BUILDING AND DELIVERING SUSTAINABILITY SOLUTIONS: INSIGHTS, METHODS, AND CASE-STUDIES

EDITED BY: Nathaniel K. Newlands, Tracy A. Porcelli, Andries B. Potgieter, Louis Kouadio, Alfredo Huete and Wei Guo

PUBLISHED IN: Frontiers in Environmental Science and Frontiers in Plant Science

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88963-039-4
DOI 10.3389/978-2-88963-039-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

BUILDING AND DELIVERING SUSTAINABILITY SOLUTIONS: INSIGHTS, METHODS, AND CASE-STUDIES

Topic Editors:
Nathaniel K. Newlands, Agriculture and Agri-Food Canada, Canada
Tracy A. Porcelli, P. Phys., Independent Researcher, Canada
Andries B. Potgieter, University of Queensland, Australia
Louis Kouadio, University of Southern Queensland, Australia
Alfredo Huete, University of Technology Sydney, Australia
Wei Guo, University of Tokyo, Japan

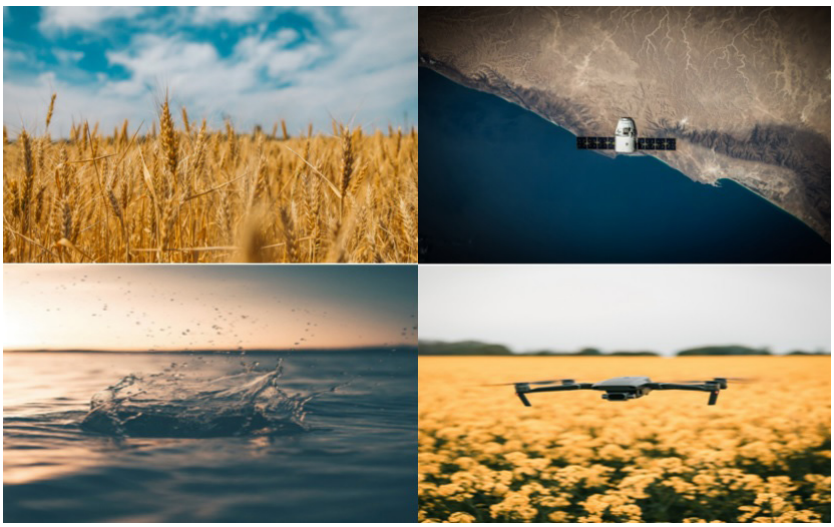


Image: Polina Rytova (TL), SpaceX (TR), Alex Perez (BL), Andreas Dress (BR). Unsplash, licenced under CC-BY (https://unsplash.com/)

Sustaining ecosystems to deliver what people need and value, while mitigating and adapting to global climate change and extreme event impacts, presents a complex set of environmental, economic, and social challenges in ensuring resilient and sustainable food production. The Climate Smart Landscape (CSL) approach has emerged as an integrated management strategy to address the increasing pressures on agricultural production, ecosystem conservation, rural livelihoods, climate change mitigation and adaptation. Deploying cheaper, more accurate, and efficient technology enables the harnessing of big data for use in solving sustainability challenges. With improved integrated analytical frameworks, statistical approaches, spatially-explicit models and indices, the CSL approach can be further developed and applied for more resilient, productive, and sustainable ecosystems. This eBook brings together original research, review, hypothesis, theory, and technology report articles, involving 87 authors from 9 countries across Asia, Europe, and North America. These articles present new methodological and technological innovation, findings, and

insights across four themes: (1) landscape productivity and crop suitability, (2) variable crop requirements for water and nutrients, (3) crop health status, phenology, and phenotyping, and (4) crop disease assessment and prediction under integrated pest management (IPM).

Citation: Newlands, N. K., Porcelli, T. A., Potgieter, A. B., Kouadio, L., Huete, A., Guo, W., eds. (2019). Building and Delivering Sustainability Solutions: Insights, Methods, and Case-Studies. doi: 10.3389/978-2-88963-039-4

Table of Contents

frontiers in Environmental Science

EDITORIAL
published: 28 May 2019
doi: 10.3389/fenvs.2019.00071

Editorial: Building and Delivering Real-World, Integrated Sustainability Solutions: Insights, Methods and Case-Study Applications

Nathaniel K. Newlands[1]*, *Tracy A. Porcelli*[2], *Andries B. Potgieter*[3], *Louis Kouadio*[4], *Alfredo Huete*[5] *and Wei Guo*[6]

[1] Agriculture and Agri-Food Canada, Science and Technology Branch, Summerland Research and Development Centre, Summerland, BC, Canada, [2] Newelli Consulting, Kelowna, BC, Canada, [3] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Toowoomba, QLD, Australia, [4] International Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, QLD, Australia, [5] School of Life Science, University of Technology Sydney, Ultimo, NSW, Australia, [6] Institute for Sustainable Agro-Ecosystem Services, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Nishitokyo, Japan

Keywords: adaptation, agriculture, big data, real-world, sustainability

Editorial on the Research Topic

Building and Delivering Real-World, Integrated Sustainability Solutions: Insights, Methods and Case-Study Applications

Sustaining ecosystems to deliver what people need and value, while mitigating and adapting to global climate change and extreme event impacts, presents a complex set of environmental, economic, and social challenges in ensuring resilient and sustainable food production (Fuhrer, 2007; Newlands, 2016; Whitfield et al., 2018). Regional climate trends drive variability in land productivity, soil water and nutrient availability, crop calendars, and the prevalence of pests and pathogens. Abrupt extreme events that defy clear prediction and attribution to climate trends cause catastrophic crop damage through water-logging, soil erosion, nutrient leaching, heat waves and drought. Sayer et al. (2013) proposed 10 core principles for an integrated landscape approach for reconciling agriculture, conservation, and other competing land uses. Building on these principles, the Climate Smart Landscape (CSL) approach has emerged as an integrated management strategy to address the increasing pressures on agricultural production, ecosystem conservation, rural livelihoods, and climate change mitigation/adaptation (Scheer et al., 2012; Salvini et al., 2018). The CSL approach is strengthened by a broad array of different science-based indicators, metrics, frameworks and modeling systems that enhance its capability to make more informed, integrated decisions. This approach, however, needs to incorporate big data, statistical, and artificial intelligence (AI) methodological improvements, new sensors, and remote-sensing technological advancements (Lee et al., 2010; Wolfert et al., 2017; Willcock et al., 2018). Such improvements offer more flexible monitoring, newer and/or higher-resolution data, better prediction methods, and tools for decision support.

This Research Topic aims to showcase research, development and technology (RDT) work toward devising and delivering integrated solutions that support and enhance the CSL-based approach. This Research Topic comprises 13 articles, including 10 Original Research articles, 1 Review, 1 Hypothesis and Theory article, and 1 Technology Report. State-of-the-art modeling approaches and sampling technologies are showcased. Contributed papers present new methodological/technological innovation, findings, and/or insights across four themes: (1) landscape productivity and crop suitability, (2) variable crop requirements for water and nutrients,

OPEN ACCESS

Edited and reviewed by:
Hayley Jane Fowler, Newcastle University, United Kingdom

***Correspondence:**
Nathaniel K. Newlands
nathaniel.newlands@canada.ca

Specialty section:
This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Environmental Science

Received: *12 April 2019*
Accepted: *10 May 2019*
Published: *28 May 2019*

Citation:
Newlands NK, Porcelli TA, Potgieter AB, Kouadio L, Huete A and Guo W (2019) Editorial: Building and Delivering Real-World, Integrated Sustainability Solutions: Insights, Methods and Case-Study Applications. Front. Environ. Sci. 7:71. doi: 10.3389/fenvs.2019.00071

(3) crop health status, phenology and phenotyping, (4) crop disease assessment and prediction under integrated pest management (IPM) and the CSL approach.

LANDSCAPE PRODUCTIVITY AND CROP SUITABILITY

Bock et al. provide a technology report on a Canadian Land Suitability Rating System (LSRS) is a rule-based algorithm integrating soil, climate, and landscape factors from accessible databases, calculating a classed suitability rating for a given landscape. It is used to support commercial field crop production, and as a spatial research tool for assessing climate change impacts. Subramanian and Crowley present a spatial-based reinforcement learning approach that uses satellite imagery to increase the predictive power of spatial dynamic models that predict, and learn better policies to manage and control spatially spreading processes. They apply this methodology in wildfire event prediction using Canadian data. Exploring landscape gradients, Xu et al. assess how measuring plant community assemblies along habitat severity gradients may improve our ability to understand and monitor community dynamics and species responses under future climate change.

VARIABLE CROP REQUIREMENTS FOR WATER AND NUTRIENTS

Neilsen et al. showcase a landscape-based water demand model for agricultural water use that regional water managers can use to better manage water demand and supply in response to climate change. This model considers high-resolution land use, soil, elevation, historical/future climate scenario data, a digital elevation model, sub-basins, aquifers, and socio-political jurisdictional boundaries. They use this model to explore future scenarios of climate change, and historical effects of crop production systems on irrigation water demand. From water to nutrient requirements, Guo et al. investigate aerially estimating nitrogen update from multi-angular hyperspectral data on winter wheat to improve the efficiency of remotely-based techniques for non-destructive, rapid detection of wheat nitrogen (N) nutrient status. A novel, modified right-side peak area index (mRPA) is benchmarked against other widely used indices, and shown to have the highest predictive power. Using this index can increase accuracy in assessing crop N status and management. Martins et al. propose a novel methodology for tracking crop micronutrient composition over time and demonstrate it for predicting maize/corn maximum requirements under variation in nutrient uptake rates, potential evapotranspiration, and micronutrient partitioning over crop growth stages.

CROP HEALTH STATUS, PHENOLOGY, AND PHENOTYPING

Watson et al. use time-lapsed "phenocam" cameras to track the phenology and identify phenological variability of native and exotic grasses across grassland areas in Australia. Their findings indicate C3/C4 species dominance to be the primary driver of phenological differences among grassland types, with the proportion of non-photosynthetic vegetation, grazing pressure, and species-dependent responses to rainfall and temperature being important biophysical drivers of grassland phenology. MODIS/Landsat satellite and field-based phenocam data were found to be in good agreement. A primary benefit of phenocam data is its higher temporal fidelity in capturing vegetation changes (i.e., increases/decreases in greenness) over periods of only 5 days, compared to coarser satellite or field measurement techniques. Using high-resolution images captured by unmanned aerial vehicles (UAVs), Guo et al. propose a two-step machine-learning based image processing method that can provide more reliable estimates of yield by detecting and counting the number of heads, than manual measurements in sorghum breeding trials, with potential for broader application in field experiments, and field production scouting operations. To better assess adaptation traits in large small-plot breeding trials, Potgieter et al. evaluate the use of a narrowband multi-spectral camera deployed on a UAV. Leaf Area Index (LAI), Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and Normalized Difference Red Edge (NDRE) were evaluated. Despite variable emergence, these indices tracked canopy cover and LAI well over a large range of plant densities, with NDVI and EVI strongly correlated with plant number per plot, canopy cover, and LAI. NDRE (i.e., leaf chlorophyll content) was found to be most useful in characterizing the leaf area dynamics and senescence patterns in contrasting genotypes. In further addressing practical constraints in genomics-assisted breeding, Watanabe et al. demonstrate UAV remote sensing (with a RGB or near-infrared green and blue (NIR-GB) camera) for measuring sorghum plant height and nitrogen availability for faster and more cost-effective throughput phenotyping. For phenotyping based on root depth distributions, Wasson et al. propose a state-of-the-art Bayesian hierarchical' nonlinear mixed statistical modeling approach to estimate root depth distributions for wheat genotypes to enable breeders to select for whole root system distributions appropriate for sustainable intensification. This approach produces de-noised profiles that exhibit rigorously discernible phenotypic traits.

CROP DISEASE ASSESSMENT, PREDICTION UNDER IPM AND CSL APPROACH

Pandey et al. provide a review of soil-borne and foliar fungal diseases of mungbean (*Vigna radiata var. radiata*), an important legume crop in South/Southeast Asia. They review pathogen characterization, economic impacts, and integrated management practices including host resistance, fungicides, biocontrol agents, natural plant products, and cultural practices. They highlight the need for longer-term studies to validate biological methods for commercial application. For wheat (yellow) stripe rust (*Puccinia striiformis f.sp. tritici*) fungal disease, the greatest global pathogen threat to wheat

production worldwide, Newlands explores the feasibility of an integrated model-based framework for predicting and controlling across large agricultural regions, using a novel spatially-explicit complex model, climate reanalysis and weather station network data.

Deploying cheaper, more accurate, and efficient technology enables the harnessing big data for use in solving sustainability challenges. With improved integrated analytical frameworks, statistical approaches, spatially-explicit models and indices, the CSL approach can be further developed and applied for more resilient, productive, and sustainable ecosystems. Smarter models will however require sufficient training data and operational frameworks assimilating data across a broad range of sampling platforms and data types. Agri-business will likely play an increasing role in sustainable landscape management by involving stakeholders and monitoring progress and outcomes (Salvini et al., 2018). Collectively, we can move faster to confront such complex interplay involved in translating scientific evidence into real-world operational or actionable solutions (Lamontagne et al., 2019).

AUTHOR CONTRIBUTIONS

NN and TP co-wrote this editorial based on contributions to the Research Topic, incorporating editorial feedback provided by AP, LK, AH, and WG.

ACKNOWLEDGMENTS

We thank all authors of contributions to this Research Topic for preparing high-quality manuscripts under a tight time schedule and the reviewers for critically evaluating the papers and for further improving them with their questions, comments, and suggestions. NKN acknowledges research funding support of the Canadian Agricultural Partnership (CAP), Agriculture and Agri-Food Canada (AAFC).

REFERENCES

Fuhrer, J. (2007). "Chapter 7: Sustainability of crop production systems under climate change," in *Agroecosystems in a Changing Climate*, eds P. C. D. Newton, R. A. Carran, G. R. Edwards, and P. A. Niklaus (Boca Raton, FL: Taylor and Francis; Chapman and Hall; CRC Advances in Agroecology Series), 167–185.

Lamontagne, J. R., Reed, P. M., Marangoni, G., and Garner, G. G. (2019). Robust abatement pathways to tolerable climate futures require immediate global action. *Nat. Clim. Chan.* 9, 290–294. doi: 10.1038/s41558-019-0426-8

Lee,W. S., Alchanatis,V.,Yang,C., Hirafuji, M., Moshou, D., and Li, C. (2010). Sensing technologies for precision specialty crop production. *Comput. Electron. Agric.* 74, 2–33. doi: 10.1016/j.compag.2010.08.005

Newlands, N. K. (2016). *Future Sustainable Ecosystems: Complexity, Risk, Uncertainty.* Boca Raton, FL: Taylor and Francis (Chapman and Hall/CRC Applied Environmental Statistics Series).

Salvini, G., Dentoni, D., Ligtenberg, A., Herold, M., and Bregt, A. K. (2018). Roles and drivers of agribusiness shaping Climate-Smart landscapes: a review. *Sustain. Dev*. 26, 533–543. doi: 10.1002/sd.1897

Sayer, J., Sunderland, T., Ghazoul, J., Pfund, J. L., Sheil, D., Meijaard, E., et al. (2013). Ten principles for a landscape approach to reconciling agriculture, conservation, and other competing land uses. *Proc. Nat. Acad. Sci. U S A*. 110, 8349–8356. doi: 10.1073/pnas.1210595110

Scheer, S. J., Shames, S., and Friedman, R. (2012). From climate-smart agriculture to climate-smart landscapes. *Agric. Food Secur.* 1:12. doi: 10.1186/2048-7010-1-12

Whitfield, S., Challinor, A. J., and Rees, R. M. (2018). Frontiers in climate smart food systems: outlining the research space. *Front. Sustain. Food Syst*. 2:2. doi: 10.3389/fsufs.2018.00002

Willcock, S., Martinez-López, J., Hooftman, D. A. P., Bagstad, K. J., Balbi, S., Marzo, A., et al. (2018). Machine learning for ecosystem services. *Ecosyst. Serv*. 33, 164–174. doi: 10.1016/j.ecoser.2018.04.004

Wolfert, S., Ge, L., Verdouw, C., and Bogaardt, M.-J. (2017). Big data in smart farming – a review. *Agric. Syst*. 153, 69–80. doi: 10.1016/j.agsy.2017.01.023

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Environmental Science

TECHNOLOGY REPORT
published: 19 July 2018
doi: 10.3389/fenvs.2018.00077

The Land Suitability Rating System Is a Spatial Planning Tool to Assess Crop Suitability in Canada

*Michael Bock[1], Pierre-Yves Gasser[2], Wayne W. Pettapiece[1], Anthony J. Brierley[1], Andy Bootsma[2], Peter Schut[2], Denise Neilsen[3] and C. A. Scott Smith[3]**

[1] Science and Technology Branch, Agriculture and Agri-Food Canada, Edmonton, AB, Canada, [2] Science and Technology Branch, Agricultural and Agri-Food Canada, Ottawa, ON, Canada, [3] Science and Technology Branch, Agriculture and Agri-Food Canada, Summerland, BC, Canada

The Land Suitability Rating System (LSRS) is a rule-based set of algorithms that integrate soil, climate and landscape factors to calculate a classed suitability rating for a given landscape to support commercial field crop production. The attributes used to define each of the factors are based on their proven ability to affect crop growth, their ability to be measured (or estimated by proxy) and their availability in accessible databases. The LSRS was first published in 1995 by Agriculture and Agri-Food Canada as a site-specific, manual calculator for spring-seeded small grains that incorporated sets of attribute point deduction curves based on expert knowledge. Since that time the system has been expanded to include additional crop modules and all data handling and calculations are automated through a set of web-based applications. The current version of LSRS (version 5) is implemented in Ruby on Rails® software as a suite of web services. The system runs against any soil map with standardized Canadian Soil Information Service soil data tables to process soil attributes and calculate limitations to crop growth. A climate factor rating is based on crop-specific agro-climatic indices and thresholds. Climatic indices have historically been calculated from 30-year climate normal periods using monthly data but LSRS can now also utilize daily data records which facilitate trend analyses within annual historic records. The use of available gridded climate datasets enables direct overlay and extraction of climate attributes to the spatial extent of soil map polygons. Lastly, the system incorporates a landscape factor related to land erodibility and constraints to management. Each of the three suitability factors is assigned a class rating between 1 (no limitations) and 7 (unsuitable) with the final overall rating being the most limiting of the three factors. Recent improvements in the ability of the system to process multiple climate datasets mean outputs from Global Circulation Models may also be useful for the LSRS model in assessing possible impacts of climate change on crop suitability. LSRS is used increasingly as a spatial research tool in assessing potential changes in crop distributions at both national and regional scales.

Keywords: land suitability, climate change, soil maps, agro-climatic indices, Global Circulation Model, agricultural capability

OPEN ACCESS

Edited by:
Tracy Ann Porcelli,
Independent researcher, Canada

Reviewed by:
Jana Mrazova,
Deakin University, Australia
Sean Smukler,
University of British Columbia, Canada

****Correspondence:***
C. A. Scott Smith
scott.smith051@gmail.com

Specialty section:
This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Environmental Science

Received: *09 March 2018*
Accepted: *26 June 2018*
Published: *19 July 2018*

Citation:
Bock M, Gasser P-Y, Pettapiece WW, Brierley AJ, Bootsma A, Schut P, Neilsen D and Smith CAS (2018) The Land Suitability Rating System Is a Spatial Planning Tool to Assess Crop Suitability in Canada. Front. Environ. Sci. 6:77. doi: 10.3389/fenvs.2018.00077

INTRODUCTION

The identification of soil landscapes suitable for production of food no doubt began with the dawn of arable agriculture. "Pedology and soil science in general have their rudimentary beginnings in attempts to group or classify soils on the basis of productivity. Early agrarian civilizations must have had some way to communicate differences and similarities among soils." Ahrens et al. (2002). In the early 1900's, German agronomist A. D. Thaer proposed a scale to describe the intrinsic fertility of soil based on attributes such as soil texture, calcium carbonate content and organic matter content (Feller et al., 2003). These early assessments were very subjective and qualitative based on general observations of features such as topography, stoniness, wetness and tree cover.

As agriculture became more mechanized and intensified in the twentieth century, there was a concomitant increase in scientific research into the agronomic requirements of cultivated crops. The evaluation of production potential became much more specific and quantitative with assessments such as the Storie Index Rating (Storie, 1933). This mathematical treatment of individual parameters such as texture, organic matter and pH was well accepted by the technical community of soil specialists. Ratings based on his procedure were incorporated into numerous soil reports produced as part of the National Cooperative Soil Survey Program in the United States (see, for example, Goodman, 1955; Arroues and Anderson, 1986). In the province of Alberta, Canada, soil survey reports from the late 1930s to the late 1950s included an interpretive land class map derived primarily from the physical characteristics of the study area. When compiling the map, soil physical data (texture, surface color, pH, soil taxon, mode of deposition, and degree of stoniness), landscape data (topographic form and relief) and climate data (rainfall and rainfall variability) were all taken into consideration (Wyatt et al., 1939). Each of these factors was assigned a numeric value for each soil area, and the multiplying together of these values gave the final index rating of the soil area. Using available pasture carrying capacity and wheat yield data, a seven-class productivity grouping of these rated areas was created. Some of the later soil survey reports produced through the end of the 1960s incorporated a stocking rate and/or wheat yield range for each group as an estimate of productivity (see, for example, Bowser et al., 1951; Odynsky et al., 1952).

By the middle of the century with greater intensification and further expansion into the less suitable fringe lands, there was recognition that some of the agricultural practices associated with the push for higher productivity were negatively affecting environmental sustainability (Standing Committee on Agriculture, Fisheries and Forestry, 1984). In Canada, this discussion had started with the wind erosion concerns in the 1920s and 30s. There was increasing competition with other land uses such as forestry, wildlife habitat, and recreation all associated with a growing population. It was into this setting that the broader concept of land capability, which included an aspect of sustainability, was introduced both in the United States (Klingebiel and Montgomery, 1961) and Canada (Agriculture and Rural Development Act, 1965).

The Canada Land Inventory (CLI) was introduced in Canada in the mid-1960s under the program of the federal Agriculture and Rural Development Act. The system used a general comparative capability approach based on severity of limitations for broadly defined land uses, specifically, agriculture, forestry, wildlife, and recreation. In the words of the authors of the first CLI report, the new system "is designed primarily for planning rather than for management. It is of a reconnaissance type, it provides information essential for land development planning at the municipal, provincial and federal levels of government. It does not provide the detailed information required for management of individual parcels of land, nor for land planning in small watersheds, local government units, etc." (Agriculture and Rural Development Act, 1965) The capability ratings were to be presented on maps at a scale of 1:250,000. At the time of its development, the then relatively new profession of land use planning was becoming important and quickly embraced this new rating system as capability was an intuitive and easily understandable concept. Municipal planners and realtors particularly liked the approach. The classification system incorporated seven classes with Class 1 being the best with no limitation for the intended use, Class 4 being marginal for the use and Class 7 being completely unsuitable. Under the umbrella of a cooperative federal-provincial program, approximately 2.5 million hectares covering all regions of Canada with multiple land use issues associated with agriculture were mapped from 1965 to 1980 using the CLI (Canada Land Inventory, 1998).

The CLI system worked very well for its intended regional development objectives; however, these regional assessments, while based on specific soil and land information still required a significant amount of subjective interpretation, extrapolation and amalgamation and climate was not considered in the ratings. In addition, the capability classes were categorical and while well-suited to land use planning functions they were less useful for more detailed analysis such as assessing the productivity function of soils (Mueller et al., 2010). Many land evaluators and managers required a much more detailed assessment at scales of 1:50,000 or larger (with units as small as several hectares), but in many areas there simply was not a more detailed soil database available. The CLI, however, was not intended for use at detailed or site-specific scales (Alberta Energy and Natural Resources, 1983). Attempts to extend the CLI approach to address the more detailed requirements provided variable results. Over the years, agencies in a number of jurisdictions either modified the original CLI system or developed entirely new systems of land capability classification for a variety of purposes. In other instances, soils specialists with different agencies within a province simply used two parallel systems following the CLI and Storie-type approaches (Alberta Soils Advisory Committee, 1983). If one person or one closely correlated group was providing the interpretations there was at least consistency within a region; however, as more people and agencies became involved, consistent evaluations became a problem (Alberta Soils Advisory Committee, 1987).

The overall result was that by the mid-1980s there were many agricultural rating systems being used across Canada. These included the systems used in the Atlantic Region (Atlantic

Advisory Committee on Soil Survey, 1988), Quebec (Mailloux et al., 1984), the Ottawa area (Marshall et al., 1979), Ontario (Brooke and Presant, 1986), Alberta (Alberta Soils Advisory Committee, 1987), and British Columbia (Kenk and Cotic, 1983). With different parameters and thresholds, the same crops could have different absolute ratings under the different systems. The original CLI was also not consistent across Canada as the ratings were based on regionally important crops that differed from one region to another. For example, an area rated as CLI Class 1 for agriculture in Saskatchewan would be based on wheat and other small grains while an area rated as CLI Class 1 for agriculture in Ontario would have to grow corn.

The use of multiple capability classification systems as well as the inconsistent manner in which the CLI was being applied nationally resulted in confusion and disagreement. The topic was raised at the meeting of the Canadian Expert Committee on Soil Survey in 1986. Agriculture and Agri-Food Canada's then Land Resource Research Centre responded by creating an Agronomic Interpretations Working Group (AIWG) in 1987 to pursue the development of a national rating system for Canada. It was asked to address several specific concerns related to the CLI-Soil Capability for Agriculture, namely:

- The proliferation of modifications to the CLI by provincial departments that had resulted in a variety of inconsistent methods used in the classification of land capability for agriculture across Canada;
- The inadequate consideration of the role of climate on land suitability for crop production;
- The omission of organic soils in the CLI, and;
- The lack of specificity in definitions and applications which lead to inconsistent ratings by those applying the system.

The result of the work was the publication in 1995 of the technical bulletin "Land Suitability Rating System for Agricultural Crops: 1. Spring-seeded small grains" (Agronomic Interpretations Working Group, 1995). The spring-seeded small grains report presented the details for the first module of LSRS. While it was anticipated by the original working group that additional crop modules would quickly follow, it would be a full decade before these would be developed.

It is the purpose of this paper then to present the story of the evolution of Canada's Land Suitability Rating System (LSRS) from its beginnings in the 1990's up to 2017. Specifically, the paper outlines the rationale associated with the initial development of LSRS, describes the model's structure and methods of implementation, describes the chronology and nature of upgrades (new versions, new modules) to the system, and presents recent examples of its application at national and regional scales.

SYSTEM BACKGROUND AND DEVELOPMENT

In describing the approach and methods used in the development of LSRS, the differences and usages in the terms land evaluation, land capability and land suitability warrant some discussion. Capability assesses the nature of limitations and degree of limitations imposed on cultivated agriculture by the physical characteristics of a land unit. In the New Zealand Land Use Capability classification system capability of land is defined as "its suitability for productive use or uses after considering the physical limitations of the land" (Lynn et al., 2009). The land capability classification for agriculture in the UK developed by the Macauley Land Research Institute described capability as "the agricultural potential of land based on the degree of limitation imposed by its biophysical properties" (Wright et al., 2006). Land capability classification involves systematically categorizing units of the landscape in a way that reflects the inherent ability to produce sustainably into the future (Lynn et al., 2009). In Canada, the Canada Land Inventory (CLI) series of capability maps have been an important reference for assessing present and potential land use activities and planning across Canada. However, in terms of capability ratings there have been issues related to specificity. The agricultural capability classification rates soils for their ability to sustain production of common crops based on their potentials and limitations. The underlying assumption is that the better the capability rating (i.e., Class 1 vs. Class 4) a wider variety of agricultural crops may be grown; however, the number of crops successfully grown varies from one region of Canada to another. For example, corn may be grown in the CLI Class 1 or 2 areas in southern Ontario, but within the Prairie region areas of equivalent capability cannot sustain equivalent corn crops.

In comparison, the term "land suitability" is an estimate of the fitness of a soil and its landscape for production of a specific agricultural crop (FAO, 1976). As plants have specific requirements related to the functional status of soil, classifications based on production limitations and crop productivity must have a certain stratification or orientation specific to a single crop or groups of related crops (Mueller et al., 2010). Early federal German systems used soil suitability classification terms like "prime wheat soil," "rye soil," or "oats soil" to describe the suitability of specific areas (Feller et al., 2003). More recently, the Muencheberg Soil Quality Rating is a soil suitability classification developed as a tool to assess the capacity of land across all scales to produce small grains (Mueller et al., 2007). Similarly, soil suitability rating systems have been developed for a variety of specific land management practices such as reduced tillage or direct drilling (Cannell et al., 1978), irrigation (Alberta Agriculture Food and Rural Development, 2000), the spreading of animal manure (Eilers and Buckley, 2002) and carcass burial (Carcass Burial Site Selection Technical Committee, 2004).

Modern land evaluation systems largely grew out of the older agricultural land capability classification. However, land evaluation is a much broader concept bringing into the assessment aspects of environmental sustainability (Smyth and Dumanski, 1995) and a range of economic and market factors that can influence land use decisions (Rossiter, 1996). Land suitability is typically one component of land evaluation. Methods of land evaluation (including land suitability) have become more sophisticated over time (Sonneveld et al., 2010) and with increased availability of large geographic and production

datasets and geographic information systems, evaluations have become more quantitative and process oriented (Triantafilis et al., 2001; Joss et al., 2008; Elsheikh et al., 2013).

Initial Concepts and Assumptions

Development of LSRS was based on an "expert system" approach that made use of existing sources of data and the collective knowledge and experience of professionals from the fields of land and plant science and those familiar with the evaluation of land suitability for crop production. The Agricultural Interpretations Working Group that developed the first version of LSRS was composed of soil surveyors, agronomists and agro-climatologists from across Canada. They examined a number of systems that were being used to rate land for the production of agricultural crops, keeping in mind the need for national consistency and the other concerns raised by the Expert Committee on Soil Survey. Climatic stratifications such as those by Chapman and Brown (1966), FAO (1976), and Williams (1983) were examined.

The initial assessment recommendations were:

1. Retain framework of the seven-class CLI system. It was well established and accepted and formed the basis for land use regulations in several provinces.
2. Rate separately elements of climate, soil and landscape, each of which can independently control land suitability for crop production.
3. Organic soils must be included and rated for suitability as is done with mineral soils.
4. Use only a limited number of key parameters to optimize the rating system (Huddleston, 1984). Using an expert system approach, each should be individually defined and explicitly rated (McCracken and Cate, 1986).
5. The individual parameter ratings should follow scientifically proven relationships and be managed in a mathematical setting leading to a composite index.
6. The system should be developed using the small seeded cereals (wheat, oats and barley) with an emphasis on barley which is the one crop with widespread production in all agricultural regions of Canada.
7. The system must use data that is available across Canada.

The first recommendation provided the overall approach that should be taken while the next four addressed the major weaknesses that had been identified in the CLI approach. The final two recommendations addressed the issue of national consistency and provided further direction for system development.

It may be noted that no mention was made of scale. The reason was that the system was meant to be scale-independent–that it should have the ability to be used at any scale appropriate to the objectives of the project and available data (see section Applications of the Rating System for description of how this applies to the current use of LSRS). Preliminary discussion identified the assumptions or guidelines that were required to provide boundary conditions for the system:

- External economic factors such as distance to market, availability of transportation and size of farm would not be criteria. These are important for municipal property assessments and taxation but would not be part of this natural resource evaluation;
- One-time costs such as clearing of trees or drainage would not be criteria but continuing annual costs. Stone removal and erosion control should be considered in the sustainability considerations;
- Exceptional skills or resources of the farm manager or specific cultural practices would not be considered;

Structure of the Land Suitability Rating System

Classes, Factors and Parameters

The basic structure of the LSRS rating output uses two hierarchical categories–classes and subclasses. Each factor is assigned a class rating. Classes are broad in scope and are based on the degree of limitation of land for production of the specified crop. Seven classes are recognized (**Table 1**). Areas assigned to the same suitability class are similar only with respect to the degree, and not the kind of limitation. Different management may be required on lands of similar class which can be composed of different soil, climate and landscape.

In LSRS, the major rating factors are related to three of the major elements that describe crop production suitability:

- Climate-controls the type and range of crops that can be grown (flexibility of production);
- Soil-controls how well the crops grow (productivity), and;
- Landscape-controls the cost to manage environmental constraints (sustainability).

LSRS assesses the climate, soil and landscape factors independently with a precisely defined index procedure that links the results to the seven-class system. This methodology required the development of clear relationships and guidelines for the assessment of the factors the lead to the assignment of class rating.

LSRS is an interpretive assessment based on the limitations controlling crop specific production. The rating factor that is most limiting ultimately determines the suitability class

TABLE 1 | The characteristics and index ratings for LSRS classes.

Suitability class	Limitation level for specified crop	General assessment	Index point rating	Comments
1	None to slight	Excellent	100–80	Prime land
2	Slight	Good	79–60	
3	Moderate	Fair	59–45	
4	Severe	Poor	44–30	Marginal land
5	Very severe	Very poor	29–20	
6	Extreme	Unsuitable	19–10	Unsustainable or unsuitable land
7	Non-arable	Unsuitable	9–0	

rating. The components and measurable parameters identified to evaluate the major rating factors were selected from those used in earlier approaches (Huddleston, 1984) and acknowledged to be critical in crop production (**Table 2**). The following criteria were used in selecting the specific parameters used to characterize the major rating factors used in the LSRS model:

- parameter is known to affect the ability to produce crops;
- parameter is known to affect the ability to respond to abiotic stress (e.g., ability to withstand drought);
- parameter must be measurable or able to be estimated from known relationships (i.e., pedo-transfer functions), and;
- parameter data must be commonly available.

Although it is recognized that there are many instances of overlap and synergy between the major factors, there are advantages with respect to simplicity, clarity and the ability to highlight specific limitations by assessing each factor separately. Having built the major climatic-soil interactions into the soil factor, any other remaining discrepancies are considered small. This decision was reasonable given the constraint of available data at that time and the need for national coverage. This approach provided the greatest flexibility to assess various environmental, crop and climatic scenarios.

Each of the major factors is assessed a rating index between 0 (most limiting) and 100 (least limiting). Initially, each factor is assigned an index rating of 100. Limitations are assessed using the specific parameters identified for each factor and point values are deducted from the initial index ratings. The final index rating (i.e., class) assigned is that of the most limiting of the three factors. Subclasses identify the parameters (**Table 2**) that have the greatest limiting influence on the final class rating. They reflect the kind of climate, soil and landscape limitations that are present. The subclass information is critical for determining conservation and management practices and for land use planning. While Classes are numbered, Subclass designations used in LSRS are letters as follows: climate—temperature/aridity (H) and moisture (A); mineral soil–water supplying ability (M), structure and consistence (D), organic matter content (F), depth of topsoil (E), reaction (V), salinity (N), sodicity (Y), organic (peaty) surface (O), and drainage (W); organic soil—soil temperature (Z), water supplying ability (M), degree of decomposition (B), reaction and nutrient status (V), salinity (N), and drainage (W), and; landscape—basic landform slope (T), stoniness (P), wood content (J), landscape pattern (K), and flooding (I).

A few examples of complete LSRS ratings would therefore be: 2A–a landscape with only slight climatic moisture limitation, or, 4N–a marginal landscape limited by severe soil salinity, or, 6T–a landscape rendered unsuited for crop production due to steep slopes. A class 1 rating (landscape without limitation) would have no subclass denoted.

TABLE 2 | The factors, components and parameters used by LSRS to assess land suitability.

Factor	Component	Measurable parameter[a]
Climate	Heat (energy) supply	Growing degree days, growing season length
	Moisture supply	Precipitation, evapotranspiration
Mineral Soils	Moisture supply	Texture, rooting depth, water table
	Nutrient supply	Organic matter content, soil reaction
	Physical conditions	Soil structure, soil bulk density
	Chemical conditions	Soil salinity, soil reaction
	Drainage	Depth to water table, drainage class
Organic Soils	Moisture supply	Fibre content, water table
	Nutrient supply	Fibre type, soil reaction
	Physical conditions	Soil structure, soil density
	Chemical conditions	Soil reaction, soil salinity
	Drainage	Depth to water table, climate
Landscape	Erodibility potential	Slope steepness, slope length, climate
	Management factors	Stoniness, drainage, pattern
	Flooding potential	Wetness, duration of flooding, landform position

[a] *see Agronomic Interpretations Working Group (1995) for definitions.*

Linkage of Deductions to the Suitability Class

As productivity is an important consideration in suitability assessments, studies of the relationships of Canada Land Inventory (CLI) classes to the yields of cereals in Alberta (Peters, 1977; Peters and Pettapiece, 1981) and apples in Ontario (van Vliet et al., 1979) provided the initial guidance in developing an assessment of the limitations-based linkage between the prescriptive numerical index ratings (0–100) and the descriptive suitability class. These studies reported a reasonably good correlation for the better classes with Class 1 (none to slight limitation) generally yielding 80% to 100% of the crop maximum. Class 3 areas (moderate limitation but still considered "Good") generally had yields about 50% of maximum or better. It was also noted that with increasing limitations-particularly landscape limitations, the yield relationships disappeared. That is, landscape features may be difficult or costly to manage but are not directly related to yield. Additional expert opinion suggested that index rating of less than approximately 1/3 of maximum (33 out of 100 points) should be considered a very severe limitation to the long-term sustainability of production.

Using these considerations as a guide, a relationship framework was formulated that provides the conceptual and mathematical linkage between the factor index ratings and LSRS classes. The calculation of LSRS class ratings required the development of the individual parameter indices. Examples of point deduction values for selected climate, soil and landscape parameters as they relate to spring seeded small grains (wheat, oats, barley) are given in **Figure 1**. The point deduction values were selected through expert opinion designed around class breaks. Originally depicted in tabular format for manual calculation, these "parameter to deduction point" relations were fit to curves to allow automated processing. Although each factor is calculated separately, LSRS is ultimately a most limiting

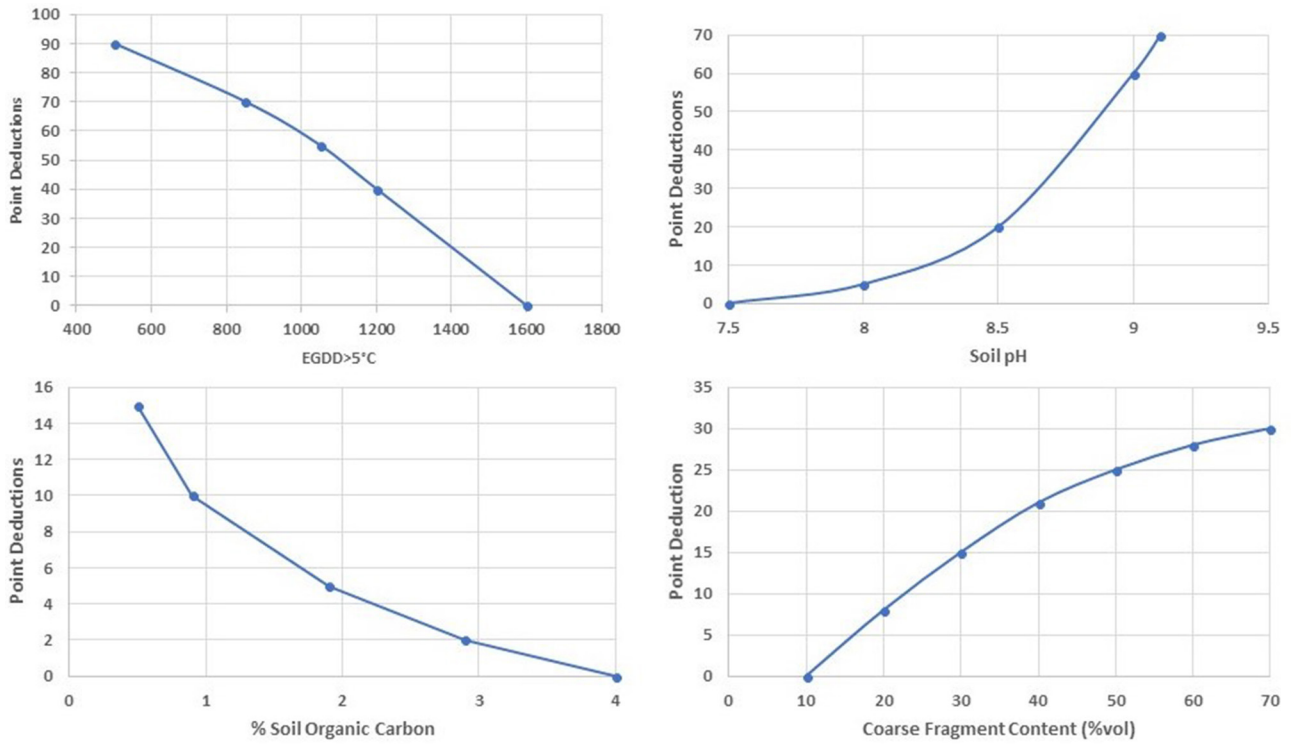


FIGURE 1 | Selected climate and soil parameter deduction curves for spring seeded small grains from Agronomic Interpretations Working Group (1995). Different parameters are set to differing deduction point ranges. Effective growing degree days (EGDD > 5°C) are a key determinant of climate suitability. A value of <900 would result in at least a 70-point deduction, enough to generate an unsuited class rating for grain production. Similarly, a soil pH > 9.0 alone would generate an unsuitable soil rating but sub-optimal soil organic carbon or coarse fragment content deductions typically result in suitability class reductions.

function tool. The final class rating given to a parcel of land is the lowest of the climate, soil or landscape rating.

Initially LSRS calculations were made site-specifically based on field observed soil and landscape properties and climate data from the closest weather station. Early calculations were based on climate normal data (30-year averages for temperature and precipitation). Early calibration was based on the period 1951–1980, later updated to 1961–1990. To facilitate the calculation of climate suitability over large areas, national scale maps of growing degree days and climatic aridity (precipitation minus potential evaporation) were constructed specifically for use with LSRS. Details of initial parameter values and the calculation of land suitability class and subclass are given in Agronomic Interpretations Working Group (1995). Over the years, validation efforts have led to many adjustments to the way soil and climate factors are assessed and rated and details of these validation efforts are given in the Supplemental Data.

Recent System Development

Over the last 20 years the LSRS has undergone developments in all aspects of the system (**Table 3**). The principle enhancements have been the introduction of additional crop modules in version 3, the transformation of the model platform to web-based processing and output (version 4), and most recently, the ability to integrate a range of gridded historic and future scenario climate datasets and calculate climate indices internally within LSRS (version 5). The requirement of the system to handle ever larger climate and soil datasets has led to implementation of more efficient data processing, storage and retrieval technologies and a more effective architecture for developing web services. The key major development areas are described below.

Multiple Crop Modules

Crop modules for corn, soybeans, forages (alfalfa, grasses), and canola are part of the current system in addition to the spring-seeded small grains module that was part of the original system described in Agronomic Interpretations Working Group (1995). The development process was similar for each module. Agronomists with crop-specific expertise were consulted to help define climate, soil and landscape parameter values for each crop. The threshold values for class limits were tested against published production figures and calculations were programmed for user access and further testing. For each new crop module, the bulk of the work lay in defining crop-specific climate indices and thresholds (**Table 4**). For grains and forages, variations of the widely used growing degree-days above 5°C index with adjustments for higher latitude growing areas (termed "effective" growing degree-days) was used to stratify climate requirements. For soybeans and corn, an adaptation of the corn heat unit which we termed "crop" heat unit was used. Some development

TABLE 3 | Outline of the version history of the land suitability rating system.

Version	Date	System	Description
–	1995	Manual	First report of the Agronomic Interpretation Working Group of AAFC. System designed for site-specific work using manual input and calculations
1	1997	Dbase	Provided an electronic access to the national soil database structures and content enabling automated LSRS calculations. Developed as a Dbase application.
2	2002	Excel	LSRS calculations developed for Alberta soil databases using spreadsheet technology (Excel). Enabled users to view spatial data and derived ratings on an individual soil map polygon basis.
3	2006	Excel	Modification of version 2 (Alberta prototype) to link national soil database information and provide the ability to generate SLC map polygon ratings across Canada.
3.5	2007	Excel	Expansion of the original LSRS program for corn, soybeans, canola and forage crops and alternate climate and soil inputs.
4	2009	Ruby on Rails with MySQL	This major revision provided the ability to run and store batch processes as a suite of web services, calculate results for multiple crops simultaneously, use a range of soil map scales and formats and accept gridded climate scenarios transmitted from other systems. Significant adjustments are made to the original algorithms as the result of regional validation efforts across Canada. System runs internal to AAFC.
4.5	2011	Ruby on Rails with MySQL	Limited calculator functionality is made available to collaborators and the public via the Internet.
5	2017	Ruby on Rails with MySQL and Redis	Addition of new climate calculator that validates climate data, calculates climate indices, and stores the calculated indices in Redis. This version enhances data management and the ability to support raster calculations

TABLE 4 | Heat indices and thresholds used in the LSRS crop modules.

Crop	Heat		Considerations
	Index	Thresholds	
Spring-seeded small grains	EGDD > 5°C,	Class 1–3:>1,050 Class 4:900–1,050 Class 5–7:<900	Calculated from date of first 5 consecutive days of GDD accumulation to date of first killing frost
Alfalfa	GDD > 5°C	Class 1:1,890 class 2–3:1,410–1,890 class 4: 930–1,410 class 5–7: <930	Calculated as total GDD accumulation between April 1 and Oct 30. Threshold values based on number of cuts possible per season
Brome	GDD > 5°C	Class 1:1,840 class 2–3: 1,380–1,840 Class 4: 880–1,380 Class 5–7:<880	Calculated as total GDD accumulation between April 1 and Oct 30. Threshold values based on number of cuts possible per season
Corn and soybeans	Crop Heat Units	Class 1: >3,500 Class 2: 2,700–3,500 Class 3–4:2,000–2,700 Class 5–7:<2,000	Crop heat units is a generic term as it applies to multiple crops but is based on the corn heat unit.
Canola	EGDD >5°C	Class 1–3:>1,050 Class 4:900–1,050 Class 5–7:<900	Same threshold values as for small grains. Added modification for number of days with temperatures > 30°C

concepts and improvisations around each of the crop modules follows. Full documentation of crop module development and testing are available as Supplemental Data.

Corn and soybeans

For both corn and soybeans, the heat index is based on crop heat units (CHU). The average daily values of crop heat units are calculated using the following formula:

$$\text{Daily CHU} = (\text{Ymax} + \text{Ymin})/2.0$$

Where Ymax and Ymin are the contributions to CHU from average daily maximum (Tmax) and minimum (Tmin) air temperatures respectively and are calculated as

$$\text{Ymax} = 3.33(\text{Tmax} - 10.0) - 0.084(Tmax - 10.0); \quad if\ Tmax < 10.0, Ymax = 0.0;$$
$$\text{Ymin} = 1.8(Tmin - 4.44); \quad if\ Tmin < 4.44, Ymin = 0.0$$

The growing season length and hence the starting and stopping dates for accumulation of the CHU values are based on temperature thresholds. The starting date is based on the average mean daily temperature (Tmean). The stopping date is based on the average mean daily minimum temperature (Tmin). Both thresholds are calibrated to the average historic dates of planting in spring and the probability of killing frost (−2°C) in the fall (Bootsma, 1991; Bootsma et al., 1999, 2005) and are presented in **Table 5**. The seasonally accumulated CHU index determined in this manner is called "CHUnorm." The original method for calculating CHUnorm used climate normal records, currently we are making the calculation from 30 years of daily climate data.

With respect to the climatic moisture assessment, the consensus was that precipitation minus potential evapotranspiration (P-PE) was a reasonable and well accepted index. Given the higher moisture demand to produce corn than cereal grains, the class limits for corn were developed by adjusting the values defined for small grains by about 100 mm.

The corn module rates soil and landscape factors and their point deductions in a similar way to that in the small grains module.

The soybean module development considerations were similar to those for corn. Both are considered warm season crops so heat requirements can be assessed within a CHU framework. Some soybean cultivars can withstand temperature limitations of relatively short, cool growing seasons but temperature requirements for soybeans are nonetheless very similar to those for corn (OMAFRA, 2002). The minimum heat requirement for the commercial production of soybeans is about 2,000 CHU.

Given the above, the agronomic consensus was that the CHU scale be used to rate the heat requirements of the crop with the same rankings as used for corn, and that P-PE should be used to characterize the moisture limitations but with the point deduction set slightly less than that for corn (**Figure 2**).

Most soil factors for soybean production are rated similarly as those for corn and the spring-seeded small grains. The only exception being the susceptibility to emergence problems relating to crusting. The soil factor therefore has an added deduction related to surface soil structure calculated as a function of the content of soil organic carbon and percentages of sand, silt and clay. The landscape factors are rated the same for soybean as for small grains as the erosion potentials and mechanical limitations are assumed to be the same as for both crops.

TABLE 5 | Definition of start and finish of the growing season for corn in eastern Canada.

Region	Spring planting date based on	Date of 10% probability of killing frost based on
	Average mean daily air temp (°C)	Average daily minimum air temp in autumn (°C)
Newfoundland	8.8	3.7
Maritime provinces	11.0	5.8
Quebec and Ontario	12.8	6.5
Prairie provinces	11.2	5.3
British Columbia (coastal)	12.7	4.6

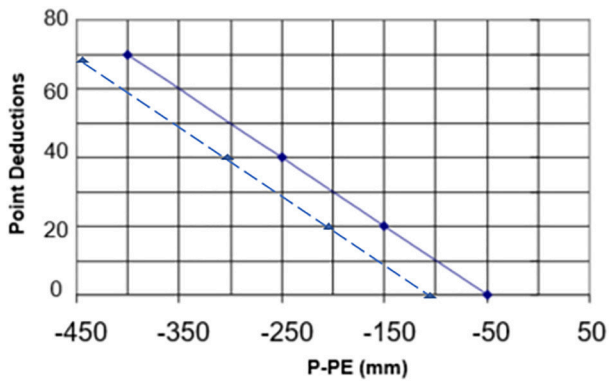


FIGURE 2 | Point deduction scheme for climatic moisture deficit for corn (solid line) and soybeans (dashed line).

Canola

In many respects, canola (*Brassica* spp.) falls within the general climatic, soils and landscape parameters used to assess small grain limitations. However, there were some refinements made in consideration of requirements for particular *Brassica* species.

Effective growing degree days (i.e., GDD base 5°C adjusted for latitudes up to 60°N) is used as the temperature index using the same scale used for small grains. The long season, higher yielding (*Brassica napus*) varieties perform best in the areas with more than about 110 frost free days which correlates with accumulations of greater than about 1,200 effective growing degree days (EGDDs). *Brassica rapa* is a shorter season variety that performs well down to about 1000 EGDDs. The baseline temperature for Brassica is approximately 4.4°C (Morrison et al., 1989). This corresponds well with the 5°C base temperature used to calculate EGDDs. However, included was a special consideration with respect to growing season heat. Flower abortion associated with temperatures greater than about 30°C is a known phenomenon (Morrison and Stewart, 2002). To address this, a heat index (HI) was calculated as the number of days where daily maximum temperature >30°C during the canola flowering period defined as between the dates of accumulation of 600 and 1100 EGDDs.

As with soybeans, the module takes into consideration the potential for soil crusting and additional deductions for soils with this potential. Otherwise, all other soil and landscape factors were rated the same for canola as for small grains.

Forages

Rating land for forage production has a fundamental difference from ratings for annual grain and oilseed crops. Typically forage crops are harvested multiple times during the growing season, often referred to as "cuts." Depending on the region of the country, forages may be harvested up to four times in a single season.

Two general types of forages (i.e., legumes and grasses) with somewhat different climatic and soil requirements are recognized within LSRS. A widely grown legume in Canada is alfalfa (*Medicago sativa* L.) and was chosen as the perennial legume to include in the system. As a perennial crop, there are no concerns with determining dates for spring seeding or for frost damage as with grain. As such, the forages have a longer growing season than the annual crops in other LSRS modules. This extended season is assumed to be the period with mean daily temperatures above 5°C (Bootsma and Boisvert, 1991).

The climatic rating for alfalfa is based on two factors, the accumulated growing degree days >5°C (GDD) and growing season length. Bootsma and Boisvert (1991) estimated about 480 GDDs to produce a first cut of alfalfa with an additional

450 GDDs for crop carryover requirements. Note that for the grass forages, LSRS uses a straight GDD >5 index that is not adjusted for high latitude day length as is the case with the effective growing degree day (EGGD) index used for small grains. Therefore, the minimum heat requirement for alfalfa is about 930 GDDs (480 GDD to obtain a harvest, 430 GDD to ensure carryover for the following year) and was used to establish the class 4-5 configuration point (**Table 6**). The climatic requirement to produce three cuts per year was used to define the Class 1 configuration point. This translates into (3 x 480) + 450 = 1,890 GDDs. The ability to support two cuts per year, [(2 × 480) + 450 = 1,410 GDDs] was set as the Class 3 threshold. The ability to produce one cut with no carry over (480 GDDs) was set as representing the Class 5-6 boundary. The maximum deduction was set at 90 points.

Length of season requirements were established using monthly climatic data and the GDD requirements for one, two and three cuts of alfalfa. Bootsma and Boisvert (1991) suggested a minimum of 45 days between cuts, 65 days to accumulate the heat to necessary to produce a first cut crop and 55 days carry-over for the crop to store enough energy to survive the winter. The relationship between growing season length and accumulated heat varies across the country. For instance, in coastal British Columbia there is a very long growing season but often insufficient heat to produce more than 2 cuts of hay.

Several soil related parameters required specific modifications for alfalfa. These include soil and subsoil pH. The most favorable soil pH range for alfalfa is between 6.5 and 7.5 (Goplen et al., 1987; Undersander et al., 1991). Therefore, a pH in this range was taken as no limitation and assigned a zero-point deduction. A pH of 5.0 is considered marginal for alfalfa and thus assigned a deduction of 70 points. The class 2 configuration point (and a deduction of 20 points) was set at pH 6.0. The class 3 configuration point (and a deduction of 40 points) was set at pH 5.5. A similar logic and set of deductions was developed for pH above 7.5.

In addition, there were modification in the module to the landscape factors of slope and stoniness. The continuous cover of perennial crops means that the risk of soil erosion is much reduced relative to annual crops. Because there is no annual cultivation requirement, internal soil coarse fragment content is not an issue from a landscape management perspective.

Climatic growth requirements and response is similar for most cool season C3 grasses (Moser et al., 1996). In humid eastern Canada, Timothy (*Phleum pretense*) is the most common forage grass. The climate of the Great Plains of western Canada is semi-arid and soils can be saline. Under these conditions smooth bromegrass (*Bromus inermis* L.) is the preferred forage grass. As with alfalfa, about 480 GDDs are required to produce a first cut of timothy with an additional 400 GDDs or so for crop carryover requirements. Carry over heat requirements for grasses are less than for alfalfa but were included for the grass forage module. The minimum heat requirement for grass forages was set at approximately 880 GDDs and was set as the boundary threshold for the class 4-5 boundary, the marginal category. Using the same approach as used for alfalfa, the ability to support two cuts per year of grass forage require 1360 GDDs and for three cuts 1840 GDDs. Point deductions for grass forage follow a similar logic as was used for alfalfa.

Also important in rating land for forage grass production is the growing season length. A literature review indicated that about 65 days of temperatures averaging >5 C are required to mature the crop for cutting and another 55 days are required for carryover into the following year. For two cuts to be achieved, it would require 65 days of initial growth, followed by 55 days of growth for the second cut and another 55 days to achieve the carryover. These counts of days were then used to determine class boundaries and deductions points. The most limiting of growing season degree-day accumulation or growing season length is used to calculate overall climatic heat suitability

Soil requirements for the brome-timothy forage model are the same as for the small grains. This is reasonable considering that these forages, like the small grains, are C3 grasses. The landscape factor deductions are the same for brome-timothy as for alfalfa.

TABLE 6 | Point deductions assigned to GDDs and growing season length for alfalfa.

GDD> 5	Growing season length (days)	Class configuration points	Point deduction
480	65	Class 5–6 boundary	80
930	120	Class 4–5 boundary	70
1410	165	Lower part of Class 3	50
1890	210	Bottom of Class 1	20

Enhanced Processing of Climate Data and Climate Indices

The current LSRS version (version 5) carries forward all the web-based processing functionality developed in version 4 and adds to it a new climate indices calculator (**Table 3**). While version 4 could accept climate indices calculated externally, version 5 incorporates the calculation of climate indices and supports raster data processing. Several major new features and improvements are incorporated into this release. Key amongst these is the use of Redis as a data store for climate data. Redis is an open source in-memory data structure store, used as a database, cache and message broker. The climate calculator validates climate data, calculates climate indices, and stores the calculated indices in Redis which simplifies access, and enhances performance and output reliability. The calculator reads climate indices directly from Redis to simplify use and to eliminate a significant source of human errors.

As the map polygon is the unit of calculation for LSRS, it is necessary to assign climate attributes (daily Tmax, Tmin, and Precipitation) to each polygon. The basic conceptual approach is the spatial union between the climate grids and soil map polygons (**Figure 3**). Depending on the scale of the grids and map polygons, when several climate grid centroids fall into one soil map polygon, these are averaged for the polygon.

The agro-climatic indices used in LSRS are calculated from mean monthly or average mean daily data for each polygon. The climate data are loaded into LSRS using a custom file

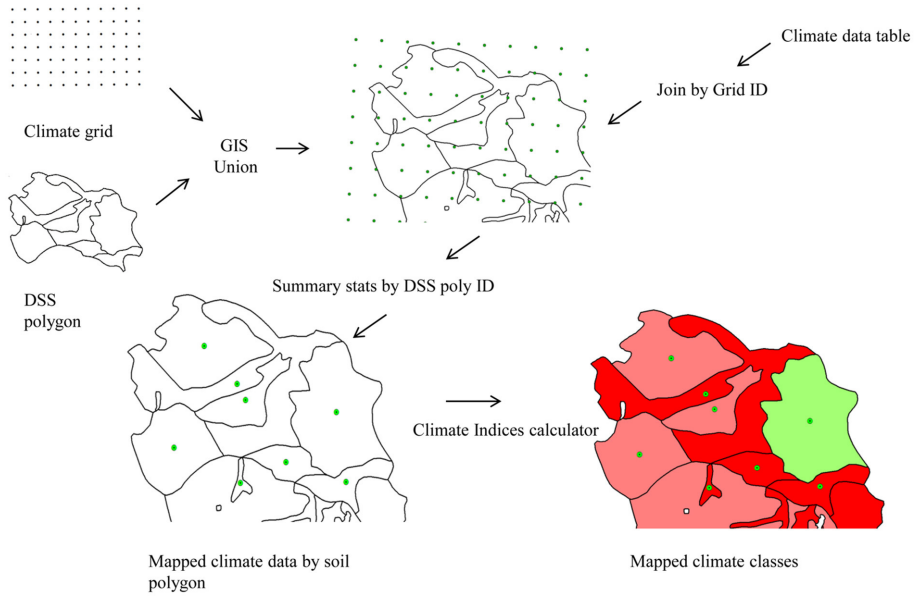


FIGURE 3 | Allocation of gridded climate data to detailed soil survey (DSS) map polygons (adapted from Gasser et al., 2016, reproduced with the permission of the copyright holder, NRC Research Press).

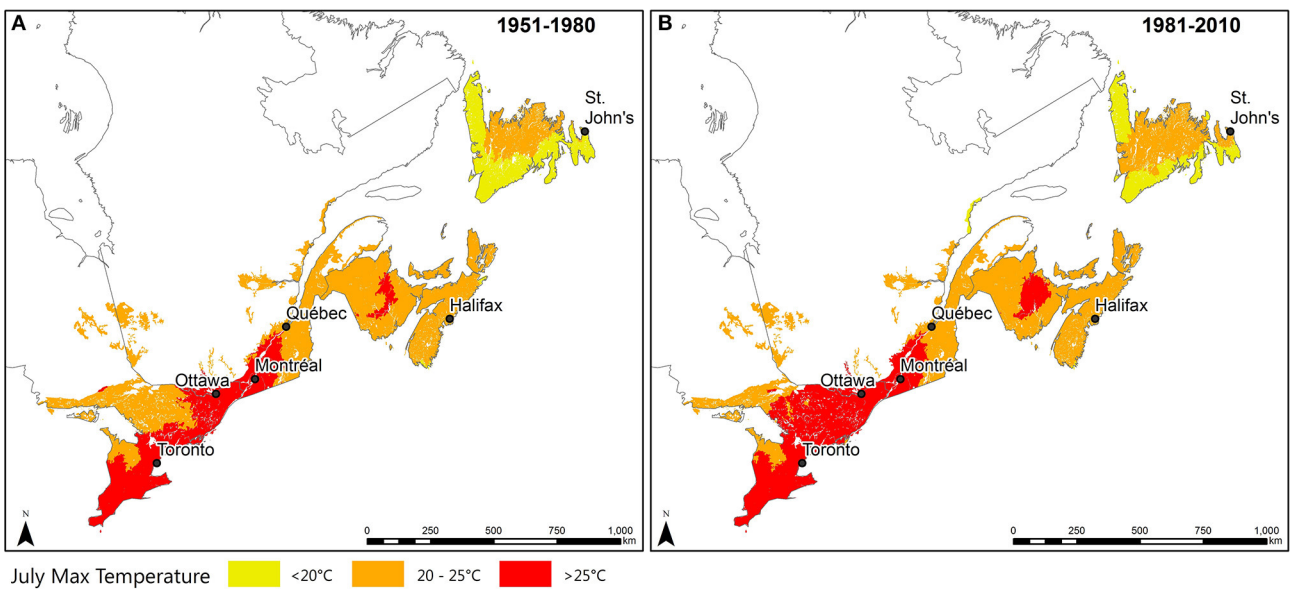


FIGURE 4 | Change in the July maximum temperature in eastern Canada between climate normal periods 1951–1980 **(A)** and 1981–2010 **(B)**. Expansion of warmer summer temperatures is visible in southern Ontario and Quebec, central New Brunswick and Newfoundland. Climate data were transposed from a national 10 km grid to 1:1,000,000 soil polygons to facilitate subsequent LSRS rating.

format that incorporates both documentation and climate data. The file is divided into two sections. The first section contains the documentation in YAML format, and the second section contains the data as a set of tab-separated values. Index calculation methods are specific to each crop. The climate indices calculator within LSRS version 5 generates the following outputs: growing season start date, growing season stop date, growing degree days >5 C, crop heat unit and P-PE.

Prior to version 5, climate indices had to be calculated outside of LSRS and then imported into the system. Early versions of the system were designed to work with monthly data. Indices requiring daily time steps were calculated by deconstructing monthly values through a curve fitting step (Bootsma and

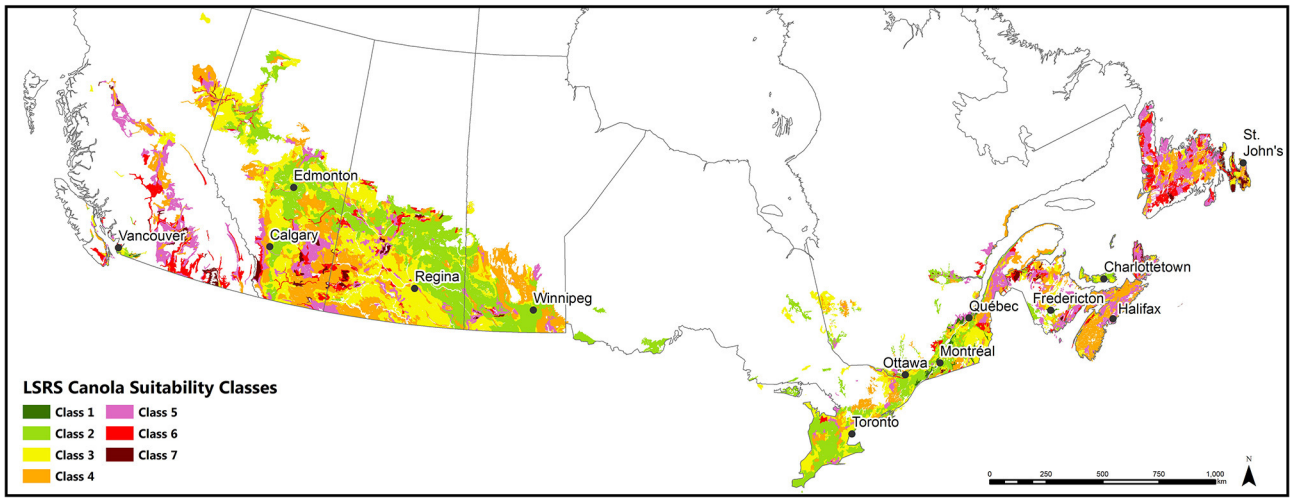


FIGURE 5 | National-scale map of land suitability classes for canola based on dominant soil types and the period 1981 to 2010. Highest suitability classes exist in Prince Edward Island (Charlottetown), in southern Quebec and Ontario and along the northern agricultural region of the Prairie provinces. Lower suitability classes result elsewhere in the country due to aridity in parts of western Canada and cool temperature in the Atlantic region.

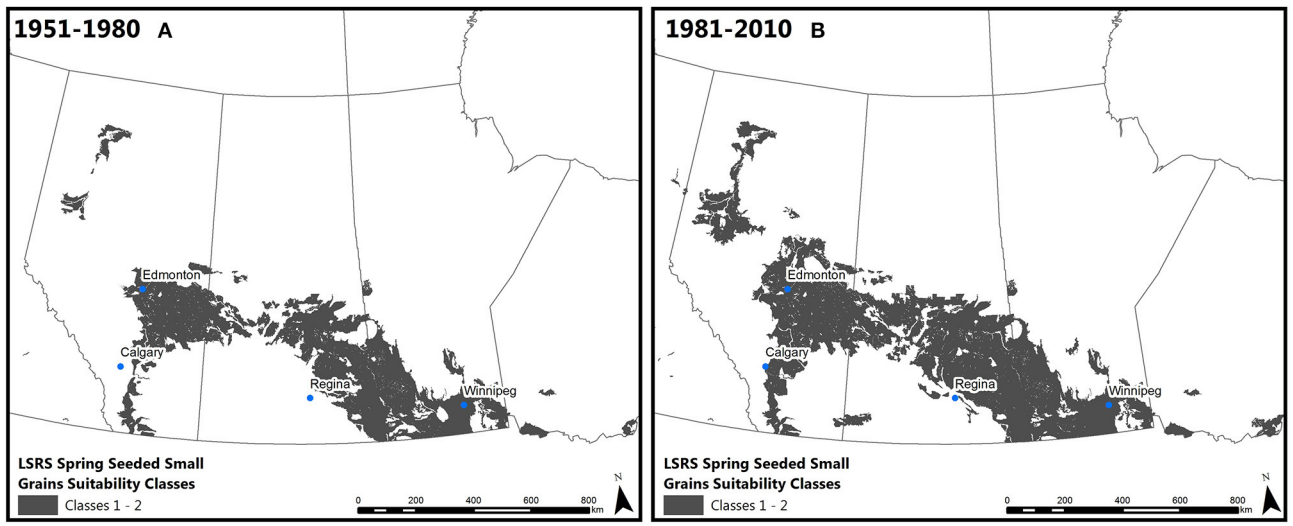


FIGURE 6 | Shaded areas show the expansion of prime land (class 1 and 2) for spring seeded small grains based on the climate normal period 1951–1980 **(A)** and 1981-2010 **(B)**.

Boisvert, 1991). To date we have not depicted results based on a single year of data (although is this possible) but have stayed with the original concept that LSRS rating are generated from climate averages over multiple years.

Computational Upgrades

Major upgrading of the LSRS system began in 2009 with the rewriting of the core LSRS calculator in Ruby as a suite of web services. The program can run and store batch processes compliant with Web Processing Services (WPS), calculate results for multiple crops simultaneously and utilize a range of map scales and data structures that exist for Canadian soil maps. The complete details of each calculation are visible as a web page (HTML) but data are also output as CSV and GDAS formats, and the system can provide thematic maps via WMS and KML. Significant adjustments were made between 2011 and 2014 to the original crop ratings algorithms as the result of regional validation efforts across Canada.

For most calculations, Version 5 generates ratings directly from class boundary (configuration) points instead of curve formulas. This approach eliminates unintended

effects that occurred in earlier version of LSRS at climate extremes, helps to improve consistency with intended class breaks, and simplifies the adjustment of deduction curves.

The internal data access mechanism and LSRS calculation requests are based on a REST (Representational State Transfer) architectural approach to simplify operations and facilitate interoperability with other systems. Our implementation creates separate access for climate, soil and landscape ratings, and stores calculation results in Redis so that the results can be reused in subsequent routines to improve the performance of raster calculations.

Finally, in order to simplify maintenance of the system, version 5 drops XSLT in favor of plain HTML for most human-readable outputs.

APPLICATIONS OF THE RATING SYSTEM

LSRS may be applied at scales based on the scale of the input soil, landscape and climate information. The scale of output is controlled by the scale of the soil map used as input. Ratings are calculated on soil profile attributes for each soil type reported to occur in a map polygon. Typically, results are depicted based on the dominant (spatially most common) soil. Climate indices are calculated from data grids at a variety of scales and intersected with the soil map polygons. LSRS may also be used on-site with attribute data coming from an observed profile, landscape parameters measured on-site and climate data obtained from a near-by weather station. In this case LSRS ratings would be site-specific. Examples of national and regional assessments are presented in the following sections.

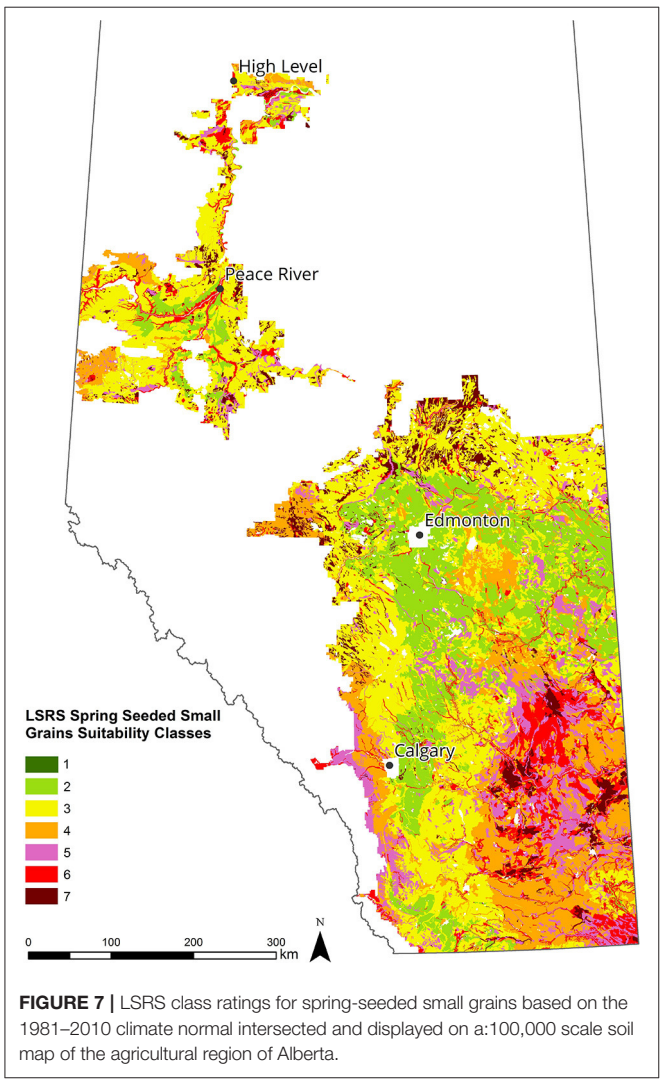


FIGURE 7 | LSRS class ratings for spring-seeded small grains based on the 1981–2010 climate normal intersected and displayed on a:100,000 scale soil map of the agricultural region of Alberta.

National Scale Assessments

LSRS can map a variety of outputs on a soil map polygon basis at regional and national scales. These include climate attributes, agro-climatic indices, and for a particular crop, climate class rating, soil class rating, landscape class rating and an overall LSRS class and subclass suitability rating. This is achieved by integrating climate data with the 1:1,000,000 Soil Landscapes of Canada mapping to produce actual LSRS ratings (Schut et al., 2011).

Outputs at the national scale help to explain historic and future projected shifts in cropping patterns and can help to interpret changes in agri-environmental performance in national policy assessments). Climate change impacts using IPCC AR3 climate scenario data downscaled to a 10 km national climate grid and linked to 1:1,000,000 Soil Landscapes of Canada polygons were assessed for the 1971 to 2000 normal period and three future 30-year periods for three regions of Canada–the Prairie provinces, central Canada (southern Quebec and Ontario) and Atlantic Canada (Chen et al., 2008). Based on suitability rating changes over time the authors speculated that grain and canola production would move onto climatically marginal crop land that is currently under forage production or livestock grazing and where aridity is not a limitation, corn and soybeans could expand into areas now mostly dominated by grain and oilseed production, especially in the northern regions of the Prairie provinces. Perhaps more importantly, was the recognition that seeding and maturity dates will shift earlier into the year as a necessary adaptation to mid-season aridity in western Canada.

LSRS has been used to examine climatic shifts within the agricultural region of the country. **Figure 4** depicts slight changes in July maximum temperatures in the agricultural portions of the country between the climate normal periods 1951–1980 and 1981 to 2010. Even though the temperature ranges are depicted at a scale of 1:1,000,000 the extent of July warming during the historical period is obvious.

Integration of climatic data with the soil and landform attributes contained with the databases linked to the Soil Landscapes of Canada map product, enables crop ratings to be depicted for the dominant soils on the landscape. **Figure 5** shows suitability ratings for canola based on climate normal data for the period 1981–2010. **Figure 6** illustrates the expansion of LSRS class 2 lands for spring seeded small grains in western Canada for two historical time periods. Due to slight climatic limitations for grain production there is no class 1 in this region of Canada.

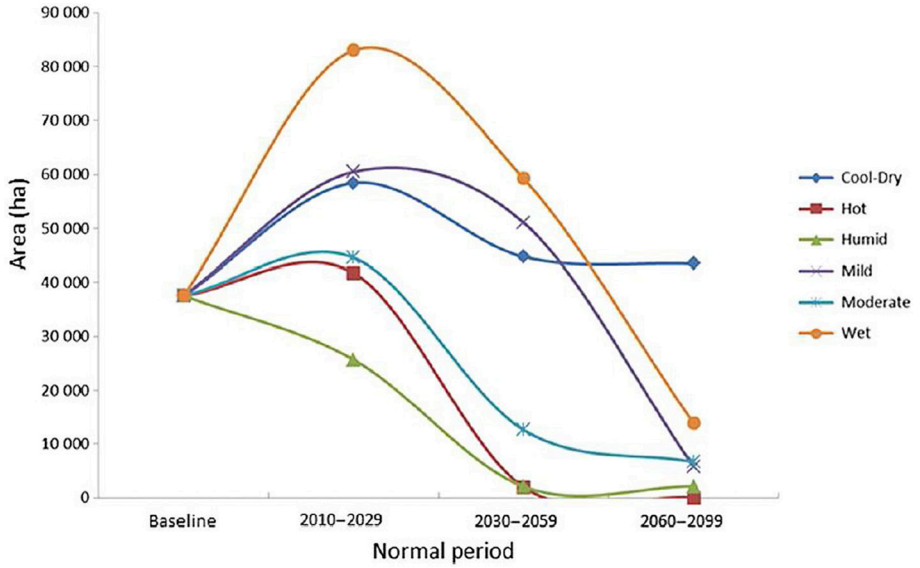


FIGURE 8 | Trends in area of suitability classes 1 and 2 for corn production in the lower Fraser Valley of British Columbia. Six different global circulation models depict a wide range of outcomes (adapter from Gasser et al., 2016, reproduced with the permission of the copyright holder, NRC Research Press).

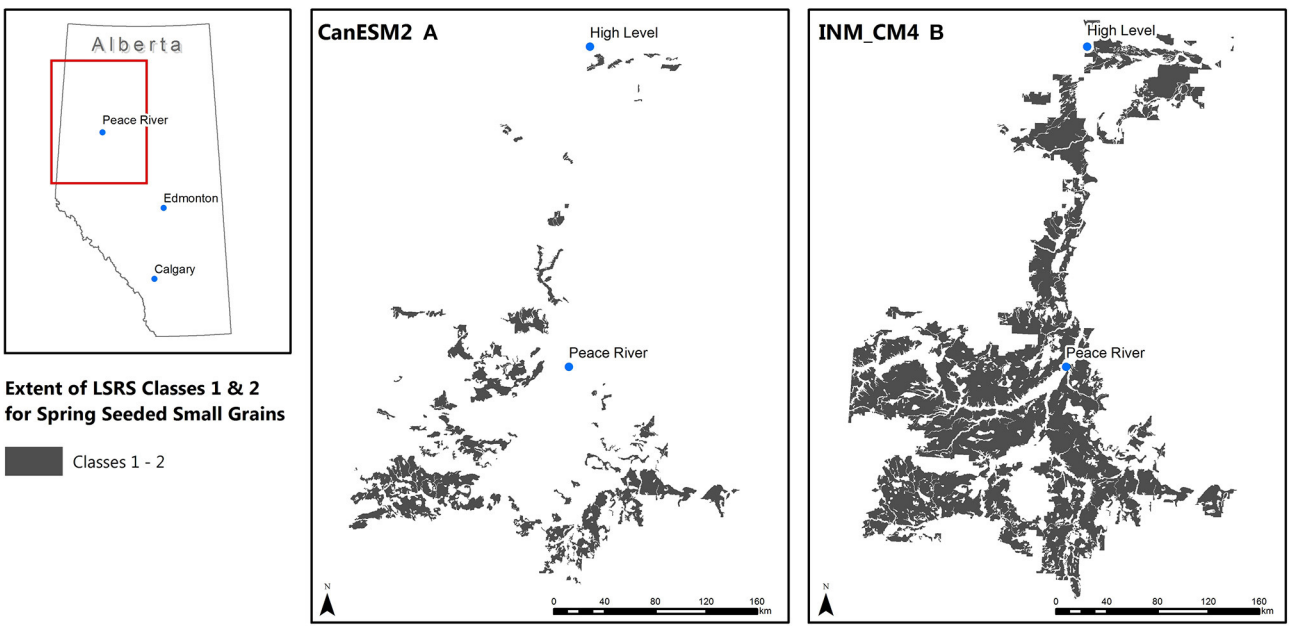


FIGURE 9 | Comparison of future prime land extents in the Peace River district of northwestern Alberta for the period 2041–2070. Shown are results based on climate outputs for RCP4.5 from the CanESM2 **(A)** and INM_CM4 **(B)** models intersected with 1:100,000 soil map polygons. The difference shown represents the degree of spatial uncertainty that exists with respect to future land suitability for grain production in the region.

Regional Climate Change Assessments

LSRS has been used locally in many provinces to replace the older CLI capability maps that were drawn based on mid-twentieth century climate conditions. Regional LSRS spatial outputs utilize regional detailed and semi-detailed soil maps as their base. In Alberta, where much of the original LSRS validation work was conducted, a provincial 1: 100,000 scale soil map is used as the base for depicting provincial-scale crop ratings. The baseline for making this comparison was the map for spring seeded small grain (**Figure 7**). Spatial validation of this type of output was used when designing the early versions of LSRS.

Gasser et al. (2016) examined the impact of a range of IPCC AR4 SRES scenarios on future production of corn in coastal British Columbia. In this instance, projected increasing

growing season aridity (a subclass of the climate rating) indicated the likely future requirement for the adoption of irrigation much more widely than is currently the practice. In all our climate change impact studies to date, a wide range of suitability outcomes are projected as a function of the model scenarios selected. This range is a measure of uncertainty. Mapped class areas for each scenario are typically summed and plotted over several time periods to see the spread in trends (**Figure 8**).

A recent application of LSRS has helped to evaluate climate change impacts on northern agricultural regions of northwest Canada. Downscaled data (Wang et al., 2016) from a subset of five climate models were selected from the IPCC CMIP5 ensemble (Cannon, 2015) to span the range of projected temperature and precipitation changes in the region. These were then used to drive LSRS for the RCP4.5 and RCP8.5 scenarios. One study area is the Peace River district of northwestern Alberta where earlier national-scale assessments indicated considerable potential change. Grain and oilseeds are the main crops in the region. We analyzed climate over both historic and future time periods. However, as with our work in British Columbia, different model outputs projected vastly different outcomes with respect to agricultural suitability by late century. **Figure 9** compares the outputs for two models, the Canadian model CanESM2 and the Russian model INM_CM4. While both models project increasing heat, CanESM2 projects significant aridity such that the extent of prime land for spring seeded grains (without irrigation) is greatly diminished relative to current condition. However, INM_CM4 projects increased precipitation along with the heating such that the extent of prime land is greatly increased relative to the current condition. The LSRS output mapping clearly demonstrates the spatial uncertainty of future land suitabilities and the range of possible crop distribution patterns.

DISCUSSION

LSRS classification remains a largely qualitative pursuit utilizing parametric scores based on expert knowledge to calculate deductions based on measurable climate, soil and landscape attributes. The derived class ratings are relative rankings rather than absolute measures. Rossiter (1996) termed this approach as a "land index," a simple rating of the goodness of fit of a particular land activity (crop type) for a given land capability. The principle purpose of agriculture land suitability assessment is to predict the potential and limitations of the land for production of a specific crop or crops. LSRS generates a measure of both potential (as a class rating) and a limitation (as a subclass modifier). LSRS contains a set of tools that allow the user to calculate a suite of agro-climatic indices integrated to a soil map to allow spatial assessment at a range of scales. LSRS has evolved to be increasingly useful to climate change studies in that it is able to depict spatially future land use possibilities. While the LSRS modules currently cover only common annual crops and some forages, the suitability concept of integrating climate, soil and landscape attributes in a spatial framework has been extended to examine climate change on potential future extent of sweet cherries in the interior of British Columbia (Neilsen et al., 2017).

Validation of the rating system remains a challenge. Two obvious independent variables, crop yield and crop presence on the landscape are two attributes against which to try to relate to suitability. However, as both of these variables are very much controlled by market/economic factors and management inputs, neither of which are considered in the LSRS computations, finding a strong correlation between an LSRS class rating and crop yield remains elusive. The early development of the system was driven by a need to have LSRS, specifically the spring seeded small grains module, emulate the CLI agriculture capability ratings. This objective was achieved but the LSRS is not a crop yield model nor necessarily a predictor of crop presence. Validation of the outputs remains an area of study.

A drawback of using a class rating structure is the inability to often detect modest changes in point ratings that may be obscured by the relatively wide range of points that constitute any one class. We encountered this problem in our climate change work in British Columbia (Gasser et al., 2016) where in some cases small point changes generated class changes, in other instances relatively large point changes did not generate any class change at all. When analyzing change, it is often better to work with point deduction values directly rather than class values.

The current version of the rating system utilizes web-based technologies that optimize data handling and storage and facilitates user access to the tools and calculators that make up the system. The efficiency in generating agro-climatic indices from both daily and monthly data makes LSRS well suited to climate change studies in Canada where many existing crop production systems are limited by climate (Campbell et al., 2014).

AUTHOR CONTRIBUTIONS

MB and CS compiled the first draft of the article and CS revised the manuscript following reviewer comments. DN was project lead at Summerland for land suitability research and provided funding and technical input to model developments. WP wrote an earlier version of the introduction and led the development of versions 1, 2, and 3 of LSRS, P-YG and MB conducted GIS work and produced the map figures, AB established the agro-climatic indices for the new crop modules and wrote an earlier version of the crop module section, AJB provided development guidance for version 3 and 4. PS built versions 4 and 5 of LSRS and currently manages the model calculators and model website.

ACKNOWLEDGMENTS

We would like to acknowledge funding from the Science and Technology Branch of Agriculture and Agri-Food Canada to support the development of the LSRS. We also acknowledge the contributions of the original Agronomic Interpretations Working Group that initiated the develop of the system. We would also like to thank David Lee and Simon Kiley for support and help with preparation of map figures for the paper.

Institutional support from the Alberta Agriculture, Food and Rural Development is gratefully acknowledged as well as the contributions of the many contractors, summer students and consultants who provided field testing, expertise, and feedback over the years.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2018.00077/full#supplementary-material

REFERENCES

Agriculture and Rural Development Act (1965). *Canada Land Inventory. Objectives, Scope and Organization*. The Canada Land Inventory Report No.1. Dept. of Forestry and Rural Development, Ottawa (Reprinted by Dept. of Regional Economic Expansion in 1970). 61.

Agronomic Interpretations Working Group (1995). *Land Suitability Rating System for Agricultural Crops: Spring-Seeded Small Grains,* ed W. W. Pettapiece. Ottawa, ON: Tech Bull. 1995-6E; Centre for Land and Biological Resources Research, Agriculture and Agri-Food Canada.

Atlantic Advisory Committee on Soil Survey (1988). "A compendium of soil survey interpretive guides used in the Atlantic Provinces," eds G. T. Patterson and H. W. Rees (Truro, NS: Agriculture Canada, Land Resources Unit), 149.

Alberta Soils Advisory Committee (1983). *Systems of rating arable land in Alberta*. 200-2-84. Edmonton, AB: Alberta Agriculture. 27.

Alberta Soils Advisory Committee (1987). "Land capability Classification for Arable Agriculture in Alberta," ed W. W. Pettapiece (Alberta Agriculture), 103. 5 maps.

Ahrens, J. R., Rice, T. J., and Eswaran, H. (2002). Soil Classification: past and present. *NCSS newsletter* 19, 1–5.

Alberta Agriculture Food and Rural Development (2000). *Procedures Manual for the Classification of Land for Irrigation in Alberta*. Lethbridge AB: Alberta Agriculture, Food and Rural Development.

Alberta Energy and Natural Resources (1983). *Soil Capability for Agriculture: A Summary and Application of Canada Land Inventory Data in Alberta*. ENR Technical Report. T/45. Edmonton, AB.

Arroues, K. D., and Anderson, C. H (1986). *Soil Survey of Kings County, California*. Washington, DC: Natural Resource Conservation Service.

Bootsma, A. (1991). *Risk Analyses of Heat Units Available for Corn Production in the Maritime Provinces*. Agriculture Canada, Research Branch Tech. Bull. 1991-8E, 49pp.

Bootsma, A., and Boisvert, J. B. (1991). *Modeling Methodology for Estimating Forage Yield Potential in Canada*. Ottawa: Tech, Bull. 1991-6-E. Research Branch, Agriculture Canada.

Bootsma, A., Gameda, S., and McKenney, D. W. (2005). Impacts of potential climate change on selected agro-climatic indices in Atlantic Canada. *Can. J. Soil Sci.* 85, 329–343. doi: 10.4141/S04-019

Bootsma, A., Tremblay, G., and Filion, P. (1999). *Risk Analyses of Heat Units Available for Corn and Soybean Production in Quebec.* Agriculture and Agri-Food Canada, Research Branch, Eastern Cereal and Oilseed Research Centre, Ottawa, Technical Bulletin, ECORC Contrib. No. 991396. 127 pp.

Bowser, W. E., Peters, T. W., and Newton, J. D. (1951). *Soil Survey of the Red Deer Sheet*. Alberta Soil Survey Report No.16. Edmonton, A. B. 86. 2 maps.

Brooke, M. A., and Presant, E. W. (1986). *Site Determination of Soil Capability for General Field Crops in the Regional Municipalities of Haldimand-Norfolk and Niagara*. Guelph, ON: Ontario Institute of Pedology Publ. No. 86-4. 20.

Canada Land Inventory (1998). *National Soils Database. Agriculture and Agri-Food Canada*. Available online at: http://sis.agr.gc.ca/cansis/nsdb/cli/index.html (Accessed March 20, 2018).

Campbell, I. D., Durant, D. G., Hunter, K. L., and Hyatt, K. D. (2014). "Food Production," in *Canada in a Changing Climate: Sector Perspectives on Impacts and Adaptation,* eds F. J. Warren and D. S. Lemmen (Ottawa, ON: Natural Resource Canada, Government of Canada), 99–134.

Cannell, R. Q., Davies, D. B., Mackney, D., and Pidgeon, J. D. (1978). The suitability of soils for sequential direct drilling of combine-harvested crops in Britain: a provisional classification. *Outlook Agric.* 9. 306–316.

Cannon, A. J. (2015). Selecting GCM scenarios that span the range of changes in a multimodel ensemble: application to CMIP5 climate extremes indices. *J. Clim.* 28, 1260–1267. doi: 10.1175/JCLI-D-14-00636.1

Carcass Burial Site Selection Technical Committee (2004). *Natural Suitability Index for Carcass Burial in Alberta.* Edmonton, AB: Alberta Agriculture. 36.

Chapman, L. J., and Brown, D. M. (1966). *The Climates of Canada for Agriculture*. The Canada Land Inventory Report No. 3. Environment Canada.

Chen, K., Hewitt, J., Brierley, T., Chipanshi, A., Hanuta, I., and Zimmer, S. (2008). *Assessment of Climate Change Impacts on Agricultural Land-Use Suitability*. Ottawa, ON: Agriculture and Agri-Food Canada and Environment Canada.

Eilers, R. G., and Buckley, K. E. (2002). *A Methodology for Evaluation Soils, Landscapes and Geology for Nutrient Management Planning in the Prairie Landscape*. Tech. Bull. 2001-1E. Ottawa: Land Resource Group, Research Branch, Agriculture and Agri-Food Canada.

Elsheikh, R., Rashid, B. A., Shariff, M., Amiri, F., Ahmad, N. B., Balasundram, S. K., et al. (2013). Agriculture Land Suitability Evaluator (ALSE): a decision and planning support tool for tropical and subtropical crops. *Comput. Electron. Agric.* 93. 98–110. doi: 10.1016/j.compag.2013.02.003

FAO (1976). *A Framework for Land Evaluation*. Rome: Soil Bull. 32. UN Food and Agriculture Organization (FAO).

Feller, C. L., Thuries, L. J.-M., Manlay, R. J., Robin, P., and Frossard, E. (2003). "The principles of rational agriculture" by Albrecht Daniel Thaer (1752–1828), An approach to the sustainability of cropping systems at the beginning of the 19th century. *J. Plant Nutr. Soil Sci.* 166, 687–698. doi: 10.1002/jpln.200321233

Gasser, P.-Y., Smith, C. A. S., Brierley, J. A., Schut, P., Neilsen, D., and, Kenney, E. A. (2016). Use of a Land Suitability Rating System to assess climate change, Fraser Valley of British Columbia. *Can. J. Soil Sci.* 96, 256–269. doi: 10.1139/cjss-2015-0108

Goodman, V. H. (1955). *Soil Survey of Waldo County, Maine. Series 1940, No. 16*. Washington, DC: Soil Conservation Service.

Goplen, B. P., Baenziger, H., Bailey, L. D., Gross, A. T. H., Hanna, M. R., Michaud, R., et al. (1987). *Growing and Managing Alfalfa in Canada*. Ottawa, ON: Agriculture Canada Publication 1705/E. Agriculture Canada.

Huddleston, J. H. (1984). Development and use of soil productivity ratings in the United States. *Geoderma* 32, 297–317.

Joss, B. N., Hall, R. J., Sidders, D. M., and Keddy, T. J. (2008). Fuzzy-logic modelling of land suitability for hybrid poplar across the Prairie Provinces of Canada. *Environ. Monit. Assess.* 141, 79–96. doi: 10.1007/s10661-007-9880-2

Kenk, E., and Cotic, I. (1983). *Land capability classification for agriculture in British Columbia. MOE Manual, I*. Victoria, BC: Ministry of Environment and Ministry of Agriculture and Food.

Klingebiel, A. A., and Montgomery, P. H. (1961). *Land Capability Classification. Agriculture Handbook No. 210*. Washington, DC: Soil Conservation Service, USDA.

Lynn, I., Manderson, A., Page, M., Harmsworth, G., Eyles, G., Douglas, G., et al. (2009). *Land Use Capability Survey Handbook–A New Zealand Handbook for the Classification of land 3rd Edn*. Hamilton, New Zealand: AgResearch Ltd.

Marshall, I. B., Dumanski, J., Huffman, E. C. and Lajoie, P. G. (1979). *Soil Capability and Land Use in 719 the Ottawa Urban Fringe.* Ottawa, ON: Research Br. Agriculture Canada.

Mailloux, A., Dube, A., and Tardif, L. (1984). Classement des sols selon leur possibilities d'utilization agricole. *Cah. Geogr. Que.* 8, 231–249.

McCracken, R. E., and Cate, R. (1986). Artificial Intellegence, cognative science and measurement theory application in soil classification. *Soil Sci. Soc. Am. J.* 50, 557–561.

Morrison, M. J., McVetty, P. B. E., and Shaykewich, C. F. (1989). The determination and verification of a baseline temperature for the growth of Westar summer rape. *Can. J. Plant Sci.* 69, 455–464.

Morrison, M. J., and Stewart, D. W. (2002). Heat stress during reproduction in summer rape. *Crop Sci.* 42, 797–803. doi: 10.2135/cropsci2002.7970

Moser, L. E., Buxton, D. R., and Casler, M. D. (1996). *Cool-season forage grasses. Agronomy series No. eds 34*. Madison: American Society of Agron.

Mueller, L., Schindler, U., Behrendt, A., Eulenstein, F., and Dannowski, R. (2007). Das Muencheberger Soil Quality Rating (SQR): ein einfaches Verfahren zur Bewertung der Eignung von Boeden als Farmland. *Mitt Dt Bodenkundl Ges*. 110, 515–516.

Mueller, L., Schindler, U., Mirschel, W., Shepherd, T. G., Ball, B. C., Helming, K., et al. (2010). Assessing the productivity function of soils. a review. *Agron. Sustain. Dev*. 30, 601–614. doi: 10.1051/agro/2009057

Neilsen, D., Smith, S., Bourgeois, G., Qian, B., Cannon, A., Neilsen, G., et al. (2017). Modelling changing suitability for tree fruits in complex terrain. *Acta Hortic*. 1160, 207–214. doi: 10.17660/ActaHortic.2017.1160.30

Odynsky, W. M., Wynnyk, A., and Newton, J. D. (1952). *Soil Survey of the High Prairie and McLennan Sheets*. Edmonton, AB: Alberta Soil Survey No. 17. 2 maps.

OMAFRA. (2002). *Agronomy Guide for Field Crops*. Guelph, ON: OMAFRA publication 811.

Peters, T. W. (1977). Relationship of yield data to agroclimate, soil capability classifications and soil of Alberta. *Can. J. Soil Sci*. 57, 341–347.

Peters, T. W., and Pettapiece, W. W. (1981). *Crop Yields in Alberta: Relationships to Soils Capability for Agriculture and Soil Type*. Alberta Institute of Pedology No. M-81-1. University of Alberta.

Rossiter, D. G. (1996). A theoretical framework for land evaluation. *Geoderma* 72, 165–190.

Schut, P., Smith, S., Fraser, W., Geng, X., and Kroetsch, D. (2011). Soil Landscapes of Canada: building a national framework for environmental information. *Geomatica* 65, 293–309. doi: 10.5623/cig2011-045

Smyth, A. J., and Dumanski, J. (1995). A framework for evaluating sustainable land management. *Can. J. Soil Sci*. 75, 401–406.

Sonneveld, M. P. W., Hack-ten Broeke, M. J. D., van Diepen, C. A., and Boogaard, H. L. (2010). Thirty years of systematic land evaluation in the Netherlands. *Geoderma* 156, 84–92. doi: 10.1016/j.geoderma.2010.02.023

Standing Committee on Agriculture, Fisheries and Forestry (1984). *Soil at Risk: Canada's Eroding Future*. Ottawa, ON. 129.

Storie, R. E. (1933). *An index for rating the agricultural value of soils. University of California Agricultural Experimental Station Bull. 556*. Berkley, CA: University of California.

Triantafilis, J., Ward, W. T., and McBratney, A. B. (2001). Land suitability assessment in the Namoi Valley of Australia, using a continuous model. *Aust. J. Soil Res.* 39, 273–290. doi: 10.1071/sr99087

Undersander, D., Martin, N., Cosgrove, D., Kelling, K., Schmitt, M., Wedberg, J., et al. (1991). *Alfalfa Management Guide*. Madison: American Society of Agronomy Crop Science Society of America Soil Science Society of America.

van Vliet, L. J., Mackintosh, E. E., and Hoffman, D. W. (1979). Effects of land capability on apple production in Southern Ontario. *Can, J. Soil Sci*. 59, 163–175.

Wang, T., Hamann, A., Spittlehouse, D., Carroll, C. (2016) Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS ONE* 11:e0156720. doi: 10.1371/journal.pone.0156720

Williams, G. D. V. (1983). Agroclimatic resource analysis-an example using an index derived and applied for Canada. Agric. *Meteorol*. 28, 31–47.

Wright, I. A., Birnie, R. V., Malcolm, A., Towers, W., and McKeen, M. (2006). *The Potential Use of the Land Capability for Agriculture Classification for Determining Support to Disadvantaged Areas of Scotland*. Aberdeen: Macaulay Institute, Craigiebuckler.

Wyatt, F. A., Bowser, W. E., and Odynsky, W. (1939). *Soil Survey of Lethbridge and Pincher Creek Sheets*. Dept of Extension, University of Alberta, Edmonton, AB. 2 maps.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in ICT

ORIGINAL RESEARCH
published: 19 April 2018
doi: 10.3389/fict.2018.00006

Using Spatial Reinforcement Learning to Build Forest Wildfire Dynamics Models From Satellite Images

*Sriram Ganapathi Subramanian and Mark Crowley**

Machine Learning Laboratory, Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON, Canada

Machine learning algorithms have increased tremendously in power in recent years but have yet to be fully utilized in many ecology and sustainable resource management domains such as wildlife reserve design, forest fire management, and invasive species spread. One thing these domains have in common is that they contain dynamics that can be characterized as a spatially spreading process (SSP), which requires many parameters to be set precisely to model the dynamics, spread rates, and directional biases of the elements which are spreading. We present related work in artificial intelligence and machine learning for SSP sustainability domains including forest wildfire prediction. We then introduce a novel approach for learning in SSP domains using reinforcement learning (RL) where fire is the agent at any cell in the landscape and the set of actions the fire can take from a location at any point in time includes spreading north, south, east, or west or not spreading. This approach inverts the usual RL setup since the dynamics of the corresponding Markov Decision Process (MDP) is a known function for immediate wildfire spread. Meanwhile, we learn an agent policy for a predictive model of the dynamics of a complex spatial process. Rewards are provided for correctly classifying which cells are on fire or not compared with satellite and other related data. We examine the behavior of five RL algorithms on this problem: value iteration, policy iteration, Q-learning, Monte Carlo Tree Search, and Asynchronous Advantage Actor-Critic (A3C). We compare to a Gaussian process-based supervised learning approach and also discuss the relation of our approach to manually constructed, state-of-the-art methods from forest wildfire modeling. We validate our approach with satellite image data of two massive wildfire events in Northern Alberta, Canada; the Fort McMurray fire of 2016 and the Richardson fire of 2011. The results show that we can learn predictive, agent-based policies as models of spatial dynamics using RL on readily available satellite images that other methods and have many additional advantages in terms of generalizability and interpretability.

Keywords: reinforcement learning, machine learning, deep learning, A3C, forest wildfire management, sustainability, spatially spreading processes

OPEN ACCESS

Edited by:
Nathaniel K. Newlands, Agriculture and Agri-Food Canada (AAFC), Canada

Reviewed by:
Saumitra Mukherjee, Jawaharlal Nehru University, India
Xinyan Huang, Hong Kong Polytechnic University, Hong Kong

***Correspondence:**
Mark Crowley
mcrowley@uwaterloo.ca

Specialty section:
This article was submitted to Environmental Informatics, a section of the journal Frontiers in ICT

Received: *25 November 2017*
Accepted: *15 March 2018*
Published: *19 April 2018*

Citation:
Ganapathi Subramanian S and Crowley M (2018) Using Spatial Reinforcement Learning to Build Forest Wildfire Dynamics Models From Satellite Images. Front. ICT 5:6. doi: 10.3389/fict.2018.00006

1. INTRODUCTION

There is a clear and growing need for advanced analytical and decision-making tools as demands for sustainable management increase and as the consequences of inadequate resources for decision-making become more profound. Artificial intelligence and machine learning methods provide ways to combine multiple modes of information such as spatial statistical ground data, weather data, and satellite imagery into a unified model for classification or prediction.

One high impact example of this potential is forest wildfire management. The risk, costs, and impacts of forest wildfires are a perennial and unavoidable concern in many parts of the world. A number of factors contribute to the increasing importance and difficulty of this domain in future years including climate change, growing urban sprawl into areas of high wildfire risk, and past fire management practices which focused on immediate suppression at the cost of increased future fire risk (Montgomery, 2014).

There are a wide range of challenging decision and optimization problems in the area of forest fire management (Martell, 2015), many of which would benefit from more responsive fire behavior models which could be run quickly and updated easily based on new data. For example, one simple decision problem is whether to allow a fire to burn or not, since burning is a natural form of fuel reduction treatment. Answering this question requires a great deal of expensive simulations to evaluate the policy options fully (Houtman et al., 2013).

These simulations are built by an active research community for forest and wildland fire behavior modeling. Data are collected using trials in real forest conditions, controlled lab burning experiments, physics-based fire modeling, and more (Finney et al., 2013). These hand crafted physics-based models' simulations have high accuracy but are expensive to create and update and computationally expensive to use. The question we ask is: **Can we learn a dynamics model from readily available satellite image data and treating wildfire as an agent spreading across a landscape in response to neighborhood environmental and landscape parameters?**

In this work, we provide evidence for an affirmative answer to this question by introducing a new approach for using reinforcement learning (RL) (Sutton and Barto, 1998) to automatically learn wildfire spread dynamics models by treating fire as an agent on the landscape taking spatial actions in reaction to its environment (Subramanian and Crowley, 2017).

Forest wildfire spread as a specific case of a more general problem which we call spatially spreading processes (SSPs), which occur when local features are changed over time by some dynamic process which is a function of properties at different locations in space and their proximity to the target location. This is to be distinguished from mere prediction of impact of a physical object moving across space as the SSP can be active in many, or all, locations at once. This is also more than merely spatial auto-correlation that measures the degree to which features at two locations are similar based on their proximity. Spatial auto-correlation can be an indicator of the presence of an SSP, but the dynamic changes over time between spatial locations may not result in values being similar or inversely correlated in a simple way. In many of these domains, the dynamics are often modeled by hand using agent-based models and geostatistics methods. In other areas, they are increasingly learned from data streams by being treated as videos. However, each of these approaches has their drawbacks. One of the goals of our research is to find novel representations of local dynamics for SSPs that can provide transparent and interpretative solutions to learning of dynamics models and decision-making across many domains. The approach proposed here offers the tantalizing possibility to represent and learn a *causal* agent-based policy representation using RL which would be a much easier model to interpret and analyze for human decision makers.

In this article, we carry out experiments on the accuracy of interpolation and forward prediction using five different RL algorithms in a new formulation of agent-based learning. We use easily accessible Landsat satellite data from the USGS satellite data portal and provide a prediction accuracy of resulting policies learned with each algorithm upon comparison with the same. The dataset is a set of satellite photos of massive wild fires in Northern Alberta in 2011 and 2016. We show that Monte Carlo Tree search and the Deep RL algorithm A3C (Mnih et al., 2013) perform the best but have advantages in different situations in this domain. For comparison, we also implemented a Gaussian Process (GP) classifier (Rasmussen and Williams, 2006) as a supervised learning model on the dataset. This classifier places a Gaussian Process prior on a latent function, which is then squashed to the domain [0, 1] through a link function to obtain the probabilistic classification.

This approach has applications in forest wildfire management and opens up possibilities for interpretative dynamics models as well as providing new challenging data sets for RL research. Our approach promises to provide another source of validation for existing wildfire models as well as gives us the opportunity to perform faster and more accurate prediction by learning patterns in the raw data. Another application of our approach would be learning transferable models from one data-rich region or time period and applying it to another region or time period where less data are available if there is evidence that fires in both regions behave similarly.

In prior work (Houtman et al., 2013), we have used standard, physics-based simulators such as Farsite (Finney, 1998) to carry out automated planning of fire control policies using Monte Carlo simulations and optimization. These simulations took on the order of 1 h to run for 100-year simulations of forest fire futures, which present a challenge when thousands of simulation trajectories are needed for statistical confidence. Thus, augmenting these detailed simulators with faster approximations learned from data could improve the ability to do automated planning.

2. REVIEW OF LITERATURE

We provide an overview of the relevant literature on the forest wildfire prediction and management problem in general and then on the previous use of machine learning algorithms for this domain.

2.1. Machine Learning on Satellite Imagery

Researchers have attempted to use machine learning in combination with satellite imagery in a series of ecological applications in the past. In Kubat et al. (1998), the authors use satellite radar images and machine learning for the detection of oil spills. Algorithms, such as C4.5 and 1-nearest neighbor, are used along with expert rules to train the classifier which the authors specify was hard and not completely accurate. Our work completely removes the need for such expert rules.

In Jean et al. (2016), satellite imagery and machine learning has been used to tackle the case of poverty. Convolutional neural nets and transfer learning have been used to derive models having good performances in terms of accuracy. This research demonstrated how machine learning tools, which are typically suited for data-rich domains, can be used for data-scarce settings too.

Coral reef research is another aspect that has been studied using satellite images and machine learning (Knudby et al., 2010). In Knudby et al. (2010), a series of statistical and machine learning models have been used on the IKONOS satellite imagery to produce spatially explicit predictions of species richness, biomass, and diversity of fish community. This research motivates the exploration of importance of different variables in the predictive models by using permutations techniques.

2.2. Forest Wildfire Prediction and Management

Montgomery (2014) discusses the increased future fire risk at the consequence of fire management practices, which focused on immediate suppression. The three core themes of externalities, incentives, and risk-based decision analysis in the case of wild fire suppression are described. The goal is to determine how the core themes contribute to the evolution of an effective future fire policy. This problem is also enumerated in Houtman et al. (2013).

The standard for wildland fire behavior and forest fire modeling is described in Finney et al. (2013), which carried out exhaustive lab experiments, real forest condition simulations, and trials to build sound and coherent fire spread theory for model reliability. The resulting models are used by the US Forest Service but are very computationally expensive to run. The model accuracy also varies widely across wild fires in different regions. Cellular automaton models are also widely used to predict wildfire spread (Yongzhong et al., 2004). Our approach is easier to apply than these methods and is shown to perform better than the same.

The work in Martell (2015) describes the need for a computationally simple and accurate model for wild fires. They highlight a number of challenging decision and optimization problems in the area of forest fire management and recent efforts to develop decision support tools to overcome them. The focus is on using methods of Operations Research to aid fire managers making complex decisions about fire suppression and resource allocation.

2.3. Machine Learning in Forest Fire Management

In Castelli et al. (2015), the authors discuss the application of an intelligent system based on genetic programming for the prediction of burned areas in a wild fire situation. They also compare the genetic programming methodology to state-of-the-art machine learning techniques in fire modeling and conclude that genetic programming techniques are better. The major machine learning techniques used are SVM with a polynomial kernel, random forests, radial basis function network, linear regression, isotonic regression, and neural networks.

Significant problems have arisen while dealing with large databases or long periods of observation (e.g., pattern recognition, geophysical monitoring, monitoring of rare events (natural hazards), etc.). The authors in Forsell et al. (2009) stipulate that the major problems in such cases are how to explore, analyze, and visualize the oceans of available information. Several important applications of machine learning algorithms for geospatial data are presented: regional classification of environmental data, mapping of continuous environmental data including automatic algorithms and optimization (design/redesign) of monitoring networks.

Machine learning algorithms use an automatic inductive approach to recognize patterns in data. Once learned, pattern relationships are applied to other similar data to generate predictions for data-driven classification and regression problems. The work in Cracknell and Reading (2014) takes a task of supervised lithology classification (geological mapping) using airborne images and multispectral satellite data and compares the application of popular machine learning techniques to the same. A 10-fold cross validation was used to select the optimal parameters to be used in all the methods. These selected parameters were used to train the machine Learning classification models on the entire set of samples.

In Sehgal et al. (2006), the authors use a machine learning approach for Geospatial Entity resolution which is the problem of consolidating data from diverse sources into a single data source referenced by location (in the form of coordinates). Several feature-based matching techniques like location name matching, coordinate matching, and location-type matching are introduced and evaluated. These feature-based matching techniques use each location feature independently. A new method integrating spatial and non-spatial features and learning a combined spatial similarity measure is introduced.

Modeling forest areas is concentrated upon in Garzón et al. (2006). The environmental variables consisting of both topographic and climatic factors are considered in this work. A modeling framework for habitat modeling is established to train, test, and validate the popular predictive machine learning methods. Neural Networks, Random Forests, and Tree-Based Classification are used as predictive models. A ROC curve (Specificity vs Sensitivity graph) analysis is done for parameter selection. Species distribution and habitat modeling are described to be complex problems much like the wild fire problem with many responsible factors and the authors admit that modeling all the factors are impossible with the current state of the art. Hence, having an agent-based approach that learns the relevant factors on its own seems to be the most suitable idea.

Machine learning is used for the spatial interpolation of environmental variables in Li et al. (2011). In this study, around 23 methods are considered including popular machine learning methods and their combinations. Along with machine learning,

the considered methods were drawn from a large pool of categories including geostatistical methods, non geostatistical methods, statistical methods, and combined methods. The dataset consists of about 177 samples of sea bed mud content in the southwest of Australian Exclusive Economic Zone, and the problem is to determine the mud content in the other points by interpolation. Several primary and secondary variables are considered to influence this decision. Random forest seems to perform the best among the methods considered. RF performance was attributed to its relative robustness to outliers and noise, its nature of not over fitting with respect to the source data, and its ability to model complex interactions.

Recknagel (2001) discusses a summary of all the different ecological modeling applications of machine learning. Artificial Neural nets, Genetic Algorithm methods, and Adaptive Agents seem to perform the best, but they seem to have advantages for different problems. The conclusions are that Artificial Neural Nets perform well in problems of non-linear ordination, visualization, multiple regression, time series modeling, and image recognition and classification. Genetic algorithms are suitable to evolve causal rules, process equations, and optimize process parameters. Adaptive agents are suggested for providing a novel framework in the aid of the discovery and forecasting of emergent ecosystem structures and behaviors in response to environmental changes. This review demonstrates that while various classical AI and ML techniques have been applied to this domain in preliminary forms, there is a gap in the literature on attempting to use modern deep learning and RL techniques.

The existing state-of-the-art methods for building fire prediction models directly satellite images include the Forest Fire risk prediction model used in China (Zhang et al., 2011), the Canadian systems like Forest Fire Weather Index (FWI) System, and the Forest Fire Behavior Prediction (FBP) System (Stocks et al., 1989; Zhang et al., 2011). These systems are still quite difficult to implement due to the requirement of large amount of heterogeneous sources of data such as fuel property and fire characteristics (Zhang et al., 2011), which requires data from high resolution close-range sensors. The work by Alkhatib (2014) explains the advantages of the satellite-based fire prediction systems but specifies that the temporal resolution of available data serves as a serious impediment. Our work specifies a method to learn from available data and reliably interpolate to fill the missing spaces to get an acceptable overall performance.

We model the forest fire domain as a Markov Decision Process (MDP). This approach is becoming more common (Mcgregor et al., 2016), but these are usually focused on finding treatment actions to be taken to reduce or alter fire spread. We take the novel approach of modeling the fire itself as a decision agent attempting to minimize prediction error. This is related to work in Forsell et al. (2009), which investigated using the state variable to represent land vegetation cover and environmental characteristics but to use the action variable to represent interaction between characteristics of nearby locations. This was not on a domain as dynamic as forest fire, however, and our approach differs entirely in implementation.

Machine learning has also been used to detect and classify burn areas in Indian forests (Saranya and Hemalatha, 2012). Saranya and Hemalatha (2012) use spatial data mining to obtain useful information from a series of datasets and subsequently apply supervised algorithms such as Artificial Neural Nets (ANN) and Sequential minimal optimization (SMO) to quantify ignition risk of different regions and hence predict the occurrence of fires. Similarly, Sitanggang and Ismail (2011) use Decision Trees and a series of IF–THEN rules to develop a classification model for forest fires in Indonesia. We prove the superiority of RL techniques to such supervised methods in our work. For instance, the results in Sitanggang and Ismail (2011) show an accuracy of about 63%, whereas our best RL models perform much better than that in similar test cases. In addition, Angayarkkani and Radhakrishnan (2009) use fuzzy logic and fuzzy membership rules to make a forest fire detection system from satellite images. The methods outlined are only capable of predicting fires at the time of satellite image availability, and it is not possible to predict forward or in-between as demonstrated in our work.

Our work also has similarities to the use of intelligent systems for predicting burned areas as suggested in Castelli et al. (2015). However, that work focused on burned area alone whereas we look at the more specific problem of prediction of actual fire spread location over the short term.

3. PROBLEM FORMULATION

The problem is formulated as a Markov Decision Process (MDP) $<S, A, P, R>$ where the set of states S describes any location on the landscape. A state $s \in S$ corresponds to the state of a cell in the landscape $(x, y, t, l, w, d, rh, r)$ where x and y are the location of the cell, t is the temperature at the particular time and location, and l is the land cover type of the cell, which could be one of water, vegetation, built up, bare land, and other (derived from satellite images), and w and d are wind speed and direction, rh is relative humidity, and r is amount of rainfall. The "agent" taking actions is a fire spreading across the landscape. The action $a \in A$ indicates the direction the fire at a particular cell "chooses" to move: north, south, east, or west or to stay put. These variables are considered to be the most contributing to fires as they are the primary variables in the Canadian Forest fire weather index as specified in Cortez and Morais (2007).

The dynamics function for any particular cell $P(s'|s, a)$ is a mapping from one state s to the next s' given an action a. Note, unlike most RL domains where the unknown dynamics of the system can be very complex, in our formulation the dynamics are actually straightforward. Most properties of the cell state do not change quickly, or at all, in response to our actions. So the action of spreading a fire into a neighboring cell directly alters p in that cell, the probability that the neighboring cell in that direction will move to a burn state.

The reward function R maps a cell state to a continuous value in the range $[-1, 1]$. Rewards are based on the land-cover, for example, a cell predominantly filled with water has less chance catching fire compared with one filled with vegetation. Thus, there is a negative reward for choosing to spread fire to a cell with high percentage of water and a small positive reward for choosing to spread to a cell covered with vegetation. Other reward function components, are derived from the ground truth

data for the training sets. Thus, in the training phase, the fire is given a positive reward for taking action choices similar to the actual scenario and negative rewards otherwise. These are applied for particular experiments and algorithms as explained in the experimental setup.

Figure 1 shows a schematic representation of the domain where some cells are currently affected by fire and have a potential to spread fire to other cells nearby which is treated as the "decision" of an agent at that location. The task then is to learn a policy for this agent, which minimizes the total error at different future steps.

The initial state will come from satellite images that correspond to the beginning of a fire. We are focusing on fire growth rather than ignition detection, so we set certain cells to have just ignited fire and assign these to the initial state. As it is impossible to predict precisely from where a "fire starts," these ignited cells are decided based on thermal image data, media reports, and approximations based on the burning areas in the first day on which the satellite image is available. Wind speed and wind direction are assumed to be a constant for a small area at a fixed point of time.

The goal is to learn a policy for this agent that recreates the spread of the fire observed in later satellite images by maximizing discounted rewards designed to reward high accuracy simulation. Actions are constrained to not cross the boundary of the domain of study.

3.1. Data Acquisition

The Richardson fire occurred in a region called Richardson Backcountry north of Fort McMurray in Northern Alberta in 2011. Fort McMurray was also affected by forest fires in 2016 during the massive Fort McMurray fires (Woo and Tait, 2016). The coordinates of Fort McMurray is (56°43′36″ N, 111°22′51″ W) and the coordinates of Richardson Backcountry is (57°22′02.3″ N, 111°19′27.1″ W). The satellite images are downloaded from the USGS Earth Explorer data portal[1] for Alberta. The Landsat Enhanced Thematic Mapper Plus (ETM +) sensor carried on Landsat 7 was the most important source of data. A summary (Kanevski et al., 2008) of the merits of various software packages and tools (such as GeoMISC, GeoKNN, GeoMLP, etc.) in relation to machine learning on geospatial data was useful in developing this work. **Figure 2** shows an example where the smoke over burning areas can be seen. A series of visual and thermal images corresponding to the duration of occurrence of the Fort McMurray fire (approximately from April 2016 to September 2016) and Richardson fire (approximately from May 2011 to October 2011) were collected for the corresponding regions of Alberta. All the images were corrected for missing values and outliers. Additional pre-processing steps were carried out as outlined in Cracknell and Reading (2014). These images gave a clear description of the ground scenario in case of fire and are a reliable source of burn areas. The pixels on fire are clearly at a far higher temperature corresponding to the pixels not on fire, and this can be easily delineated from the thermal images. **Figure 3** shows the areas in darker shades of black are on fire. The spatial resolution of the satellite images was 30 m. This means that a pixel

[1]https://earthexplorer.usgs.gov/ (Accessed: December 5, 2016).

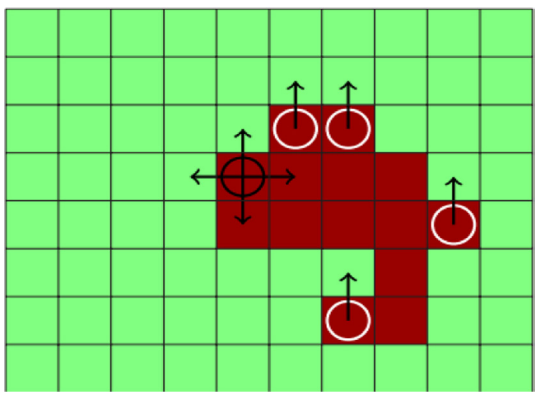


FIGURE 1 | Forest wildfire satellite data domain: a schematic of the wildfire motion domain at a particular state and timestep. The red (dark) cells are on fire, the green cells (light) are not on fire, and the dark circle indicates the current cell or agent being spread by the policy. The arrows around the dark circle indicate the action choices possible. The white circle indicates that other cells will be considered for spread. The arrows from the white circle indicate that there is a strong wind blowing toward north, and the north action is the most likely action choice for these cells (effect of wind).

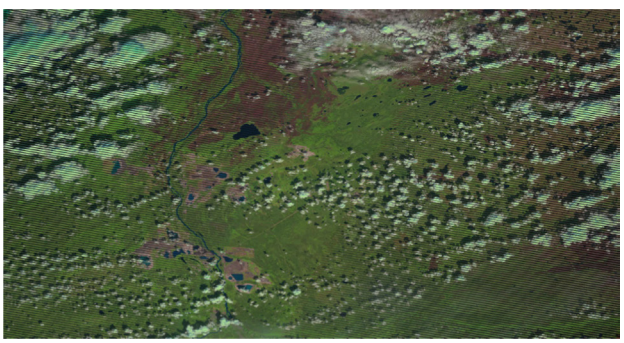


FIGURE 2 | Raw color image. Images obtained from USGS/NASA Landsat Program.

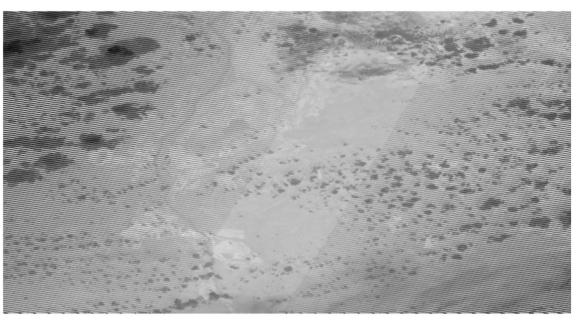


FIGURE 3 | Thermal image. Images obtained from USGS/NASA Landsat Program.

in the satellite image corresponds to 30 m × 30 m on the ground. This also means that two objects, at least 30 m long or wide, sitting side by side, can be separated (resolved) on a Landsat image. The temporal resolution of this satellite is 16 days. Thus a 16-day periodic information of a particular area during the course of the fires are available. The Landsat Enhanced Thematic Mapper Plus (ETM +) sensor is capable of sensing in 8 different spectral bands, which are made available in its datasets. Only the bands in the visible spectrum (red, blue, and green) and the thermal band are used in this work. Landsat 8 imagery has a radiometric resolution of 12-bits (16-bits when processed into Level-1 data products).

The land-cover value is obtained by processing the satellite images in an open source geoprocessing software (Multispec). Temperature is obtained from processing thermal images from the same data source. Wind speed, rain, and wind direction are obtained from the historical datasets present in the Canada Information Portal[2] and World Clim datasets for the region of study. Relative humidity is obtained from processing satellite datasets from the USGS as explained in Peng et al. (2006).

The resolution for all the inputs was fixed to be 30 m in accordance with the spatial resolution of the satellite images. As all the other information came from data sources that have lesser resolution that 30 m all the inputs were comfortably generalized to 30 m. For example, the initial state of the Richardson fire contained the following values for a particular spatial location (fixed x, y) and calendar day. Date = May 8, 2011, geographic coordinates (x, y) = (57°24′11.2″ N, 111°13′20.8″ W), temperature (t) = 10.1 C, land cover (l) = vegetation (black spruce), wind speed (w) = 8.05 mph, wind direction (d) = NW, rainfall (r) = 35.1 mm, and relative humidity (rh) = 50%.

4. DEFINITION OF ALGORITHMS

We compare five very different, widely used RL algorithms on this domain to see the range and application of the idea. We begin with three classic approaches for iteratively solving MDPs, value iteration, policy iteration, and Q-learning. We then consider the recent and very different approaches of Monte Carlo Tree Search and A3C, a recent policy gradient RL algorithm that utilizes deep Learning for value function representation. These algorithms will be reviewed briefly with modifications needed for our problem highlighted.

For all of the following methods each cell in the target area that is visited has the resulting value estimate stored in a hash data structure so it expands only as new states are encountered.

4.1. Asynchronous Value Iteration (VI)

The optimal value of the state $V^*(s)$ under the greedy fire spread policy is given by the following Bellman equation:

$$V^*(s) = R(s) + \max_a \gamma \sum_{s'} P(s' \mid s,a) V^*(s') \tag{1}$$

where s' is the successor state and γ denotes the discount factor which we set to 0.9. States are randomly sampled to update the value function according to equation (1) and updates are stopped when the change in value made by the successive iterations is less than 0.1. To provide a signal for training, a pixel in the center of the region of consideration that is currently burning in the next time step is given a high positive reward. This can be considered the goal state to reach for the burning cell as in a classic navigation problem using RL. This approach is utilized for training both value and policy iteration algorithms.

4.2. Policy Iteration (PI)

We begin with an initial random policy for acting in each state and iteratively improve it through alternating policy evaluation 2 and improvement 3 steps defined as follows:

$$V_\pi(s) = R(s, \pi_{(s)}) + \gamma \sum_{s' \in S} T(s, \pi(s), s') V_\pi(s') \tag{2}$$

$$\pi_{k+1}(s) = \text{argmax}_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V_i^{\pi_k}(s')] \tag{3}$$

The discount factor γ was fixed to be 0.9. For this algorithm, the value function $V_\pi(s)$ is an array of **Net Policy Worth** values for all states, which is the highest cumulative reward that could be obtained for an optimal policy passing through that state. This is required to maintain single-step dynamics since each cell could take more than one action at a time as fire spreads in multiple directions.

4.3. Q-Learning (QL)

This algorithm (Watkins, 1989) performs off-policy exploration and uses temporal differences to estimate the optimal policy. In Q-learning, the agent maintains a state-action value function $Q(S,A)$ instead of a state value function. This is updated as follows:

$$\hat{Q}(s,a) = (r + \gamma(\max Q(s',a'))) \tag{4}$$

and

$$Q(s,a) = Q(s,a) + \alpha(\hat{Q}(s,a) - Q(s,a)) \tag{5}$$

where α denotes the learning rate and γ denotes the relative value of delayed vs immediate rewards. S' is the new state after action a. a and a' are actions taken in states s and s', respectively. $\max_a Q(s', a')$ denotes the estimate of maximum discounted future reward expected.

The learning rate for Q-learning was chosen to be 0.9 as we need the most recent information to have a higher impact in a continuous spatial environment like that of forest fire. There is a high positive spatial auto-correlation in spatial datasets corresponding to fire as the similar pixels(on fire or not fire) tend to cluster together. The Q-learning is made to exploit this property using the appropriate learning rate. The discount factor is determined to be 0.9 as the long-term rewards are more important than the short-term rewards in our model and we want the model to converge. Using a lower discount rate decreases the level of exploration and the risk of falling into a local optimum becomes high. Each state-action pair is considered as a step taken in the real world. For every valid state, the highest Q value (utility value) for the state is recorded. The burned area is determined to be all the states having utility values over the threshold. Further, for each successive time step in the real world, a suitable number of

[2]https://www.canada.ca/en/services/environment/weather.html (Accessed: December 5, 2016).

steps taken by the fire are determined approximately using the difference between the total number of cells being burnt between successive time steps and is used to cap the total number of steps (state—action pairs analyzed) in the Q-learning implementation for the time step at which the learning is taking place. This is done to obtain a faster convergence.

4.4. Monte Carlo Tree Search (MCTS)

These are a class of algorithms that perform approximate, but confidence-bounded, RL by performing short roll-outs of the current policy to obtain enough statistical information about a state to keep it or discard it as belonging to the optimal path. A good survey is provided in Browne et al. (2012).

In our implementation of MCTS, each node in the search tree is a valid cell state in the fire model. From each state, any possible action could be taken, which is modeled as branches in the tree. The nodes in the tree are made to be selected by the **Upper Confidence Bound for trees (UCT)** (Kocsis and Szepesvári, 2006) method to minimize the cumulative regret. Each "step" in the MCTS tree is defined as a possible action taken from any valid state. A correlation between number of steps a fire can take in a 16-day period has been worked out empirically to be about 1,000. So, the MCTS roll-out policy is forced to stop after taking 1,000 steps. Note that there is a stay action defined for every fire location. Thus, in this implementation, it is possible to go a few levels down the tree without that corresponding to any real action in the world. Rewards are given based at each step of the roll-out and when complete the combined reward is used to update the value for the initial state at the root of the roll-out tree. For computational simplicity, during the roll-out a simpler state representation is used which focuses only on number of cells burning.

4.5. Deep RL

The final comparison is using the **Asynchronous Advantage Actor-Critic (A3C)** algorithm (Mnih et al., 2016), which represents the state-action value function using a Deep Q-Network (DQN) (Mnih et al., 2013). This algorithm defines a global network in addition to multiple worker agents with individual sets of parameters. The input to this global network was formalized as a grid of 100 × 100 cells with each cell having state values, which is an average of the state values of several pixels derived from satellite images. For our problem, A3C has the advantage of defining multiple worker agents. Each separate instance of a fire (unconnected to other fires) in a neighborhood is given its environment as an input, and the fire is defined as an individual worker. In our data, there are 96 instances of fire (thus, 96 worker agents) considered for training and testing. Each worker would then update the global environment, and we have plotted the result obtained. The deep network used is based on the DQN network given in Mnih et al. (2013), which uses an input layer of 100 × 100 pixel windows from the satellite image for the start date. Then there is a convolution layer of 16, 8 × 8 filters with a stride of 4 followed by a rectifier nonlinear unit (ReLU) layer. The second hidden layer involves 32 4 × 4 filters using a stride of 2 also followed by a ReLU layer. The final fully connected layer uses 256 ReLUs, which is output to a layer of threshold functions for the output of each of the five possible actions.

5. SUPERVISED CLASSIFICATION—GAUSSIAN PROCESS

In this supervised classification method of performing the experiments, the Gaussian process algorithm (Rasmussen and Williams, 2006) is used to estimate if a particular cell is burning or not based on the burning conditions of other cells in the scene. For each particular cell, the attributes of temperature, rain, relative humidity, land-cover type, wind speed, and wind direction are considered to influence the decision.

A distribution over function f can be specified by the Gaussian process as given by:

$$p(f) = GP(f \mid 0, k(x, x')),$$

where the mean function is 0, and the covariance is defined by some kernel function. We use the RBF kernel in our experiments because we have an assumption of 2D proximity of features being relevant in any direction.

The GP prior is a multivariate normal given by:

$$p(y) = Normal(y \mid 0, K),$$

where K is a covariance matrix given by evaluating $k(x_n, x_m)$ for each pair of inputs. The mean of the GP prior is assumed to be 0 in accordance with the above equations.

To obtain a probability the output response is condensed into the range [0, 1], which is an appropriate choice for classification. The probability is given by the condensed value, and a Bernoulli distribution is used to determine the label.

The likelihood of an observation (x_n, y_n) is given by:

$$p(y_n \mid z, x_n) = Bernoulli(y_n \mid logit^{-1}(z_n)).$$

Gaussian processes were chosen for this problem as they are known to work well for modeling non-linear relationships between variables especially in spatial domains. For the GP classifier, we use the logistic (sigmoid) function, whose integral is approximated for the binary case to map output to a probability. For the covariance function, we use an RBF kernel (stationary kernel) because we have an assumption of 2D proximity of features being relevant in any direction. The hyper-parameters of the kernel are optimized by maximizing the log-marginal-likelihood of the optimizer. A Cholesky decomposition was used to decompose the kernel matrix. The result is a binary output response, which can be used for classification.

6. EXPERIMENTAL SETUP

To evaluate this idea of agent-based RL for SSP prediction, we performed a cross-comparison of five RL algorithms and a supervised spatial GP approach using a two-stage training process. In the first stage, we learn the cell-based fire spread policy directly using the satellite images from the start of the fire and the image at its next successive time step. This is the policy from the MDP, which describes spread of fire from one cell to another based on

local conditions. The training data images are spaced 16 days apart, so many successive calls to this policy are required to reach the fire spread for 16 days. Thus, the second stage of training is to choose a cap *C* for the number of calls to this policy to obtain the correct spread area. We use another future satellite image to choose this cap based on maximum number of cells the policy got correct. The cap is used to set the upper limit of the total number of moves in any experiment.

In experiment (A), we test the ability of the algorithms to learn the spread of fire in an intermediate time step if the previous and next time step are known. Thus, we optimize *C* based on the ground truth data for time step 3. Now, the learned policy and the cap are used to determine the fire spread at the intermediate step 2 given that this is halfway between the state reached in step 3.

In experiment (B), we start with the initial state of the fire, we provide rewards based on the time step 2 of the fire, and we ask the algorithm to predict time step 3. This experiment is similar to asking the question: Where will the fire spread in the next 16 days, given its position currently? In this case, the cap value *C* is tuned for a 16-day duration.

In experiments (C)–(F), we apply the learned policy from the Richardson fire to the Fort McMurray fire over four 16-day time steps but tune the cap value *C* using the first transition in the Fort McMurray data. This was a fire that happened in Northern Alberta 5 years after the Richardson fire used for training. As the regions are similar and very near each other, the general properties that the model encapsulates should remain relevant. The initial state for the Fort McMurray fire is provided in a similar way to the Richardson fire.

After limiting the number of calls in any experiment based on the cap, we need a policy that determines if a fire should spread from one burning cell to a nearby non-burning cell. For the RL algorithms, this is done by applying a threshold to the value function to determine spread or non-spread. This threshold is determined for each algorithm to balance true positives and false positives based on the training data. For all the experiments, the result from the algorithms in the form of burned areas is compared against the actual scenario, using the satellite images corresponding to the predicted time step.

All experiments were run on an Intel core i7-6700 CPU with 32GB RAM.

7. ANALYSIS OF RESULTS

The task in these experiments is essentially to classify each cell at each time step correctly in terms of burning or not burning. So we can use the true positive rate and false positive rate to compute a receiver operating curve (ROC) and a corresponding area under the curve (AUC) metric. The AUC for experiment (A) is shown in **Table 1** for all algorithms and shows that the A3C algorithm provides an excellent threshold while classic value iteration is the least favorable. This seems to augur well with the overall accuracy for experiment (A) seen in **Table 2**.

Figures 4 and **5** show the visualization of some of the results for the different experiments. For all the images, the red pixels correspond where fire was classified correctly, blue pixels represent those which were classified as burning but were not yet at that point. White pixels represent false negatives where the policy predicted no fire but fire was indeed present. Black pixels represent true negatives, the pixels that were not on fire and correctly classified.

TABLE 1 | AUC for all the methods.

Method	AUC
VI	0.6806
PI	0.8983
QL	0.7138
MCTS	0.8256
A3C	0.9294

TABLE 2 | Average accuracy for each algorithm on the different test scenarios.

Method	Rich. fire (A)	Rich. fire (B)
GP	62.4%	50.8%
VI	72.2%	25.4%
PI	73.3%	38.2%
QL	67.2%	10.4%
MCTS	61.3%	**60.2**%
A3C	**87.3**%	53.2%

Bold font indicates the best performing algorithms in the corresponding experiments.

From all the algorithms considered, A3C seems to be the best overall, which is not surprising as it has the most flexible state-action value function representation. Policy Iteration in general comes second best in most tests. Q-learning a model free method of learning gives a lesser accuracy than model-based approaches like Value and Policy Iteration. In a spatial domain, the environment has a high influence on the spread of the agent. Thus, the strong consideration of the model of the spatial environment helps the model-based approaches.

The output images in **Figures 5G,H** show the two most successful algorithms MCTS and A3C on the region on which they were trained, filling in fire predictions between the start state and the reward target. Looking at the output images in **Figures 4E,F**, we see these two algorithms applied to the Fort McMurray fire (similar region, different year, and not part of the training data) at the same number of days forward after the start state. We can see that the model trained on one set images applies quite well to a different start state it has never seen, but A3C has much fewer false positives than MCTS. **Figure 5** visualizes the results of all the RL algorithms on predicting the next time step after the training data on the Richardson fire, here A3C clearly produces the most accurate prediction.

Somewhat surprisingly, MCTS outperforms A3C in experiment (B) for forward prediction of the next 16-day period on the same location as the training data. It seems that the MCTS rollouts have fit a better model to the sense of real world "time" to predict the extent of a fire in the next time period. As can be seen for the results for (C) and (D) in **Table 3**, the A3C algorithm outperforms all others on prediction when the learned policy for one location is applied to another location in the same region. MCTS does significantly worse in this challenge, meaning its model is not as good at generalizing the essentials of the policy for transfer to a new location even though it seems to have a better model of dynamics.

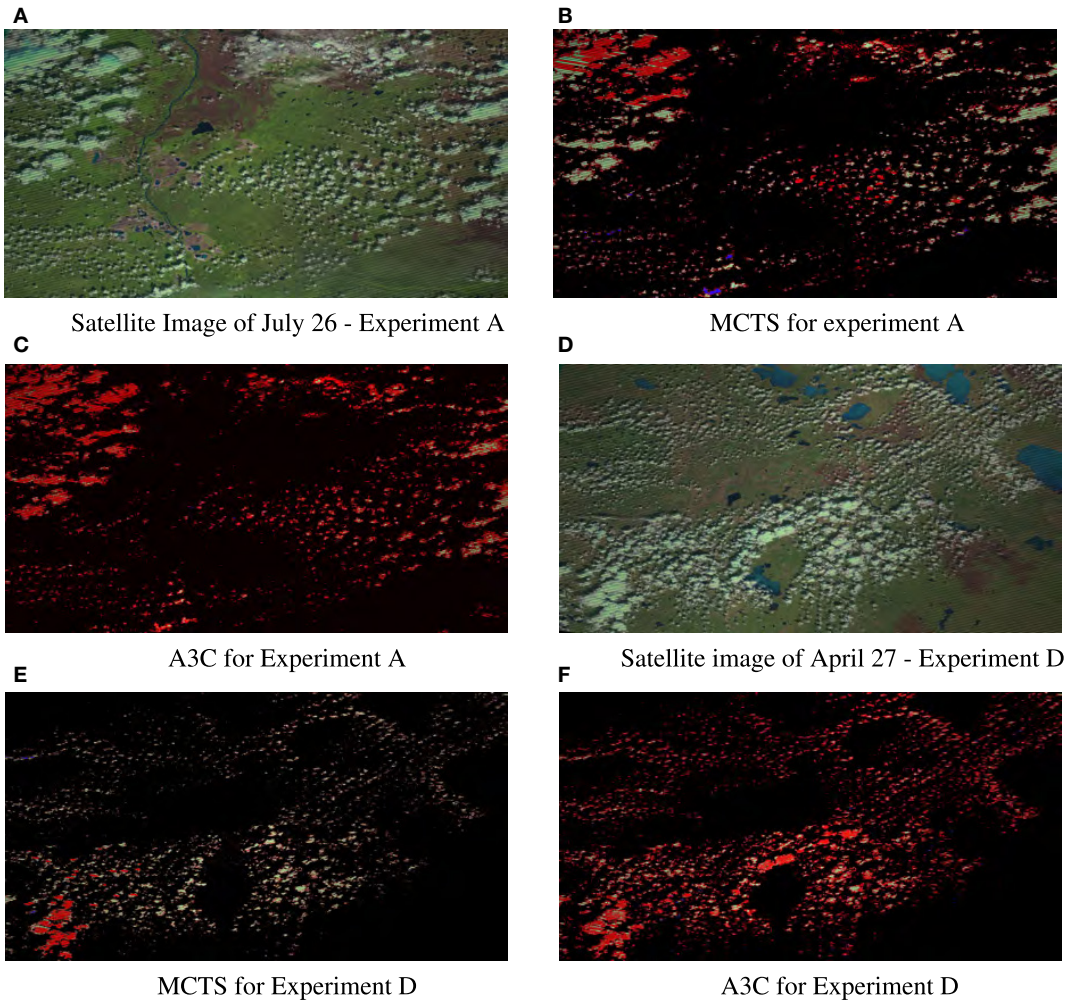


FIGURE 4 | Results for experiments from the best two algorithms MCTS and A3C for filling in an (intermediary state) in the training region (A) and applying that policy in a different region (D). Red pixels were on fire and classified correctly (true positives), blue pixels were incorrectly classified as burning (FP), white pixels were incorrectly classified as not burning (FN), and black pixels were correctly classified as not burning (TN). **(A)** Satellite image of July 26—experiment (A). **(B)** MCTS for experiment (A). **(C)** A3C for experiment (A). **(D)** Satellite image of April 27—experiment (D). **(E)** MCTS for experiment (D). **(F)** A3C for experiment D. Images obtained from USGS/NASA Landsat Program.

Note that the accuracy of all algorithms reduces with increasing time after the start of a fire as can be see by looking at experiments (C)–(F) in **Table 3**. This is unsurprising, as predicting multiple steps into the future is inherently harder. However, it is also partially due to the fact that later fires are larger, more intense, and spread faster, making learning more difficult. In the first time step, all algorithms do very well as it is easier to predict the first few moves of the fire. The accuracy goes down rapidly, and in the last time step it is only around 10%.

It is also interesting to note that even though MCTS starts of with a low accuracy compared with the other methods in the first time step, its cautious approach causes its accuracy to decrease much slower than other methods so that it seems tied with A3C when predicting three steps ahead. Similarly, the supervised GP method that loses out in most comparisons does far better for the hardest problem of predicting four time steps (about 2 months) ahead after all other methods have degraded. The GP approach is purely spatial pattern learning, with no element of dynamics. So one possible explanation is that, GP's with only a spatial model and no notion of time loses out when changes over time are relevant, but once we are so far ahead after the start of the fire a notion of time actually proves to be detrimental as the behavior of a very intense fire is difficult to correlate with time and the GP's more accurate spatial model wins out.

Turning our attention to the running times, we ran all experiments on similar scale problems running with the same system resources and training took between 5 and 7.5 h for all algorithms. Q-learning was the fastest to finish execution and obtained reasonably results even when even training time was subjected to a threshold cutoff. So Q-learning is a good candidate algorithm for fast approximations for testing purposes. One unique aspect of A3C is that it can exploit multithreading of CPU cores rather than GPU acceleration. Each fire in our dataset could be run as its own thread allowing A3C use similar training time to obtain

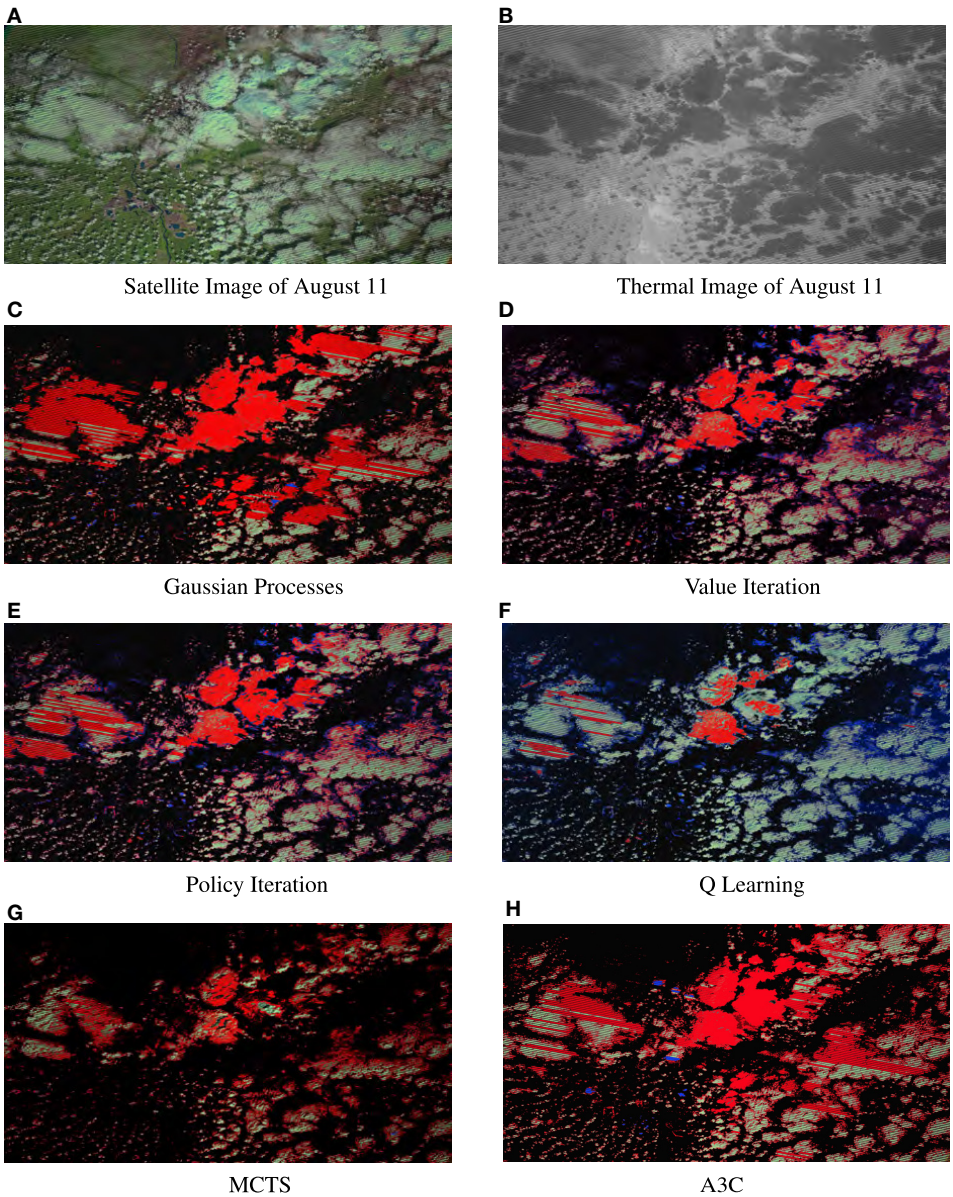


FIGURE 5 | Results for experiment (B) from all algorithms showing performance on the prediction of the next state directly after the training data. **(A)** Satellite image of August 11. **(B)** Thermal image of August 11. **(C)** Gaussian processes. **(D)** Value iteration. **(E)** Policy iteration. **(F)** Q-learning. **(G)** MCTS. **(H)** A3C. Images obtained from USGS/NASA Landsat Program.

superior results. MCTS was the slowest algorithm we tried since it is not multithreaded and requires extra roll-out simulations and back propagation at every iteration.

There are several reasons why the best RL algorithms are more suited to such domains than supervised learning algorithms. The first reason is that, RL can model the spatial dynamics along with time in such domains. This enables RL to predict action choices using a policy tuned to a particular time of fire spread test as compared with supervised learning which estimates a model based on inputs and outputs only. The second reason being RL prepares a policy for the agent that takes actions which model the underlying causal fire behavior. The supervised learning algorithms do not have such a state-action mapping. Thus this RL approach should be able to learn a reasonable policy in data-scarce scenarios by focusing on the reachable state-action space only.

8. CHALLENGES AND FUTURE WORK

As expounded in Malarz et al. (2002), forest fire prediction requires additional information consisting of firefighting intervention (such as fire fighting strategy and time elapsed), which are not taken into consideration in this study, as we chose a study region having very minimal fire fighting. In future work, we aim to incorporate this kind of information as well as enriching the model by including more land characteristics such as moisture,

TABLE 3 | Average accuracy of each algorithm trained on the Richardson fire but applied on the Fort McMurray fire for different time durations.

Method	(C)	(D)	(E)	(F)
GP	60.5%	47.9%	45.3%	**20.5**%
V.I	88.5%	68.4%	30.1%	6.4%
P.I	89.3%	67.8%	35.8%	8.9%
Q.L	84.2%	61.4%	26.4%	5.3%
MCTS	65.3%	55.7%	49.7%	5.8%
A3C	**90.1**%	**81.8**%	**50.8**%	13.4%

They are the best performing algorithms in the corresponding experiments.

slope, and directional aspect as state variables in individual cells. We will also perform a wider comparison against different existing wildfire models algorithms such as those in Castelli et al. (2015).

We also plan to investigate improvements to the structure of the Deep Neural Network policy representation to tailor it more closely to this kind of spatially spreading problem. For example, the relatively better behavior of the supervised GP approach on distance predictions where the dynamics approximation may be hindrance, suggest trying a hybrid approaches, learning a pair of models, one temporal and atemporal to achieve the best of both.

A severe challenge in the wildfire domain is the relative paucity of data compared with some other image analysis problems. While there are vast databases of satellite imagery, they are not all readily available and find the very small proportion which contains forest wildfires is not straightforward. This study has looked at a pair of fires in a well known region but future work will use data from more locales and automate the process of data collection. We will also look at flooding as a similar but more data-rich domains. This limited amount of image data is one reason, we chose to use a state feature extraction approach rather than learning on the entire image directly. There simply may not be enough images to learn effectively. However, a filter-based approach such as Convolutional Neural Networks (CNNs) applied directly to the images would be possible if we use small filters on the same scale as the local neighborhood we used in this study. We are currently trying this approach using a combination of CNNs and Recurrent Neural Networks to better encapsulate the effect of time on the fire.

9. CONCLUSION

In this work, we presented a novel approach for utilizing RL for learning forest wildfire spread dynamics directly from readily available satellite images. Our approach inverts the usual RL setup so that the dynamics of the MDP is a simple function of the fire spread actions being explored while the agent policy is a learned model of the dynamics of a complex spatially spreading process. Our results indicate that A3C is better at predicting spread dynamics at intermediate time steps and MCTS performs better while predicting the future spread. As the test data diverge from the training data and temporal changes become less relevant, such as when the intensity of fire increases, the fully supervised Gaussian process approach performs better than the RL algorithms.

The intersection between the decision-making tools of Artificial Intelligence, the pattern recognition tools of machine learning, and the challenging datasets of sustainability domains offers a rich area for research. For the machine learning community, our approach opens up new set of challenging and plentiful datasets for learning patterns of spatial change over time in the form of spatially spreading wildfires and a platform for experimenting with new Deep RL approaches on a challenging problem with high social impact.

We hope this can lead to development of a comprehensive way of integrating deep learning and RL approaches to support the tasks of prediction, dynamics model learning, and decision making in problems with SSP structure. This would also involve new representations for spatial data and policies, which can benefit theoretical as well as applied practitioners. Finally, the algorithmic approach demonstrated here could lead to more effective modeling and decision-making tools for domain practitioners in forest wildfire management which we are exploring with collaborators.

AUTHOR CONTRIBUTIONS

SS was responsible for primary design of experiments, collection of data and results, implementation of the reinforcement learning algorithms, and writing. MC carried out high level design of the algorithms, experimental design, and writing.

FUNDING

This project has been funded by internal funds provided for research by the Electrical and Computer Engineering Department at the University of Waterloo. No external granting agency funding was used for this project. This project was also partially funded by a Mitacs Globalink Graduate Fellowship Award.

REFERENCES

Alkhatib, A. A. (2014). A review on forest fire detection techniques. *Int. J. Distrib. Sens. Netw.* 10, 597368–597380. doi:10.1155/2014/597368

Angayarkkani, K., and Radhakrishnan, N. (2009). Efficient forest fire detection system: a spatial data mining and image processing based approach. *Int. J. Comput. Sci. Netw. Secur.* 9, 100–107.

Browne, C. B., Powley, E., Whitehouse, D., Lucas, S. M., Cowling, P. I., Rohlfshagen, P., et al. (2012). A survey of Monte Carlo tree search methods. *IEEE Trans. Comput. Intell. AI Games* 4, 1–43. doi:10.1109/TCIAIG.2012.2186810

Castelli, M., Vanneschi, L., and Popovič, A. (2015). Predicting burned areas of forest fires: an artificial intelligence approach. *Fire Ecol.* 11, 106–118. doi:10.4996/fireecology.1101106

Cortez, P., and Morais, A. (2007). "A data mining approach to predict forest fires using meteorological data," in *Volume Proceedings of the 13th Portuguese Conference on Artificial Intelligence*, Guimarães, Portugal.

Cracknell, M. J., and Reading, A. M. (2014). Geological mapping using remote sensing data: a comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* 63, 22–33. doi:10.1016/j.cageo.2013.10.008

Finney, M. A. (1998). *Farsite: Fire Area Simulator-Model Development and Evaluation*. Technical Report RMRS-4. Missoula, MT: USDA Forest Service, Rocky Mountain Research Station.

Finney, M. A., Cohen, J. D., McAllister, S. S., and Jolly, W. M. (2013). On the need for a theory of wildland fire spread. *Int. J. Wildland Fire* 22, 25–36. doi:10.1071/WF11117

Forsell, N., Garcia, F., and Sabbadin, R. (2009). "Reinforcement learning for spatial processes," in *18th World IMACS/MODSIM Congress*, Cairns, Australia, 755–761.

Garzón, M. B., Blazek, R., Neteler, M., De Dios, R. S., Ollero, H. S., and Furlanello, C. (2006). Predicting habitat suitability with machine learning models: the potential area of *Pinus sylvestris* l. In the Iberian Peninsula. *Ecol. Model.* 197, 383–393. doi:10.1016/j.ecolmodel.2006.03.015

Houtman, R. M., Montgomery, C. A., Gagnon, A. R., Calkin, D. E., Dietterich, T. G., McGregor, S., et al. (2013). Allowing a wildfire to burn: estimating the effect on future fire suppression costs. *Int. J. Wildland Fire* 22, 871–882. doi:10.1071/WF12157

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., and Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science* 353, 790–794. doi:10.1126/science.aaf7894

Kanevski, M., Pozdnukhov, A., and Timonin, V. (2008). "Machine learning algorithms for geospatial data applications and software tools," in *4th International Congress on Environmental Modelling and Software*, Barcelona, Spain.

Knudby, A., LeDrew, E., and Brenning, A. (2010). Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using Ikonos imagery and machine-learning techniques. *Remote Sens. Environ.* 114, 1230–1241. doi:10.1016/j.rse.2010.01.007

Kocsis, L., and Szepesvári, C. (2006). "Bandit based Monte-Carlo planning," in *European Conference on Machine Learning* (Berlin, Heidelberg: Springer), 282–293.

Kubat, M., Holte, R. C., and Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Mach. Learn.* 30, 195–215. doi:10.1023/A:1007452223027

Li, J., Heap, A. D., Potter, A., and Daniell, J. J. (2011). Application of machine learning methods to spatial interpolation of environmental variables. *Environ. Model. Software* 26, 1647–1659. doi:10.1016/j.envsoft.2011.07.004

Malarz, K., Kaczanowska, S., and Kuakowski, K. (2002). Are forest fires predictable? *Int. J. Mod. Phys. C* 13, 1017–1031. doi:10.1142/S0129183102003760

Martell, D. L. (2015). A review of recent forest and wildland fire management decision support systems research. *Curr. For. Rep.* 1, 128–137. doi:10.1007/s40725-015-0011-y

Mcgregor, S., Houtman, R., Buckingham, H., Montgomery, C., Metoyer, R., and Dietterich, T. G. (2016). "Fast simulation for computational sustainability sequential decision making problems," in *Proceedings of the 4th International Conference on Computational Sustainability*, New York, 5–7.

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., et al. (2016). "Asynchronous methods for deep reinforcement learning," in *International Conference on Machine Learning*, New York, 1928–1937.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., et al. (2013). Playing Atari with Deep Reinforcement Learning. *CoRR*, abs/1312.5602.

Montgomery, C. (2014). "Chapter 13: fire: an agent and a consequence of land use change," in *The Oxford Handbook of Land Economics*, eds J. M. Duke and J. Wu (Oxford University Press), 281–301.

Peng, G., Li, J., Chen, Y., Norizan, A. P., and Tay, L. (2006). High-resolution surface relative humidity computation using Modis image in peninsular Malaysia. *Chin. Geogr. Sci.* 16, 260–264. doi:10.1007/s11769-006-0260-6

Rasmussen, C. E., and Williams, C. K. (2006). *Gaussian Processes for Machine Learning*, Vol. 1. Cambridge: MIT Press.

Recknagel, F. (2001). Applications of machine learning to ecological modelling. *Ecol. Model.* 146, 303–310. doi:10.1016/S0304-3800(01)00316-7

Saranya, N. N., and Hemalatha, M. (2012). "Integration of machine learning algorithm using spatial semi supervised classification in FWI data," in *2012 International Conference on Advances in Engineering, Science and Management (ICAESM)* (Nagapattinam: IEEE), 699–702.

Sehgal, V., Getoor, L., and Viechnicki, P. D. (2006). "Entity resolution in geospatial data integration," in *Proceedings of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems* (Arlington, Virginia: ACM), 83–90.

Sitanggang, I. S., and Ismail, M. H. (2011). Classification model for hotspot occurrences using a decision tree method. *Geomatics Nat. Hazards Risk* 2, 111–121. doi:10.1080/19475705.2011.565807

Stocks, B. J., Lynham, T., Lawson, B., Alexander, M., Wagner, C. V., McAlpine, R., et al. (1989). Canadian forest fire danger rating system: an overview. *For. Chron.* 65, 258–265. doi:10.5558/tfc65258-4

Subramanian, S. G., and Crowley, M. (2017). "Learning forest wildfire dynamics from satellite images using reinforcement learning," in *Conference on Reinforcement Learning and Decision Making* (Ann Arbor, MI).

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Watkins, C. (1989). *Learning from Delayed Rewards*. PhD thesis, King's College, Cambridge, UK.

Woo, A., and Tait, C. (2016). Up to 90,000 Evacuated from Fort McMurray. *The Globe and Mail*.

Yongzhong, Z., Feng, Z.-D., Tao, H., Liyu, W., Kegong, L., and Xin, D. (2004). "Simulating wildfire spreading processes in a spatially heterogeneous landscapes using an improved cellular automaton model," in *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, Vol. 5 (Anchorage, AK: IEEE), 3371–3374.

Zhang, J.-H., Yao, F.-M., Liu, C., Yang, L.-M., and Boken, V. K. (2011). Detection, emission estimation and risk prediction of forest fires in china using satellite sensors and simulation models in the past three decades – an overview. *Int. J. Environ. Res. Public Health* 8, 3156–3178. doi:10.3390/ijerph8083156

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 February 2018
doi: 10.3389/fpls.2018.00180

Shifts in Plant Community Assembly Processes across Growth Forms along a Habitat Severity Gradient: A Test of the Plant Functional Trait Approach

Jinshi Xu[1,2], *Yongfu Chai*[1,2], *Mao Wang*[1,3], *Han Dang*[1,2], *Yaoxin Guo*[1,2], *Yu Chen*[1], *Chenguang Zhang*[1], *Ting Li*[1,2], *Lixia Zhang*[1,2] *and Ming Yue*[1,2*]

[1] *Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, Northwest University, Xi'an, China,* [2] *School of Life Sciences, Northwest University, Xi'an, China,* [3] *College of Grassland and Environment Sciences, Xinjiang Agricultural University, Urumchi, China*

Species respond to changes in their environments. A core goal in ecology is to understand the process of plant community assembly in response to a changing climate. Examining the performance of functional traits and trait-based assembly patterns across species among different growth forms is a useful way to explore the assembly process. In this study, we constructed a habitat severity gradient including several environment factors along a 2300 m wide elevational range at Taibai Mountain, central China. Then we assessed the shift on functional trait values and community assembly patterns along this gradient across species among different growth forms. We found that (1) although habitat-severity values closely covaried with elevation in this study, an examined communities along a habitat severity gradient might reveal community dynamics and species responses under future climate change. (2) the occurrence of trait values along the habitat severity gradient across different growth forms were similar, whereas the assembly pattern of herbaceous species was inconsistent with the community and woody species. (3) the trait-trait relationships of herbaceous species were dissimilar to those of the community and woody species. These results suggest that (1) community would re-assemble along habitat severity gradient through environmental filtering, regardless of any growth forms and that (2) different growth forms' species exhibiting similar trait values' shift but different trait-trait relationship by different trait combinations.

Keywords: functional traits, habitat severity, community assembly, climate change, functional structure, trait-trait relationships, woody species, herbaceous species

OPEN ACCESS

Edited by:
Louis Kouadio, University of Southern Queensland, Australia

Reviewed by:
John Morgan, La Trobe University, Australia
Georgina Conti, Instituto Multidisciplinario de Biologia Vegetal, Argentina

***Correspondence:**
Ming Yue
yueming@nwu.edu.cn
orcid.org0000-0002-1601-7988

Specialty section:
This article was submitted to Plant Biophysics and Modeling, a section of the journal Frontiers in Plant Science

Received: *28 April 2017*
Accepted: *31 January 2018*
Published: *15 February 2018*

Citation:
Xu J, Chai Y, Wang M, Dang H, Guo Y, Chen Y, Zhang C, Li T, Zhang L and Yue M (2018) Shifts in Plant Community Assembly Processes across Growth Forms along a Habitat Severity Gradient: A Test of the Plant Functional Trait Approach. Front. Plant Sci. 9:180. doi: 10.3389/fpls.2018.00180

INTRODUCTION

The climate acts as a filter for the species pool on a regional scale (Southwood, 1988), as it shifts the interactions of plant species (Chapin et al., 1998), as well as the community assembly process. Plant functional traits and their value's distribution patterns have become proxies for examining the plant community assembly process (Grabherr et al., 1994; Díaz et al., 1999, 2007; Mcgill et al., 2006; Vittoz et al., 2008; Walther et al., 2009; Mason and de Bello, 2013; Yablon, 2013). Such approaches might

reveal species responses to climate change (Weiher and Keddy, 1995; Woodward and Cramer, 1996; Díaz and Cabido, 1997; Cornwell and Ackerly, 2009; Götzenberger et al., 2012; Mason et al., 2012; Spasojevic and Suding, 2012; May et al., 2013).

Plant functional traits are the outcome of a history of species adaptation (Southwood, 1988). They represent the primary strategy that plants utilize to adapt to a changing environment (Lamanna et al., 2014). Plants respond to variable habitat conditions by adjusting their metabolism (Pappas et al., 2016) and performance (Keddy, 1992). Meanwhile, habitats filter species according their particular combination of traits (Keddy, 1992). Thus, plant functional traits, especially "response traits" which are measurable characteristics of plants, are assumed to reflect a plant's response to changes in its habitat (McIntyre et al., 1999; Lavorel et al., 2007; Meng et al., 2009; Borchardt et al., 2013). Response traits also provide information on the physiological adaptations of vegetation to various environmental gradients (McIntyre et al., 1999; de Bello et al., 2005; May et al., 2013; Purcell, 2016). These processes are believed to shape the range of functional trait values within communities (Cornwell and Ackerly, 2009) through habitat filtering or interspecific competition. Convergence of a trait value suggests co-occurring species often appeared in similar abiotic conditions, leading to habitat filtering (Grime, 2006; Cornwell and Ackerly, 2009). In contrast, interspecific competition is expected to exclude species with high trait similarity, resulting in trait divergence (MacArthur and Wilson, 1967; Weiher and Keddy, 1998; Stubbs and Wilson, 2004; Kraft et al., 2009). These functional traits distribution patterns can be described by using standardized effect sizes (SES) of traits. The SES of traits observed value to null expectation value (Kraft et al., 2009; Kraft and Ackerly, 2010).

As communities' functional trait values would change under the assembly process, species' trait value would change within community. It is necessary to examine the relationship between functional trait values and the associated biotic and abiotic conditions where the plant community is established to predict assembly patterns with climate change. Plant functional traits are assumed to be adaptively differentiated with habitats differing in some key factors (e.g., disturbance) or resource availability (Brouillette et al., 2014). Previous studies have focused on the trait-habitat relationships along various gradients in different regions and scales (Díaz and Cabido, 1997; Fonseca et al., 2000; Westoby and Wright, 2002; Wright et al., 2004; Fynn and Kirkman, 2005; Lambrecht and Dawson, 2007; Cornwell and Ackerly, 2009; Qi et al., 2009; Maharjan et al., 2011; Violle et al., 2011; Lawson and Weir, 2014). However, since plant functional traits vary among life forms (e.g., woody vs. herbaceous plants in one community, Meng et al., 2015), the concept of plant growth form is important for community dynamics (Meng et al., 2015). Nevertheless, previous studies have largely focused on woody species or particular species in forest communities (Cunningham et al., 1999; Fonseca et al., 2000; Qi et al., 2009), rather than herbaceous species (Oyarzabal et al., 2008). Studies on functional trait values might reveal whether woody or herbaceous species respond differently to environmental gradients (Yablon, 2013). Moreover, analyzing the assembly process in woody/herbaceous level and in community level separately is a useful way to detect the actual community assembly mechanisms comprehensively. These studies may be helpful to reveal the most focused ecological question of plant community, while they are seldom to be involved.

Examining different growth forms' responses seems necessary to explore the community assembly which respond to changes in environmental gradients (Keddy, 1992; Westoby, 1998). However, there are few studies examining the response of functional trait/functional trait distribution to the combination of these environment factors (Weiher and Keddy, 1995; de Bello et al., 2006) or plant functional traits (Smith and Wilson, 1994). We developed a habitat severity matrix representing the level of environmental stress in the habitat. We measured functional traits (leaf morphologic traits, leaf chemometrical traits, plant height, and seed mass) of each species in the communities based on the intrinsic dimensions theory (Laughlin, 2014; Laughlin and Messier, 2015). Additionally, we assessed the functional traits values and their distributions changes along the environmental severity gradient across varying levels (community, woody species, and herbaceous species) at Taibai Mountain, central China.

We expected that (1) the community assembly pattern would show convergence in more stressful habitat, and (2) economic spectrum-related functional trait-trait relationships and functional trait distributions would demonstrate different patterns across different growth forms.

MATERIALS AND METHODS

Study Site

The study was carried on the Taibai Mountain Nature Reserve, central China, located on 33°59′45″N−34°05′12″N, 107°41′18″E−107°48′22″E. This temperate zone is expected to permit a wider range of trait combinations than the tropical zone (Lamanna et al., 2014) since there are more heterogenous micro niches. The elevations of the reserve region vary from 940 to 3,767 m. In the present study, we selected the core area of forest which elevations range from 1,140 to 3,481 m. Such a wide-range altitudinal gradient provides opportunities to study various responses of plant species to changes in habitat conditions (Cordell and Handley, 1999). The climate of the study site is dominated by continental monsoon, mean annual temperature (MAT) varies from 0.9 to 12.3°C related to elevation, and annual precipitation is 640–1,000 mm (Tang and Fang, 2004). Forest coverage is over 82%, with high species diversity. In our study region, 389 woody species potentially exist which are recorded in the literature; the vegetation distributes along zonal zone in Taibai Mountain (Xu et al., 2017). The zonal vegetation of the Taibai Mountain region is highly heterogeneous in the warm temperate-zone deciduous broad-leaved forest and coniferous forest resulting from differences in hydrothermal conditions along such a wide-range altitudinal gradient (Zhu, 1981).

Plot Setup and Sampling

Along the altitudinal gradient, we established 39 plots; the area of each plot is 20 × 30 m. We selected more than 3 plots per

200 m elevation range, which were set as far as possible in order to represent the whole study region integrally. These selected plots commonly have different topographic factors or species composition in order to maximize variation in the environmental factors (sampling maps see Appendix Figure 1). Main field work was conducted in July 2014.

We collected soil samples dug from the 10–20-cm layer below ground surface and litter at four corners of the plots. We collected soil samples within same date for making sure the weather condition's consistency of samples. Soil samples were sealed in plastic bags, and they were naturally dried with 3 weeks for experimental analysis. In addition, we also recorded the habitat information, including elevation, location (by HOLUX EZ-Tour GPS recorder, HOLUX Technology Inc.), topographic slope (by a specific compass), and canopy coverage.

All species within each plot were identified, the abundance and coverage of the species in each plot were documented. The abundance of woody species (DBH $>$ 8 cm) was counted as the number of stems. The abundance of herbaceous species was calculated by its relative coverage via assignment: We established eight 0.3 $\times$ 0.3 m quadrats within each plot for recording the abundance and coverage of each herbaceous species. Based on these data, we estimated the total abundance of herbaceous species within each plot (Marteinsdóttir and Eriksson, 2014). We recorded plant height of each woody species individual. For individuals $<$2 m height, height was measured using a tape, while individuals $>$ 2 m height were measured using a height indicator; thereafter, the maximum plant height of each woody species was determined. For herbaceous species, the tallest individual was chosen to measure plant height. Because leaves are most exposed to habitat conditions and the changes in their traits have been interpreted as adaptations to specific environments (Fahn and Cutler, 1992), we collected 18–20 fully expanded sun-exposed leaves at top of crown from various directions of each mature species individuals within each plot. For each species, we selected at least five individuals if possible. The leaf samples were preserved in moist filter paper until analysis (Pérez-Harguindeguy et al., 2013).

Traits Measuring

We measured eight functional traits. Apart from the highest plant height (Hmax) mentioned above, we measured leaf morphologic and chemometrical traits by using leaf samples we collected. Leaf area (LA) is a trait determined by gradients in available moisture and temperature (Meng et al., 2009), Specific leaf area (SLA) is key trait reflecting species resource acquisition strategies (Cheng et al., 2016). Leaf dry matter content (LDMC) is another basic important leaf morphologic traits linking to habitat conditions (Yan et al., 2012). These leaf morphologic traits were measured following standard methods (Cornelissen et al., 2003). We also measured leaf chemometrical traits. Leaf nitrogen content (LNC), leaf carbon content (LCC) were measured by elemental analyzer (EA3000, EuroVector Inc.), and the leaf carbon-nitrogen ratio (C:N) was calculated afterwards. We also obtained seed mass (SM) data by weighing seed specimen preserved in the specimen museum. Missing data on seed mass of some species were compensated by literature review or website information (http://data.kew.org/sid/sidsearch.html). It is fine to assess the community assembly patterns for using these traits, as it would minimize the number of traits but maximize the number of dimensions (Laughlin, 2014).

Although trait values are often weighted by its relative abundance (Garnier et al., 2004; Violle et al., 2007) which calculated by its abundance divided by the sum of total abundance numbers within a community. In our study, there were species belonging different growth forms, relative abundance may not reflect the information of their actual biomass. Here we weighted the trait values of each species by its importance value (IV). The importance value of each species in each community was calculated as the sum of its relative abundance, relative height, and relative coverage and then divided by 3 (Xu et al., 2017). Note that the importance value of each species was calculated separately for all species combined (community level), only canopy species (woody species level), and only undergrowth herbaceous species (herbaceous species level). We calculated the weighted mean trait values of community, woody species and herbaceous species of each plot afterwards (method see Díaz et al., 2007).

Environmental Factors

Traits are related to temperature, light, moisture availability, soil pH and nutrients. (Landolt et al., 2010). Here we developed a "habitat-severity value" as an agency of all these factors. We measured the soil pH by the pH indicator (PB-100, Sartorius Inc.). Soil fertility factors nitrate nitrogen content (NN), ammonium nitrogen content (AN), total nitrogen content (TN), and rapidly available phosphorus (RAP) were analyzed by discontinuous chemical analyzer (CleverChem 2000, Dechem-Tech Inc.). Soil water content (SWC) which calculated by the ratio of water mass to total soil mass was measured as an indicator of moisture availability. We used woody species canopy coverage degree (WCD) as habitat cover. Furthermore, topographical slope was measured in our study, as it showed correlation with species leaf traits and height (Ackerly and Cornwell, 2007). MAT and air relative humidity (RH) were calculated by empirical equations adapted to Taibai Mountain (Tang and Fang, 2004):

$$\mathrm{MAT} = -0.0049 \times \mathrm{ALT} + 17.9 \ (r^2 = 0.99,\ P < 0.001) \quad (1)$$

$$\mathrm{RH} = 0.4 \times 10^{-6} \times (\mathrm{ALT})^2 - 0.0153 \times \mathrm{ALT} + 83.7 \ (r^2 = 0.95,\ P < 0.01) \quad (2)$$

In the equations, ALT indicates elevation of plots.

Constructing Habitat Severity Matrix

As there are numerous traits, we used principal component analysis (PCA) of all these traits for dimensionality reduction (Shipley, 2015). The PC1 axis of community weighted mean trait values captured 55.56% of the total variance of traits, while the PC1 axes of woody species and herbaceous species weighted mean trait values captured 52.09 and 36.89%, respectively. The PC1 axes of community, woody species and herbaceous species seemed to have statistical correlation with almost all the traits (Appendix Table 1). We used the PC1 scores of three levels as the

trait variables to perform variance decomposition. We input PC1 scores as dependent variables, while input environmental factors as independent variables for variance decomposition, which was performed in "hier.part" package of R 3.1.1 program. Variance decomposition (Pappas et al., 2016) was used to identify the relative roles in variance of traits in community level and in woody or herbaceous species level. We obtained the relative roles of factors in every habitat for trait variations in community, woody species and herbaceous species level respectively. The results of variance decomposition were in form as percentage (Appendix Table 2).

Thereafter, we constructed the habitat-severity values for above three levels. All the habitat factors were normalized to [0, 1], the more stressful for plant growth, the value is more nearby 1. Therefore, we used these scores as the habitat-severity values, in other words, as habitat severity gradient.

Data Analysis

We performed linear regression between weighted mean traits values and habitat-severity values in community, woody species and herbaceous species levels. In addition, we assessed the functional trait distribution of each plot in community, woody species and herbaceous species level as well by using SES of traits. The SES which describes functional trait distribution is the degree of discrepancy of trait observed value to null expectation value (Cornwell and Ackerly, 2009; Kraft et al., 2009; Kraft and Ackerly, 2010). Before calculating SES, mean functional distance (MFD) within a community should be calculated (Webb, 2000). MFD describes the mean difference between two species.

$$SES = \frac{MFD_{observed} - MFD_{randomized}}{sdMFD_{randomized}}$$

In the equation, $MFD_{observed}$ is actual MFD which calculated by observed functional traits values; $MFD_{randomized}$ is calculated by null model approach which is run in R 3.1.1 program for 999 times; sd means standard deviation. If SES < 0, means functional structure convergence; in contrast, SES > 0 indicates functional structure divergence. Owing to convergence or divergence of functional structures mirrored community assembly processes, we could assess performances of woody or herbaceous species comparing to community total species along habitat severity gradient and understand more detailed assembly processes (**Figure 1**).

We also showed the trait-trait relationships among above levels in order to determine differentiation of trait-response among different levels. We tested the Pearson correlation coefficient of each trait in community, wood species, and herbaceous level (**Tables 2**, **3**) by using the weighted values, and then compared with each other so that we could find out patterns which were consistent or not among different components in common community. As we found our habitat-severity values was dominated by MAT and RH which calculated by elevation (Appendix Table 2, Appendix Table 3), we also compared the Pearson correlation coefficient of elevations/habitat-severity values to weighted mean functional trait values in above three levels for assessing the habitat severity matrix (**Table 1**).

RESULTS

Performances of Habitat-Severity Values

Habitat severity values (HV) were dominated by MAT and RH, and MAT and RH were calculated using the elevations of plots. The correlation analysis between HV and elevation confirmed the notion that other environmental factors might correlate with elevation, even across three different levels (Appendix Table 3). We found nearly equal correlations between HV and traits and elevation and traits, regardless of level (**Table 1**).

Functional Trait Distributions Vary along Habitat Severity Gradient: Different Patterns across Different Growth Forms

Functional trait distribution of community and woody species exhibited almost complete similarity (no significant difference by U-test). Without exception, the functional distributions of LA, SLA, Hmax, and SM across community and woody species levels showed convergent tendencies with higher habitat severity. For herbaceous species, although there were convergent tendencies on LA and SLA along habitat severity, the inflection points of divergence-convergence were inconsistent with community and woody species. The SES of herbaceous LNC showed a divergent tendency at higher habitat severity. The SES of other traits did not show significant correlation with the habitat severity gradient (**Figure 1**).

Functional Traits Vary along Habitat Severity Gradient

Variation of traits among the community, woody species, and herbaceous species along habitat severity gradients displayed similar patterns to each other (**Figure 2**). For Hmax and SM, patterns among different levels were almost consistent; Hmax and SM decreased along the habitat severity gradient, though the slope of fit lines in herbaceous species were relatively flat. Chemometrical traits among the community, woody species, and herbaceous species exhibited interesting patterns; the LCC of these levels were consistent with each other (increasing with habitat severity), while LNC demonstrated a significant increase in only herbaceous species. Due to the varying performances of LNC, C:N showed similar patterns along habitat seventy. For leaf morphological traits, the LA of community, woody species, and herbaceous species exhibited decreased tendencies along habitat severity gradient. SLA and LDMC of woody species did not demonstrate significant relationships to habitat severity; however, community and herbaceous species exhibited increased patterns.

Trait-Trait Relationships Show Different Patterns across Different Growth Forms

The Pearson correlation semi-matrices of woody species and herbaceous species were compared with the semi-matrix of community (**Tables 2**, **3**). Trait-trait relationships of community and woody species were similar, except for very few trait-trait relationships, such as the relationships which involving Hmax. However, there were relatively different trait-trait patterns between the community and herbaceous species.

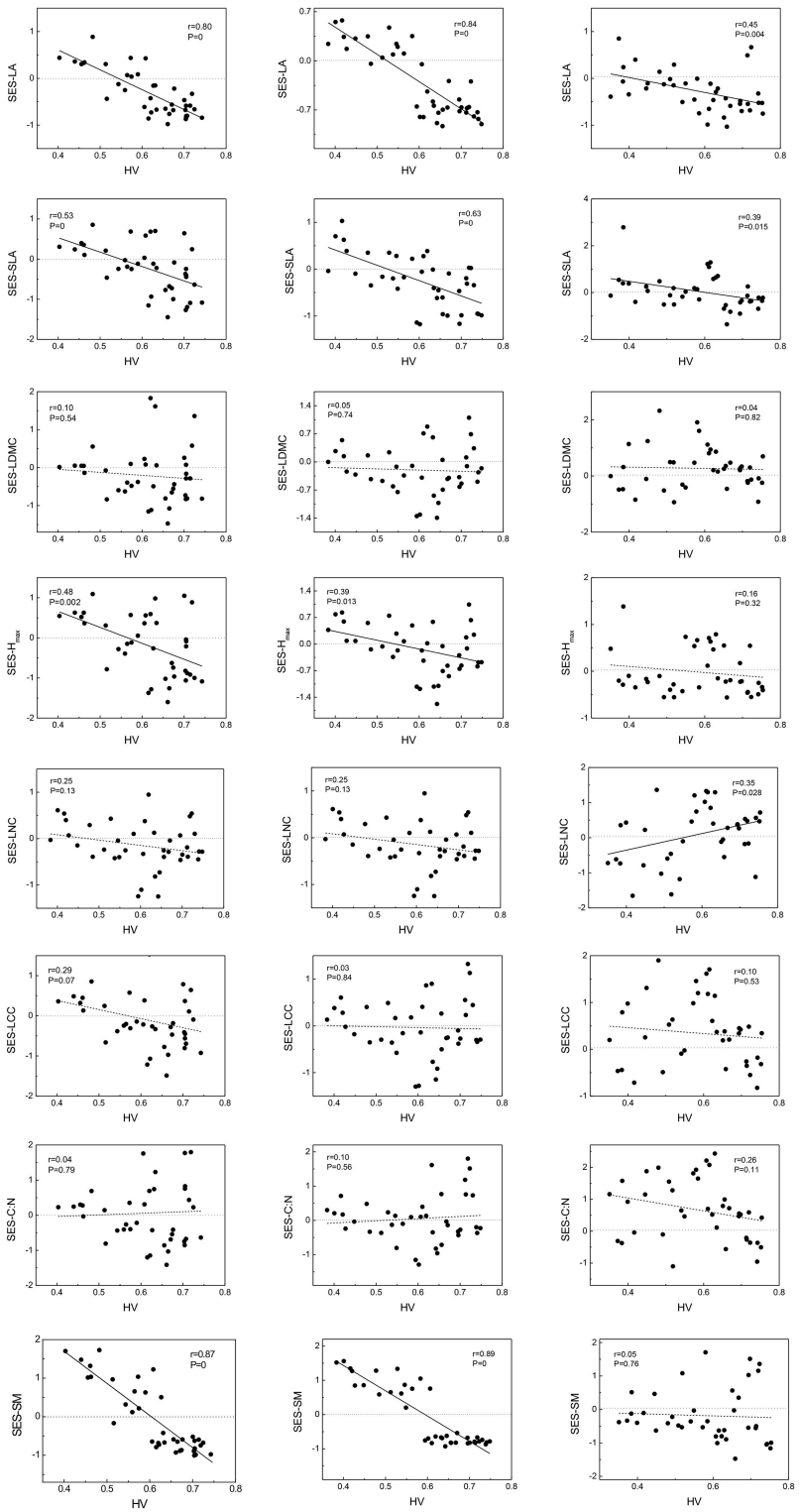


FIGURE 1 | Standardized effect sizes (SES) of traits vary along habitat-severity values (HV) gradient. Graphs on left column represent woody species level, graphs on right column represent herbaceous species level, and graphs on central column represent community level. Solid straight lines in graphs mean fit lines which $p < 0.05$; dash lines in graphs mean fit lines which $p > 0.05$. Dot lines locate on 0-level represent patterns of null expectation. LA, leaf area; SLA, specific leaf area; LDMC, leaf dry mass content; Hmax, the max plant height; LNC, leaf Nitrogen content; LCC, leaf Carbon content; C:N, leaf Carbon-Nitrogen ratio; SM, seed mass, same in below.

TABLE 1 | The correlation coefficient of mean weighted trait values and elevation/habitat-severity values (HV).

Trait	HV	Elevation	Trait	HV	Elevation	Trait	HV	Elevation
C-LA	**−0.797**	**−0.822**	W-LA	**−0.819**	**−0.839**	H-LA	**−0.393**	**−0.459**
C-SLA	**−0.513**	**−0.699**	W-SLA	**−0.294**	**−0.589**	H-SLA	**−0.675**	**−0.734**
C-LDMC	**0.464**	**0.534**	W-LDMC	**0.336**	**0.482**	H-LDMC	**0.479**	**0.429**
C-Hmax	**−0.692**	**−0.622**	W-Hmax	**−0.685**	**−0.597**	H-Hmax	**−0.360**	**−0.342**
C-LCC	**0.652**	**0.672**	W-LCC	**0.438**	**0.464**	H-LCC	**0.594**	**0.592**
C-LNC	−0.113	−0.143	W-LNC	−0.143	−0.287	H-LNC	**0.392**	**0.505**
C-C:N	**0.460**	**0.444**	W-C:N	**0.463**	**0.536**	H-C:N	−0.114	−0.206
C-SM	**0.888**	**−0.911**	W-SM	**−0.854**	**−0.901**	H-SM	**−0.333**	**-0.442**

C-, W-, and H- trait represents mean weighted trait values of community, woody species, and herbaceous species. The Pearson correlation coefficient which has statistical significance is highlighted in bold.

TABLE 2 | The trait-trait relationship semi-matrix of woody species.

	LA	SLA	LDMC	Hmax	LCC	LNC	C:N	SM
LA		**0.522**	**−0.553**	**0.633**	**−0.651**	0.135	**−0.426**	**0.811**
SLA	**0.454**		**−0.658**	**0.386**	**−0.479**	**0.475**	**−0.520**	**0.572**
LDMC	**−0.458**	**−0.653**		**−0.420**	**0.557**	**−0.355**	**0.490**	**−0.511**
Hmax	**0.629**	0.155	−0.225		**−0.632**	0.179	**−0.477**	**0.758**
LCC	**−0.497**	**−0.397**	**0.555**	−0.301		0.025	**0.321**	**−0.720**
LNC	0.191	**0.563**	**−0.355**	0.043	0.155		**−0.814**	0.100
C:N	**−0.466**	**−0.563**	**0.436**	−0.289	0.152	**−0.820**		**−0.490**
s SM	**0.841**	**0.455**	**−0.445**	**0.742**	**−0.466**	0.199	**−0.518**	

Mind that the semi-matrix at top right corner is the trait-trait relationship within community level. The Pearson correlation coefficient which has statistical significance is highlighted in bold.

Although the patterns of traits in community and herbaceous species levels along habitat severity gradient shifted in a similar manner, the relationships between traits each other of community was dissimilar to those relationships of herbaceous species.

DISCUSSION

Habitat Severity Gradient Applications in Plant Response to More Stressful Condition

The plant responses to environmental gradients seems to be a result of habitat severity, which links to the environmental filtering effect (Díaz et al., 1999). In the past, ecological studies elevations that varied in temperature, moisture availability, and other environment factors (Guittar et al., 2016) for community assembly processes. We developed habitat-severity values (HV) to assess the severity of habitat conditions. Our HV is strongly correlated with elevation (Appendix Table 3), and the correlations between HV and traits are similar to those between elevation and traits; this suggests that elevation might replace habitat severity in regions with a wide range of elevations but short geographic distances. Some studies revealed that the filtering process occurred at higher elevations (Pottier et al., 2012; Hulshof et al., 2013), similar to the pattern observed at high habitat severity in our study. However, even in regions without an obvious gradient, we suggest applying this framework to explore the shift in functional traits and their distributions. This framework could reflect plant response to more stressful conditions in changing world.

Assembly Processes Is Different across Different Growth Forms

The process of community assembly is related to environmental conditions. Generally speaking, plant traits tend to be clustered under stressful conditions, suggesting the occurrence of habitat filtering. Under less extreme environmental conditions, plant traits exhibit high differentiation due to interspecific competition (Kluge and Kessler, 2011). In our study, functional trait distributions of community and woody species were covariant in all examined traits along habitat severity gradient, whereas the functional trait distribution of herbaceous species showed dissimilar patterns. These results are consistent with other studies (Cornwell and Ackerly, 2009; Yablon, 2013) and suggest differing responses to habitat change across different growth forms. However, it should be noted that the robustness of HV for herbaceous plants is unclear, as the PC1 of herbaceous species' traits for structuring the HV only explained 35% of the total variance in traits. These results might be explained by environmental factors which were not measured in our study, such as light availability or herbivore

TABLE 3 | The trait-trait relationship semi-matrix of herbaceous species.

	LA	SLA	LDMC	Hmax	LCC	LNC	C:N	SM
LA		**0.522**	**−0.553**	**0.633**	**−0.651**	0.135	**−0.426**	**0.811**
SLA	**0.340**		**−0.658**	**0.386**	**−0.479**	**0.475**	**−0.520**	**0.572**
LDMC	**−0.375**	**−0.620**		**−0.420**	**0.557**	**−0.355**	**0.490**	**−0.511**
Hmax	−0.061	0.216	0.029		**−0.632**	0.179	**−0.477**	**0.758**
LCC	−0.259	**−0.520**	**0.542**	−0.113		0.025	**0.321**	**−0.720**
LNC	−0.024	−0.152	−0.297	**−0.388**	0.035		**−0.814**	0.100
C:N	−0.074	−0.090	**0.460**	**0.335**	**0.343**	**−0.874**		**−0.490**
SM	0.278	**0.435**	**−0.346**	0.112	−0.265	−0.117	−0.058	

Mind that the semi-matrix at top right corner is the trait-trait relationship within community level. The Pearson correlation coefficient which has statistical significance is highlighted in bold.

activity. These factors might have a stronger influence on herbaceous species than on woody species, potentially resulting in inconsistency compared to the community and woody species.

Community Assembly Pattern Changes with Habitat

The functional trait-based approach is a good approach for understanding plant community assembly patterns. Functional trait distributions change along the habitat severity gradient in this study and seems worthwhile to discuss. With increasingly stressful environmental conditions, the functional distributions of LA, SLA, Hmax, and SM of the community and woody species showed convergent patterns. In other words, abiotic filters restricted the range of viable strategies, thus creating a similar suite of traits (Woodward and Diament, 1991; Thuiller et al., 2004; de Bello et al., 2005; Fraser et al., 2016. These results suggest the filtering effect of habitat severity (Díaz et al., 1998; May et al., 2013). In fact, functional trait distributions will change along several environmental gradients which were included in our habitat severity matrix. For example, SLA is restricted with temperature (Joshi, 2013) and soil moisture (Cornwell and Ackerly, 2009); Hmax of shade-tolerant species is filtered by WCD (Cornwell and Ackerly, 2009). It would be interesting to discuss the convergent tendency of SM along the habitat severity gradient, as it seldom exhibited this pattern (Garnier and Navas, 2012). Due to the abundant species diversity in our study, highly-conserved traits—such as seed mass (Cavender-Bares et al., 2006)—would exhibit greater divergence at low to middle altitudinal regions than at higher elevations. As a result, seed mass demonstrated a convergent tendency along the habitat severity gradient, which was closely related to elevation. Therefore, the community assembly pattern would change as habitat changed.

Functional Traits Vary along Habitat Severity Gradient in Similar Pattern across Different Growth Forms

Our results demonstrated similar patterns of trait-habitat relationships among community, woody species, and herbaceous species, suggesting species within a community respond to environmental conditions similarly. Leaf traits, plant height, and seed mass correlated with environment conditions (Díaz et al., 1998; Cornelissen, 1999). Habitat filtering might cause these shifts in mean weighted trait values along environmental gradients (Ackerly and Cornwell, 2007; Cornwell and Ackerly, 2009; May et al., 2013). Species located in more arid sites had significantly lower LA, confirming a functional trade-off between stress tolerance and productivity in leaves (Thuiller et al., 2004). SLA is lower at regions with low rainfall and/or temperature, owing to thicker leaves and/or denser tissues (Tranquillini, 1964; Hadley and Smith, 1986; Westoby and Wright, 2002). Such habitats lead to higher LDMC and LCC in species as defense against stressful conditions (Oyarzabal et al., 2008). Plant height was also strongly correlated with stress (usually decreased, Lamanna et al., 2014), particularly when directly related to resource competition (e.g., temperature and light availability; Schwinning and Weiner, 1998). Smaller seed mass is often observed in species from more stressful conditions (Cornelissen, 1999; Luo et al., 2014). Such adaptations might allow plants to disperse to better environmental conditions (Körner, 1999). Our findings are consistent with the previously mentioned studies. In our study, higher habitat severity values mirrored lower temperature, water availability, and poorer soil conditions. At the community level, as habitat severity increased, community mean weighted trait values of LA, SLA, Hmax, and SM decreased, as well as findings in herbaceous' level. However, SLA and LDMC of woody species did not decrease along habitat severity gradient significantly; this phenomenon might be caused by the morphological differences between broad-leaved species and needle-leaved species, which often appear at middle to high elevations (middle to high HV in this study). We also observed that LNC of herbaceous species increased with the habitat severity gradient; this is consistent with a study in high-mountain grasslands, representing a greater investment in photosynthetic nitrogen (Díaz and Cabido, 1997).

It should be noted that we only focused on interspecific trait differences, which are the primary source of variability in trait values. The traits we selected for this study often show low phenotypic plasticity (especially for leaf stoichiometry, SLA, LDMC, and seed mass).

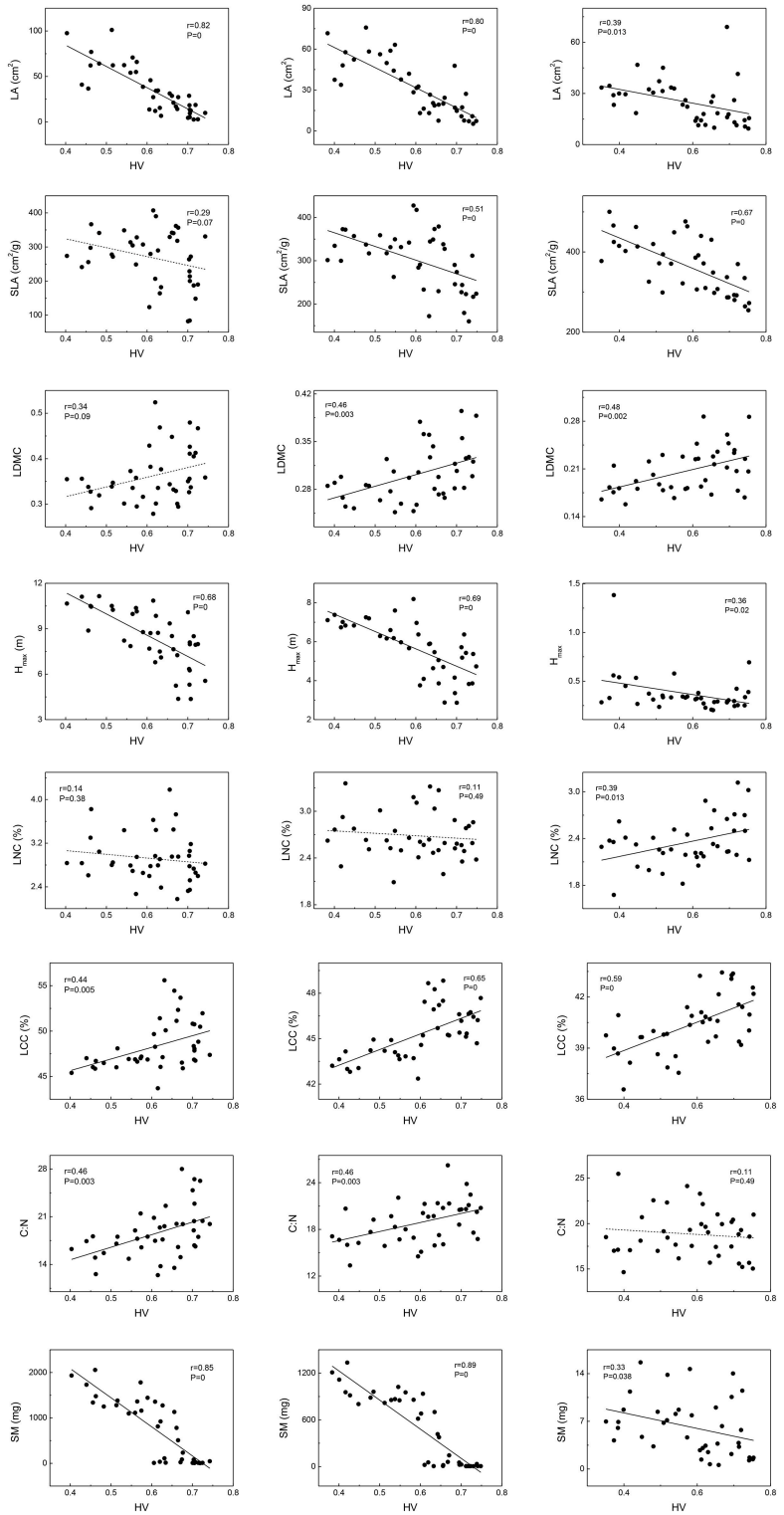


FIGURE 2 | Mean weighted trait values vary along habitat-severity values (HV) gradient. Graphs on left column represent woody species level, graphs on right column represent herbaceous species level, and graphs on central column represent community level. Solid straight lines in graphs mean fit lines which $p < 0.05$; dash lines in graphs mean fit lines which $p > 0.05$.

Trait-Trait Relationships of Herbaceous Species Are Unique Compared to Community

Analyses of intra-community trait-trait relationships can reveal the functional trade-offs operating along gradients of environmental stress. Dwyer and Laughlin (2017) proposed a novel insight that, with habitat becoming more stressful, trait covariance would become more significant. This phenomenon could be inferred from the patterns observed in functional distributions. In more stressful habitats, functional traits often show convergent patterns. Thus, two traits are more likely to exhibit significant covariance. For the community and woody species in our study, LA, SLA, Hmax, and SM showed significant tendencies to decrease along the habitat severity gradient (**Figure 1**). Combining the patterns of these traits along habitat severity (**Figure 2**), the more covariant relationship among these traits each other would be inferred. However, as herbaceous species did not demonstrate significant convergent patterns for most traits at more stressful habitats (**Figure 1**, right), we cannot infer the pattern that trait covariance might present along habitat severity. Our Pearson analysis results revealed the difference between the community and woody species and herbaceous species. Significant correlations among leaf morphological traits are not surprising. LA, SLA, and LDMC are directly related to each other. We found dissimilar patterns of trait-trait relationships between the community and herbaceous species. However, the trait-trait relationships between the community and woody species were very similar. One study explored the relationship between SM and plant height, concluding that plants with more biomass can afford to allocate more energy to seed development, thus producing larger seeds (Thuiller et al., 2004). These results illustrate that the SM of herbaceous species is not covariant with Hmax, suggesting that there are unique strategies of dispersal for herbaceous species compared to the community and woody species. Other trait-trait relations of herbaceous species comparing with community further highlighted the specificity of herbaceous species. Considering that (1) traits change similarly along a habitat stress gradient across growth forms, and (2) trait-trait relationships exhibit different patterns across growth forms, we conclude that the differing community assembly mechanisms between community and woody species and herbaceous species are likely caused by different trait combinations that are filtered by the environment.

AUTHOR CONTRIBUTIONS

JX and MY: Conceived and designed the experiments; YFC and MW: Analyzed the data; HD: Wrote the paper; YG: Conducted field work; YC, TL, and LZ: Performed the experiments and collected the data; CZ: Helped completing the manuscript.

ACKNOWLEDGMENTS

We are grateful to F. de Bello for the critical discussion. This study was funded by the National Natural Science Foundation of China (41571500, 31600337), Shaanxi Science and Technology Co-ordination and Innovation Project (2013KTZB03-01-03), and Shaanxi Scientific Research Project of Key Laboratory (13JS095).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.00180/full#supplementary-material

REFERENCES

Ackerly, D. D., and Cornwell, W. K. (2007). A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecol. Lett.* 10, 135–145. doi: 10.1111/j.1461-0248.2006.01006.x

Borchardt, P., Oldeland, J., Ponsens, J., and Schickhoff, U. (2013). Plant functional traits match grazing gradient and vegetation patterns on mountain pastures in sw Kyrgyzstan. *Phytocoenologia* 43, 171–181. doi: 10.1127/0340-269X/2013/0043-0542

Brouillette, L. C., Mason, C. M., Shirk, R. Y., and Donovan, L. A. (2014). Adaptive differentiation of traits related to resource use in a desert annual along a resource gradient. *New Phytol.* 201, 1316–1327. doi: 10.1111/nph.12628

Cavender-Bares, J., Keen, A., and Miles, B. (2006). Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* 87, S109–S122. doi: 10.1890/0012-9658(2006)87[109:PSOFPC]2.0.CO;2

Chapin, F. S. III., Sala, O. E., Burke, I. C., Grime, J. P., Hooper, D. U., Lauenroth, W. K., et al. (1998). Ecosystem consequences of changing biodiversity. *Bioscience* 48, 45–52. doi: 10.2307/1313227

Cheng, J., Chu, P., Chen, D., and Bai, Y. (2016). Functional correlations between specific leaf area and specific root length along a regional environmental gradient in inner Mongolia grasslands. *Funct. Ecol.* 658, 259–280. doi: 10.1111/1365-2435.12569

Cordell, S., and Handley, L. L. (1999). Allocation of nitrogen and carbon in leaves of metrosideros polymorpha regulates carboxylation capacity and δ13c along an altitudinal gradient. *Funct. Ecol.* 13, 811–818. doi: 10.1046/j.1365-2435.1999.00381.x

Cornelissen, J. H. (1999). A triangular relationship between leaf size and seed size among woody species: allometry, ontogeny, ecology and taxonomy. *Oecologia* 118, 248–255. doi: 10.1007/s004420050725

Cornelissen, J. H. C., Lavorel, S. B., Garnier, E. B., Díaz, S. M., Buchmann, N., Gurvich, D. E. C., et al. (2003). Handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Austr. J. Bot.* 51, 335–380. doi: 10.1071/BT02124

Cornwell, W., and Ackerly, D. (2009). Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecol. Monogr.* 79, 109–126. doi: 10.1890/07-1134.1

Cunningham, S. A., Summerhayes, B., and Westoby, M. (1999). Evolutionary divergences in leaf structure and chemistry, comparing rainfall and soil nutrient gradients. *Ecol. Monogr.* 69, 569–588. doi: 10.1890/0012-9615(1999)069[0569:EDILSA]2.0.CO;2

de Bello, F., Leps, J., and Sebastià, M. T. (2005). Predictive value of plant traits to grazing along a climatic gradient in the mediterranean. *J. Appl. Ecol.* 42, 824–833. doi: 10.1111/j.1365-2664.2005.01079.x

de Bello, F., Leps, J., and Sebastià, M. T. (2006). Variations in species and functional plant diversity along climatic and grazing gradients. *Ecography* 29, 801–810. doi: 10.1111/j.2006.0906-7590.04683.x

Díaz, S., and Cabido, M. (1997). Plant functional types and ecosystem function in relation to global change. *J. Veg. Sci.* 8, 463–474. doi: 10.1111/j.1654-1103.1997.tb00842.x

Díaz, S., Cabido, M., and Casanoves, F. (1998). Plant functional traits and environmental filters at a regional scale. *J. Veg. Sci.* 9, 113–122. doi: 10.2307/3237229

Díaz, S., Cabido, M., Zak, M., Carretero, E. M., and Araníbar, J. (1999). Plant functional traits, ecosystem structure and land-use history along a climatic gradient in central-western Argentina. *J. Veg. Sci*. 10, 651–660. doi: 10.2307/3237080

Díaz, S., Lavorel, S., de Bello, F., Quétier, F., Grigulis, K., and Robson, T. M. (2007). Incorporating plant functional diversity effects in ecosystem service assessments. *Proc. Natl. Acad. Sci. U.S.A*. 104, 20684–20689. doi: 10.1073/pnas.0704716104

Dwyer, J. M., and Laughlin, D. C. (2017). Constraints on trait combinations explain climatic drivers of biodiversity: the importance of trait covariance in community assembly. *Ecol. Lett*. 20, 872–888. doi: 10.1111/ele.12781

Fahn, A., and Cutler, D. F. (1992). Xerophytes Berlin: Gebruder Borntraeger.

Fonseca, C. R., Overton, J. M., Collins, B., and Westoby, M. (2000). Shifts in trait-combinations along rainfall and phosphorus gradients. *J. Ecol*. 88, 964–977. doi: 10.1046/j.1365-2745.2000.00506.x

Fraser, L. H., Garris, H. W., and Carlyle, C. N. (2016). Predicting plant trait similarity along environmental gradients. *Plant Ecol.* 217, 1297–1306. doi: 10.1007/s11258-016-0628-3

Fynn, R. W. S., and Kirkman, K. P. (2005). Plant strategies and trait trade-offs influence trends in competitive ability along gradients of soil fertility and disturbance. *J. Ecol*. 93, 384–394. doi: 10.1111/j.0022-0477.2005.00993.x

Garnier, E., Cortez, J., Billès, G., Navas, M. L., Roumet, C., and Debussche, M., et al. (2004). Plant functional markers capture ecosystem properties during secondary succession. *Ecology* 85, 2630–2637. doi: 10.1890/03-0799

Garnier, E., and Navas, M. L. (2012). A trait-based approach to comparative functional plant ecology: concepts, methods and applications for agroecology. A review. *Agron. Sustain. Dev*. 32, 365–399. doi: 10.1007/s13593-011-0036-y

Götzenberger, L., de Bello, F., Bråthen, K. A., Davison, J., Dubuis, A., and Guisan, A., et al. (2012). Ecological assembly rules in plant communities—approaches, patterns and prospects. *Biol. Rev*. 87, 111–127. doi: 10.1111/j.1469-185X.2011.00187.x

Grabherr, G., Gottfried, M., and Paull, H. (1994). Climate effects on mountain plants. *Nature* 369, 448–448. doi: 10.1038/369448a0

Grime, J. P. (2006). Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *J. Veg. Sci*. 17, 255–260. doi: 10.1111/j.1654-1103.2006.tb02444.x

Guittar, J., Goldberg, D., Klanderud, K., Telford, R. J., and Vandvik, V. (2016). Can trait patterns along gradients predict plant community responses to climate change? *Ecology* 97, 2791–2801. doi: 10.1002/ecy.1500

Hadley, J. L., and Smith, W. K. (1986). Wind effects on needles of timberline conifers: seasonal influence on mortality. *Ecology* 67, 12–19. doi: 10.2307/1938498

Hulshof, C. M., Violle, C., Spasojevic, M. J., McGill, B., Damschen, E., and Harrison, S., et al. (2013). Intra-specific and inter-specific variation in specific leaf area reveal the importance of abiotic and biotic drivers of species diversity across elevation and latitude. *J. Veg. Sci*. 24, 921–931. doi: 10.1111/jvs.12041

Joshi, C. (2013). *Inferential Tests and Modelling of Functional Trait Convergence along Environmental Gradients*. International Statistical Institute

Keddy, P. A. (1992). Assembly and response rules: two goals for predictive community ecology. *J. Veg. Sci*. 3, 157–164. doi: 10.2307/3235676

Kluge, J., and Kessler, M. (2011). Phylogenetic diversity, trait diversity and niches: species assembly of ferns along a tropical elevational gradient. *J. Biogeogr*. 38, 394–405. doi: 10.1111/j.1365-2699.2010.02433.x

Körner, C. (1999). "Alpine plants: stressed or adapted?," in *Physiological Plant Ecology*, eds M. C. Press, J. D. Scholes, and M. G. Barker (Oxford, UK: Blackwell), 297–312.

Kraft, N. J. B., and Ackerly, D. D. (2010). Functional trait and phylogenetic tests of community assembly across spatial scales in an amazonian forest. *Ecol. Monogr*. 80, 401–422. doi: 10.1890/09-1672.1

Kraft, N. J. B., Valencia, R., and Ackerly, D. D. (2009). Functional traits and niche-based tree community assembly in an amazonian forest. *Science.* 324, 580–582. doi: 10.1126/science.1169885

Lamanna, C., Blonder, B., Violle, C., Kraft, N. J. B., Sandel, B., Šímová, I., et al. (2014). Functional trait space and the latitudinal diversity gradient. *Proc. Natl. Acad. Sci. U.S.A*. 111, 13745–13750. doi: 10.1073/pnas.1317722111

Lambrecht, S. C., and Dawson, T. E. (2007). Correlated variation of floral and leaf traits along a moisture availability gradient. *Oecologia* 151, 574–583. doi: 10.1007/s00442-006-0617-7

Landolt, E., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F., Lämmler, W., et al. (2010). *Flora Indicativa - Ecological Inicator Values and Biological Attributes of the Flora of Switzerland and the Alps: Ökologische Zeigerwerte und Biologische Kennzeichen zur Flora der Schweiz und der Alpen*. Bern: Haupt.

Laughlin, D. C. (2014). The intrinsic dimensionality of plant traits and its relevance to community assembly. *J Ecol*. 102, 186–193. doi: 10.1111/1365-2745.12187

Laughlin, D. C., and Messier, J. (2015). Fitness of multidimensional phenotypes in dynamic adaptive landscapes. *Trends in Ecology & Evolution* 30: 487–496. doi: 10.1016/j.tree.2015.06.003

Lavorel, S., Díaz, S., and Cornelissen, J. H. C. (2007). "Plant functional types: are we getting any closer to the Holy Grail?," in *Terrestrial Ecosystems in a Changing World,* eds J. Canadell, L. F., and Pataki Deds (Berlin:Springer-Verlag), 149–164. doi: 10.1007/978-3-540-32730-1_13

Lawson, A. M., and Weir, J. T. (2014). Latitudinal gradients in climatic-niche evolution accelerate trait evolution at high latitudes. *Ecol. Lett*. 17, 1427–1436. doi: 10.1111/ele.12346

Luo, Y., Widmer, A., and Karrenberg, S. (2014). The roles of genetic drift and natural selection in quantitative trait divergence along an altitudinal gradient in *Arabidopsis thaliana*. *Heredity* 114, 220–228. doi: 10.1038/hdy.2014.89

MacArthur, R. H., and Wilson, E. O. (1967). *The Theory of Island Biography*. Princeton, NJ.

Maharjan, S. K., Poorter, L., Holmgren, M., Bongers, F., Wieringa, J. J., and Hawthorne, W. D. (2011). Plant functional traits and the distribution of west African rain forest trees along the rainfall gradient. *Biotropica*, 43, 552–561. doi: 10.1111/j.1744-7429.2010.00747.x

Marteinsdóttir, B., and Eriksson, O. (2014). Plant community assembly in semi-natural grasslands and ex-arable fields: a trait-based approach. *J. Veg. Sci*. 25, 77–87. doi: 10.1111/jvs.12058

Mason, N. W. H., and de Bello, F. (2013). Functional diversity: a tool for answering challenging ecological questions. *J. Vegetat. Sci.* 24, 777–780. doi: 10.1111/jvs.12097

Mason, N. W. H., Richardson, S. J., Peltzer, D. A., de Bello, F., Wardle, D. A., and Allen, R. B. (2012). Changes in coexistence mechanisms along a long-term soil chronosequence revealed by functional trait diversity. *J. Ecol.* 100, 678–689. doi: 10.1111/j.1365-2745.2012.01965.x

May, F., Giladi, I., Ristow, M., Ziv, Y., and Jeltsch, F. (2013). Plant functional traits and community assembly along interacting gradients of productivity and fragmentation. *Perspect. Plant EcolEvolut. Syst*. 15, 304–318. doi: 10.1016/j.ppees.2013.08.002

Mcgill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends Ecol. Evol*. 21, 178–185. doi: 10.1016/j.tree.2006.02.002

McIntyre, S., Lavorel, S., Landsberg, J., and Forbes, T. D. A. (1999). Disturbance response in vegetation – towards a global perspective on functional traits. *J. Veg. Sci*. 10, 621–630. doi: 10.2307/3237077

Meng, T. T., Ni, J., and Harrison, S. P. (2009). Plant morphometric traits and climate gradients in northern china: a meta-analysis using quadrat and flora data. *Ann. Bot.* 104, 1217–1229. doi: 10.1093/aob/mcp230

Meng, T. T., Wang, H., Harrison, S. P., Prentice, I. C., Ni, J., and Wang, G. (2015). Responses of leaf traits to climatic gradients: adaptive variation versus compositional shifts. *Biogeosciences*, 12 1–14. doi: 10.5194/bg-12-5339-2015

Oyarzabal, M., Paruelo, J. M., Pino, F. D., Oesterheld, M., and Lauenroth, W. K. (2008). Trait differences between grass species along a climatic gradient in south and north America. *J. Veg. Sci.* 19, 183–192. doi: 10.3170/2007-8-18349

Pappas, C., Fatichi, S., and Burlando, P. (2016). Modeling terrestrial carbon and water dynamics across climatic gradients: does plant trait diversity matter? *New Phytol.* 59, 137–151. doi: 10.1111/nph.13590

Pérez-Harguindeguy, N., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M. S., et al. (2013). New handbook for standardised measurement of plant functional traits worldwide. *Austr. J. Bot*. 61, 167–234. doi: 10.1071/BT12225

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C. F., et al. (2012). The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Glob. Ecol. Biogeogr*. 22, 52–63. doi: 10.1111/j.1466-8238.2012.00790.x

Purcell, A. (2016). *Functional Trait Variation along a Hydrological Gradient and Trait-Based Predictions of the Composition of a Wetland Plant Community. A Thesis Submitted in Partial Fulfillment of the Requirements for the degree of Masters of Science*, The University of Waikato.

Qi, J., Ma, K., and Zhang, Y. (2009). Leaf-trait relationships of quercus liaotungensis, along an altitudinal gradient in dongling mountain, beijing. *Ecol. Res*. 24, 1243–1250. doi: 10.1007/s11284-009-0608-3

Schwinning, S., and Weiner, J. (1998). Mechanisms determining the degree of size asymmetry in competition among plants. *Oecologia* 113, 447–455. doi: 10.1007/s004420050397

Shipley, B. (2015). Describing, explaining and predicting community assembly: a convincing trait-based case study. *J. Veg. Sci*. 26, 615–616. doi: 10.1111/jvs.12294

Smith, B., and Wilson, J. B. (1994). Vegetation texture as an approach to community structure: community-level convergence in a new Zealand temperate rainforest. *N. Zeal. J. Ecol.* 18, 41–50.

Southwood, T. R. E. (1988). Tactics, strategies and templets. *Oikos* 52, 3–18. doi: 10.2307/3565974

Spasojevic, M. J., and Suding, K. N. (2012). Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *J. Ecol*. 100, 652–661. doi: 10.1111/j.1365-2745.2011.01945.x

Stubbs,W. J., and Wilson, J. B. (2004). Evidence for limiting similarity in a sand dune community. *J. Ecol*. 92, 557–567. doi: 10.1111/j.0022-0477.2004.00898.x

Tang, Z. Y., and Fang, J. Y. (2004). Patterns of woody plant species diversity along environmental gradients on Mt.Taibai, Qinling Mountains. *Chin. Biodiversity*, 12, 115–122

Thuiller, W., Lavorel, S., Midgley, G., Lavergne, S., and Rebelo, T. (2004). Relating plant traits and species distributions along bioclimatic gradients for 88 leucadendron taxa. *Ecology*, 85, 1688–1699. doi: 10.1890/03-0148

Tranquillini, W. (1964). The physiology of plants at high altitudes. *Plant Biol*. 15, 345–362. doi: 10.1146/annurev.pp.15.060164.002021

Violle, C., Bonis, A., Plantegenest, M., Cudennec, C., Damgaard, C., Marion, B., et al. (2011). Plant functional traits capture species richness variations along a flooding gradient. *Oikos* 120, 389–398. doi: 10.1111/j.1600-0706.2010.18525.x

Violle, C., Lecoeur, J., and Navas, M. L. (2007). How relevant are instantaneous measurements for assessing resource depletion under plant cover? A test on light and soil water availability in 18 herbaceous communities. *Funct. Ecol*. 21, 185–190. doi: 10.1111/j.1365-2435.2006.01241.x

Vittoz, P., Bodin, J., Ungricht, S., Burga, C. A., and Walther, G. R. (2008). One century of vegetation change on Isla Persa, a nunatak in the Bernina massif in the Swiss Alps. *J. Vegetat. Sci.* 19, 671–680. doi: 10.3170/2008-8-18434

Walther, G. R., Beißner, S., and Burga, C. A. (2009). Trends in upward shift of alpine plants. *J. Veg. Sci.* 16, 541–548. doi: 10.1111/j.1654-1103.2005.tb02394.x

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am. Nat*. 156, 145–155. doi: 10.1086/303378

Weiher, E., and Keddy, P. A. (1995). Assembly rules, null models, and trait dispersion: new questions from old patterns. *Oikos* 74, 159–164. doi: 10.2307/3545686

Weiher, E., and Keddy, P. A. (1998). Community assembly rules, morphological dispersion, and the coexistence of plant species. *Oikos* 81, 309–322. doi: 10.2307/3547051

Westoby, M. (1998). A leaf-height-seed (lhs) plant ecology strategy scheme. *Plant Soil* 199, 213–227. doi: 10.1023/A:1004327224729

Westoby, M., and Wright, I. J. (2002). Plant ecological strategies: some leading dimensions of variation between species. *Ecol. Evol. Syst*. 33, 125–159. doi: 10.1146/annurev.ecolsys.33.010802.150452

Woodward, F. I., and Cramer, W. (1996). Plant functional types and climatic change: introduction. *J. Veg. Sci*. 7, 306–308. doi: 10.1111/j.1654-1103.1996.tb00489.x

Woodward, F. I., and Diament, A. D. (1991). Functional approaches to predicting the ecological effects of global change. *Funct. Ecol*. 5, 202–212. doi: 10.2307/2389258

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., et al. (2004). The worldwide leaf economics spectrum. *Nature* 428, 821–827. doi: 10.1038/nature02403

Xu, J., Chen, Y., Zhang, L., Chai, Y., Wang, M., Guo, Y., et al. (2017). Using phylogeny and functional traits for assessing community assembly along environmental gradients: a deterministic process driven by elevation. *Ecol. Evol*. 7, 5056–5069. doi: 10.1002/ece3.3068

Yablon, E. (2013). *Functional Traits, Environmental Gradients and Community Assembly in a Temperate Forest*. Senior Honors thesis, Washington University in St. Louis.

Yan, B., Zhang, J., Liu, Y., Li, Z., Huang, X., Yang, W., et al. (2012). Trait assembly of woody plants in communities across sub-alpine gradients: identifying the role of limiting similarity. *J. Veg. Sci*. 23, 698–708. doi: 10.1111/j.1654-1103.2011.01384.x

Zhu, Z. C. (1981). The regulation and characteristic of dominant type of forest in Taibai Mountain, Qinling Mountains. *Shaanxi Forest Sci. Technol.* 5, 29–39.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 03 August 2018
doi: 10.3389/fenvs.2018.00074

Landscape Based Agricultural Water Demand Modeling—A Tool for Water Management Decision Making in British Columbia, Canada

Denise Neilsen[1], Matthew Bakker[2], Ted Van der Gulik[3], Scott Smith[1], Alex Cannon[4], Istvan Losso[1] and Anna Warwick Sears[5]*

[1] Agriculture and Agri-Food Canada, Summerland, BC, Canada, [2] School of Resource and Environment Management, Simon Fraser University, Burnaby, BC, Canada, [3] Partnership for Water Sustainability, Abbotsford, BC, Canada, [4] Climate Research Division, Environment and Climate Change Canada, Victoria, BC, Canada, [5] Okanagan Basin Water Board, Kelowna, BC, Canada

OPEN ACCESS

Edited by:
Louis Kouadio, University of Southern Queensland, Australia

Reviewed by:
Nidhi Nagabhatla, United Nations University Institute for Water Environment and Health, Canada
Carmen De Jong, Université de Strasbourg, France

***Correspondence:**
Denise Neilsen
denise.neilsen@agr.gc.ca

Specialty section:
This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Environmental Science

Received: *02 March 2018*
Accepted: *20 June 2018*
Published: *03 August 2018*

Citation:
Neilsen D, Bakker M, Van der Gulik T, Smith S, Cannon A, Losso I and Warwick Sears A (2018) Landscape Based Agricultural Water Demand Modeling—A Tool for Water Management Decision Making in British Columbia, Canada. Front. Environ. Sci. 6:74. doi: 10.3389/fenvs.2018.00074

Regional water managers face challenges managing water demand and supply in response to climate change. In British Columbia (BC), <5% of the total land surface is suitable for cultivation and consequently, urban development and agriculture co-exist on lower elevation sites and compete for water. Surface and ground water supply all uses, affecting in-stream habitats and aquifer levels. To assess water needs, we used a GIS-based water demand model for agricultural water use, with layers for detailed land use, soils, a digital elevation model, sub-basins, aquifers, and socio-political jurisdictional boundaries. The model was driven by gridded daily minimum and maximum temperatures and precipitation at a 500 m spatial resolution using historical data (1961–2010) and downscaled climate scenarios (1961–2100) derived from five CMIP5 climate models under two greenhouse gas concentration scenarios (RCP4.5 and RCP8.5). Two case studies were examined:

1. Changes in water use were determined for scenarios of climate change and expanded cropping within the BC Agricultural Land Reserve for 17 agricultural regions. Potential increases in irrigation water demand (IWD) in response to climate change ranged from 21 to 58 and 30 to 114% under low and high emissions scenarios respectively. Land use change scenarios resulted in very large potential increases in water demand, up to 2,400%. Output from this work forms the basis for a web-based agricultural water license calculator.
2. Effects of crop production systems on IWD were examined in the Okanagan region. Combinations of pasture and forage crops with inefficient irrigation systems were most vulnerable to drought as indicated by two indices: relative vulnerability index (IWD/ETo) and allocation vulnerability index (IWD/Maximum annual water allocation). In drought years, up to 70% of the irrigated area was vulnerable. A comparison of detailed land use surveys made in 2006 and 2014 indicated a large shift to highly efficient irrigation systems in the horticultural sector from 38 to 68% of the irrigated

acreage. Similar shifts for other agricultural sectors may require financial support. On-going development of regional drought management will require collaborative decision making by water suppliers and users.

Keywords: irrigation water demand (IWD), climate change, agricultural land reserve, water management tool, land use

INTRODUCTION

Concerns about freshwater water availability and use have risen in recent years as regions have experienced ongoing droughts and water shortages. Notable droughts have occurred in California, northern China, Australia, and the Canadian prairies. However, water stress is not limited to these areas. It was estimated that 36% of the world population lived in water stressed areas in 2011 (Grebmer et al., 2012). These conditions are likely to persist, with the proportion of the world population living in water stressed areas predicted to increase to as much as 66% by 2025 (Kjellen and McGranahan, 1997). This trend is being driven by continued population growth, environmental degradation, and increasing overall water demand (Arnell, 1999).

The challenges around water availability and use will be further complicated by future climate change as variation in temperature and precipitation, decease the reliability of water resources (Jiménez Cisneros et al., 2014) and significantly affect water use (Wang et al., 2014). Broadly speaking, there is a consensus that some regions will receive greater precipitation while others less, and long-term seasonal, or short-term water shortages will likely result from greater variability of precipitation and reduced snow and ice storage, but how these changes will interact with demographic and economic factors is still unclear (Jiménez Cisneros et al., 2014).

Changes to water availability are particularly important to agriculture. Globally, agriculture is the largest user of water and that which is required to satisfy irrigation requirements, or irrigation water demand (IWD), accounts for ∼70% of global water use (Raskin et al., 1997), and 90% of global consumptive water use (Shiklomanov and Rodda, 2003). This demand is projected to increase in response to population growth and greater demand for agricultural production but at the same time, research suggests that agriculture will face a disproportionate level of water stress resulting from climate change compared to other water uses (Wang et al., 2014). As precipitation becomes more unreliable, rain-fed agricultural regions will become more vulnerable to poor production, increasing the need for irrigation. In general, irrigated systems are more efficient than rain-fed systems as ∼40% of the world's food production is estimated to come from the 17% of agricultural land that is irrigated (Postel, 2000). However, the poor quality of some irrigation water may lead to problems with salination both of soils and groundwater resources in some parts of the world.

The far-reaching implications of climate change for agriculture have led to a substantial body of research into future IWD. Most notably, modeling efforts that employ Global Climate Models (GCMs) and Earth Systems Models (ESMs) have contributed significantly to our understanding of future water demand. These studies examine a range of climate change scenarios that vary in projected global warming. Assessments at the global scale range from projections of a slight decrease of IWD (Zhang and Cai, 2013), to no change (Konzmann et al., 2013), to increases of 14–45% (Fischer et al., 2007; Wada et al., 2013) by 2080. While this variation in projections highlights substantial uncertainty at a global scale, the majority of scenarios predict that certain regions will experience >20% increase in IWD (Jiménez Cisneros et al., 2014). For example, there is high confidence that IWD will increase in Europe, USA, and parts of Asia, while other regions, such as India and Pakistan, will likely experience decreases in IWD.

The interregional variation seen in global IWD models is supported by studies at smaller spatial scales. Regional IWD models based on GCMs and ESMs have projected a wider range of future IWD. Annual IWD has been projected to increase by 9–24% by 2100 in the Arkansas River Basin (Elgaali et al., 2007), by 54–645% by 2095 in Bhadra, India (Rehana and Mujumdar, 2013), and by 2,000% by 2080 in North Rhine-Westphalia, Germany (Kreins et al., 2015), although current irrigation water use is very low in this last region. The interregional variability in projected IWD seen in global studies and supported by the projections of regional studies highlights the significant influence of localized climatic conditions on IWD. As a result, while global and regional models offer insights into future IWD, they are often of limited benefit to regional water management efforts outside the area of study.

The Province of British Columbia (BC) has an abundance of water resources but the amount of agricultural land (∼4.5%) and access to water are limited by the complex, mountainous terrain (Walker and Sydneysmith, 2008). Agricultural land in the province is protected by an Agricultural Land Reserve (ALR) which comprises 4.599 m ha of which 53% falls within agricultural capability class 1–4, considered suitable for crop production (https://www.alc.gov.bc.ca.) Agricultural water supply includes surface flows from rainfall, snowmelt, glacier melt and groundwater. Annual precipitation ranges from >3,000 mm in coastal regions to <300 mm in Interior rain-shadow valleys. Summers are dry and low summer flows restrict water availability for both withdrawals and ecological requirements. British Columbia has experienced increasing stress on its water resources in recent years with seasonal droughts in 2003, 2009, 2015, and 2017 resulting in extremely low late summer and fall stream flows in southern interior and coastal regions. Additionally, many municipalities implement watering restrictions for all outdoor uses in response to decreasing water availability and increasing demand. Water stress in BC is likely to worsen in response to climate change. Seasonal

mean air temperatures at most observing stations in B.C. have increased by at least +0.5 to +2.5°C since 1950 (Bush et al., 2014). Over B.C. as a whole, climate model projections indicate further increases of +1.0 to +3.8°C for the summer season and +1.1 to +3.9°C for the winter season by the middle of the twenty-first century (ECCC, 2016). Accompanying regional increases in air temperatures, changes in the intensity and seasonal distribution of precipitation are also projected to occur, although there is large uncertainty on the sign and magnitude of changes by mid-century; over B.C., projected changes range from −2.8% to +17.8% in summer and +1.5% to +17.9% in winter precipitation. Southern and northern BC are likely to experience greater rates of warming, while central BC is likely to experience a greater increase in precipitation (Pacific Climate Impacts Consortium, 2013). Taken together, changes in temperature and precipitation are projected to affect regional freshwater hydrologic regimes (Zwiers et al., 2011). In western Canada, projections show that existing glaciers will shrink 60–80% in volume, with some disappearing by 2100 (Clarke et al., 2015). Nivo-glacial hydrologic regimes will shift toward pluvio-glacial, altering the timing and quantity of downstream freshwater supply. Along the coast, minimum losses of 50% of glacial volume are anticipated for even the lowest emissions scenario. Snow-melt dominant regimes will shift to be rainfall-dominant, leading to reductions to both winter snowpack and late summer flows and increases in winter flooding and summer drought, especially in the south (Merritt et al., 2006; Mantua et al., 2010; Shrestha et al., 2012).

Withdrawals from groundwater and the limits to groundwater supply are not well-recorded in the province. Until the advent of the Water Sustainability Act in 2016 (https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/laws-rules/water-sustainability-act), licensing of groundwater was not required. Also under the Water Sustainability Act, there is increased emphasis on maintaining adequate stream flows to meet eco-system requirements which were not in place when many agricultural water licenses were granted. Moreover, there are additional pressures on water resources from fracking operations within the province (Parfitt, 2017). In fact, there is a lack of reliable information on water use data from all sectors which greatly undermines management of water resources (Parfitt et al, 2012). Adding to these challenges overall water demand in southern BC is projected to increase as the population is expected to grow by 1.4 million by 2025 (BC Ministry of Environment, 2008).

The projected changes to water availability in BC will have serious implications for agriculture. Currently, agriculture accounts for 70% of consumptive water use in some areas of the province.) Reduced water availability may lead to crop failures and place an extreme financial burden on farmers (Crawford and Beveridge, 2013). Similarly, changes in climate may necessitate a transition to crops better suited to the climate, the costs of which can be prohibitive. These challenges arise at a time when the emphasis on BC agriculture has never been higher. Approximately 50% of BC's agricultural products are supplied by local farmers and the importance placed on growing local food is increasing. At the same time, agriculture faces pressures from competing land uses and volatile international markets (Crawford and Beveridge, 2013). There is an immediate need to understand future water availability and water demand in order to support BC agriculture, and the integration of climate science, climate model projections and agricultural water management has been recommended as an important step toward a resilient agricultural sector (Crawford and McNair, 2012). It is clear then that greater understanding of climate change and its impacts on water demand is required to effectively manage BC's agricultural water resources for the future.

The most detailed studies on water supply and demand for multiple uses have been undertaken in the Okanagan Basin in S. British Columbia. Consequently, in this paper we examine the Okanagan Basin in more detail. It is one of the few watersheds that is not glacier fed and that also does not have any major hydro-electricity developments. The Okanagan R. is a tributary of the Columbia with 8,000 km^2 of the basin lying in Canada. The climate in the lower elevations is semi-arid and crops require irrigation. The region has a mixture of high–value horticultural crops (wine-grapes and tree fruits) and beef and dairy production. There is a growing urban and rural-residential population, and earlier studies have indicated that outdoor water use through irrigation of agricultural and non-agricultural lands will potentially increase in response to climate change and urban development. Estimates for the period 1995–2006 indicated that agricultural irrigation accounted for 54%, non-agricultural irrigation for 24% of total water use and the remaining 22% accounted for domestic, industrial, commercial and institutional indoor use (Van der Gulik et al., 2010). Around 72% of irrigation water came from surface supplies, 23% from groundwater and 5% from reclaimed water (Okanagan Basin Water Board, 2011). Agricultural water use was predicted to increase by 50% between 2010 and 2040 in response to climate change and agricultural expansion and urban water use by 300% in response to climate change and growth (http://obwb.ca/wsd/wp-content/uploads/2014/07/OWSD_Phase3_Scenarios.pdf). In addition, the Okanagan provides a large portion of the breeding habitat for the Columbia River sockeye salmon population, and there has been considerable effort by the Okanagan Nation Alliance (ONA, www.syilx.org) and Partners to restore the sockeye salmon run including the development of a decision support system for managing Okanagan lake flows for fish (Hyatt et al., 2015).

Supplying water to all of these needs plus those for other ecosystem requirements is potentially conflict ridden considering that supply is dependent on highly variable annual snowpack (https://www2.gov.bc.ca/gov/content/environment/air-land-water/water/water-science-data/water-data-tools/snow-survey-data). Since the drought of 2003, there have been several initiatives to avoid conflicts among water users (particularly agricultural, domestic users and government departments protecting environmental flow needs) and to plan for the future through dialogue (Cohen et al., 2006; Melnychuk et al., 2016) and scientific assessment of water resources (Merritt et al., 2006; Neilsen et al., 2006; Okanagan Basin Water Board, 2011). Much of the dialogue has been led by the Okanagan Basin Water Board (OBWB) and its Stewardship Council, with a commitment to

provide scientifically informed decision making (Melnychuk et al., 2016). As a result there has been an increase in drought planning, although the efforts are varied. Of the 19 major water purveyors in the basin, 16% have a full sub-watershed plan taking into account environmental flow requirements with staged water restriction; 21% have a drought plan, water conservation plan and staged water restrictions; 16% have a drought plan with staged water restriction, 42% have stage water restriction and 5% have no plan at all. Staged water restrictions are frequently imposed as a percentage water reduction to all users yet for agricultural users vulnerabilities differ depending on crop, soil and irrigation system type. However, there is some potential for adaptation to water shortages by changes in crop and irrigation practices.

The BC/Canada irrigation water demand model (IWDM) addresses the need for greater information regarding climate change and BC's water resources. This model defines the spatial boundaries of agricultural production taking into account the physical landscape—soil and terrain characteristics, and the managed landscape—agricultural land use. The IWDM was initially developed to measure and predict agricultural IWD for the Okanagan Basin and has since been extended, to incorporate non-agricultural outdoor water use and indoor water use (Van der Gulik et al., 2010). The results from the IWDM can be generated for an entire study region or for any number of sub-regions, for periods ranging from daily to annual. IWD projections can be assembled for agencies such as municipal governments and water purveyors, as well as geographic areas such as watershed sub-basins. This allows for the monitoring of future water demand over a range of spatial scales, acknowledges and presents the influence of local climatic conditions, and ultimately provides local decision-makers with site-specific IWD projections. The success of the Okanagan Basin Water Supply and Demand Study (OBWB, 2010) has lead to requests from local governments for the IWDM to be run in other BC regions. To date, the IWDM has been used to calculate current water demand in a 17 regions across BC (The Partnership for Water Sustainability in BC, 2015). With this information local water managers and stakeholders are better able to plan strategically for the future of BC's water.

Given the need for a regional understanding of the implications of climate change on future IWD, two case studies are presented:

1. A comprehensive review of projected IWD in BC's agricultural regions from 2000 to 2100, including the uncertainty surrounding projected IWD and the implications of developing additional land for agricultural uses.
2. An analysis of vulnerabilities and adaptation in water use in the Okanagan Basin in response to agricultural land use and irrigation management.

METHODS

Study Areas

The study focuses on 17 regions in BC where agriculture is prominent or has the potential to be prominent in the future (**Figure 1**). The dimensions of the study regions examined in this paper evolved through previous applications of the IWDM. It was initially developed for the Okanagan Basin and its successful application in the Okanagan Basin Water Supply and Demand Study resulted in requests from other municipalities and regional districts. The nature of subsequent studies and the dimensions of the modeling regions were thus defined by the requesting administrative or civic organization. Thus some regions conform to regional district boundaries and others to river basins (**Figure 1**; **Table 1**). The Cariboo region as represented in the current study is just a small fraction of the Cariboo Regional District area. Regardless of the motivations behind their origins, the study regions are all host to significant agricultural activity. The Agricultural Land Reserve (ALR) in the study regions ranges between ∼2,200 ha (S. Gulf Islands) to ∼195,000 ha (S. Thompson) of which only a small fraction is irrigate, between <1% (Cariboo) and 32% (Cowichan) (**Table 1**). Agricultural capability class (1–4) land, considered suitable for crop cultivation, also varies among regions. Generalized information (https://www.alc.gov.bc.ca) indicates that, within the ALR, coastal regions have ∼74%, and interior regions have ∼55–60% class (1–4) land, with the exception of the Cariboo region which has around 37%. Much of the remaining ALR area is in pasture and range. The source of water for irrigation was identified by linking agricultural lots to surface water licenses (**Table 1**). Lots not linked to surface water licenses were, by default, assigned to groundwater sources, and where possible specific aquifers. Details on specific subbasins and aquifers sources are available in the IWDM reports for specific regions (https://www2.gov.bc.ca/gov/content/industry/agriculture-seafood/agricultural-land-and-environment/water/water-management/agriculture-water-demand-model). The Okanagan region also has a small supply ∼5% of reclaimed water. However, it should be noted that the production of fresh fruits and vegetables requires potable water, so reclaimed water is only a possibility for non-food crops (https://www.canadagap.ca/). The regions cover most of Southern British Columbia which has a widely varied landscape, characterized by coastal valleys, high mountain ranges and inter-montane plateaus and basins. Climate is also highly variable resulting in vegetation types ranging from coastal temperate rainforests to dry interior forests, sage brush and grasslands. Growing season effective precipitation (EP) (rainfall > 5 mm) and potential evapotranspiration (ETo) (**Table 1**) were derived from the climate data sets used to run the IWDM (see below for details). Growing season mean EP ranged between 59 mm (N. Thompson) and 265 mm (Fraser Valley) and mean ETo between 400 mm (Cariboo) and 634 mm (Okanagan) with the largest moisture deficits (ETo-EP) occurring in the Okanagan region.

IWDM

The IWDM is driven by detailed land use data and gridded precipitation and temperature data. Land use information comprised of mapped crop type, irrigation system, and soil type, was collected for all study regions and stored in a GIS database (Van der Gulik et al., 2010). Land use was represented as cadastral parcels that denoted property boundaries. These

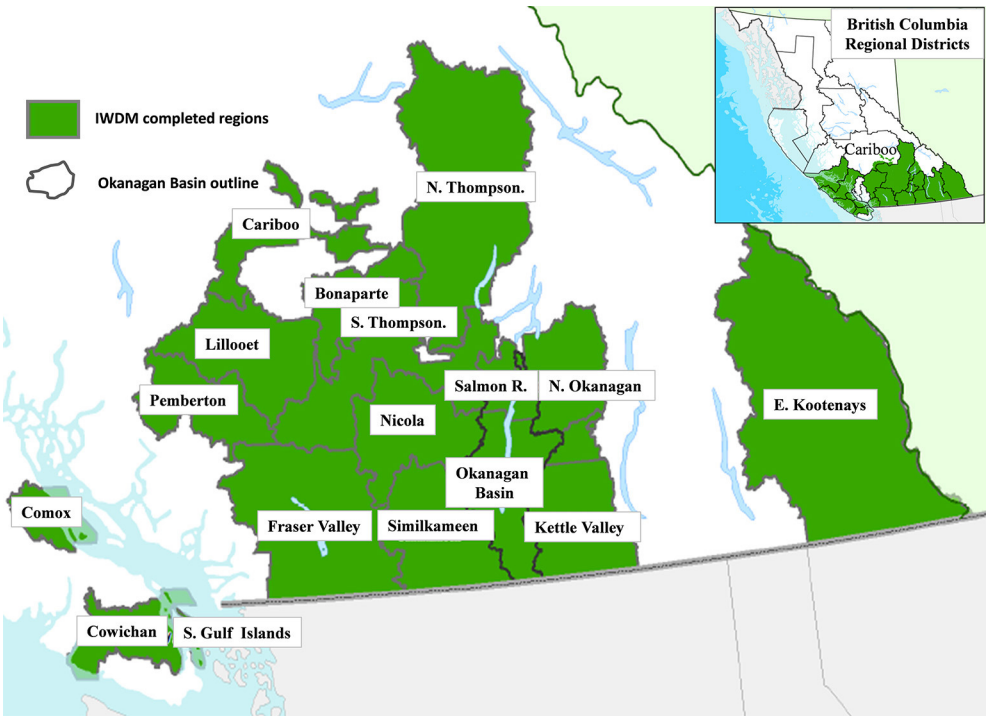


FIGURE 1 | Map of British Columbia regional districts and IWDM study regions.

TABLE 1 | Regional characteristics: agricultural land reserve (ALR), irrigation water source, mean growing season effective precipitation (EP) and potential evapotranspiration (ETo).

Study region	Region type	ALR (Ha)	Irrigated area (Ha)	Water source[Z] surface (%)	EP[Y] (mm)	ETo[Y] (mm)	ETo-EP[Y]	
							Mean (mm)	Max (mm)
Bonaparte	(I) Basin[X]	39,626	2,976	92	94	486	392	532
Cariboo	(I) RD[W]	84,295	586	88	114	400	286	428
Comox	(C) RD	22,626	1,691	43	164	557	393	553
Cowichan	(C) RD	7,865	2,504	66	136	584	448	616
E. Kootenay	(I) RD	91,517	5,770	73	116	528	412	574
Fraser Valley	(C) RD	187,735	15,455	66	265	603	338	479
Kettle Valley	(I) Basin	45,820	3,988	61	106	604	498	664
Lillooet	(I) RD	17,360	1,369	89	77	546	469	618
Nicola	(I) Basin	126,402	6,537	78	137	588	451	691
N. Okanagan	(I) RD	68,676	14,966	52	117	580	463	667
N. Thompson	(I) RD	64,863	6,639	82	59	485	426	596
Okanagan	(I) Basin	81,671	20,548	75	107	634	526	706
Pemberton	(C) RD	9,027	811	57	124	610	485	636
Salmon River	(I) Basin	29,248	4,835	85	120	556	436	678
Similkameen	(I) Basin	35,937	4,533	56	85	581	496	648
S. Gulf Is.	(C) RD	2,213	108	11	91	491	400	506
S. Thompson	(I) RD	194,743	6,911	71	90	582	493	679

[Z] Water supply was identified by linkage of agricultural lots to surface water licenses (private or water purveyor). With the exception of the Okanagan, which has 5% water supply from reclaimed sources, remaining irrigation water is considered to be supplied from groundwater.

[Y] Calculated for 1961–2000 for agricultural regions during the growing season (does not include non-agricultural and high elevation areas).

[X] Area defined as a watershed (I) interior or (C) coastal environments.

[W] Area defined as a Regional District; Cariboo sample area is only a small part of the Cariboo Regional District.

parcels were divided into polygons according to crop, irrigation and soil texture. The IWDM calculates IWD for each polygon for each day within the modeling period (Van der Gulik et al., 2010; Neilsen et al., 2015). Daily temperature data were used to calculate potential evapotranspiration in each climate grid cell using a modified Penman–Monteith method (Allen et al., 1998). Daily crop water use was calculated from potential ET modified by specific crop coefficients, crop rooting depth and rooting efficiency. After accounting for surface evaporation, available water stored in the soil and daily effective precipitation (>5 mm), IWD was calculated based on the daily crop water use, modified by irrigation system efficiency and losses to deep water percolation. Available water storage capacity and surface evaporation were determined by soil texture and losses to deep percolation by combinations of soil texture and irrigation system type. Details of coefficients used in the model are found in Van der Gulik et al. (2010). The IWDM was validated using pump data from a major water purveyor in the Okanagan Region, South East Kelowna Irrigation District. Because both agricultural irrigation and all other out water uses (domestic outdoor, parks and recreation outdoor, domestic, commercial, industrial and institutional indoor), were supplied through the same distribution system, output from the full version of the water demand model, which accounts for all of these uses (Van der Gulik et al., 2010), was used in the validation (**Supplementary Figure 1**). Weekly modeled and measured water use for the years 1995 to 2006 were well-correlated ($R^2 = 0.91$; $p < 0.01$).

IWDM Specifications

Agricultural Land Development

With a changing climate and a predicted increase in demand for agricultural products in the future, there is the potential for the expansion of land used for agricultural purposes. Land currently regarded as unsuitable may become suitable with warming temperatures and changing precipitation patterns (Neilsen et al., 2017). Further, increased demand for agricultural products may make agriculture more economically attractive. Accordingly, land that was determined to be suitable for irrigated agriculture was included in the simulations. Criteria included slope, elevation, current use, distance to water source, and presence in the Agricultural Land Reserve with an agricultural capability class of 1–4 (**Supplementary Table 1**). Depending on local factors, parcels of suitable land were assigned alfalfa or apple as their crop type.

Seasonality of Irrigation Water Demand

Within the IWDM, the start and end of each crop's growing season are determined by internal temperature-based rules driving crop phenology (Van der Gulik et al., 2010). This allows the start, end, and length of the growing season to change in relation to the influences of climate change. For example, the growing season may begin earlier or end later as the result of warmer temperatures. However, it must be noted that this approach also potentially challenges the current legislated irrigation season boundaries developed in response to previous climate normals.

Climate Data

The IWDM was run with both historical and future climate data. Observed daily historical climate data was based on the 300 arc second (∼10 km) gridded ANUSPLIN dataset (1961–2010) (McKenney et al., 2011). To model past and project future IWD for each region in BC, simulated temperature and precipitation (1961–2100) from an ensemble of climate models under historical and two future global socio economic development scenarios were used as input to the IWDM. Daily temperature and precipitation data were obtained from archived Coupled Model Intercomparison Project Phase 5 (CMIP5) simulations by GCMs and ESMs driven by scenarios representing potential future paths of global socio-economic development. These scenarios, Representative Concentration Pathways (RCPs), differ according to the concentration of greenhouse gas emissions (GHG) that may result from different development paths (Meinhausen et al., 2011). Four RCP scenarios, 2.6, 4.5, 6.0, and 8.5 indicate possible radiative forcing values ($W \cdot m^{-2}$) in 2100 relative to pre-industrial values. For the purposes of this study, the RCP2.6 scenario was excluded as it seemed unlikely that its assumptions would be borne out. It assumes that global GHG emissions will peak between 2010 and 2020, with emissions declining after that. Additionally, RCP6.0 was omitted as it represents a middle-ground between RCP4.5 and RCP8.5, and offered comparatively less insight into the range of potential changes in IWD.

Previous research has recommended using a multi-model and multi-climate forcing approach to represent uncertainty in IWD projections (Gosling et al., 2011). Accordingly, data from a subset of five climate models (**Table 2**), selected from the CMIP5 ensemble to span the range of projected temperature and changes in the region (Cannon, 2015), was used to drive the IWD for the RCP4.5 and RCP8.5 scenarios, with outputs from each scenario condensed into an ensemble mean. To complement the ensemble mean, which reduces variability, smoothing out projected values, the range, standard deviation, and coefficient of variation of projected values were also calculated. This workflow is described in **Supplementary Figure 2**.

Downscaling

To accommodate the effects of complex terrain in the region on climate conditions, fine-scaled (500 m) gridded datasets of

TABLE 2 | Global Climate Models (GCMs) and Earth System Models (ESMs) used in this study.

Model name	Abbreviated name	Model type
Australian Community Climate and Earth-System Simulator	ACCESS	ESM
Canadian Earth System Model	CanESM2	ESM
Centre National de Recherches Meteorologiques Climate Model	CNRM-CM5	GCM
CSIRO Mk3.5	CSIRO Mk3.5	GCM
Institute of Numerical Mathematics Climate Model	INMCM4	GCM

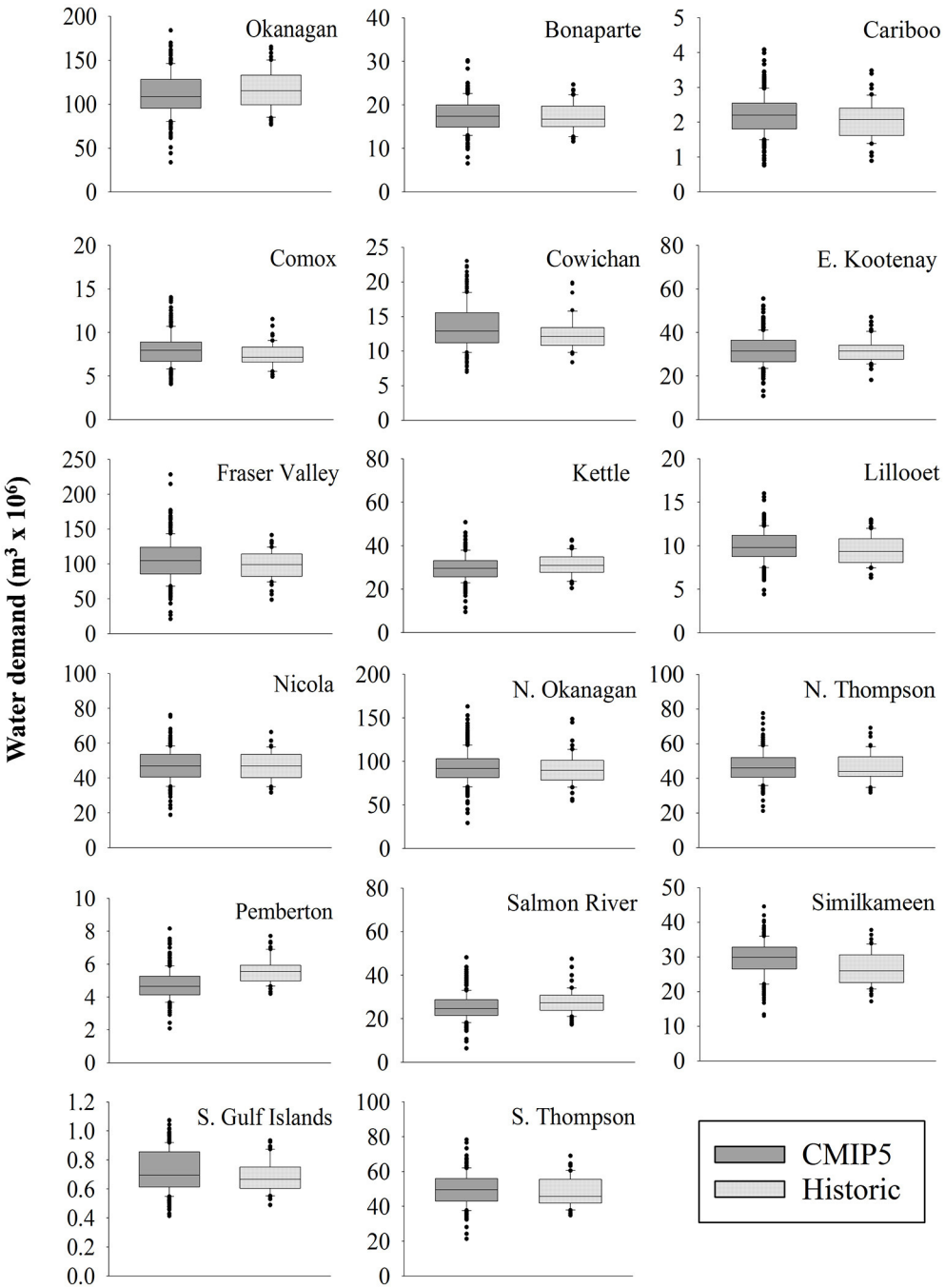


FIGURE 2 | Boxplots of modeled annual IWD for 16 agricultural regions in British Columbia in response to greenhouse gas emissions (RCP 8.5) and historical climate data for the period 1961–2010. Greenhouse gas emission scenarios data include output from five climate models ACCESS, CanESM2, CNRM-CM5, CSIRO Mk3.5, INMCM4.

daily minimum, maximum temperature and precipitation were downscaled from the historical ANUSPLIN and future GCM climate data. The 300 arc second (∼10 km) historical ANUSPLIN dataset for 1951–2010 (McKenney et al., 2011) was spatially-disaggregated using a local lapse rate upsampling algorithm (Wang et al., 2006) and then bias corrected so that climatological monthly means matched the Pacific Climate Impact Consortium PRISM climatology https://www.pacificclimate.org/data/prism-climatology-and-monthly-timeseries-portal. This produced historical daily climate surfaces for 1961–2010. GCM and ESM outputs for British Columbia (**Table 2**) were statistically downscaled onto the ANUSPLIN grid using the Bias-Corrected

Constructed Analogs with Quantile mapping reordering (BCCAQ) algorithm (Werner and Cannon, 2016) to produce daily climate surfaces for 1961–2100. http://www.pacificclimate.org/data/statistically-downscaled-climate-scenarios. Outputs were then spatially disaggregated and bias corrected in the same manner as the historical ANUSPLIN dataset.

Data Analysis

The ensemble means of projected IWD were summarized as annual totals and presented for 2000–2100 for each region. Simple linear regression using the least squares approach was applied to the projected IWD to show the average rate of increase over the study period for each region. Additionally, the range of potential IWD-values for each year in each region is presented. Uncertainty was captured through absolute and relative uncertainty. Absolute uncertainty was calculated as the standard deviation of the five model outputs for each year, for each scenario. As with the IWD, simple linear regression was applied to the standard deviation to show the trend in standard deviation over the study period. Relative uncertainty was represented as the coefficient of variation for the first and last year of the study period for each region, and each scenario. This illustrates the change in uncertainty over the study period. Finally, IWD that may result from the development of agricultural land was also summarized annually and presented for 2010–2040.

Okanagan Basin Case Study

Two detailed land use data sets were assembled for the Okanagan Basin, one surveyed in 2005–2006 (Van der Gulik et al., 2010) and the second in 2013–2014 according to the same methodology. Both land use datasets (labeled 2006 and 2014), were run through the IWDM using the historical climate data allowing comparisons of land use, irrigation system use and water use. Thirty-two crop production systems were identified in 2006 and 40 in 2014. In each case there were 18 different types of irrigation management system resulting in a potential 536 crop × irrigation combinations in 2006 and 720 combinations in 2014. Because of the large number of combinations, results are presented for grouped data: 6 crop classes reflecting rooting depth and function; 5 soil texture classes reflecting available water storage capacity and 4 irrigation system classes reflecting efficiency (**Supplementary Table 2**). The region has the highest in-season water deficits (ETo-EP) in Southern BC (**Table 1**). Water shortages can occur because of (1) low spring mountain snowpack (SWE) causing low reservoir storage; (2) low in-season precipitation (EP) and (3) high in-season potential ET (ETo) (**Supplementary Figure 3**) These do not necessarily coincide with provincially declared hydrologic droughts associated mainly with low late summer flows in unregulated streams. Mountain snow pack is highly variable, ranging between 212 and 775 mm in April and 13–775 mm in May (**Supplementary Figure 3A**). The most notable water shortages have occurred when low in-season EP coincides with high in-season ETo as in 2003 for example (**Supplementary Figure 3B**). These occurrences lead to high in-season water deficits (707 mm in 2003) and the low EP also results in low stream flows and an increased dependence on storage resulting in greater vulnerability for agriculture. Two indices of vulnerability were created. The first index was a ratio of IWD to potential evapotranspiration (IWD/ETo) which indicated the relative vulnerability of crop/irrigation system/soil combinations to water demand. The second index was a ratio of IWD to the maximum allocations for selected water purveyors which usually would be reduced in response to drought. In all cases, the selected purveyors have water storage in high elevation reservoirs which supply irrigation needs in the growing season. Drought management tools are based on "trigger" graphs related to reservoir levels throughout the year usually with five stages, 0–4. Restriction on agricultural producers vary and may call for a reduction in water use in July and August to 50–70% of maximum allocation for stage 3 to a complete restriction of irrigation for stage 4. The water allocation vulnerability index compared estimated water use with maximum allocation values (550, 686, and 800 mm/yr respectively), for three irrigation districts from the North, Central and South Okanagan.

RESULTS

Irrigation Water Demand (IWD) in British Columbia Agricultural Regions

Water use from 1961 to 2010 was calculated using both historical climate data and the suite of downscaled climate models (**Table 2**). Historical IWD varied widely among region, ranging between median values of 0.6 MCM ($m^3 \times 10^6$) for the Southern Gulf Islands region and 115 MCM for the Okanagan Region

TABLE 3 | Historical average IWD (1961–2000) and projected increase in demand (2000–2100) in response to two greenhouse gas scenarios.

Region	IWD				
	Historical 1961–2000	RCP 4.5		RCP 8.5	
	$m^3 \times 10^6$	$m^3 \times 10^6$	%	$m^3 \times 10^6$	%
Bonaparte	17.30	7.71	45	12.02	69
Cariboo	2.06	1.40	68	2.25	109
Comox	7.38	2.26	31	3.34	45
Cowichan	12.46	3.81	31	4.53	36
E. Kootenay	31.16	11.13	36	24.60	79
Fraser Valley	103.65	35.59	34	51.49	50
Kettle Valley	33.50	12.86	38	19.49	58
Lillooet	9.49	3.95	42	6.06	64
Nicola	46.83	21.31	46	33.72	72
N. Okanagan	90.83	35.20	39	54.56	60
N. Thompson	46.13	17.34	38	27.52	60
Okanagan	116.52	42.43	36	63.50	54
Pemberton	5.57	1.60	29	2.34	42
Salmon River	28.06	11.75	42	18.59	66
Similkameen	26.74	10.28	38	15.29	57
S. Gulf Islands	0.68	0.16	23	0.23	33
S. Thompson	48.32	28.36	59	46.05	95

TABLE 4 | Uncertainty in projected IWD 2000–2100.

Region	RCP4.5				RCP8.5			
	2000		2100		2000		2100	
	σ (m^3 × 10^6)	CV	σ (m^3 × 10^6)	CV	σ (m^3 × 10^6)	CV	σ (m^3 × 10^6)	CV
Bonaparte	3.30	0.18	5.00	0.19	2.55	0.13	6.73	0.22
Cariboo	0.57	0.23	0.84	0.22	0.50	0.2	1.16	0.25
Comox	1.49	0.18	2.11	0.2	1.43	0.17	2.59	0.22
Cowichan	2.75	0.2	3.63	0.2	2.64	0.18	3.91	0.2
Fraser Valley	27.09	0.24	31.96	0.22	25.52	0.23	34.49	0.21
Kettle Valley	5.39	0.17	7.67	0.17	4.58	0.14	9.07	0.17
Kootenays	16.61	0.19	19.27	0.16	13.78	0.17	26.94	0.19
Lillooet	1.58	0.15	2.56	0.18	1.30	0.12	3.26	0.19
Nicola	7.90	0.16	12.39	0.17	6.37	0.12	16.07	0.19
N. Okanagan	17.29	0.17	23.48	0.17	14.37	0.14	30.64	0.2
N. Thompson	7.77	0.16	10.26	0.15	5.54	0.11	14.68	0.19
Okanagan	22.26	0.19	28.46	0.18	19.67	0.16	33.76	0.18
Pemberton	0.83	0.16	1.09	0.16	0.70	0.14	1.39	0.19
Salmon River	5.46	0.2	7.61	0.19	4.23	0.15	10.15	0.22
Similkameen	4.15	0.13	5.77	0.14	3.56	0.11	6.85	0.14
S. Gulf Islands	0.13	0.17	0.15	0.17	0.12	0.16	0.16	0.16
S.Thompson	7.54	0.17	11.13	0.15	6.66	0.16	15.14	0.18

TABLE 5 | Current and potential (additional) land base and projected additional IWD in response to 100% increase in agricultural land development and two greenhouse gas scenarios for BC agricultural regions by 2100.

Region	Land base			IWD				
	Current	Potential	Total	Historical 1961–2000	RCP4.5		RCP8.5	
	Ha	Ha	%ALR	m^3 × 10^6	m^3 × 10^6	%	m^3 × 10^6	%
Bonaparte	2,976	2,418	14	17.30	16.72	97	17.63	102
Cariboo	586	10,012	13	2.06	48.86	2371	51.04	2477
Comox	1,691	4,759	29	7.38	24.16	328	25.82	350
Cowichan	2,504	4,431	88	12.46	22.74	182	24.25	195
E. Kootenays	5,770	8,933	37	31.16	56.97	183	60.7	195
Fraser Valley	15,455	18,650	17	103.65	130.44	56	135.86	58
Kettle	3,988	3,843	8	33.50	32.01	96	34.13	102
Lillooet	1,369	191	9	9.49	1.66	17	13.98	147
Nicola	6,537	9,673	13	46.83	68.79	147	72.08	154
N. Okanagan	14,966	20,573	52	90.83	139.03	153	146.77	162
N, Thompson	6,639	18,365	39	46.13	127.1	276	131.65	285
Okanagan	20,548	5,800	32	116.52	35.33	30	38.04	33
Pemberton	811	2,423	36	5.57	15.51	278	16.39	294
Salmon River	4,835	6,779	40	28.06	38.88	139	41.36	147
Similkameen	4,533	1,827	18	26.74	2.68	10	11.53	43
S. Gulf Islands	108	460	4	48.32	9.55	20	9.78	20
S. Thompson	6,911	1,119	26	0.68	2.48	362	2.65	387

(**Figure 2**). In general, similar median results were found for the climate model simulations over the historical time period (1961–2010), although the range was greater for the ensemble model outputs due to sampling multiple realizations of natural climate variability by the five different CMIP5 climate models. The inter-annual range for the historic data set falls within the extreme values generated by the five CMIP5 models and within the 95% confidence interval of the ensemble mean. There is a slight upward trend in the ensemble mean that is not evident in the historic data.

Trends in IWD between 2000 and 2100 were derived from linear regression lines fitted to the annual ensemble means. In all study regions IWD was projected to increase between 2000 and 2100 (**Table 3**). For the RCP4.5 scenario increases in IWD ranged from 21.3 to 63.7%, with 14 of the 17 regions expected to see increases >30%. Under the RCP8.5 scenario, every region was projected to have greater IWD relative to the RCP4.5 scenario. Projected increases in IWD under RCP8.5 ranged from 29.8 to 114.1%, with 14 of the 17 regions experiencing >40% increase in IWD. Increases in IWD were highly variable among regions. In both scenarios the Cariboo and S. Thompson regions were projected to see the greatest increases in IWD. Conversely, the regions on or adjacent to Vancouver Island (Comox, Cowichan, and S. Gulf Islands) were projected to see the smallest increases in IWD. Increases in IWD were also highly variable among scenarios in each region. The difference between the RCP4.5 and RCP8.5 scenarios was most pronounced for the Cariboo, Kootenay, and S. Thompson regions. Those regions near the south coast (Comox, Cowichan, S. Gulf Islands, Fraser Valley) had the least variation between scenarios.

To clarify the reasons for the differences between coastal and interior regional responses, changes in ETo and seasonal water deficits (ETo-EP) were examined for CanESM2 RCP8.5 scenarios for the S. Gulf Islands and Cariboo regions. In the coastal example, there was relatively little change in either ETo or ETo-EP throughout the century (**Supplementary Figure 4A**). The larger spread in ETo-EP than ETo was due to the range in EP data. In the Interior example, both ETo and ETo-EP increased linearly over time, and again the range in EP data increased the spread for the water deficit calculation (**Supplementary Figure 4B**). In both cases, there was no trend in precipitation data.

Uncertainty

Absolute uncertainty (σ) increased over the study period (2000–2100) for all regions in both scenarios (**Table 4**). It was also greater in RCP8.5 than RCP4.5 scenarios in all regions at the end of the study period. In some regions, the difference in absolute uncertainty between the two scenarios in 2100 was substantial (e.g., Kootenays, North Okanagan) while in others the difference was minimal (Southern Gulf Islands).

A comparison of relative uncertainty (CV) at the beginning and end of the study period shows that uncertainty in the RCP4.5 scenario was generally stable. Relative uncertainty remained the same or slightly decreased in 12 of 17 regions in the RCP4.5 scenario. In contrast, relative uncertainty in the RCP8.5 scenario increased in every region other than the Fraser Valley and S. Gulf Islands. Comparing the two scenarios, RCP8.5 had lower relative uncertainty in all 17 regions in 2000, but a slightly greater relative uncertainty in 11 of 17 regions in 2100.

Agricultural Land Development

As would be expected, projected IWD increased as additional agricultural land was developed (**Table 5**). Projected increases for scenario RCP4.5 and RCP8.5 were nearly identical in response to development relative to their respective baselines. Just as in the non-development scenarios, interregional variation in IWD was considerable, ranging from a projected ∼10–2,477% increase in both scenarios. It is important to note that a region's increased IWD is dependent on the amount of potential agricultural land available in conjunction with the effects of climate change. Thus, in this case interregional variation was less influenced by local climatic conditions and more by differences in the amount of agricultural land available.

TABLE 6 | Areal distribution of grouped soil texture, irrigation system, and crop rooting attributes for the Okanagan basin in 2006 and 2014.

Attribute	% Total basin area					
Soil texture	**V. coarse**	**Coarse**	**Medium**	**Fine**	**V. fine**	
2006 land use data	0.07	10	56	18	16	
2014 land use data	0.05	10	58	18	15	
Irrigation system	**Drip**	**Micro**	**Sprinkler**	**Gun**	**Flood**	
2006 land use data	13	9	66	12		
2014 land use data	23	13	55	9	0.20	
Crop rooting	**V. deep**	**Deep**	**Grape**	**Medium**	**Shallow**	**V. Shallow**
2006 land use data	16	14	15	19	36	0.56
2014 land use data	14	22	19	13	31	0.53

Okanagan Basin Case Study

Land Use Characteristics

Due to slight shifts in crop area and crop type, there were slight differences in the texture of irrigated soils between the 2006 and 2014 surveys (**Table 6**) and over 50% of the irrigated area had medium textured soils which were predominantly sandy loams (**Supplementary Table 2**). Use of sprinkler irrigation systems were dominant in both land use surveys but had decreased by 10% in 2014 with a concomitant increase in more efficient micro-irrigation and drip systems. The area under grapes and other deep rooted crops (primarily forage crops) increased in the 2014 survey while the area under medium and shallow rooted crops decreased. Efficient irrigation systems (drip and micro-irrigation) were associated with horticultural production as they are best adapted to crops grown in rows (**Figure 3A**). The choice of irrigation systems is more limited for pasture and forage crops although there was some evidence of change from less efficient gun to sprinkler systems in 2014. Between 2006 and 2014 there was a large shift to efficient irrigation systems in horticultural crops with 68% of the horticultural area under drip and micro-irrigation in 2014 compared with 38% in 2006 (**Figure 3B**).

Relative Vulnerability Index (IWD/ETo)

There was considerable annual variation in the overall relative vulnerability index (**Figure 4A**) which was related to growing season temperatures and precipitation. The magnitude of the

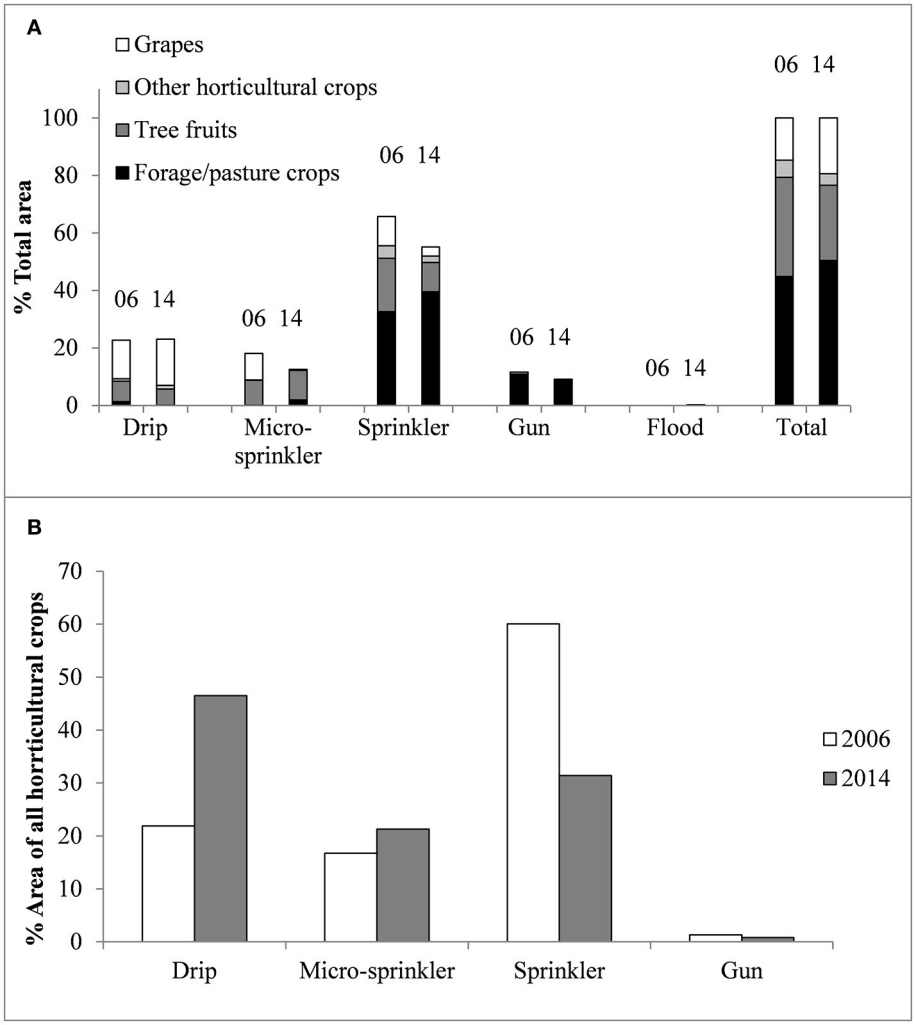


FIGURE 3 | Areal distribution of **(A)** Okanagan Basin major cropping and irrigation systems and **(B)** horticultural irrigation systems in 2006 and 2014.

relative vulnerability index varied annually and was largest in 1998, 2002, and 2003 (hottest and driest years) and lowest in 1997 (coolest and wettest year). In hot dry years, between 60 and 70% of the basin irrigated area had a relative vulnerability index>1.0 which fell to <5% in wet cool years (**Figure 4B**). Crops grown on finer textured soils had lower relative vulnerability indices than those grown on coarser soils, and in general shallow rooted crops grown on coarse textured soils the highest relative vulnerability indices (**Figures 5A,B**). Irrigation system also greatly affected the relative vulnerability index with gun and flood irrigation giving rise to high relative vulnerability indices, even for deep rooted crops, and drip irrigation resulting in low relative vulnerability indices (**Figures 5C,D**).

Water Allocation Vulnerability Index (IWD/Maximum Water Allocation)

Estimated annual IWD was compared with maximum allocation values (550, 686, and 800 mm/yr., respectively), for three irrigation districts from the North, Central and South Okanagan. All three regions had over 50% of their area in medium textured soils, mainly sandy loams (**Table 7**). The Central Okanagan had more coarse textured soils than the other two regions. The three irrigation districts had different cropping profiles, with the N. Okanagan dominated by pasture and forage crops, which can be both deep and shallow rooted (**Supplementary Table 2**), and inefficient irrigation systems, when compared to the Central and S. Okanagan. Between 2006 and 2014, the area in wine grapes had doubled in all three regions and the amount of drip irrigation had increased five-fold in the Central Okanagan. Maximum allocations were exceeded in all regions with the area affected depending on year and region (**Figure 6**). There were slight differences in vulnerability for the two land use surveys. For the 1996 survey, between 1997 and 2010, allocations were exceeded in more than 50% of the area in 11 years for the N. Okanagan, in 5 years for the Central Okanagan and in 1 year for the S. Okanagan. For the 2014 survey, allocations

were exceeded in more than 50% of the area in 8 years (N. Okanagan), 4 years (Central Okanagan) and 0 years (S. Okanagan) In 2003, one of the hottest, driest years, maximum allocations for the 2006 land use data were exceeded for 93%), 84% () and 53% (of the irrigated area respectively North, Central and South Okanagan districts and for the 2014 land use data they were exceeded for 91, 85, and 47% of the irrigated area for the three districts respectively. These results indicate an imbalance between expected maximum irrigation requirements and maximum allocation. The magnitude of the water allocation vulnerability index ranged between 0.55 and 2.26 for the North Okanagan, between 0.46 and 1.94 for the Central Okanagan and between 0 and 1.38 for the South Okanagan. In all locations, crops irrigated with inefficient systems (gun and sprinkler) were most affected regardless of rooting depth and those irrigated with drip and micro sprinklers were usually minimally affected. Wine grape crops were rarely affected, mainly because the IWD model accounts for the lower water requirements for their production. The effects of crop production system, soil texture and irrigation system on the magnitude of the water allocation vulnerability index are illustrated for the North Okanagan in 2003 (**Figure 7**). Crops grown on very coarse or coarse textured soils had the highest vulnerability indices regardless of rooting depth (**Figures 7A,B**). Greatest vulnerability was associated with crops grown with inefficient irrigation systems, flood and gun, and least vulnerability was associated with crops grown with drip and micro-irrigation (**Figures 7C,D**).

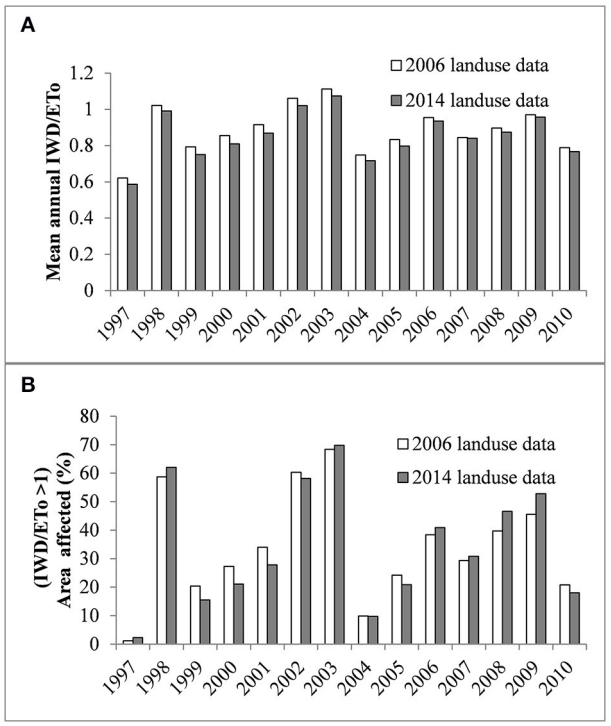


FIGURE 4 | Annual climatic effects on **(A)** the relative vulnerability index (IWD/ETo) and **(B)** the area with IWD/ETo>1 in the Okanagan Basin for land use data collected in 2006 and 2014.

DISCUSSION

Irrigation Water Demand (IWD) in British Columbia Agricultural Regions

The findings of this paper indicated that climate change alone will likely lead to significant increases in IWD in BC. Under the RCP4.5 scenario, which assumes emissions will peak around 2040 and decline thereafter, IWD in every region was projected to increase >20%, while for the majority of regions a >30% increase in IWD is expected. In the higher GHG emissions scenario (RCP8.5) every region in BC was projected to experience a >30% increase in IWD, while the majority will face a >40% increase. These findings project significant increases in agricultural IWD in BC, and align with previous research on the Okanagan Basin (Neilsen et al., 2006). Past studies at the regional scale have projected dramatic increases in IWD as a result of climate change (Fischer et al., 2007; Rehana and Mujumdar, 2013; Wada et al., 2013). The increases projected in IWD for BC reiterate the need to enhance water infrastructure and decision-making processes and the projections included in this paper can be incorporated into future decision-making to address these challenges. However, the uncertainty associated with future water supply in the region (Merritt et al., 2006; Shrestha et al., 2012; Clarke et al., 2015) highlights the need to integrate climate science and future water demand with changes in water supply at the basin and sub-basin scale in order to develop sound decisions. This is currently of increasing importance as BC moves to license groundwater. The results of these regional water demand model runs have been incorporated into a new web-based tool which allows applicants for either groundwater or surface water licenses to calculate the agricultural water demand for their properties (http://www.bcagriculturewatercalculator.ca/), While substantial in all regions, the projected increases in IWD were not uniform across the province. Our findings illustrate that the relationship between GHG emissions and IWD is unique to each region. Interregional variation in projected IWD highlights the significant influence of local climatic conditions. Under both scenarios the range of projected increases was large (RCP4.5 ∼20–60%, RCP8.5 ∼30–115%). Changes in temperature affecting potential ETo appeared to be the major driver of increased water demand, as there were no apparent trends in precipitation data. These temperature effects were greater for Interior than for Coastal regions. This stresses the need for a localized and regional approach to water management in BC (Cohen et al., 2006; Melnychuk et al., 2016) as the policies and regulations put in place for one region may not work in another.

The course of global socio-economic development will also have an uneven impact on BC's IWD. Projected IWD did not increase uniformly between the RCP4.5 and RCP8.5 scenarios in all regions. This suggests that different regions have different sensitivities to climate change drivers. In certain regions, a future where GHG emissions grow exponentially may lead to a near doubling of IWD relative to a future where emissions decline after 2040. However, in other regions the trend in GHG emissions has a limited effect on future IWD. Whether future

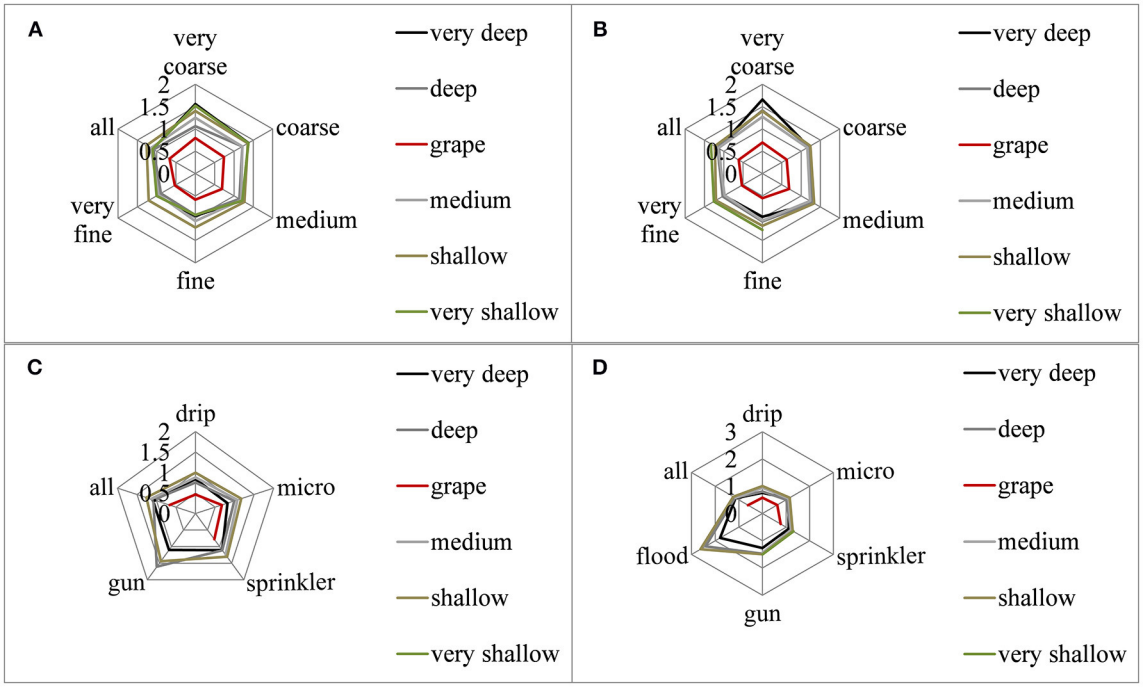


FIGURE 5 | Relative vulnerability index (IWD/ETo) for Okanagan Basin cropping systems in 2003, in response to crop rooting depth and **(A)** soil texture for 2006 landuse data, **(B)** soil texture for 2014 land use data, **(C)** irrigation systems for 2006 land use data and **(D)** irrigation system for 2014 land use data.

GHG emissions follow a RCP4.5 or RCP8.5 path will have serious implications for the climate sensitive regions in the province. Accordingly, the integration of climate science and water resource planning should be an ongoing process. Only through continued research and reporting on rates of GHG emissions and IWD will local decision-makers have access to the most up-to-date information necessary for sound decision-making.

Ongoing research into future water demand will also help to address the uncertainty surrounding projected IWD. While relative uncertainty can be regarded as low, absolute uncertainty would be regarded as considerable for regional water purveyors. In the majority of regions the absolute uncertainty surrounding projected IWD in 2100 was comparable to the projected increase in IWD. Further, the uncertainty grew when comparing RCP8.5 to RCP4.5. Uncertainty in projected IWD is related to the level of understanding of our climate's response to increasing GHG emissions (Figure SPM.10 of https://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_SPM_FINAL.pdf; Figure SPM.10 of https://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_SPM_FINAL.pdf). As the climate moves away from historical norms, our understanding of the climate's response becomes weaker, and uncertainty increases. This explains the greater increase in both absolute and relative uncertainty found in the RCP8.5 scenario. While we can say with relative confidence that IWD will increase in the coming decades, the uncertainty of projected IWD emphasizes the need for a flexible and adaptive approach to water resource planning.

Similar to projected IWD, the contribution of agricultural land development to IWD will vary region to region. For some regions it could far exceed the increases brought by climate change alone, while for others agricultural development would contribute relatively little to increased IWD. This is contingent on the level of development, the amount of land available for development, and the region's sensitivity to climate change. Some regions, for example the Cariboo Region, have a large potential for increases in irrigated agriculture as much of the Agricultural Land Reserve is currently used for dry land ranching. However, the region is very remote, sparsely populated and poorly served by transportation networks, which may provide a substantial barrier to development. On the other hand, there may be increased opportunities for the production of high value horticultural crops in this region, particularly as minimum winter temperatures moderate. Neilsen et al. (2017), using phenology modeling and terrain analysis, showed that areas of the S. Cariboo region, particularly in the terraces adjacent to the Fraser R, would become suitable for sweet cherry production by the 2040s. At the other extreme, a region such as the S. Gulf Islands, has little potential available agricultural land and is also constrained by ground-water availability. These relationships have important implications for land use management: first, they again highlight the unevenness of projected IWD and the need for regional decision-making; second, they stressed that the benefits that may be brought by expanding the agricultural land base must be weighed against the potential costs and risks associated with increasing IWD. These might include the costs of

new infrastructure; the costs of monitoring and modeling surface and groundwater source; environmental risks from reducing in-stream flows and aquifer levels; risks to producers if planting crops at their climatic limits etc. Accordingly, water resource planning should be undertaken at a regional level and adopt a long-term perspective.

Okanagan Basin Case Study

Considerable attention has been focused on incorporating science based decision making into water resource planning in the Okanagan Basin (Melnychuk et al., 2016) and there is on-going interest in pursuing both basin-wide drought and flood management strategies. The Okanagan water stewardship council, which is a long-standing committee of water professionals includes representives of As much of the region is supplied by large water purveyors, who have storage licenses for large upland reservoirs and/or withdrawal licenses from large valley bottom lakes, drought management strategies are largely aimed at retaining sufficient storage for winter and early spring withdrawals for non-irrigation purposes and instream flow requirements. Thus the likelihood of "drought" under these circumstances is increased by a combination of low winter snowpack, high in-season ET and low in-season precipitation. This type of drought differs from a hydrologic drought, declared by the Province, when low river flows threaten environmental flow needs in un-regulated streams. Projected increases in agricultural water demand due to climate change (up to 54%), while likely within the licensed volume for agricultural purposes (Neilsen et al., 2006) may not be met by future water supply (Merritt et al., 2006). Moreover, the findings of this study indicated that even within the recent historical climatic range, allocation of water resources may not be adequate to meet agricultural needs in some areas. In the N. Okanagan irrigation district, the area potentially vulnerable to under-allocation was >50% in all but the coolest and wettest years which was due to both a low maximum allocation value and to the crop profile in the region. Forage and pasture production systems which rely on less efficient irrigation systems (sprinklers, irrigation guns and flood irrigation) are the most vulnerable to under-allocation and comprise 73% of the irrigated area in the N. Okanagan, compared to around 50% in the basin as a whole. In contrast, the irrigation district in the S. Okanagan which had a much lower potential for under-allocation was characterized by a higher allocation value and only around 25% of the irrigated area in pasture and forage crops.

Unfortunately, the majority of irrigation districts and water purveyors in the basin have not defined maximum water allocations, so this type of analysis was not possible. However, there was a strong relationship between the relative vulnerability index (IWD/ETo) and the water allocation vulnerability index (IWD/max water allocation) with R^2-values varying between 0.82 and 0.92 for the three irrigation districts (n = 10,000–124,722 land use polygons × 4 years). Taking this into consideration, if maximum allocations are imposed across the basin as drought planning continues, there is potential for considerable vulnerability to insufficient water availability for some production systems. That these vulnerabilities are already evident for historic climatic conditions before the 34–54% increase projected in response to climate change, or potential 30% potential increase in irrigable land, emphasizes the need for adaptive measures both in water use and drought planning.

TABLE 7 | Areal distribution of soil texture groups, irrigation system groups and crop rooting groups for three Okanagan irrigation districts (North, Central, South) in 2006 and 2014.

Region	% Irrigation district area					
	Soil texture					
	V. coarse	**Coarse**	**Medium**	**Fine**	**V. fine**	
North 2006	0.1	0.3	54	25	21	
North 2014	0.1	0.4	58	26	16	
Central 2006	0.4	25	61	3	11	
Central 2014	0.1	25	64	3	9	
South 2006	0	8	65	27	0	
South 2014	0.1	8	65	27	0	
	Irrigation system					
	Drip	**Micro**	**Sprinkler**	**Gun**	**Flood**	
North 2006	11	3	57	29	0	
North 2014	9	6	69	17	0.2	
Central 2006	5	39	51	6	0	
Central 2014	25	32	41	1	0	
South 2006	25	12	62	0.4	0	
South 2014	32	16	48	4	0	
	Crop rooting					
	V. deep	**Deep**	**Grape**	**Medium**	**Shallow**	**V. Shallow**
North 2006	26	10	0.2	7	56	0.6
North 2014	35	22	0.6	13	29	0.2
Central 2006	6	28	7	31	28	0
Central 2014	2	26	15	19	38	0
South 2006	0.6	9	7	34	50	0
South 2014	4	11	16	28	41	0

Adaptation to potential shortfalls in water supply due to current production methods can best be achieved through changes in irrigation management. Conversion to low pressure irrigation systems (drip, micro-sprinkler, LP-center pivot) increases irrigation efficiency to 78% and greater. Such changes are already underway in the horticultural sector in the region where water was supplied to 38% of the irrigated area through micro-sprinklers or drip in 2006 and to 68% of the irrigated area in 2014. This provides a number of benefits to producers including the ability to target water spatially and temporally to meet plant demand; to produce plants which are tolerant of partially dry soils and to allow maintenance of the same irrigation practice under mild drought conditions if water use is well below maximum allocations. There are also options to increase efficiency within pasture and forage crops. Low pressure center pivot irrigation has been widely adopted in many regions. A recent study of Alberta irrigation districts indicated that 76%

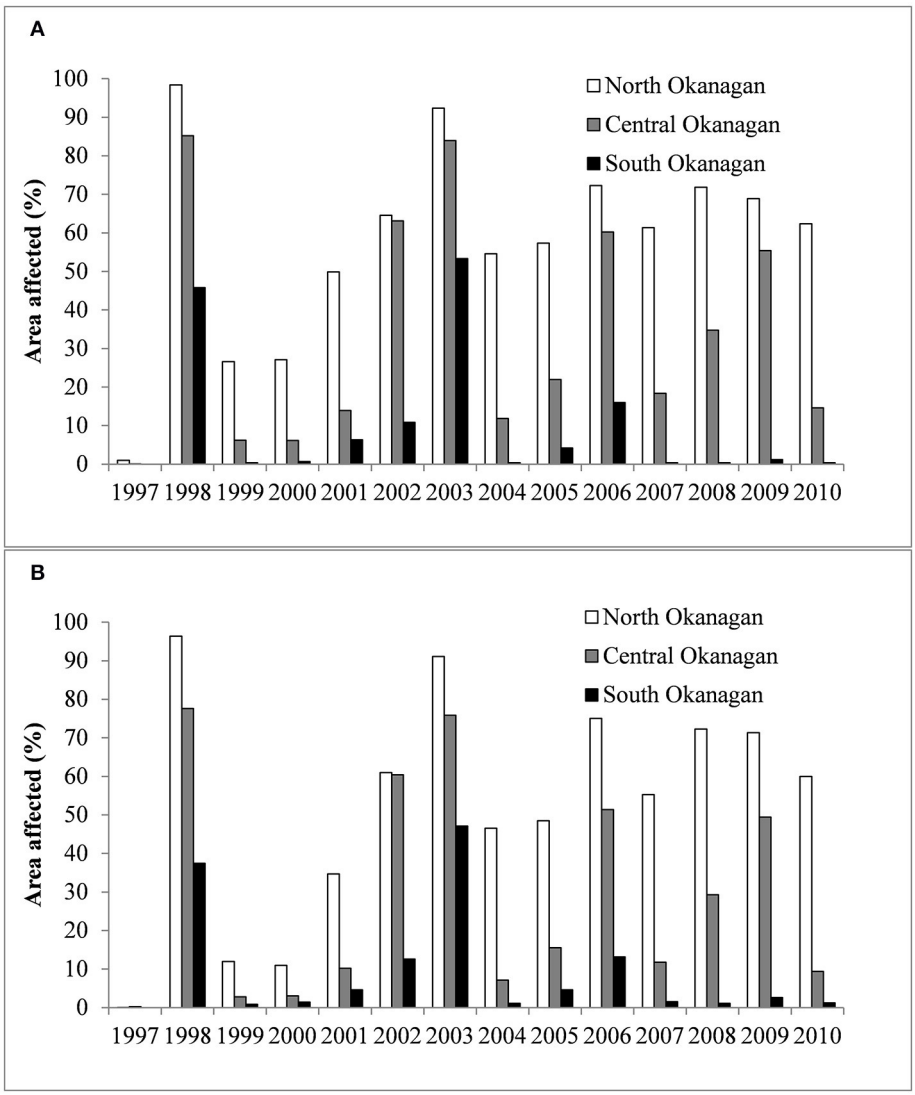


FIGURE 6 | Area of irrigated land where IWD exceeded annual allocations in three irrigation districts from the Okanagan basin in response to **(A)** 2006 and **(B)** 2014 land use data.

of the irrigated area was in the process of change to more efficient management and that mostly was due to the use of LP-center pivot systems (Wang et al., 2015). In the Okanagan basin, the small area of pivot irrigation increased from 442 to 484 ha between 2006 and 2014 with LP-center pivot from 1 to 67%. The extent to which the size of land holdings and topography limit the adoption of center pivot technology in this region is unclear and requires further study. The use of sub-surface drip (95% efficient) has also been adopted for forage corn, alfalfa and grass pasture in regions with severe water shortages including the US Great Plains (Lamm et al., 2012a) and Australia (Wood and Finger, 2006) and is gaining in popularity due to improved technology and crop production (Lamm et al., 2012b), despite some concerns about the economics of installation (Heard et al., 2012). There has been little adoption of this technology in the Okanagan region with a decrease from 43 to 3.2 ha between 2006 and 2014.

Other irrigation adaptation strategies include the use of deficit irrigation. In addition to the inherent efficiency found in deep-rooted grape vines, water use for wine-grape production is often controlled through planned deficits which reduce canopy development and enhance fruit quality (Keller, 2010). Use of small deficits, without detrimental effect has been reported for tree fruits (Marsal et al., 2009, 2010; Neilsen et al., 2016), but there is insufficient evidence for adoption as practical management tool to enhance fruit value and further research is required.

Adaptation to water shortages may also be addressed by drought planning, which requires the development of watershed supply information and models in addition to water demand modeling. Forecasting droughts on a seasonal basis is aided to

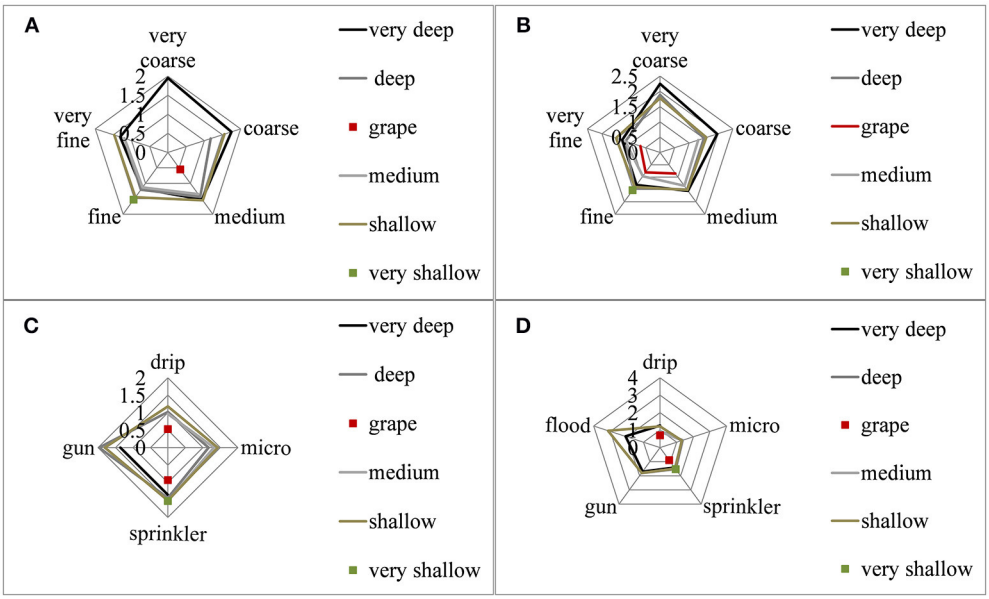


FIGURE 7 | Allocation vulnerability index (IWD/Maximum water allocation) for cropping systems in 2003 in the North Okanagan, in response to crop rooting depth and **(A)** soil texture for 2006 landuse data, **(B)** soil texture for 2014 land use data, **(C)** irrigation systems for 2006 land use data, and **(D)** irrigation system for2014 land use data.

some degree in regulated watersheds by understanding the role of storage, in particular the switch from stream to storage use and the requirements until the next freshet. Early use of stored water will often determine the imposition of watering restrictions and with early warnings producers may have some options to offset water shortages (Marsal et al., 2009, 2010; Neilsen et al., 2016). Drought planning is frequently approached by universal application of percentage reductions in water allocations in response to declared drought stages and this is often interpreted as a percentage reduction in current producer use. Cropping systems which have a low vulnerability index are already using less than the maximum allocation and likely have little room for further reductions in water use. Similarly, some systems, particularly in pasture/forage production may be chronically under-supplied unless the most efficient irrigation technology is adopted. Thus drought planning may need to take into account both a realistic development of water allocation thresholds and incentives for changes in irrigation management that go beyond the use of water pricing as the major mechanism for change (Melnychuk et al., 2016). A further option is the development of more water storage if there is sufficient water supply to fill new reservoirs.

In summary, future IWD will vary regionally across the province and in response to different GHG emission scenarios. These findings stress the importance of regional, adaptive decision-making undertaken with a long-term perspective that incorporates the latest climate science. A detailed analysis of the Okanagan Region illustrates the complexity of drought management issues. Vulnerability to a potential under-supply of water has been identified, particularly for the pasture/forage systems underpinning high value animal production which are typical of many agricultural regions in the province and may require economic support for farmers in the adoption of more efficient irrigation systems. In addition more research is required into the adoption of efficient water use practices by all water users Boland et al. (2005) documented considerable resistance to new irrigation practices in Australia for example, but the comparative land use surveys in the Okanagan region documented in this paper have shown widespread adoption of efficient irrigation systems by some sectors.

The current paper deals only with changes in annual demands in response to climate and expanded agricultural land use. To add to our understanding of future IWD, the IWDM can be used to track seasonal shifts in IWD as changes in the timing and volume of water demand have been highlighted as important not only to BC farmers, but also to the maintenance of stream flows for ecological requirements. Additionally, the IWDM could be used to examine the shifting composition of BC agriculture in response to climate change. It is unlikely that this will remain static over the coming decades as temperature and precipitation patterns shift.

The scope of this study was restricted to current and future agricultural IWD in relation to environmental changes brought about by climate change. These cannot be separated from climate change effects on water supply and other ecosystem requirements. Any decision regarding water management will also need to address requirements to provide ecosystem services as well as economic, demographic, and social drivers of water demand. Additionally, the impacts of non-agricultural water demand for both outdoor irrigation and indoor uses should also be factored into water management decisions. Such discussions necessarily will require input from both stakeholders

and decision makers. A successful model for collaborative decision making is the Okanagan Water Stewardship Council (Melnychuk et al., 2016), which has a membership of local, regional, provincial, federal and First Nations planners and technical experts. In addition there are members from academia, professional associations and sector groups (forestry, agriculture, tourism). The group provides technical support, develops policy and participates in water science projects for the regional Okanagan Basin Water Board, a regional political entity with local government members. A recent collaborative project of this group with the Okanagan Nation Alliance on environmental flow needs highlights the importance of incorporating First Nations knowledge in collaborative decision making. Beyond the regions we have reported on in this paper, there is also major competition for both agricultural land and water resources from other industries. For example, the Peace River region in Northern BC will lose land to hydro-electric power generation after the construction of the site-C dam and considerable water resources are also being used by the oil and gas industry.

AUTHOR CONTRIBUTIONS

DN and MB wrote the first draft of the article. DN and TV developed the IWDM. DN developed the vulnerability indices and analysis. SS and MB provided the GIS support. MB and IL ran the models and summarized the data. AC provided the downscaled climate data and interpretation. AW supported the land-use data collection and coordinated the Okanagan Basin Water Supply and Demand study.

ACKNOWLEDGMENTS

We would like to acknowledge funding from AAFC for supporting the development of the IWDM and co-op student support. We also acknowledge the contributions of the BC Ministry of Agriculture, the Partnership for Water Sustainability and The Okanagan Basin Water Board for funding land use data collection and climate analysis. Wenda Mason BC Ministry of Environment oversaw the funding and execution of the Okanagan Basin Water Supply and Demand study. Ron Fretwell, RHF Systems Ltd. coded the IWDM.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2018.00074/full#supplementary-material

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). *Crop Evapotranspiration Guidelines For Computing Crop Water Requirements*. FAO Irrigation and Drainage Paper 56. United Nations Food and Agriculture Organization, Rome.

Arnell, N. W. (1999). Climate change and global water resources. *Global Environ. Change* 9, S31–S49. doi: 10.1016/S.0959-3780(99)00017-5

BC Ministry of Environment (2008). *Living Water Smart: British Columbia's Water Plan*. Victoria, BC: British Columbia Ministry of Environment. Available online at: http://www.livingwatersmart.ca/docs/livingwatersmart_book.pdf

Bush, E. J., Loder, J. W., James, T. S., Mortsch, L. D., and Cohen, S. J. (2014). An overview of Canada's changing climate. Canada in a Changing Climate," in *Sector Perspectives on Impacts and Adaptation*, eds F. J. Warren and D. S. Lemmen (Ottawa, ON: Government of Canada), 23–64.

Cannon, A. J. (2015). Selecting GCM scenarios that span the range of changes in a multimodel ensemble: application to CMIP5 climate extremes indices. *J. Clim.* 28, 1260–1267. doi: 10.1175/JCLI-D-14-00636.1

Clarke, K. C., Jarosch, A. H., Anslow, F. S., Radic, V., and Menounos, B. (2015). Projected deglaciation of western Canada in the twenty-first century. *Nat. Geosci.* 8, 372–377. doi: 10.1038/ngeo2407

Cohen, S., Neilsen, D., Smith, S., Neale, T., Taylor, B., Barton, M., et al. (2006). Learning with local help: expanding the dialogue on climate change and water management in the Okanagan Region, British Columbia, Canada. *Clim. Change* 75, 331–358. doi: 10.1007/s10584-006-6336-6

Crawford, E., and Beveridge, R. (2013). *Strengthening BC's Agriculture Sector in the Face of Climate Change*. Victoria, BC: Pacific Institute for Climate Solutions. Available online at: http://pics.uvic.ca/sites/default/files/uploads/publications/Strengthening%20BC's%20Agriculture%20Sector_0.pdf

Crawford, E., and McNair, E. (2012). *BC Agriculture Climate Change Adaptation Risk and Opportunity Assessment – Provincial Report*. Victoria, BC: BC Agriculture and Food Climate Action Initiative. Available online at: http://pics.uvic.ca/sites/default/files/uploads/publications/BC%20Agriculture_summary.pdf

Environment and Climate Change Canada (ECCC) (2016). *Climate Data and Scenarios for Canada: Synthesis of Recent Observation and Modelling Results*. Gatineau, QC.

Elgaali, E., Garcia, L. A., and Ojima, D. S. (2007). High resolution modeling of the regional impacts of climate change on irrigation water demand. *Clim. Change* 84, 441–461. doi: 10.1007/s10584-007-9278-8

Fischer, G., Tubiello, F. N., Velthuizen, H., and Wiberg, D. A. (2007). Climate change impacts on irrigation water requirements: effects of mitigation, 1990–2080. *Tech. Forecast. Soc. Change* 74, 1083–1107. doi: 10.1016/j.techfore.2006.05.021

Gosling, S. N., Taylor, R. G., Arnell, N. W., and Todd, M. C. (2011). A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models. *Hydrol. Earth Syst. Sci.* 15, 279–294. doi: 10.5194/hess-15-279-2011

Grebmer, K., Ringler, C., Rosegrant, W., Olofinbiyi, T., Wiesmann, D., Fritchel, H., et al. (2012). *2012 Global Hunger Index. The Challenge of Hunger: Ensuring Sustainable Food Security Under Land, Water, and Energy Stress.* Washington, DC: International Food Policy Research Institute (IFPRI). Available online at: http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/127150/filename/127

Heard, J. W., Porker, M. J., Armstrong, D. P., Finger, L., Ho, C. K. M., Wales, W. J., et al. (2012). The economics of subsurface drip irrigation on perennial pastures and fodder production in Australia Agric. *Water Manage.* 11, 68–78. doi: 10.1016/j.agwat.2012.05.005

Hyatt, K. D., Alexander, C. A. D., and Stockwell, M. M. (2015). A decision support system for improving "fish friendly" flow compliance in the regulated Okanagan Lake and River System of British Columbia. *Can. Water Resour. J.* 40, 87–110. doi: 10.1080/07011784.2014.98551

Jiménez Cisneros, B. E. T., Oki, N. W., Arnell, G., Benito, J. G., Cogley, P., Döll, T., et al. (2014). "Freshwater resources," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, et al. (Cambridge; New York, NY: Cambridge University Press), 229–269.

Keller, M. (2010). Managing grapevines to optimise fruit development in a challenging environment: a climate change primer for viticulturists. *Aust. J. Grape Wine Res.* 16, 56–69. doi: 10.1111/j.1755-0238.2009.00077.x

Kjellen, M., and McGranshan, G. (1997). "Comprehensive assessment of the freshwater resources of the world," in *Urban Water - Towards Health and Sustainability*. Available online at: https://books.google.ca/books/about/Urban_Water.html?id=uygoAQAAMAAJ&redir_esc=y

Konzmann, M., Gerten, D., and Heinke, J. (2013). Climate impacts on global irrigation requirements under 19 GCMs, simulated with a vegetation and hydrology model. *Hydrol. Sci. J.* 58, 88–105. doi: 10.1080/02626667.2013.746495

Kreins, P., Henseler, M., Anter, J., Herrmann, F., and Wendland, F. (2015). Quantification of climate change impact on regional agricultural irrigation and groundwater demand. *Water Resourc. Manage.* 29, 3585–3600. doi: 10.1007/s11269-015-1017-8

Lamm, F. R., Bordovsky, J. P., Schwankl, L. J., Grabow, G. L., Enciso-Medina, J., Peters, R. T., et al. (2012b). Subsurface drip irrigation: status of the technology in 2010. *Trans. ASABE* 55, 483–491.

Lamm, F. R., Harmoney, K. R., Aboukheira, A. A., and Johnson, S. K. (2012a). Alfalfa production with subsurface drip irrigation in the Central Great Plains. *Trans. ASABE* 55, 1203–1212. doi: 10.13031/2013.42258

Mantua, N., Tohver, I., and Hamlet, A. (2010). Climate change impacts on streamflow extremes and summertime stream temperature and their possible consequences for freshwater salmon habitat in Washington State. *Clim. Change* 102, 187–223. doi: 10.1007/s10584-010-9845-2

Marsal, J., Lopez, G., Arbones, A., Mata, M., Vallverdu, X., and Girona, J. (2009). Influence of post-harvest deficit irrigation and pre-harvest fruit thinning on sweet cherry (cv. New Star) fruit firmness and quality. *J.Hortic. Sci. Biotech.* 84, 273–278. doi: 10.1080/14620316.2009.11512516

Marsal, J., Lopez, G., del Campo, J., Mata, M., Arbones, A., and Girona, J. (2010). Postharvest regulated deficit irrigation in 'Summit' sweet cherry: fruit yield and quality in the following season. *Irrig. Sci.* 28, 181–189. doi: 10.1007/s00271-009-0174-z

McKenney, D. W., Hutchinson, M. F., Papadopol, P., Lawrence, K., Pedlar, J., and Campbell, J. et al. (2011). Customized spatial climate models for North America. *Bull. Amer. Meteor. Soc.* 92, 1611–1622. doi: 10.1175/2011BAMS3132.1

Meinhausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., et al. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim. Change* 109, 213–241. doi: 10.1007/s10584-011-0156-z

Melnychuk, N., Jatel, N., and andWarwick Sears, A. L. (2016). Integrated water resource management and British Columbia's Okanagan Basin Water Board. *Int. J. Water Resour. Dev.* 33, 408–425. doi: 10.1080/07900627.2016.121490

Merritt, W. S., Alila, Y., Barton, M., Taylor, B., Cohen, S., and Neilsen, D. (2006). Hydrologic response to scenarios of climate change in sub watersheds of the Okanagan Basin, British Columbia. *J. Hydrol.* 326, 79–108. doi: 10.1016/j.jhydrol.2005.10.025

Neilsen, D., Neilsen, G. H., Guak, S., and Forge, T. (2016). Consequences of deficit irrigation and crop load reduction on plant water relations, yield and quality of 'Ambrosia' apple. *Hortscience* 51, 98–106.

Neilsen, D., Smith, S., Bourgeois, G., Qian, B., Cannon, A., Neilsen, G., et al. (2017). Modelling changing suitability for tree fruits in complex terrain. *Acta Hortic.* 1160, 207–214. doi: 10.17660/ActaHortic.2017.1160.30

Neilsen, D., Van Der Gulik, T., Cannon, A., Taylor, B., and Fretwell, R. (2015). Modeling future water demand for current and future climate in the Okanagan basin, B.C, Canada. *Acta Hort.* 1068, 211–218.

Neilsen, D., Smith, S., Frank, G., Koch, W., Alila, Y., Merritt, W., et al. (2006). Potential impacts of climate change on water availability for crops in the Okanagan Basin, British Columbia. *Can. J. Soil Sci.* 86, 909–924. doi: 10.4141/S05-113

Okanagan Basin Water Board (2011). *Okanagan Basin Water Supply and Demand Project*. Final summary report. Available online at: http://www.obwb.ca/wsd/

Pacific Climate Impacts Consortium (2013). *Regional Climate Summaries.* Available online at: https://www.pacificclimate.org/resources/publications/summaries?tid%5B%5D=54&keys

Parfitt, B. (2017). *Fracking, First Nations and Water*. Vancouver, BC: Canadian Centre for Policy Alternatives. Available online at: www.policy alternatives.ca

Parfitt, B., Baltutis, J., and Brandes, O. M. (2012). *From Stream to Steam: Emerging Challenges for BC's Interlinked Water and Energy Resources.* Vancouver, BC: Canadian Centre for Policy Alternatives. Available online at: www.policy alternatives.ca

Postel, S. L. (2000). Entering an era of water scarcity: the challenges ahead. *Ecol. Appl.* 10, 941–948. doi: 10.1890/1051-0761(2000)010[0941:EAEOWS]2.0.CO;2

Raskin, P., Gleick, P., Kirshen, P., Pontius, G., and Strzepek, K. (1997). *Water Futures: Assessment of Long-Range Patterns and Problems. Background Report for the Comprehensive Assessment of the Freshwater Resources of the World.* Stockholm: Stockholm Environment Institute. Available online at: http://www.sei-nternational.org/publications?pid=360

Rehana, S., and Mujumdar, P. P. (2013). Regional impacts of climate change on irrigation water demands. *Hydrol. Process.* 27, 2918–2933. doi: 10.1002/hyp.9379

Shiklomanov, I. A., and Rodda, J. C. (2003). *World Water Resources at the Beginning of the Twenty-First Century*. Cambridge: Cambridge University Press. Available online at: http://catdir.loc.gov/catdir/samples/cam034/2002031201.pdf

Shrestha, R. R., Schnorbus, M. A., Werner, A. T., and Berland, A. J. (2012). Modelling spatial and temporal variability of hydrologic impacts of climate change in the Fraser River basin, British Columbia, Canada. *Hydrol. Process.* 26, 1840–1860. doi: 10.1002/hyp.9283

The Partnership for Water Sustainability in BC (2015). *Agriculture Water Demand Model*. Available online at: http://waterbucket.ca/aw/2015/03/05/agriculture-water-demand-model/

Van der Gulik, T., Neilsen, D., and Fretwell, R. (2010). *Agriculture Water Demand Model: Report for the Okanagan Basin.* Kelowna, BC: Okanagan Basin Water Board. Available online at: http://www2.gov.bc.ca/assets/gov/farming-natural-resources-and-industry/agriculture-and-seafood/agricultural-land-and-environment/water/agriculture-water-demand-model/500300-3_agric_water_demand_model-okanagan_report.pdf

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., et al. (2013). Multimodel projections and uncertainties of irrigation water demand under climate change. *Geophys. Res. Lett.* 40, 4626–4632. doi: 10.1002/grl.50686

Walker, I. J., and Sydneysmith, R. (2008). "British Columbia," in *From Impacts to Adaptation: Canada in a Changing Climate,* eds D. S. Lemmen, F. J. Warren, J. Lacroix, and E. Bush (Ottawa, ON: Government of Canada), 334–391.

Wang, J., Klein, K. K., Bjornlund, H., Zhang, L. and Zhang, W. (2015) Adoption of improved irrigation scheduling methods in Alberta: An empirical analysis. *Can. Water Resourc. J.* 40, 47–61. doi: 10.1080/07011784.2014.975748

Wang, T., Hamann, A., Spittlehouse, D. L., and Aitken, S. N. (2006). Development of scale-free climate data for Western Canada for use in resource management. *Int. J. Climatol.* 26, 383–397. doi: 10.1002/joc.1247

Wang, X., Zhang, J., Shamsuddin, S., Guan, E., Wu, Y., Gao, J., et al. (2014). Adaptation to climate change impacts on water demand. *Mitigation Adapt. Strat. Global Change* 21, 1–19. doi: 10.1007/s11027-014-9571-6

Werner, A. T. and Cannon, A. J. (2016) Hydrologic extremes - an intercomparison of multiple gridded statistical downscaling methods. *Hydrol. Earth Syst. Sci.* 20, 1483–1508. doi: 10.5194/hess-20-1483-2016

Wood, M. L., and Finger, L. (2006). Influence of irrigation method on water use and production of perennial pastures in northern Victoria. *Aust. J. Exp. Agric.* 46, 1605–1614. doi: 10.1071/EA05197

Zhang, X., and Cai, X. (2013). Climate change impacts on global agricultural water deficit. *Geophys. Res. Lett.* 40, 1111–1117. doi: 10.1002/grl.50279

Zwiers, F. W., Schnorbus, M. A., and Maruszeczka, G. D. (2011). *Hydrologic Impacts of Climate Change on BC Water Resources*. Summary Report for the Campbell, Columbia and Peace River Watersheds, Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 May 2018
doi: 10.3389/fpls.2018.00675

Remotely Estimating Aerial N Uptake in Winter Wheat Using Red-Edge Area Index From Multi-Angular Hyperspectral Data

Bin-Bin Guo, Yun-Ji Zhu, Wei Feng*, Li He, Ya-Peng Wu, Yi Zhou, Xing-Xu Ren and Ying Ma*

State Key Laboratory of Wheat and Maize Crop Science, National Engineering Research Centre for Wheat, Henan Agricultural University, Zhengzhou, China

OPEN ACCESS

Edited by:
Tracy Ann Porcelli, Independent Researcher, Canada

Reviewed by:
Sung Kyeom Kim, Rural Development Administration, South Korea
Andrew Davidson, Government of Canada, Canada

***Correspondence:**
Yun-Ji Zhu
hnndzyj@126.com
Wei Feng
fengwei78@126.com

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *30 November 2017*
Accepted: *03 May 2018*
Published: *25 May 2018*

Citation:
Guo B-B, Zhu Y-J, Feng W, He L, Wu Y-P, Zhou Y, Ren X-X and Ma Y (2018) Remotely Estimating Aerial N Uptake in Winter Wheat Using Red-Edge Area Index From Multi-Angular Hyperspectral Data. Front. Plant Sci. 9:675. doi: 10.3389/fpls.2018.00675

Remote sensing techniques can be efficient for non-destructive, rapid detection of wheat nitrogen (N) nutrient status. In the paper, we examined the relationships of canopy multi-angular data with aerial N uptake of winter wheat (*Triticum aestivum* L.) across different growing seasons, locations, years, wheat varieties, and N application rates. Seventeen vegetation indices (VIs) selected from the literature were measured for the stability in estimating aerial N uptake of wheat under 13 view zenith angles (VZAs) in the solar principal plane (SPP). In total, the back-scatter angles showed better VI behavior than the forward-scatter angles. The correlation coefficient of VIs with aerial N uptake increased with decreasing VZAs. The best linear relationship was integrated with the optimized common indices DIDA and DDn to examine dynamic changes in aerial N uptake; this led to coefficients of determination (R^2) of 0.769 and 0.760 at the -10° viewing angle. Our novel area index, designed the modified right-side peak area index (mRPA), was developed in accordance with exploration of the spectral area calculation and red-edge feature using the equation: mRPA $= (R_{760}/R_{600})^{1/2} \times (R_{760}\text{-}R_{718})$. Investigating the predictive accuracy of mRPA for aerial N uptake across VZAs demonstrated that the best performance was at -10° [$R^2 = 0.804$, $p < 0.001$, root mean square error (RMSE) = 3.615] and that the effect was relatively similar between -20° to $+10^\circ$ ($R^2 = 0.782$, $p < 0.001$, RMSE = 3.805). This leads us to construct a simple model under wide-angle combinations so as to improve the field operation simplicity and applicability. Fitting independent datasets to the models resulted in relative error (*RE*, %) values of 12.6, 14.1, and 14.9% between estimated and measured aerial N uptake for mRPA, DIDA, and DDn across the range of -20° to $+10^\circ$, respectively, further confirming the superior test performance of the mRPA index. These results illustrate that the novel index mRPA represents a more accurate assessment of plant N status, which is beneficial for guiding N management in winter wheat.

Keywords: winter wheat, multi-angular hyperspectral, vegetation indices, aerial N uptake, monitoring model

INTRODUCTION

Nitrogen (N) is a vital element for higher photosynthetic functioning; N resource management is a major factor that can enhance plant growth and influence the quality of plant crops (Woodard and Bly, 1998; Smil, 2002). To ensure productivity, crop producers commonly supply plants in the field with N fertilizers. N supply generally appreciably surpasses plant N uptake, leading to the loss of nitrate through soil leaching, increased greenhouse gas (N_2O) emission, and ground water pollution (Sehy et al., 2003; Ju et al., 2006). To minimize potential N losses, N fertilizer should be applied at the correct time and according to the requirements of the crops. The development, therefore, of techniques which insure higher yield and good product while reducing ecological environment pollution attributed to unsuitable N application is essential.

In currently, remote sensing technique is among the most promising approach that has been shown to rapidly predict the spatial-temporal variability of crops and monitor crop growth status (Hansen and Schjoerring, 2003; Ecarnot et al., 2013). Many indices have been constructed by extracting characteristic spectral information for evaluating biochemical properties in crop plants (Hatfield et al., 2008; Ecarnot et al., 2013). Several researchers have demonstrated a close relationship between NDVI- and RVI-like spectral indices and aboveground N uptake (Mistele and Schmidhalter, 2008; Li et al., 2013), chlorophylls and carotenoids (Blackburn, 1998), canopy leaf biomass (Le Maire et al., 2008). An additional group of vegetation indices are constructed by the forms of three band combination. For instance, Wang et al. (2012) added the $2\times R_{423}$ band to the NDVI (R_{703}, R_{924}) and effectively improved the sensitivity of leaf nitrogen concentration (LNC) estimation in rice and wheat. Feng et al. (2015) showed that the three-band spectral index (R_{759}-1.8 × R_{419})/(R_{742}-1.8 × R_{419}) was a good indicator of above ground N uptake in wheat. A third group vegetation indices were developed by area-based algorithm. These include the triangle vegetation index (TVI) and modified TVI for green leaf area index (LAI) (Broge and Leblanc, 2000; Haboudane et al., 2004); the red-edge reflectance curve area for green biomass (Ren et al., 2011); and the adjusted TVI for aerial N uptake (Li et al., 2013). Additionally, remote sensing technology was also applied in the field of phenotype. Rothamsted Research reported that the data collecting from Unmanned Aerial Vehicle (UAV) based remote sensing could rapidly and accurately measure the wheat plant height and growth rate (Holman et al., 2016). Andradesanchez et al. (2014) investigated that the tractor-based phenotyping system could acquire and record data for canopy temperature, height and reflectance of cotton plants at much higher rates. However, the canopy spectral reflectance was sampled only from the vertical observation angle in prior researches, and the nadir observation were difficult to extract spatial structure from the middle and lower layers of plants (Thenkabail et al., 2000; Erdle et al., 2011).

Compared with the nadir observations, multi-angle observations contain more detailed and reliable canopy structure information that permits effective monitoring of crop N status in the middle and lower layers and provides a novel approach for quantitative remote sensing (Pocewicz et al., 2007; Huang et al., 2011). To date, many studies have shown that multi-angle measurements could improve the performance of indices when estimating the structural characteristics of ground objects (Shibayama and Wiegand, 1985; Diner et al., 1999). For instance, Galvão et al. (2009) showed that the varieties of soybean could be distinguished with the best predictive ability in the backward scattering direction. Gemmell and McDonald (2000) showed that the performance of indices under off-nadir angle can effectively discriminate forest cover and LAI. Furthermore, some studies have used to multi-angular datasets to assess plant variables, particularly biochemical components (Hasegawa et al., 2010; Huang et al., 2011). The effect of indices in estimating agronomic parameters and yields changes with the LAI and VZAs (Gemmell and McDonald, 2000; Inoue et al., 2008). Stagakis et al. (2010) focused on using satellite spectral data to estimate chlorophyll a (chl a), chl b, and carotenoids of semi-deciduous shrubs by utilizing different viewing angles and narrow-band indices. He et al. (2016) developed a multi-angular VI to enhance the estimation stability and accuracy of leaf nitrogen concentrations. No matter what the understory vegetation was (green or senesced) the relationships between canopy or total LAI and NDVI or Enhanced Vegetation Index (EVI) varied little across VZAs in pine forests (Pocewicz et al., 2007). However, only a few researches systematically tested the ability of multi-angle spectral data for predicting aerialN uptake of wheat. Taken together, these studies report the construction of a range of novel indices that use multi-angle datasets to enhance the precision and robustness of prediction indices for plant biophysical traits.

The specific aims of the paper were: to (1) study the performance of ground-based spectra and common indices to detect the aerial N uptake of winter wheat under different VZAs; (2) construct an improved, novel VI for aerial N uptake prediction; (3) compare the aerial N uptake predictive ability of the novel model with published VIs; and (4) establish the best observation angle and the best estimation model for aerial N uptake. The results of this study provide technical knowledge and a theoretical basis for monitoring N status by remote sensing technology; the information obtained from these techniques can then be used to help guide appropriate N fertilization application of wheat.

Abbreviations: ASD, Analytical Spectral Devices; CARI, Chlorophyll absorption ratio index; Chl a, Chlorphyll a; CIred-edge3, Red-edge chlorophyll index-3; DDn, New double difference index; DIDA, Double-peak areas; DVI, Difference vegetation index; EVI, Enhanced Vegetation Index; LAI, Leaf area index; LNC, Leaf N concentration; LSDR, Left-side peak area; MCARI-1, Modified chlorophyll absorption in reflectance index; mRER, Modified red-edge ratio; mRPA, Modified right-side peak area index; MSR, Modified simple ratio; mSR705, Modified red-edge ratio; MTVI1, Modified triangular vegetation index; N, Nitrogen; NAOC, Normalized area over reflectance curve index; NDDA, Normalized difference of the double-peak areas; NDRE, Normalized difference red-edge index; NDVI, Normalized Difference Vegetation Index; NIR, Near infrared region; PSSRb, Pigment specific simple ratio chlorophyll b; R^2, Coefficients of determination; RDVI, Re-normalized difference vegetation index; RE, Relative error; REFCA, Red edge reflectance curve area; RMSE, Root mean square error; RSDR, Right-side peak area; SAVI, Soil-adjusted vegetation index; SPAD, Soil and Plant Analyzer Development; SPP, Solar principal plane; TCI, Triangular chlorophyll index; TVI, Triangle vegetation index; UAV, Unmanned Aerial Vehicle; VI, Vegetation index; VIopt, Optimal vegetation index; VZA, View zenith angle.

MATERIALS AND METHODS

Experimental Fields

The field experiments were designed over a 4-year period in Zhengzhou and Shangshui city, China. The different locations, N fertilizer rates, wheat cultivars, and growth seasons were used (**Table 1**). Urea as N sources was divided into two equal doses, one administered before seeding and the rest at jointing period. Before seeding, 150 kg ha^{-1} P_2O_5 [as $Ca(H_2PO_4)_2$] and 90 kg ha^{-1} K_2O (as KCl) were used to all treatments. The N treatments with triplicates were assigned as completely random blocks in the experiment. The density of the seedlings was 3.0×10^6 plants ha^{-1}.

Data Acquisition

Measurement of Canopy Multi-Angular Hyperspectral Reflectance

Canopy reflectance spectra were obtained in a 1 m^2 area in each plot under sunny and windless conditions between11:00 to 13:00 using an ASD (Analytical Spectral Devices Inc., Boulder, CO, USA) FieldSpec Handheld spectrometer. This spectrum instrument was equipped with 25°-field-of-view optics fiber, sampling interval of 1.6 nm and spectral resolution of 3.5 nm from 325 to 1,075 nm. The multi-angle data are obtained with a Field Goniometer System, which was designed based on the system developed by Sandmeier and Itten (1999). The goniometer is a device used to position a sensor at these different angles and azimuths (**Figure 1**). The observed azimuth was fixed relative to the direction of the sun, and the measured plane was defined as the SPP (Myneni et al., 1995). The VZA was divided into backward direction (the observation direction same to the sun, −) and the forward direction (the observation direction against to the sun, +); the nadir position was defined as 0°. The VZA from backward to forward direction is −60, −50, −40, −30, −20, −10, 0, 10, 20, 30, 40, 50, 60°. The VZA become larger from 0° to ±60°, regardless of the backward and forward direction. The 10 sites were averaged to a single spectral sample of each plot. The black and base-line reflectance was calculated by a 40 × 40 cm $BaSO_4$.

Plant Measurements

The areas of 0.20 m^2 of plant samplings were randomly uprooted from each test area almost simultaneously with the canopy spectral acquisition. The plant samples were weighed after desiccation in an oven at 70°C to a constant weight. The ovendried plants were grinded into powder (1 mm) for N content analysis in laboratory. The aerial nitrogen concentration was

TABLE 1 | The experimental conditions, N fertilizer levels, and measured stages.

Exp. No.	Season and site	Cultivar	Soil characteristics	Treatments	Stages
Exp. 1	2012–2013 Zhengzhou city	Yumai 49–198 Zhengmai9694	Type: fluvo-aquic soil, Organic-M: 17.47 kg^{-1}, Soil pH (CaCl2): 7.9, Total N: 0.84 g kg^{-1}, NO_3-N: 8.1 mg kg^{-1}, Available N:78.4 mg kg^{-1}, Available P: 18.83 mg kg^{-1}, Available K: 252.56 mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1(120), N2(240), N3(360), and 50% prior to seeding and 50% at jointing	Jointing Booting Heading Anthesis
Exp. 2	2013–2014 Zhengzhou city	Yumai 49–198 Zhengmai 9694	Type: fluvo-aquic soil, Organic-M: 16.8 g kg^{-1}, Soil pH (CaCl2): 7.8, Total N: 0.89 g kg^{-1}, NO_3-N: 9.3 mg kg^{-1},Available N:113.0 mg kg^{-1}, Available P: 19.20 mg kg^{-1}, Available K: 252.30 mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1(120), N2(240), N3(360), and 50% prior to seeding and 50% at jointing	Jointing Booting Heading Anthesis
Exp. 3	2013–2014 Shangshui city	Zhoumai 27	Type: lime concretion black soil, Organic-M: 20.8 g kg^{-1}, Soil pH (CaCl2): 7.1, Total N: 1.36 g kg^{-1}, NO3-N:14.1 mg kg^{-1}, Available N: 93.2 mg kg^{-1}, Available P: 4.92 mg kg^{-1}, Available K: 176.1 mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1(180), N2 (240), N3(300), and 50% prior to seeding and 50% at jointing	Jointing Booting Anthesis
Exp. 4	2014–2015 Zhengzhou city	Yumai 49–198 Zhengmai9694	Type: fluvo-aquic soil, Organic-M: 9.7 g kg^{-1}, Soil pH (CaCl2): 8.01, Total N: 0.71 g kg^{-1} NO_3-N: 7.2 mg kg^{-1}, Available N: 64.6 mg kg^{-1}, Available P: 28.8 mg kg^{-1}, Available K: 101.7 mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1(120), N2 (240), N3(360), N4(450), and 50% prior to seeding and 50% at jointing	Jointing Booting Heading Anthesis
Exp. 5	2014–2015 Shangshui city	Yumai 49–198	Type: lime concretion black soil, Organic-M: 21.7 g kg^{-1}, Soil pH (CaCl2): 8.06, Total N: 1.13 g kg^{-1}, NO_3-N: 10.6 mg kg^{-1}, Available N: 85.7 mg kg^{-1}, Available P: 13.1 mg kg^{-1}, Available K: 111.3 mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1 (120), N2 (240), N3(360), and 50% prior to seeding and 50% at jointing	Jointing Booting Anthesis
Exp. 6	2011–2012 Zhengzhou city	Yumai 49–198 Zhengmai9694	Type: fluvo-aquic soil, Organic-M: 10.6 g kg^{-1}, Soil pH (CaCl2): 7.9, Total N: 0.91 g kg^{-1}, NO_3-N: 8.4 mg kg^{-1}, Available N: 82.0 mg kg^{-1}, Available P: 25.6 mg kg^{-1}, Available K: 124.5mg kg^{-1}	N rate (kg ha^{-1}): N0(0), N1(120), N2 (240), N3(360), and 50% prior to seeding and 50% at jointing	Booting Heading Anthesis Initial-fillingMid-filling

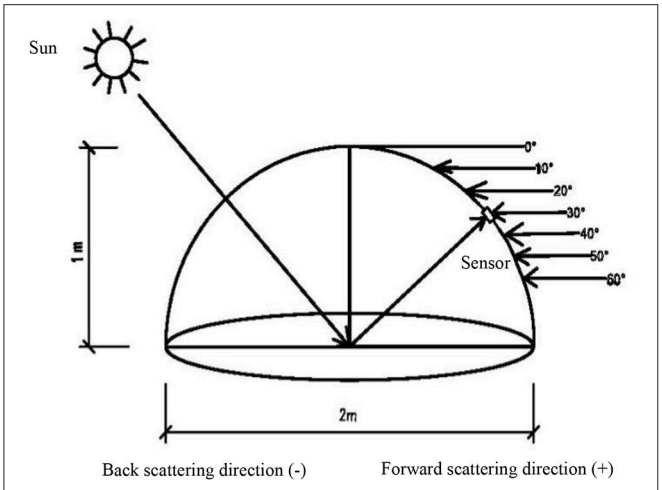


FIGURE 1 | Dimensions and design of Field Goniometer System.

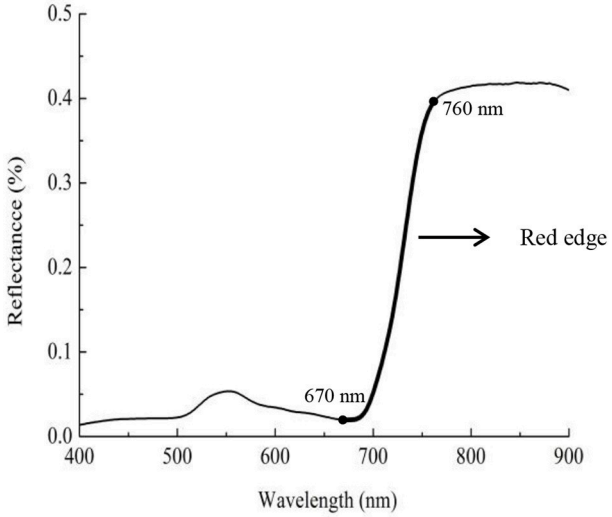


FIGURE 2 | Schematic representation of red edge reflectance curve.

determined in line with the micro-Kjeldahl method (Isaac and Johnson, 1976).

Construction of the New VI

The red edge is a region of steep variations in spectral reflectance, and this value may provide some useful information on crop growth and N status (**Figure 2**) (Sims and Gamon, 2002; Cho and Skidmore, 2006; Hatfield et al., 2008; Feng et al., 2015). To date, few area-based optimized indices have been constructed using red-edge information for non-destructive, rapid assessment of plant N status. Our preliminary research found that the double-peak area parameters constructed based on analysis of the red-edge double-peak characteristics could be effective for assessing Leaf N concentration (LNC) (Feng et al., 2014). There are several techniques to divide the red edge double-peak area into the right-side peak area (RSDR) and the left-side peak area (LSDR) (**Figure 3**). In this study, the datasets obtained using a 0° observation angle in Exp. 1–5 were used to analyze the relationship between LSDR, RSDR(from different splitting methods), and aerial N uptake. The results showed that RSDR (R_{760}-R_{718}) divided by characteristic wavelength method had the best performance ($R^2 = 0.740$, $p < 0.001$; **Figure 4**), which suggested RSDR as potential indicator for estimating aerial N uptake. The previous research showed that the ratios of two or more bands (such as R_{801}/R_{670}, R_{801}/R_{550}) could increase sensitivity to crop physiological traits and reduce variation because of external influencial factors (Daughtry et al., 2000). Haboudane et al. (2008) inserted the $(R_{700}/R_{670})^{1/2}$ into the TVI formula to decrease the combined impacts of the soil background reflectance. We, therefore, inserted a coefficient [in the form of $(\lambda_1/\lambda_2)^{1/2}$] into RSDR ($R_{760}$-$R_{718}$) to construct a novel VI called the mRPA. The spectral region of the above λ_1 and λ_2 were located within 400 and 900 nm. **Figure 5** gave a comprehensive overview of the correlation coefficients for any two band combinations (λ_1, λ_2), and this is valid for selecting the sensitive bands of aerial N uptake. The wavebands λ_1 and λ_2 ranged between 750–900 and 550–650 nm, respectively. This area

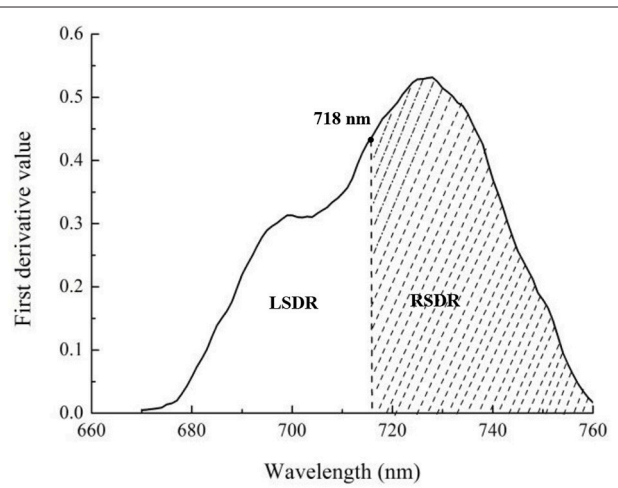


FIGURE 3 | The division of the whole red-edge double-peak area into left and right single peak areas.

had the highest precision, with R^2 values above 0.78, especially R^2 performed best when $\lambda_1 = 760$ nm, $\lambda_2 = 600$ nm. Therefore, the final formula of mRPA was:

$$\text{mRPA} = (R_{760}/R_{600})^{1/2} \times (R_{760} - R_{718}) \quad (1)$$

Model Calibration and Validation

The datasets from Exp. 1–5 were used to construct the evaluating models, and the dataset from Exp. 6 were used to validate these aerial N uptake evaluating models. The correlation of aerial N uptake and VIs was examined by MATLAB 7.0. In this study, 18 VIs were selected and summarized in **Table 2**. The quantitative relationship between the optimal VI and aerial N uptake could be established based on the highest R^2. The model's behavior was

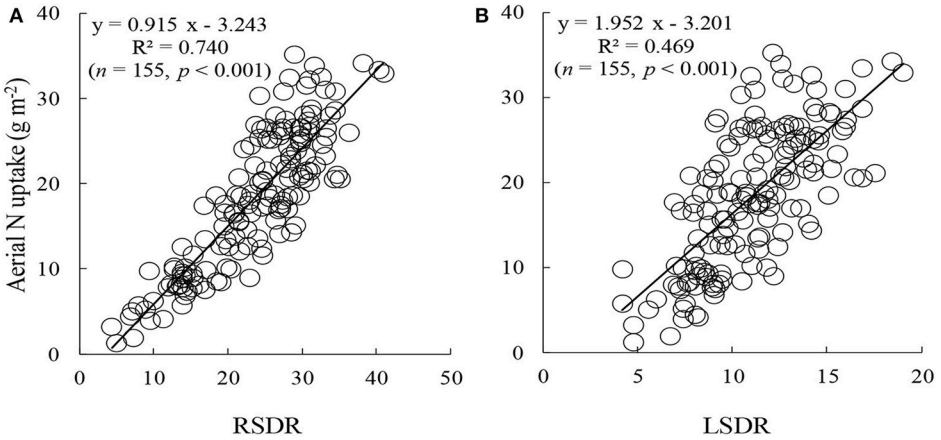


FIGURE 4 | The relationships of aerial N uptake to RSDR **(A)** and LSDR **(B)** at 0°view zenith angle.

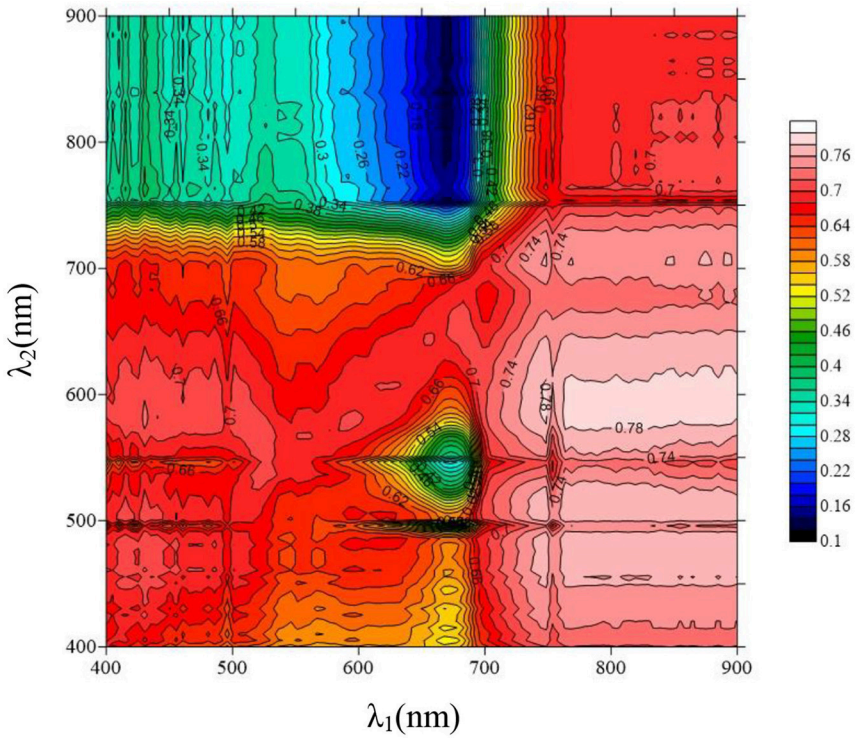


FIGURE 5 | Contour maps of coefficients of determination between aerial N uptake and the mRPA with formula $(R_{\lambda 1}/R_{\lambda 2})^{1/2} \times (R_{760}\text{-}R_{718})$ ($n = 155$, $p_{0.001} = 0.262$).

evaluated by employing the R^2, RMSE, and *RE* (%). Among the indices, those with the highest R^2 and the lowest RMSE and *RE* were considered as the best. RMSE and *RE* were calculated from actual and predicted values of samples according to Equations (2) and (3), respectively:

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (Pi - Oi)^2} \quad (2)$$

$$RE(\%) = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left(\frac{Pi - Oi}{Oi}\right)^2} \times 100\% \quad (3)$$

Here, P_i, and O_i represented the estimated and measured values, respectively, and n represented the sampling number. The prediction was considered to be excellent if $RE < 10\%$, good if $RE = 10$–20%, fair if $RE = 20$–30%, and poor if $RE > 30\%$ (Feng et al., 2014).

TABLE 2 | Summary of selected vegetation indices published in the literature.

Spectral index	Definition or equation	Reference
TWO BANDS		
Optimal vegetation index (VIopt)	$(1+0.45)\times[(R_{800})^2+1)/(R_{670}+0.45)]$	Reyniers et al., 2006
Difference vegetation index (DVI)	$R_{810}-R_{560}$	Richardson and Wiegand, 1977
Re-normalized difference vegetation index (RDVI)	$(R_{800}-R_{670})/(R_{800}+R_{670})^{1/2}$	Roujean and Breon, 1995
Pigment specific simple ratio chlorophyll b (PSSRb)	R_{800}/R_{635}	Blackburn, 1998
Modified simple ratio (MSR)	$[R_{800}/R_{670}-1]/[(R_{800}/R_{670})^{1/2}+1]$	Chen, 2014
Normalized difference red-edge index (NDRE)	$(R_{790}-R_{720})/(R_{790}+R_{720})$	Fitzgerald et al., 2006
Soil adjusted vegetation index (SAVI)	$(1-0.08)\times(R_{825}-R_{735})/(R_{825}+R_{735}-0.08)$	Huete, 1988
Red-edge chlorophyll index-3 (CIred-edge3)	$R_{790}/R_{720}-1$	Gitelson et al., 2005
THREE BANDS		
Difference index of the double-peak areas (DIDA)	$(R_{755}+R_{680}-2\times R_{718})/(R_{755}-R_{680})$	Feng et al., 2014
New double difference index (DDn)	$2\times R_{710}-R_{660}-R_{760}$	Le Maire et al., 2008
Modified chlorophyll absorption in reflectance index (MCARI-1)	$1.2\times[(2.5\times(R_{800}-R_{670})-1.3\times(R_{800}-R_{550})]$	Haboudane et al., 2004
Modified triangular vegetation index (MTVI1)	$1.2\times[1.2\times(R_{800}-R_{550})-2.5\times(R_{670}-R_{550})]$	Haboudane et al., 2004
Modified red-edge ratio (mRER)	$(R_{759}-1.8\times R_{419})/(R_{742}-1.8\times R_{419})$	Feng et al., 2015
Modified red-edge ratio (mSR705)	$(R_{750}-R_{455})/(R_{705}-R_{445})$	Sims and Gamon, 2002
Modified right-side peak area index (mRPA)	$(R_{760}/R_{600})^{1/2}\times(R_{760}-R_{718})$	This study
OVER THREE BANDS		
SDr-SDb	$\int_{680}^{760}\frac{dR\lambda}{d\lambda}d\lambda-\int_{490}^{530}\frac{dR\lambda}{d\lambda}d\lambda$	Feng et al., 2008
Triangle vegetation index (TVI-3)	$60\times(R_{nir}-R_{green})-100\times(R_{red}-R_{green})$	Broge and Leblanc, 2000
Red-edge reflectance curve area (REFCA)	$SUM(R_i/R_{780})$ i = 680–780	Ren et al., 2011

R is the reflectance at a given wavelength. R_{800}, R_{670}, R_{635},...and R_{680} are the spectral reflectance values at 800, 670, 635..., and 680 nm, respectively. R_λ is the spectral reflectance at wavelength λ.

RESULTS

Variability of Wheat Aerial N Uptake Under Different Growth Stages

The datasets from Exp. 1 are shown in **Figure 6** to illustrate the general distribution of aerial N uptake. The aerial N uptake of the two wheat cultivars increased in the vegetative period because of increasing biomass. Across the different applied N rates, the aerial N uptake of Yumai 49–198 ranged from 6.0–23.5, 7.7–28.6, 8.1–32.2, and 8.9–35.1 g kg^{-1} in the jointing, booting, heading, and anthesis stage, respectively. The aerial N uptake of Zhengmai 9694 varied from 6.9–24.3, 9.3–25.1, 10.6–27.1, and 13.4–31.2 g kg^{-1}, respectively, in these stages. With the progression of the growth stages, coefficients of variation for aerial N uptake increased. Thus, it can be seen that the aerial N uptake was significantly influenced by the different wheat cultivars and growth period.

Relationship Between Canopy Reflectance and Aerial N Uptake at Different VZAs

We plotted the correlation between canopy reflectance and aerial N uptake under different VZAs with the data from Exp. 1–5 (**Figure 7**). In the 13 VZAs, a negative relationship was detected between aerial N uptake and the reflectance in the 400–720 nm. The minimum correlation coefficient was under 560–710 nm ($r<-0.57$), caused by red valley and chlorophyll absorption. The highest *r*-value was presented in the near infrared region (NIR), and increased with decreasing VZA both in backward and forward scattering. No matter in the backward and forward scattering, the r sharply changed in the red band region (690–760 nm), and it tended to 0 near 720 nm; this indicates that *r* from this region was not sensitive to the VZA.

Relationship Between the Aerial N Uptake and Spectral Indices at Different VZAs

Canopy spectral data of different VZAs were influenced by numerous comprehensive factors including soil background, meteorological conditions, leaves, stems, and spectral noise. In this study, 18 spectral indices, including the new index and 17 published indices are presented in **Table 2**. Among them, the common indices were chose in accordance with a comprehensive literature investigation of red edge characteristics. To further describe the ability of the spectral indices in assessing aerial N uptake, we used the datasets from Exp. 1–5 to compare the predictive ability. As shown in **Table 3** the r from backward scattering directions were higher than forward scattering directions for most two-band indices (except for DVI and RDVI). To the performance of 13 single VZAs, the relatively high correlation coefficient within -40° to $+30^\circ$ was present. For the well-performance VIs, the mRPA, DIDA, and DDn were advantageous at a viewing angle of -10°, with r scores of 0.896, −0.877, and −0.872, respectively;$CI_{red-edge3}$ and mSR705 were most sensitive at the -20° viewing angle, with r scores of 0.776 and 0.771, respectively; PSSRb and MSR demonstrated the best performance at the 0° viewing angle, with r scores of 0.808 and 0.789, respectively; SDr-SDb and TVI-3 had the higher

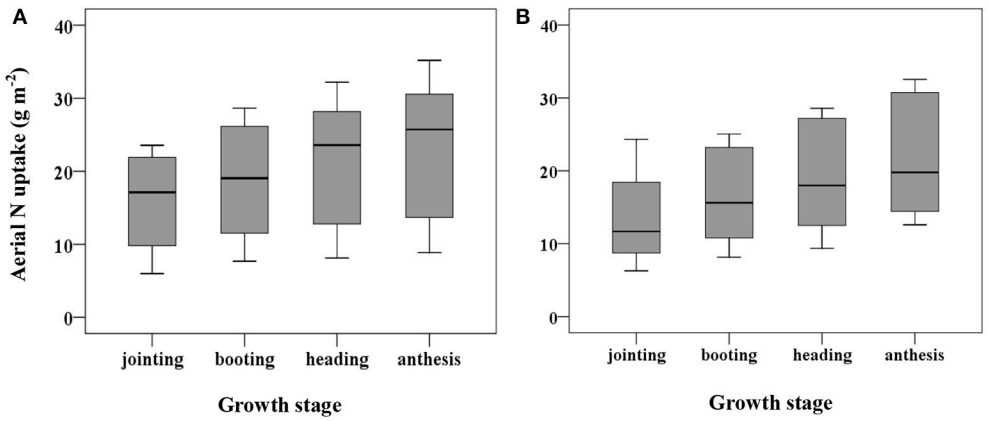


FIGURE 6 | Variation in plant N uptake over jointing-anthesis growth stages in wheat cultivars of Yumai 49-198 **(A)** and Zhengmai 9694 **(B)**.

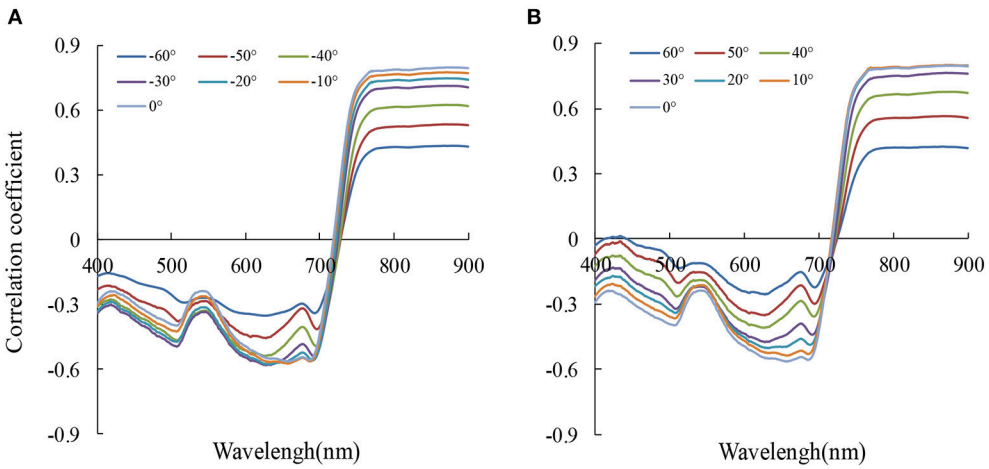


FIGURE 7 | The correlation coefficient (r) between reflectance (R_λ) and aerial N uptake at 13 viewing zenith angles (**A**: backward scattering, **B**: forward scattering) (n = 155, $p_{0.001}$ = 0.262).

correlations at the +10° viewing angle, with r scores of 0.814 and 0.809, respectively. Notably, eight of 18 indices produced the best correlations at the −10° viewing angle. These results illustrate that the around −10° VZA may be the most suitable for aerial N uptake estimation.

The four indices [mRPA, the two best-performing common indices (DIDA and DDn, and average (corresponding average value of 18 VIs at different VZAs, shown as average)] based on the R^2 values of the correlations between VIs and aerial N uptake were plotted in **Figure 8**. The results demonstrated that R^2 increased with decreasing VZA in both backward and forward scattering directions, and the highest R^2 were obtained under −20° to +10° VZAs. DIDA, DDn, and mRPA had strong correlations (R^2 > 0.72) to aerial N uptake in this region. Nevertheless, the average did not show any strong correlations with aerial N uptake (R^2 < 0.66). Compared with the average from −20° to +10° VZAs, the R^2 of mRPA, DIDA, and DDn was increased by 21.3–23.9, 15.3–18.8, and 14.4–17.6%, respectively.

Suitable Combined Angles for Aerial N Uptake Assessment Using VIs

To ascertain the suitable range of VIs to the mRPA, R^2 and RMSE were selected to compare different angle combinations. As showed in **Figure 9,** the performance of the R^2 and RMSE in the back scattering direction (R^2 = 0.426–0.804, p < 0.001) were superior to those in the forward view angles (R^2 = 0.337–0.775, p < 0.001). Among the 13 VZAs, −20° to +10° VZAs showed significantly higher predictive ability (high R^2 and small RMSE), and the most significant viewing angle was found to be −10°, with an R^2 and RMSE of 0.804 and 3.615, respectively. **Figure 10** revealed that these combinations including large VZAs combinations generated relatively poor correlations. There was a dominant region in range of −10° to 0° (R^2 = 0.796), and the predictive ability of mRPA in −20° to +10° combination was also relatively high (R^2 = 0.782). When the VZA was out of −40° to +20°, the predictive accuracy was relatively low (R^2 < 0.740). A comparison of VIs (**Figure 11**) demonstrated that

TABLE 3 | The correlation coefficient (r) for the relationships of vegetation indices with aerial N uptake at different viewing zenith angles ($n = 155$, $p_{0.001} = 0.262$).

	−60°	−50°	−40°	−30°	−20°	−10°	0°	10°	20°	30°	40°	50°	60°
TWO BANDS													
Vlopt	0.560	0.644	0.732	0.795	0.831	0.858	0.855	0.829	0.764	0.712	0.643	0.531	0.390
DVI (810,560)	0.548	0.627	0.711	0.779	0.806	0.832	0.828	0.823	0.811	0.791	0.719	0.624	0.493
RDVI (800,670)	0.535	0.591	0.672	0.743	0.781	0.827	0.825	0.808	0.803	0.763	0.694	0.599	0.475
PSSRb	0.490	0.641	0.715	0.765	0.795	0.802	0.808	0.753	0.710	0.679	0.624	0.539	0.412
MSR	0.460	0.571	0.655	0.718	0.761	0.787	0.789	0.752	0.685	0.627	0.540	0.423	0.306
NDRE	0.602	0.683	0.732	0.759	0.774	0.778	0.779	0.760	0.738	0.715	0.689	0.646	0.583
SAVI (825,735)	0.565	0.660	0.706	0.742	0.765	0.764	0.758	0.749	0.734	0.715	0.687	0.643	0.587
Clred-edge3	0.558	0.670	0.726	0.758	0.776	0.766	0.753	0.748	0.728	0.702	0.674	0.626	0.552
THREE BANDS													
DIDA	−0.699	−0.736	−0.792	−0.828	−0.853	−0.877	−0.871	−0.869	−0.842	−0.818	−0.789	−0.752	−0.672
DDn	−0.656	−0.715	−0.786	−0.827	−0.850	−0.872	−0.867	−0.865	−0.838	−0.812	−0.784	−0.724	−0.614
MTVI1	0.411	0.484	0.577	0.671	0.722	0.756	0.781	0.785	0.771	0.724	0.635	0.519	0.387
mRER	0.592	0.601	0.603	0.632	0.705	0.780	0.777	0.775	0.712	0.679	0.674	0.616	0.558
mSR705	0.504	0.643	0.719	0.753	0.771	0.769	0.749	0.713	0.674	0.653	0.612	0.534	0.407
mRPA	0.653	0.738	0.816	0.857	0.879	0.896	0.893	0.880	0.853	0.823	0.783	0.711	0.578
OVER 3 BANDS													
DD	0.656	0.689	0.759	0.807	0.835	0.868	0.862	0.857	0.847	0.821	0.772	0.701	0.588
SDr-SDb	0.486	0.573	0.661	0.740	0.775	0.787	0.798	0.814	0.802	0.762	0.678	0.570	0.434
TVI-3	0.457	0.546	0.637	0.723	0.761	0.789	0.808	0.809	0.800	0.761	0.676	0.565	0.426
REFCA	−0.609	−0.655	−0.703	−0.736	−0.754	−0.750	−0.748	−0.741	−0.704	−0.695	−0.661	−0.621	−0.566

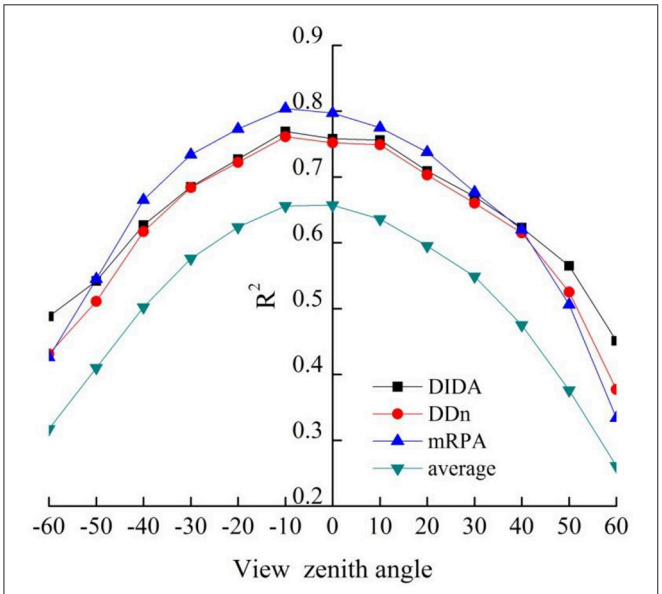


FIGURE 8 | Relationship between aerial N uptake and DIDA, DDn, mRPA, and average (corresponding average value of 18 VIs at different VZAs, shown as average) at different VZAs ($n = 155$, $p_{0.001} = 0.262$).

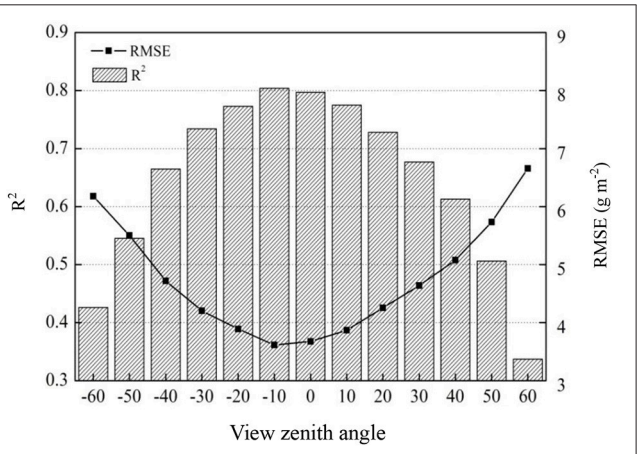


FIGURE 9 | Comparison of the prediction power of mRPA at 13 VZAs in terms of aerial N uptake ($n = 155$, $p_{0.001} = 0.262$).

the mRPA ($R^2 = 0.782$ and 0.734) at −20° to +10° and −30° to +20° VZAs were more sensitive than the two best-performing published index DIDA ($R^2 = 0.740$ and 0.712) and DDn ($R^2 = 0.726$ and 0.701). **Figure 12** showed the quantitative relationship between aerial N uptake and mRPA. The R^2 increased from 0.734 (from −30° to +20° combination) to 0.782 (from −20° to +10° combination). Compared to the −10° VZA having the highest R^2 value, mRPA in −20° to +10° combination only had a slightly decreased R^2 (2.7%) and an increased RMSE (5.2%). As a result, the novel mRPA model is the most forceful index for assessing aerial N uptake because of its insensitivity to VZAs of −20° to +10°, increasing the practicality of mRPA in actual production process.

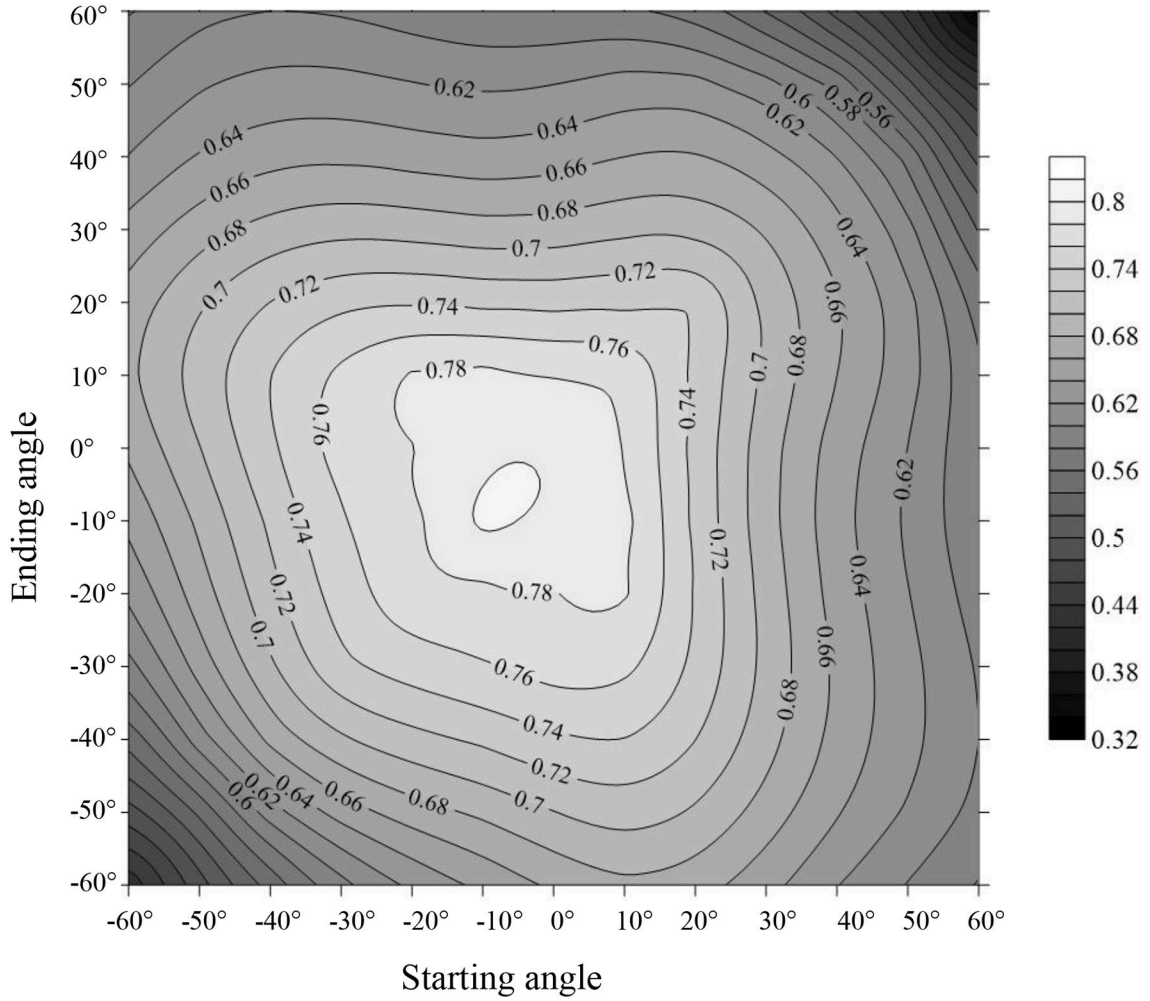


FIGURE 10 | Comparison of the predictive ability (R^2) of the indices within different view zenith angles combinations in terms of aerial N uptake.

Testing Aerial N Uptake Estimation Models

The relationship between aerial N uptake and the spectral indices (across −20° to +10° VZAs) discussed above were measured utilizing data from Exp. 6 using R^2, RMSE, and *RE* to evaluate the accuracy and applicability between observed and predicted values. Data analysis was carried out on the common best-performing VIs DIDA and DDn, and on the novel index mRPA (**Figure 13**). The DIDA and DDn showed acceptable performance in the tests, with R^2 of 0.804, *RE* of 14.1%, and a RMSE of 2.464 for DIDA and R^2 of 0.792, *RE* of 14.9%, and a RMSE of 2.554 for DDn; this also indicates that DIDA is a better indicator than DDn. Compared with these two common indices, the mRPA demonstrated the superior predictive ability of aerial N uptake, with R^2 of 0.825, *RE* of 12.1% and RMSE of 2.190. In summary, mRPA seems to be the best index for predicting aerial N uptake of winter wheat under different management conditions.

DISCUSSION

Remote sensing technology is widely used in agricultural production. It mainly includes the following aspects: crop growth measuring (e.g., biomass, N content, and yield), agricultural disaster monitoring (e.g., plant diseases and insect pests, droughts and floods) and crop phenotyping (e.g., crop height, leaf size, shape, and canopy longevity), and so on (Le Maire et al., 2008; Cao et al., 2015; He et al., 2016; Holman et al., 2016; Virlet et al., 2017), which could provide technical support for production management. However, spectral reflectance has previously been shown to be significantly affected by canopy structure, planting density, the wind, the angle of the sun, and various other factors (Rondeaux et al., 1996). To extract exact information for different characteristics and improve the detection precision, area calculations have been introduced to reduce the background effects (Broge and Leblanc, 2000; Ren et al., 2011; Li et al., 2013). Delegido et al. (2010) developed

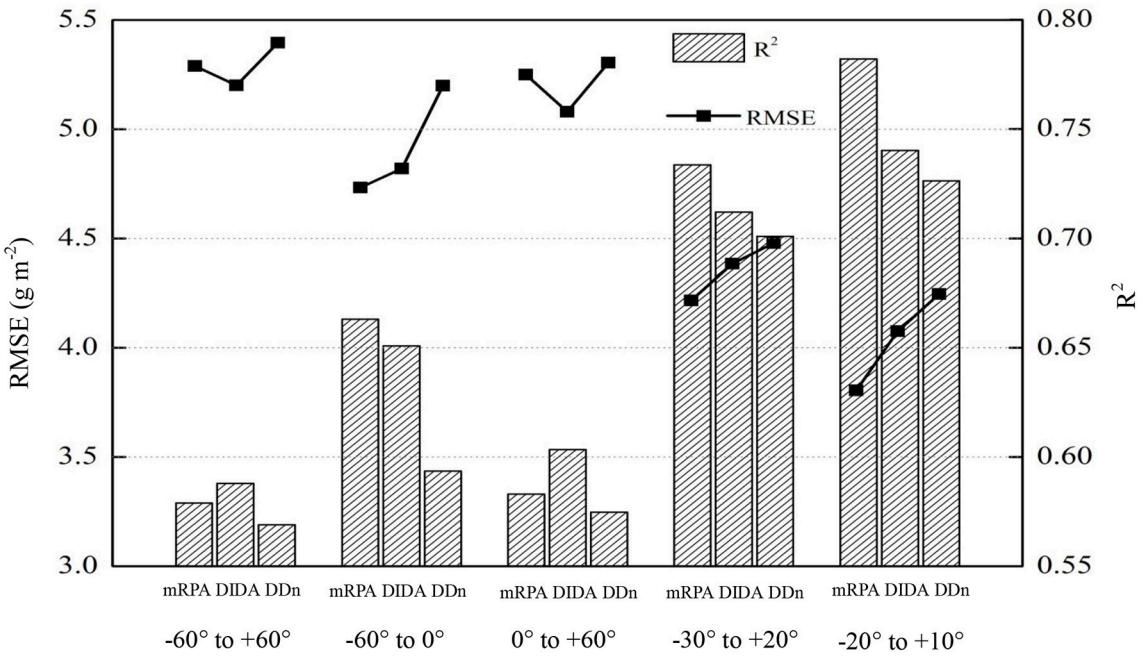


FIGURE 11 | Comparison of the predictive ability of the indices within five kinds of view zenith angles combinations in terms of aerial N uptake.

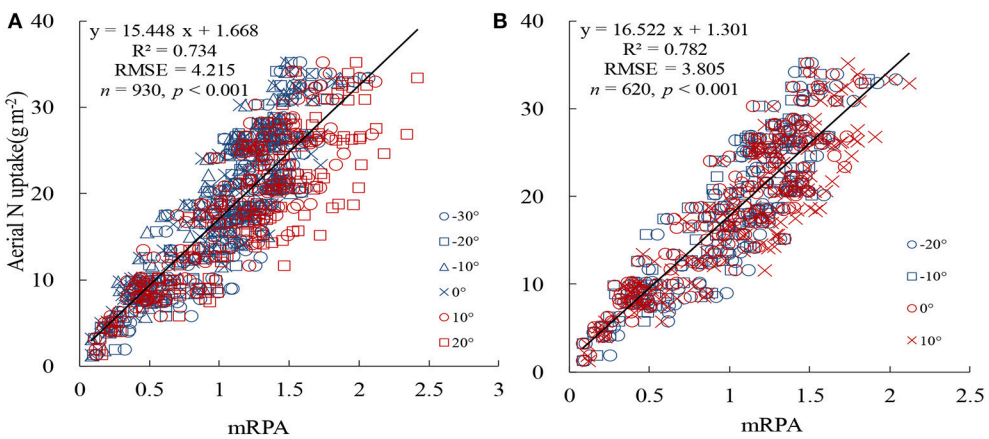


FIGURE 12 | Comparison of the prediction power of mRPA at different VZA combinations in terms of aerial N uptake. **(A)** −30° to +20° and **(B)** −20° to +10°.

the normalized area over reflectance curve index (NAOC) and this index showed a linear relation with chlorophyll content. The normalized difference of the double-peak areas (NDDA) was calculated and discovered to correlate strongly with LNC (Feng et al., 2014). Broge and Leblanc (2000) found that the TVI (constructed by the area under the concave curve of red light absorption) could effectively assess the chlorophyll content and LAI. The chlorophyll absorption integral infers chlorophyll concentration through calculating the surrounding area between a connecting line of 600 and 735 nm and the red edge (Oppelt and Mauser, 2004). In this study, five of 17 VIs have previously been constructed using the area-based algorithm. Among them, DIDA, SDr-SDb, and TVI had significant correlations with aerial N uptake, with r-values of −0.877, 0.814, and 0.809, respectively, at their advantageous viewing angles (−10, 10, and 10°, respectively). The best-performing common index DIDA had strong correlations (R^2 > 0.74) to aerial N uptake within −20° to +10° VZAs. In addition, DIDA showed acceptable performance in the tests, with R^2 of 0.804, *RE* of 14.1% and a RMSE of 2.464. The above results showed that the VIs constructed by area algorithm could potentially be used to precisely predict aerial N uptake.

Many vegetation indices have been also constructed, including those NDVI-, RVI-, and DVI-like spectral indices, or other

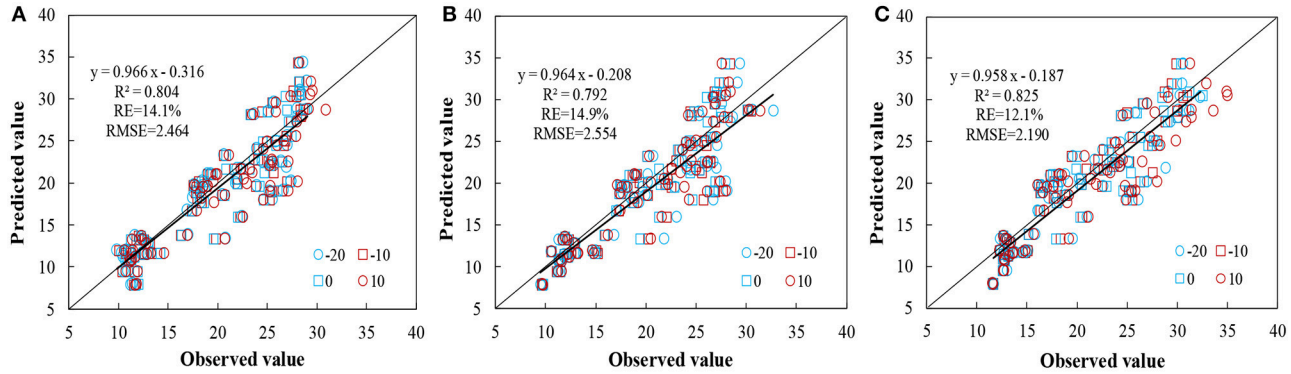


FIGURE 13 | Comparison between estimated and measured aerial N uptake based on DIDA **(A)**, DDn **(B)**, and mRPA **(C)** for −20° to +10° combinations ($n = 252$, $p < 0.001$).

derived functions to enhance accuracy of estimating models (Huete, 1988; Wang et al., 2012). Baret and Guyot (1991) and Rondeaux et al. (1996) constructed the transformed soil-adjusted vegetation index (TSAVI) and optimized SAVI (OSAVI) by adding soil line parameters into NDVI to decrease the sensitivity of the soil background reflectance at low LAI. In order to reduce the combined influences of the canopy non-photosynthetic materials and increase the sensitivity of chlorophyll concentration determination, Daughtry et al. (2000) added the R_{700}/R_{670} to the chlorophyll absorption ratio index (CARI) to obtain modified CARI. Haboudane et al. (2008) brought the $(R_{700}/R_{670})^{1/2}$ into the TVI to construct the triangular chlorophyll index (TCI) to increase its sensitivity of chlorophyll changes. These studies inspired us to develop a novel index by adding a coefficient. We attempted to derive the coefficient by combining different bands in a square root form. Finally, R_{760} and R_{600} were selected from different combinations of bands and changed into the form of $(R_{760}/R_{600})^{1/2}$. This was integrated with RSDR (760,718) to construct the mRPA, and this novel VI had a high correlation coefficient within −20° to +10° VZAs ($R^2 = 0.782$).The mRPA makes the best of area algorithms and red-edge information, and effectively improves the monitoring accuracy of aerial N uptake.

The vegetation indices displayed anisotropy depending on the canopy structural development, shadowing, the view angles of the sensors, the inherent viewing geometry of sensors, and in some respects the underlying soil (Kimes et al., 1985). Vegetation indices merging into multiple VZAs had the potency to further enhance anisotropy estimation. Pocewicz et al. (2007) took full advantage of the hotspot effect in the backscatter direction to improve quantitative estimation of LAI. Galvão et al. (2009) highlighted that the back-scattering direction was suited to predict the yield of soybean. He et al. (2016) showed that the novel VI constructed by a four-band VI from two angles (−20° and +10°) was sensitive to the change of the LNC in wheat. In this study, the back scattering direction improved indices performance for aerial N uptake compared with the forward scattering direction. The main reason may be that back scattering observations contain more signals from sunlit branches or leaves with higher reflectance values, while forward-scatter observations derive mostly from shady branches/leaves with lower reflectance values (Stagakis et al., 2010). In addition, our results showed that the R^2 increased with decreasing view zenith angles in back and forward-scatter direction. This was mainly because that the spectral data obtained at small angles mainly includes the total plant characteristics (lower, middle, and upper) of wheat. In general, sampling involving upper, middle, and lower wheat leaf layers could determine the aerial N uptake of the target region. In summary, further survey analysis on the variations of relationships between canopy attributes and remote sensing observations and on the availability off-nadir in fetching information is recommended.

Multi-angular remote sensing is able to obtain three-dimensional vegetation structure information, and thus it is better than vertical measurement for monitoring the canopy structural properties and the biochemical component of ground objects (Pocewicz et al., 2007). Rautiainen et al. (2008) demonstrated that high VZAs are the best fit for detecting over-story LAI values because of the quite limited influence of bottom layer on the whole signal. The value of canopy chlorophyll inversion index(CCII) at ±50 and ±60°, ±30 and ±40°, and nadir, ±20 and ±30°, VZAs were selected for inverting the chlorophyll at the upright upper, middle and bottom layer (Huang et al., 2011). Song et al. (2016) showed that −40° VZA was suitable to effectively monitor the LNC of wheat. This study showed that the R^2 of mRPA changed strongly with VZAs, which the highest R^2 value was found at −10° VZA ($R^2 = 0.804$). The mRPA within −10° could improve the monitoring accuracy of aerial N uptake. However, it is not convenient to accurately control measuring angle under off-nadir conditions in the field operation. Our research found that mRPA was relatively insensitive between −20° and +10° VZAs region for R^2 changes. A comparison among spectral parameters demonstrated that performance of mRPA in −20° to +10° combined dataset was superior to the two better published indices, with R^2 of 0.782 for mRPA, 0.750 for DIDA, and 0.736 for DDn, respectively. Compared to the most sensitive VZA (−10°), mRPA within −20° to +10° VZAs only had a slightly reduced R^2 (2.7%). This allowed

us to construct a unified model in variable VZAs range to assess aerial N uptake in wheat, which decrease the influence of the VZAs and increase the field operation simplicity and applicability in a wide-angle region using portable monitors. Therefore, it is vital to select appropriate VI formulas and VZAs, which could reduce the variance due to soil background and crop canopy structure. In summary, the novel index plays an important role in predicting aerial N uptake of wheat and could be utilized to more precisely regulate N fertilization rate for different cultivation sites and plant types.

In this study, the novel index mRPA had the higher predictive ability with range of -20° to $+10^\circ$. This can not only provide optimized parameters for the development of the portable monitor, but also offer dynamic information for guiding precise N budgeting. In order to obtain higher yield and avoid wasting resources, it is important to consider crop-N demand as well as soil-N supply to optimum N fertilizer strategy (Ju et al., 2009; Hartmann et al., 2015). The N fertilizer requirement was calculated using the following formula: $N_{req} = (N_{target} - N_{uptake} - N_{soil})/f_{NUE}$, where N_{target} is the total crop-N demand for a target yield and grain protein, calculated according to Angus (2001), N_{uptake} is the aerial N uptake, N_{soil} is the potential soil-N supply for the rest of the growing period, and f_{NUE} is the fertilizer-N use efficiency. The mRPA model developed in this study could effectively estimate N_{uptake}, which will contribute to managing the N application in winter wheat. The prediction power of VIs was affected by cultivation factors. Only if the VI was seldom influenced by the factors of cultivation, the applicability of the model was strong. We synthesized dataset from the vegetative growth stages to develop a unified model that could be easily used to assess the N status. However, this research was designed only on winter wheat in Henan province, the dependability and adaptability of this novel model ought to be tested in other crops and areas.

CONCLUSIONS

Timely assessment of aerial N uptake is important to diagnose crop N status, maximizing yields and minimizing disadvantageous environmental impacts. In this study, we compared the use of 18 VIs, including 17 common VIs and a novel index constructed in this study, to estimate aerial N uptake of wheat. The results demonstrated that back scattering observation angles improved the ability to predict aerial N uptake compared with forward-scatter viewing angles VZAs. To decrease the restrictions on the environmental conditions and to further explore the superiority of spectral information, we combined the advantages of red-edge characteristics and area-based algorithms to construct a novel index mRPA to illuminate dynamic changes in aerial N uptake. The novel VI have the characteristic of simplicity and reliability and could be developed according to the formula: $mRPA = (R_{760}/R_{600})^{1/2} \times (R_{760}\text{-}R_{718})$. Compared with the best-performing traditional indices DIDA and DDn, the predictive ability of mRPA at -10° view angle was effectively enhanced by 4.6–5.8%. Further systematic analysis of VZA combinations showed that mRPA had the best forecasting ability when compared with the traditional indices, with small difference between combinations of -20° to $+10^\circ$. This has guided us in the development of a unified model for forecasting the aerial N uptake of wheat across a wide angle range; this will increase the precision of N predictions under a range of angles using portable monitors. The integrated index mRPA was shown to be practical and exact for aerial N uptake evaluation of winter wheat. This result will be also beneficial for choosing appropriate VZA and for the construction of more precise sensors for ecosystem monitoring. Nonetheless, it is also need to further validate the reliability and stability of the novel VI and to examine its effectiveness under the condition of different cultivation environment.

AUTHOR CONTRIBUTIONS

B-BG, WF, and Y-JZ conceived the research. B-BG, Y-PW, YZ, and X-XR performed the experiments. B-BG and WF wrote the paper. LH and YM contributed to the results analysis and discussion.

FUNDING

This work was supported by the Thirteenth Five-year Plan of National Key Research Project of China [grant number 2016YFD0300604], the National Natural Science Foundation of China [grant number 31671624; 31571607], the Program of Science&Technology Innovation Talents in Universities of Henan Province, China [grant number 17HASTIT036].

REFERENCES

Andradesanchez, P., Gore, M. A., Heun, J. T., Thorp, K. R., Carmosilva, A. E., French, A. N., et al. (2014). Development and evaluation of a field-based high-throughput phenotyping platform. *Funct. Plant Biol.* 41, 68–79. doi: 10.1071/FP13126

Angus, J. F. (2001). Nitrogen supply and demand in Australian agriculture. *Aust. J. Exp. Agric.* 41, 277–288. doi: 10.1071/EA00141

Baret, F., and Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35, 161–173. doi: 10.1016/0034-4257(91)90009-U

Blackburn, G. A. (1998). Quantifying chlorophylls and caroteniods at leaf and canopy Scales. *Remote Sens. Environ.* 66, 273–285. doi: 10.1016/S0034-4257(98)00059-5

Broge, N. H., and Leblanc, E. (2000). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* 76, 156–172. doi: 10.1016/S0034-4257(00)00197-8

Cao, X., Luo, Y., Zhou, Y., Fan, J., Xu, X., West, J. S., et al. (2015). Detection of powdery mildew in two winter wheat plant densities and prediction of grain yield using canopy hyperspectral reflectance. *PLoS ONE* 10:e121462. doi: 10.1371/journal.pone.0121462

Chen, J. M. (2014). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Can. J. Remote Sens.* 22, 229–242. doi: 10.1080/07038992.1996.10855178

Cho, M. A., and Skidmore, A. K. (2006). A new technique for extracting the red edge position from hyperspectral data: the linear extrapolation method. *Remote Sens. Environ.* 101, 181–193. doi: 10.1016/j.rse.2005.12.011

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., Colstoun, E. B. D., and Iii, M. M. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239. doi: 10.1016/S0034-4257(00)00113-9

Delegido, J., Alonso, L., González, G., and Moreno, J. (2010). Estimating chlorophyll content of crops from hyperspectral data using a normalized area over reflectance curve (NAOC). *Int. J. Appl. Earth Obs. Geoinf.* 12, 165–174. doi: 10.1016/j.jag.2010.02.003

Diner, D. J., Asner, G. P., Davies, R., Knyazikhin, Y., Muller, J. P., Nolin, A. W., et al. (1999). New directions in earth observing: scientific applications of multiangle remote sensing. *Bull. Am. Meteorol. Soc.* 80, 2209–2228. doi: 10.1175/1520-0477(1999)080<2209:NDIEOS>2.0.CO;2

Ecarnot, M., Compan, F., and Roumet, P. (2013). Assessing leaf nitrogen content and leaf mass per unit area of wheat in the field throughout plant cycle with a portable spectrometer. *Field Crops Res.* 140, 44–50. doi: 10.1016/j.fcr.2012.10.013

Erdle, K., Mistele, B., and Schmidhalter, U. (2011). Comparison of active and passive spectral sensors in discriminating biomass parameters and nitrogen status in wheat cultivars. *Field Crops Res.* 124, 74–84. doi: 10.1016/j.fcr.2011.06.007

Feng, W., Guo, B.-B., Wang, Z. J., He, L., Song, X., Wang, Y. H., et al. (2014). Measuring leaf nitrogen concentration in winter wheat using double-peak spectral reflection remote sensing data. *Field Crops Res.* 159, 43–52. doi: 10.1016/j.fcr.2014.01.010

Feng, W., Guo, B. B., Zhang, H. Y., He, L., Zhang, Y. S., Wang, Y. H., et al. (2015). Remote estimation of above ground nitrogen uptake during vegetative growth in winter wheat using hyperspectral red-edge ratio data. *Field Crops Res.* 180, 197–206. doi: 10.1016/j.fcr.2015.05.020

Feng, W., Zhu, Y., Tian, Y. C., Cao, W. X., Yao, X., and Li, Y. X. (2008). Monitoring leaf nitrogen accumulation with hyper-spectral remote sensing in wheat. *Acta Ecol. Sin.* 28, 23–32. doi: 10.1016/S1872-2032(08)60018-9

Fitzgerald, G. J., Rodriguez, D., Christensen, L. K., and Belford, R. (2006). Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments. *Precis. Agric.* 7, 233–248. doi: 10.1007/s11119-006-9011-z

Galvão, L. S., Roberts, D. A., Formaggio, A. R., Numata, I., and Breunig, F. M. (2009). View angle effects on the discrimination of soybean varieties and on the relationships between vegetation indices and yield using off-nadir Hyperion data. *Remote Sens. Environ.* 113, 846–856. doi: 10.1016/j.rse.2008.12.010

Gemmell, F., and McDonald, A. J. (2000). View zenith angle effects on the forest information content of three spectral indices. *Remote Sens. Environ.* 72, 139–158. doi: 10.1016/S0034-4257(99)00086-3

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., and Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* 32, 93–114. doi: 10.1029/2005GL022688

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90, 337–352. doi: 10.1016/j.rse.2003.12.013

Haboudane, D., Tremblay, N., Miller, J. R., and Vigneault, P. (2008). Remote estimation of crop chlorophyll content using spectral indices derived from hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 46, 423–437. doi: 10.1109/TGRS.2007.904836

Hansen, P. M., and Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* 86, 542–553. doi: 10.1016/S0034-4257(03)00131-7

Hartmann, T. E., Yue, S. C., Schulz, R., He, X. K., Chen, X. P., Zhang, F. S., et al. (2015).Yield and N use efficiency of a maize–wheat cropping system as affected by different fertilizer management strategies in a farmer's field of the North China Plain. *Field Crops Res.* 174, 30–39. doi: 10.1016/j.fcr.2015.01.006

Hasegawa, K., Matsuyama, H., Tsuzuki, H., and Sweda, T. (2010). Improving the estimation of leaf area index by using remotely sensed NDVI with BRDF signatures. *Remote Sens. Environ.* 114, 514–519. doi: 10.1016/j.rse.2009.10.005

Hatfield, J. L., Gitelson, A. A., Schepers, J. S., and Walthall, C. L. (2008). Application of spectral remote sensing for agronomic decisions. *Agron. J.* 100, 117–131. doi: 10.2134/agronj2006.0370c

He, L., Zhang, H. Y., Zhang, Y. S., Song, X., Feng, W., Kang, G. Z., et al. (2016). Estimating canopy leaf nitrogen concentration in winter wheat based on multi-angular hyperspectral remote sensing. *Eur. J. Agron.* 73, 170–185. doi: 10.1016/j.eja.2015.11.017

Holman, F. H., Riche, A. B., Michalski, A., Castle, M., Wooster, M. J., and Hawkesford, M. J. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sens.* 8, 1–23. doi: 10.3390/rs8121031

Huang, W., Wang, Z., Huang, L., and Lamb, D. W. (2011). Estimation of vertical distribution of chlorophyll concentration by bi-directional canopy reflectance spectra in winter wheat. *Precis. Agric.* 12, 165–178. doi: 10.1007/s11119-010-9166-5

Huete, A. R. (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* ,25, 295–309. doi: 10.1016/0034-4257(88)90106-X

Inoue, Y., Penuelas, J., Miyata, A., and Mano, M. (2008). Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO_2 flux measurements in rice. *Remote Sens. Environ.* 112, 156–172. doi: 10.1016/j.rse.2007.04.011

Isaac, R. A., and Johnson,W. C. (1976). Determination of total nitrogen in plant tissue using a bloc digestor. *Assoc. Offic. Analyt. Chem.* 59, 98–100.

Ju, X. T., Kou, C. L., Zhang, F. S., and Christie, P. (2006). Nitrogen balance and groundwater nitrate contamination: comparison among three intensive cropping systems on the North China Plain. *Environ. Pollut.* 143, 117–125. doi: 10.1016/j.envpol.2005.11.005

Ju, X. T., Xing, G. X., Chen, X. P., Zhang, S. L., Zhang, L. J., Liu, X. J., et al. (2009). Reducing environmental risk by improving N management in intensive Chinese agricultural systems. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3046–3051. doi: 10.1073/pnas.0813417106

Kimes, D. S., Newcomb, W. W., Tucker, C. J., Zonneveld, I. S., Wijngaarden, W. V., Leeuw, J. D., et al. (1985). Directional reflectance factor distributions for cover types of Northern Africa. *Remote Sens. Environ.* 18, 1–19. doi: 10.1016/0034-4257(85)90034-3

le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J. Y., Bréda, N., et al. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sens. Environ.* 112, 3846–3864. doi: 10.1016/j.rse.2008.06.005

Li, F., Mistele, B., Hu, Y., Chen, X., and Schmidhalter, U. (2013). Comparing hyperspectral index optimization algorithms to estimate aerial N uptake using multi-temporal winter wheat datasets from contrasting climatic and geographic zones in China and Germany. *Agric. For. Meteorol.* 180, 44–57. doi: 10.1016/j.agrformet.2013.05.003

Mistele, B., and Schmidhalter, U. (2008). Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. *Eur. J. Agron.* 29, 184–190. doi: 10.1016/j.eja.2008.05.007

Myneni, R. B., Maggion, S., Iaquinta, J., Privette, J. L., Gobron, N., Pinty, B., et al. (1995). Optical remote sensing of vegetation: modeling, caveats, and algorithms. *Remote Sens. Environ.* 51, 169–188. doi: 10.1016/0034-4257(94)00073-V

Oppelt, N., and Mauser, W. (2004). Hyperspectral monitoring of physiological parameters of wheat during a vegetation period using AVIS data. *Int. J. Remote Sens.* 25, 145–159. doi: 10.1080/0143116031000115300

Pocewicz, A., Vierling, L. A., Lentile, L. B., and Smith, R. (2007). View angle effects on relationships between MISR vegetation indices and leaf area index in a recently burned ponderosa pine forest. *Remote Sens. Environ.* 107, 322–333. doi: 10.1016/j.rse.2006.06.019

Rautiainen, M., Lang, M., Mõttus, M., Kuusk, A., Nilson, T., Kuusk, J., et al. (2008). Multi-angular reflectance properties of a hemiboreal forest: an analysis using CHRIS PROBA data. *Remote Sens. Environ.* 112, 2627–2642. doi: 10.1016/j.rse.2007.12.005

Ren, H., Zhou, G., and Zhang, X. (2011). Estimation of green aboveground biomass of desert steppe in Inner Mongolia based on red-edge reflectance curve area method. *Biosyst. Eng.* 109, 385–395. doi: 10.1016/j.biosystemseng.2011.05.004

Reyniers, M., Walvoort, D. J. J., and Baardemaaker, J. D. (2006). A linear model to predict with a multi-spectral radiometer the amount of nitrogen in winter wheat. *Int. J. Remote Sens.* 27, 4159–4178. doi: 10.1080/01431160600791650

Richardson, A. J., and Wiegand, C. L. (1977). Distinguishing Vegetation from Soil background information. *Photogramm. Eng. Rem. Sens.* 43, 1541–1552.

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. doi: 10.1016/0034-4257(95)00186-7

Roujean, J. L., and Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* 51, 375–384. doi: 10.1016/0034-4257(94)00114-3

Sandmeier, S. R., and Itten, K. I. (1999). A field goniometer system (FIGOS) for acquisition of hyperspectral BRDF data. *IEEE Trans. Geosci. Rem. Sens.* 37, 978–986. doi: 10.1109/36.752216

Sehy, U., Ruser, R., and Munch, J. C. (2003). Nitrous oxide fluxes from maize fields: relationship to yield, site-specific fertilization, and soil conditions. *Agric. Ecosyst. Environ.* 99, 97–111. doi: 10.1016/S0167-8809(03)00139-7

Shibayama, M., and Wiegand, C. L. (1985). View azimuth and zenith, and solar angle effects on wheat canopy reflectance. *Remote Sens. Environ.* 18, 91–103. doi: 10.1016/0034-4257(85)90040-9

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Smil, V. (2002). Nitrogen and food production: proteins for human diets. *Ambio* 31, 126–131. doi: 10.1579/0044-7447-31.2.126

Song, X., Xu, D. Y., He, L., Feng, W., Wang, Y. H., Wang, Z. J., et al. (2016).Using multi-angle hyperspectral data to monitor canopy leaf nitrogen content of wheat. *Precis. Agric.* 17, 721–736. doi: 10.1007/s11119-016-9445-x

Stagakis, S., Markos, N., Sykioti, O., and Kyparissis, A. (2010). Monitoring canopy biophysical and biochemical parameters in ecosystem scale using satellite hyperspectral imagery: an application on a *Phlomis fruticosa* Mediterranean ecosystem using multiangular CHRIS/PROBA observations. *Remote Sens. Environ.* 114, 977–994. doi: 10.1016/j.rse.2009.12.006

Thenkabail, P. S., Smith, R. B., and Pauw, E. D. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sens. Environ.* 71, 158–182. doi: 10.1016/S0034-4257(99)00067-X

Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., and Hawkesford, M. J. (2017). Field scanalyzer: an automated robotic field phenotyping platform for detailed crop monitoring. *Funct. Plant Biol.* 44, 143–153. doi: 10.1071/FP16163

Wang, W., Yao, X., Yao, X. F., Tian, Y. C., Liu, X. J., Ni, J., et al. (2012). Estimating leaf nitrogen concentration with three-band vegetation indices in rice and wheat. *Field Crops Res.* 129, 90–98. doi: 10.1016/j.fcr.2012.01.014

Woodard, H. J., and Bly, A. (1998). Relationship of nitrogen management to winter wheat yield and grain protein in South Dakota. *J. Plant Nutr.* 21, 217–233. doi: 10.1080/01904169809365397

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

HYPOTHESIS AND THEORY
published: 01 September 2017
doi: 10.3389/fpls.2017.01482

Preliminary Studies to Characterize the Temporal Variation of Micronutrient Composition of the Above Ground Organs of Maize and Correlated Uptake Rates

Karla Vilaça Martins[1], Durval Dourado-Neto[1], Klaus Reichardt[2], Quirijn de Jong van Lier[2], José Laércio Favarin[1], Felipe Fadel Sartori[1], Guilherme Felisberto[1] and Simone da Costa Mello[1]*

[1] Crop Science Department, Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ), Universidade de São Paulo, Piracicaba, Brazil, [2] Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, Brazil

OPEN ACCESS

Edited by:
Nathaniel Newlands,
Government of Canada, Science and Technology, Canada

Reviewed by:
Yinsuo Zhang,
Agriculture and Agri-Food Canada, Canada
Louis Kouadio,
University of Southern Queensland, Australia
Kurt A. Rosentrater,
Iowa State University, United States

***Correspondence:**
Durval Dourado-Neto
ddourado@usp.br

Specialty section:
This article was submitted to Plant Biophysics and Modeling, a section of the journal Frontiers in Plant Science

Received: 05 January 2017
Accepted: 10 August 2017
Published: 01 September 2017

Citation:
Martins KV, Dourado-Neto D, Reichardt K, de Jong van Lier Q, Favarin JL, Sartori FF, Felisberto G and Mello SC (2017) Preliminary Studies to Characterize the Temporal Variation of Micronutrient Composition of the Above Ground Organs of Maize and Correlated Uptake Rates. Front. Plant Sci. 8:1482. doi: 10.3389/fpls.2017.01482

The improvement of agronomic practices and the use of high technology in field crops contributes for significant increases in maize productivity, and may have altered the dynamics of nutrient uptake and partition by the plant. Official recommendations for fertilizer applications to the maize crop in Brazil and in many countries are based on critical soil nutrient contents and are relatively outdated. Since the factors that interact in an agricultural production system are dynamic, mathematical modeling of the growth process turns out to be an appropriate tool for these studies. Agricultural modeling can expand our knowledge about the interactions prevailing in the soil-plant-atmosphere system. The objective of this study is to propose a methodology for characterizing the micronutrient composition of different organs and their extraction, and export during maize crop development, based on modeling nutrient uptake, crop potential evapotranspiration and micronutrient partitioning in the plant, considering the production environment. This preliminary characterization study (experimental growth analysis) considers the temporal variation of the micronutrient uptake rate in the aboveground organs, which defines crop needs and the critical nutrient content of the soil solution. The methodology allowed verifying that, initially, the highest fraction of dry matter, among aboveground organs, was assigned to the leaves. After the R_1 growth stage, the largest part of dry matter was partitioned to the stalk, which in this growth stage is the main storage organ of the maize plant. During the reproductive phase, the highest fraction of dry matter was conferred to the reproductive organs, due to the high demand for carbohydrates for grain filling. The micronutrient (B, Cu, Fe, Mn, and Zn) content follows a power model, with higher values for the initial growth stages of development and leveling off to minimum values at the R_6 growth stage. The proposed model allows to verify that fertilizer recommendations should be related to the temporal variability of micronutrient absorption rates, in contrast to the classic recommendation based

on the critical soil micronutrient content. The maximum micronutrient absorption rates occur between the reproductive R_4 and R_5 growth stages. These evaluations allowed to predict the maximum micronutrient requirements, considered equal to respective stalk sap concentrations.

Keywords: *Zea mays*, micronutrient content, micronutrient partition, productivity, agricultural modeling

INTRODUCTION

Maize (*Zea mays* L.) is the most cultivated cereal worldwide, mainly because of the different ways of consumption, as human food and animal feed, as well as many byproducts for the high technology industry (Edwards, 2009). Brazil is the world's third largest maize producer, behind the United States and China. In Brazil, maize is general grown in two cropping seasons, and covered an area of 15,627,300 ha in the 2014/2015 season. The national average yield triplicated from 1,632 kg ha^{-1} to 5,382 kg ha^{-1} over the last 40 years (CONAB, 2015). Such increases in productivity are due to the development of agriculture in relation to the breeding of plants and management practices, including the correction and fertilization of soils (Bender et al., 2013; Ciampitti et al., 2013). However, the application of micronutrients also played an important role, although the information on their absorption and partition by the maize plant rely on older literature as stated by Ciampitti et al. (2013). The most recent studies on the subject have been carried out mainly in the United States. Few studies were performed on the absorption and nutrient partitioning in modern maize hybrids used in Brazil (Von Pinho et al., 2009). The Brazilian lime and fertilizer recommendations are mainly based on 15–20-year-old studies such as van Raij et al. (1996), Ribeiro et al. (1999), Oliveira (2003), SBCS (2004) and Cantarutti et al. (2007).

In addition, agricultural production systems have also changed, with higher plant densities, reduced seed spacing, use of agrochemicals for crop protection and use of transgenic hybrids (Bender et al., 2013).

The use of increasingly growing high-tech crops may have changed the dynamics of absorption and partition of nutrients by the maize crop. Therefore, studies on the current absorption patterns and partition of micronutrients are welcome to update official fertilizer recommendations, which are still based on soil chemical analyses. This is essentially a static approach, whereas processes during crop development are dynamic. For example, using only soil chemical analyses does not allow to consider variations of the critical micronutrient content among phenological growth stages, expected productivity and soil and climatic interactions.

The proposed model will be useful for the improvement of the traditional fertilizer methodology based on soil analysis, giving emphasis to the plant as a nutrient extractor. The model considers that the fertilizer recommendation should be based on the temporal variability of the nutrient absorption rate, in comparison to the classic recommendation based on the critical soil nutrient content.

This study is based on the following hypotheses: (i) the maximum micronutrient concentration in the sap depends on productivity and transpiration, (ii) the micronutrient content in the different organs is characterized by a power function and does not depend on productivity, and (iii) the nutrient with concentration in the sap equal to the required critical concentration of the crop, limits productivity (Liebig Law).

This preliminary study aims to propose a methodology for characterizing the composition of different organs and extraction, distribution and export of the micronutrients boron (B), copper (Cu), iron (Fe), manganese (Mn), and zinc (Zn) during maize plant development. Based on modeling micronutrient uptake, crop potential evapotranspiration and micronutrient partition in the plant, taking into account the micronutrient uptake rate in a given production system, it contributes to an improvement in the recommendation of these micronutrients.

MATERIALS AND METHODS

Environmental Conditions

A field experiment with maize was carried out in Piracicaba, state of São Paulo, Brazil (22° 41′ S; 47° 38′ W, 546 m above sea level) to characterize the temporal variation of above ground dry matter accumulation and micronutrient contents from sowing until physiologic maturity.

The climate is of the Köppen Cwa type (Alvares et al., 2013), with a rainy summer and dry winter, annual average air temperature 21.4°C and yearly rainfall 1,257 mm. The reference evapotranspiration (ETo, mm d^{-1}) was calculated by the Penman-Monteith method (Allen et al., 1998), and the water balance was established according to Thornthwaite and Mather (1955).

The soil was classified as a typical Hapludox as defined by the USDA Soil Taxonomy (Soil Survey Staff, 1975). The micronutrients B (determined in hot-water-soluble method developed by Berger and Truog, 1940), Cu, Fe, Mn and Zn (determined in DTPA pH 7.3 method developed by Lindsay and Norvell, 1978) concentrations in the soil were, respectively, 0.32, 4.0, 15.0, 16.8, and 2.5 mg dm^{-3}.

Cropping System Characterization

Maize was sown on 26 March 2013, using a population of 65,000 plants ha^{-1} (spacing between rows of 0.45 m). The maize simple hybrid DKB 390 VT PRO™ 2 was chosen due to its favorable features, specifically: (i) high potential productivity, (ii) *YieldGard* technology (tolerance to *Spodoptera frugiperda*, *Helicoverpa zea,* and *Diatraea saccharalis*) and (iii) RR technology (tolerance to the glyphosate herbicide).

For dry matter composition characterization, a homogeneous single plot of 5,000 m^2 was sown and managed in the same way

applying nitrogen (30 kg ha^{-1} of N), phosphorus (80 kg ha^{-1} of P_2O_5) and potassium (40 kg ha^{-1} of K_2O). An additional of 90 kg ha^{-1} of N was applied at the V_4 phenological stage.

The seed treatment consisted of insecticide and fungicide applications (Fipronil, Pyraclostrobin and Thiophanate-methyl at a rate of 200 mL per 100 kg of seeds).

Sampling Description

The plot was subdivided into 315 parcels used for sampling of the above ground plant parts. Each parcel of 12.6 m^2 consisted of four maize lines 7 m long, the central ones used for plant sampling. With this large number of parcels, it was possible to randomly sample only two plants per parcel during the complete cycle of the crop.

Samplings consisted of plant collection at times according to the growth stages defined by Ritchie et al. (1996), as follows: V_2, V_4, V_6, V_8, and V_{10}, which occurred at 14, 21, 28, 35, and 42 days after seeding (*t, d*), respectively (**Table 1**). Sixty plants were collected at each sampling date, two per plot, one of each central line, using 30 parcels chosen randomly over the whole plot. The 60 sampled plants were subdivided, also randomly, into six replicates (composed samples) of 10 plants each for dry matter and chemical analyses.

At V_{12}, V_{15}, R_1, R_2, R_2, R_3, R_4, R_4, R_5, R_5, and R_6, which occurred at 50, 56, 70, 77, 84, 91, 104, 111, 118, 127, and 139 days, respectively (**Table 1**), the number of harvested plants was reduced to 30 plants per sampling. Therefore, the resulting composed samples consisted of five plants.

In this way, during the experimental period, 630 plants were sampled, corresponding to 1.93% of all plants. To determine the crop development growth stage, phenological characterization was performed every 2 days during the crop cycle, according to Ritchie et al. (1996).

Plant organ samples were dried at 65°C to characterize leaf, stalk, tassel, ear, straw, style-stigma and total dry matter. Subsamples were used for micronutrient analyses.

The leaf area was evaluated with a LI-COR® sensor (model Li-3100C, Lincoln, Nebraska, USA) allowing leaf area index (LAI, m^2 m^{-2}) estimation during crop development. LAI was determined in all growth stages, with six replicates, using the same leaves for dry matter and chemical analyses.

Harvest was performed at physiologic maturity (R_6 growth stage), collecting all plants of the central two lines of 7 m, discarding 0.5 m at each border. Grain yield was estimated from the weight of 1,000 seeds at 13% water content.

Basic Hypothesis of the Micronutrient Absorption Model

Considering that at a given time t (*d*) within the crop cycle, plants have accumulated a mass of dry matter per area *D* (kg ha^{-1}), with a nutrient content *N* (mg kg^{-1}), the cumulative nutrient absorption *A* (kg ha^{-1}) is given by the product of *D* and *N*.

Using observed data for parameterization, *D* (sigmoid function) and *N* (power function) were modeled as a function of time (*t*) using the following empirical equations:

$$D = k_1 + \frac{k_2 {k_4}^2}{{k_4}^2 + \left(t - k_3\right)^2} \quad (1)$$

$$N = k_5 . t^{k_6} \quad (2)$$

in which k_1 (kg ha^{-1}), k_2 (kg ha^{-1}), k_3 (d), k_4 (d), k_5 (mg kg^{-1} d^{-k_6}) and k_6 are empirical fitting parameters calibrated from experimental data of *D* and *N* by minimizing the sum of square errors.

The development of the general model is based on the growth curve of the maize plant given by the accumulation of the total dry matter, the sigmoidal Equation (1). For the temporal changes of the micronutrient content in the plant, the power function 2 was chosen.

Based on the dry matter production curve *D* and micronutrient content *N* of the above ground plant, a model was

TABLE 1 | Description of sampling (S) date (D), growth stages (GS), accumulated degree-days (DD, °Cd) and relative development (R_d, %) based on DD of the maize crop (hybrid DKB 390 VT PRO 2), during vegetative and reproductive phases, as a function of time (t, d), from March 26 to August 12, 2013.

S	Vegetative phase					S	Reproductive phase				
	t	D	GS[a]	DD[b]	R_d		t	D	GS[a]	DD[b]	R_d
–	0	March 26	Seeding	–	–	8	70	June 4	R_1	863	54.1
–	7	April 2	V_E	0	0.0	9	77	June 11	R_2	943	59.1
1	14	April 9	V_2	215	13.5	10	84	June 18	R_2/R_3	1,021	64.0
2	21	April 16	V_4	307	19.2	11	91	June 25	R_3	1,097	68.7
3	28	April 23	V_6	385	24.1	12	104	July 8	R_4	1,249	78.3
4	35	April 30	V_8	474	29.7	13	111	July 15	R_4/R_5	1,325	83.0
5	42	May 7	V_{10}	567	35.5	14	118	July 22	R_5	1,411	88.4
6	50	May 15	V_{12}	653	40.9	15	127	July 31	R_5/R_6	1,462	91.6
7	56	May 21	V_{15}	728	45.6	16	139	August 12	R_6	1,596	100.0

[a] *Ritchie et al. (1996).*

[b] *DD: 10°C as the lower base temperature and 35°C as the upper base temperature.*

proposed to characterize the nutrient absorption A:

$$A = D.N = k_1 k_5 . t^{k_6} + \frac{k_2 {k_4}^2 k_5 . t^{k_6}}{{k_4}^2 + (t - k_3)^2} \quad (3)$$

The shape of the sigmoidal curve represents a positive increase in dry matter accumulation with increasing rates in the vegetative growth stages and with decreasing rates in the reproductive growth stages.

Application of the Proposed Model

To calibrate the absorption rate curve (λ, mg ha^{-1} d^{-1}) (**Figure 1**) for each of the micronutrients, the following equation was used:

$$\lambda = \frac{dA}{dt} = k_1 k_5 . k_6 t^{k_6 - 1} + k_2 {k_4}^2 k_5 \left\{ \frac{k_6 t^{k_6 - 1}}{\left[{k_4}^2 + (t - k_3)^2 \right]} - \frac{2t^{k_6} (t - k_3)}{\left[{k_4}^2 + (t - k_3)^2 \right]^2} \right\} \quad (4)$$

Starting from the first derivative of the absorption rate of a micronutrient λ (or the second derivative of the absorption march – A, Equation 3):

$$f(t) = \frac{d\lambda}{dt} = \frac{d^2 A}{dt^2} \quad (5)$$

The maximum micronutrient absorption rate (λ_{max}) can be determined according to $f(t) = 0$. It is assumed to be related to the critical micronutrient content in the soil solution. The maximum absorption rate λ_{max} (maximum maize crop demand at $t = t_{max} = t_{i+1}$) can be found at time t, in the iteration "i+1" (**Figure 1**), minimizing the sum of square errors, corresponding to $d^2A/dt^2 = 0$ using the iterative Newton-Raphson method:

$$t_{i+1} = t_i - \frac{f(t_i)}{f'(t_i)} \quad (6)$$

$$f(t) = C_1 t^{k_6 - 2} + C_2 t^{k_6 - 2} \left\{ \frac{k_6}{T_2} - \frac{2t^2 T_1}{{T_2}^2} \right\} + \frac{C_2}{(k_6 - 1)} \left\{ \frac{(-2k_6 - 4t) T_1 - 2t^2}{{T_2}^2} - \frac{8t^2 {T_2}^2}{{T_2}^3} \right\} t^{k_6 - 1} \quad (7)$$

where C_1 [$C_1 = k_1 k_5 k_6 (k_6 - 1)$], C_2 [$C_2 = k_2 {k_4}^2 k_5 (k_6 - 1)$], T_1 [$T_1 = t - k_3$] and T_2 [$T_2 = {k_4}^2 + (t - k_3)^2$] are auxiliary variables.

$$f'(t) = C_1 (k_6 - 2) t^{k_6 - 3} + C_2 (k_6 - 2) t^{k_6 - 3} \left(\frac{\frac{k_6}{T_2} - \frac{2t^2 T_1}{{T_2}^2}}{} \right) + C_2 t^{k_6 - 2} \left\{ \frac{\frac{-2k_6 T_1}{{T_2}^2} - \frac{(4tT_1 + 2t^2) {T_2}^2 - 8t^2 {T_1}^2 T_2}{{T_2}^4}}{} \right\} + C_2 t^{k_6 - 2} \left\{ \frac{\frac{(-2k_6 - 4t) T_1 - 2t^2}{{T_2}^2} - \frac{8t^2 {T_1}^2}{{T_2}^3}}{} \right\} + \frac{C_2}{(k_6 - 1)} t^{k_6 - 1} \left\{ \frac{2\left[-2T_1 - k_6 - 2t - t^2\right] T_2 + 8(k_6 + 2t) {T_1}^2 + 8t^2 T_1}{{T_{2_i}}^3} - \frac{16\left[t{T_1}^2 + t^2 T_1\right] T_2 - 48t^2 {T_1}^3}{{T_2}^4} \right\} \quad (8)$$

If $|f(t_{i+1})| < \varepsilon$ then $t_{max} = t_{i+1}$, where ε is the maximum allowed error (10^{-7}). The maximum absorption rate λ_{max} corresponds to the soil supplying capacity (λs, mg ha^{-1} d^{-1}) that meets the micronutrient demand during the entire crop cycle (**Figure 1**), so that:

$$\lambda_{max} = \lambda (t_{max}) = \lambda s = k_1 k_5 . k_6 {t_{max}}^{k_6 - 1} + k_2 {k_4}^2 k_5 \left\{ \frac{k_6 {t_{max}}^{k_6 - 1}}{{k_4}^2 + (t_{max} - k_3)^2} - \frac{2{t_{max}}^{k_6} (t_{max} - k_3)}{\left[{k_4}^2 + (t_{max} - k_3)^2 \right]^2} \right\} \quad (9)$$

To estimate the maximum concentration Cc of a micronutrient in the maize stalk sap at time t ($t = t_{i+1}$) of λ_{max} (when $\lambda = \lambda s$), first the water flow rate absorbed by the roots is taken equal to the sum of the actual transpiration (Ta, mm d^{-1}) and the absorbed water needed for daily plant growth (α, mm d^{-1}).

Since micronutrient flux, here considered as λ_{max} for each micronutrient, is the product of the water flux and the maximum sap concentration (Cc, mg L^{-1}), we have:

$$Cc = \frac{\lambda_{max}}{10^4 (Ta + \alpha)} \quad (10)$$

where α is the daily absorbed water by the plant to form dry matter and the daily water retained by the plant (kg m^{-2} d^{-1}).

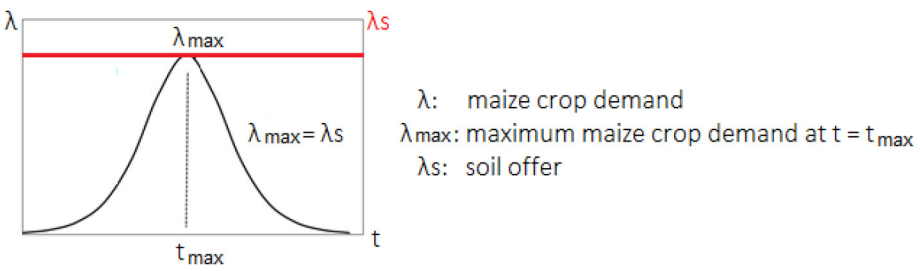


FIGURE 1 | Micronutrient absorption rate (λ, mgha^{-1}d^{-1}) as a function of time (t, d), for the determination of the maximum absorption rate (λ_{max}, mgha^{-1}d^{-1}) corresponding to the rate of the soil supplying capacity (λs, mgha^{-1}d^{-1}), considered sufficient during the crop cycle for a given productivity.

The maximum sap concentration ($Cc*$) can also be expressed in mmol L^{-1} dividing the term $\lambda_{max}/[10^4(Ta+\alpha)]$ by M, where M is the molecular mass (g mol^{-1}) of the considered micronutrient (B, Cu, Fe, Mn, and Zn).

Statistical Analyses

All statistical analyses (regressions, model fitting) in this study were performed using the software *Table Curve 2D*, version 5.01 (Systat Software, 2000).

RESULTS

Field Experiment

From March 26 to August 12, 2013, total precipitation was 409.6 mm (**Figure 2**). From seeding until the growth stage V_4, there was no significant drought and from V_4 to V_{15} growth stages (20–56 days) a drought period occurred. The grain productivity of the maize crop was 10,335 kg ha^{-1} (13% of seed water content).

Climatic Conditions and Soil Water Storage

During the crop cycle, minimum (Tn), average (Td), and maximum (Tx) air temperatures oscillated, respectively, between 6.1 and 21.7°C (mean value: 14.6°C), 11.2 and 26.3°C (mean value: 20.7°C) and 14.6 and 35.7°C (mean value: 28.4°C). The actual evapotranspiration (ETa) varied between 0.1 and 1.9 mm d^{-1} most of the time (mean value: 0.96 mm d^{-1}) (**Figure 2**).

During the dry spell, soil water storage reduced from 46.4 mm (soil water holding capacity—SWHC) to about 12.0 mm (**Figure 2**). From visual field observations, it was concluded that the crop water stress was not severe. Between 57 and 121 days, water supply by rain allowed a normal development of the crop (V_{15} to R_5 growth stage, **Figure 2**). After 121 days, there was no more rain until maturity (R_6 growth stage) and the harvest could be performed under excellent conditions.

Leaf Area Index and Dry Matter Accumulation

Positive increments of dry matter were observed from the onset of growth and development until the end of the vegetative phase (R_1 growth stage) (**Table 1**), when the total dry matter reached $7.3{\cdot}10^3$ kg ha^{-1} (**Table 2**).

In relation to leaf area index (LAI, m^2 m^{-2}), at the start of the crop cycle (day 14, V_2 growth stage), its value was 0.06 m^2 m^{-2} and at 70 days (R_1 growth stage), it expanded to 3.98 m^2 m^{-2}. After this, at flowering, the leaf area index continued practically constant until 127 days (R_5/R_6 growth stage), with a significant drop at 139 days (R_6 growth stage) presenting 2.90 m^2 m^{-2} (**Table 2**).

Total dry matter D increased until the physiologic maturity (R_6 growth stage), at 139 days with 23,069 kg ha^{-1} (354.9 g $plant^{-1}$) for a population of 65,000 plants ha^{-1} (**Table 2**). The

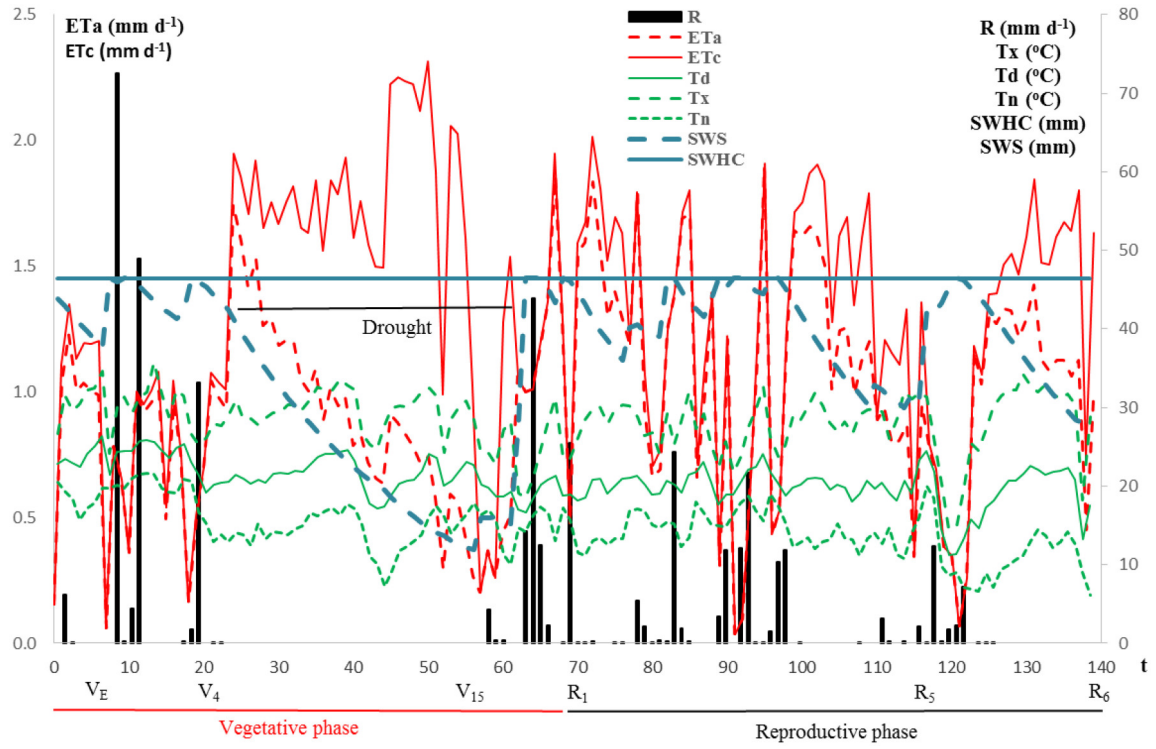


FIGURE 2 | Minimum (*Tn*), average (*Td*) and maximum (*Tx*) air temperatures (°C), rainfall (R, mm d^{-1}), potential (*ETc*, mm d^{-1}), and actual (*ETa*, mm d^{-1}) evapotranspiration, soil water holding capacity (SWHC, mm), and soil water storage (SWS, mm) as a function of time (t, d), from March 26 to August 12, 2013 (Time zero corresponds to sowing).

TABLE 2 | Average values of leaf area index (LAI, m^2m^{-2}) and dry matter ($kgha^{-1}$) of leaf, stalk, tassel, ear, style-stigma, straw and total of the maize crop (hybrid DKB 390 VT PRO 2) as a function of time (t, d).

t	LAI	Dry matter (kg ha^{-1})						
		Leaf	Stalk	Tassel	Ear	Style-stigma	Straw	Total
14	0.06	18.2	7.8	.	.	.	.	26.0
21	0.28	100.1	59.2	.	.	.	.	159.3
28	0.75	286.0	203.5	.	.	.	.	489.5
35	1.56	655.2	1,007.5	.	.	.	.	1,662.7
42	2.47	1,302.6	1,106.3	.	.	.	.	2,408.9
50	3.35	2,031.9	2,779.4	.	.	.	.	4,811.3
56	3.93	2,782.0	3,773.3	.	.	.	.	6,555.3
70	3.98	2,429.7	4,153.5	418.0	37.1	22.8	234.0	7,295.0
77	3.86	2,329.6	4,982.9	226.9	133.9	65.0	535.6	8,273.9
84	3.85	2,302.3	5,635.5	158.6	468.7	96.2	969.2	9,630.4
91	3.69	2,191.8	4,961.5	152.8	1,120.0	122.2	1,253.9	9,802.0
104	3.47	2,103.4	4,828.9	146.3	3,396.9	65.7	1,494.4	12,035.4
111	3.86	2,374.5	6,926.4	169.7	6,019.7	252.9	1,996.8	17,739.8
118	3.59	2,589.6	6,630.7	153.4	8,457.2	85.2	2,148.9	20,064.9
127	3.79	2,810.6	6,496.1	145.6	10,030.2	55.3	2,217.2	21,754.9
139	2.90	2,555.8	6,479.9	152.8	11,659.1	152.8	2,069.0	23,069.2

fit of total dry matter accumulation to the growth model resulted in an R^2 of 0.97.

The Pearson correlation coefficient value (0.985) was larger than the critical value, at 10% of significance level, for the dry matter sigmoid model of the maize crop (hybrid DKB 390 VT PRO 2) as a function of time (t, d). The empirical parameters are $k_1 = -5{,}776.147\,\text{kg ha}^{-1}$, $k_2 = 30{,}474.954\,\text{kg ha}^{-1}$, $k_3 = 151{,}749$ days and $k_4 = 68.397$ days.

Micronutrient Content

Observed micronutrient contents (B, Cu, Fe, Mn, and Zn) in leaf and stalk were initially high, decreasing to a stable value near the end of the cycle (**Figure 3A**). In most cases, nutrient contents in the leaf were higher than in the stalk, except for Mn, which showed a higher concentration in the stalk until day 112 (between R_4/R_5 and R_5 growth stages). The concentration of Zn was slightly lower in leaves at the very early stages, but from 60 days (between V_{15} and R_1 growth stages), its concentration in leaf and stalk decreased continuously.

In the stalk, Mn contents practically did not vary over time. Iron was the micronutrient with highest contents in the leaf and in the stalk, being higher in leaf.

For the tassel and style-stigma, in general, all micronutrients presented a concentration increase over time, except B in the style-stigma. In the straw, B, Cu, and Zn concentrations decreased over time, and Fe and Mn increased over time (**Figure 3B**).

The Mn content in straw was constant. Concentrations of Cu, Fe, Mn, and Zn in the tassel were higher during the final growth stages of the crop. B followed by Cu are the nutrients that present highest contents in the style-stigma. Furthermore, when comparing the three organs, one can observe greater contents of B and Cu in the style-stigma and of Fe and Mn in the tassel (**Figure 3B**).

In relation to the tassel, the micronutrient contents are lower at beginning and increase with crop development, with highest levels between days 111 and 118 (between R_4/R_5 and R_5 growth stages).

In relation to the straw, the Fe and Mn contents are constant during the time of collection and, the Zn contents are lower in the final development growth stages of maize plant (**Figure 3B**). When comparing the three organs, higher micronutrient contents in the style-stigma and tassel in comparison to straw are observed (**Figure 3B**).

From 104 days on (R_4 growth stage), it was observed that the micronutrient contents are higher in the ear when compared to contents in the grain (**Figure 3C**).

The mineral elements in the ear (cob + grain) were analyzed from 70 (R_1 growth stage) to 91 days (R_3 growth stage) (**Figure 3C**).

At 104 days (R_4 growth stage), the analysis was made in separate for the cob and the grain (**Figure 3C**). The content of micronutrients is higher in the early ear development and decreases until the end of the evaluation period, 91 days (R_3 growth stage).

As in the case of the ear, the levels of micronutrients in the cobs and grain decrease toward the end of the crop cycle (**Figure 3C**). However, the Fe and Mn levels are constant in cob and the levels of Fe and Zn are also fairly constant for the grain.

After the R_4 growth stage, at 104 days, the Fe contents were higher in the grain when compared to the cob (**Figure 3C**).

At the R_1 growth stage (70 days), higher B, Fe, and Mn contents were found in the ear.

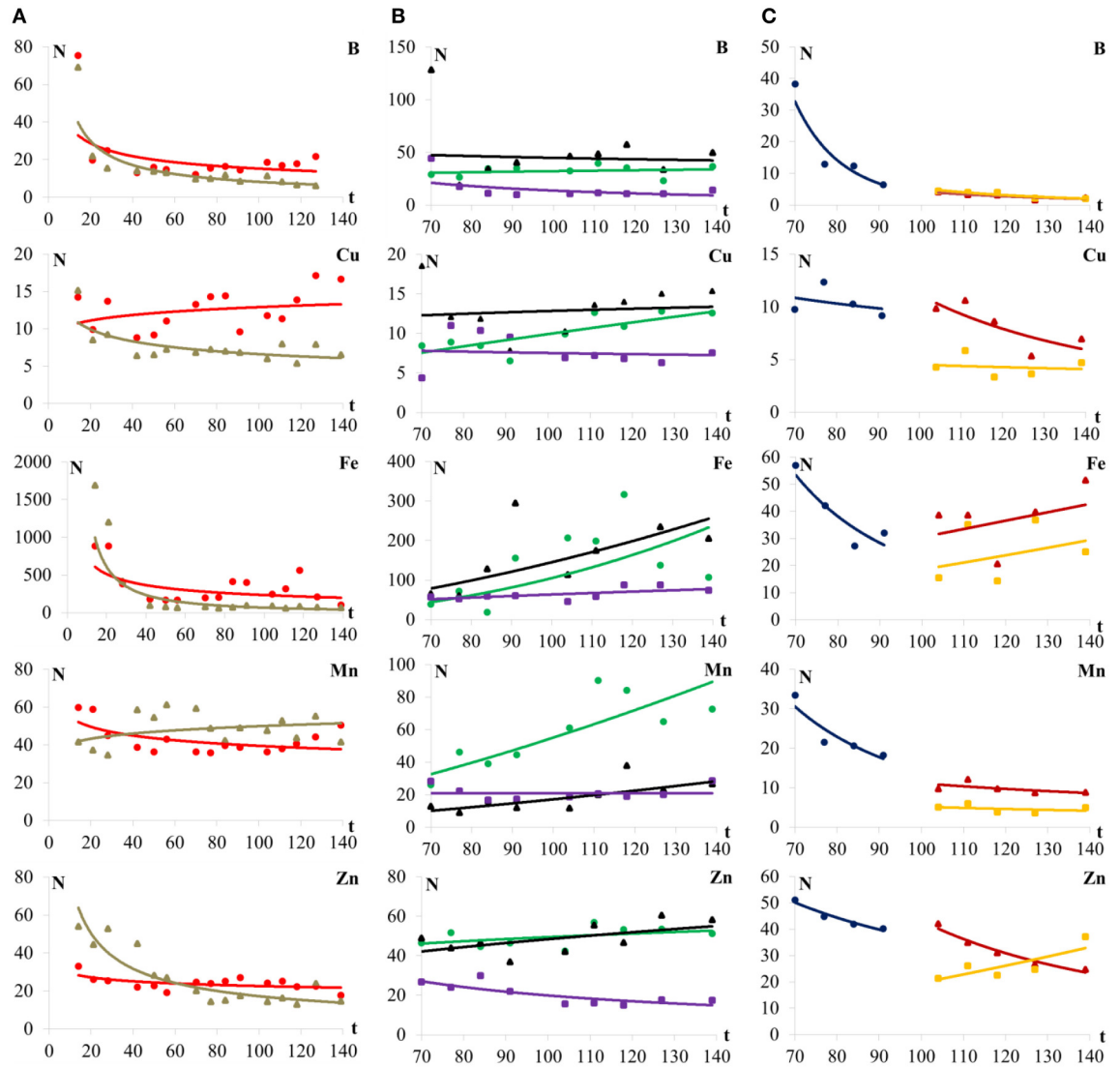


FIGURE 3 | B, Cu, Fe, Mn, and Zn contents (*N*, mgkg^{-1}) in **(A)** leaf (●), stalk (▲), **(B)** tassel (●), style-stigma (▲), straw (■), **(C)** ear (●), cob (▲), and grain (■) of maize plants (hybrid DKB390VTPRO2) as a function of time (t, d).

In grain, there are higher levels of B, Cu, and Mn at 104 (R_4 growth stage) and Fe and Zn at 140 days (R_6 growth stage) (**Figure 3C**).

In 31 out of 45 cases, the Pearson correlation coefficient values were larger than the critical values (69%) (leaf: 80%, stalk: 80%, tassel: 60%, style-stigma: 60%, straw: 60%, ear: 80%, cob: 60%, grain: 40%, and plant: 100%), at 10% of significance level, for micronutrient content power model (**Table 3**).

Micronutrient Absorption *A* and Absorption Rate λ for the Above Ground Plant

The total micronutrient absorption increased with crop growth and development. At 14 days, the Cu, Fe, Mn, and Zn absorptions were, respectively, 3, 430, 10, and 10 g ha^{-1}, and at the end of the cycle at 139 days, 180, 1,040, 930 and 430 g ha^{-1}. The absorption rate was lower at the beginning of the cycle, 14 days (V_2 growth stage) it was, respectively for Cu, Fe, Mn, and Zn, 1, 20, 3, and 3 g ha^{-1} d^{-1}, the absorption peak was observed at 111 days (R_4/R_5 growth stage) with a value of 2, 90, 12, and 5 g ha^{-1} d^{-1} (**Table 4**).

Equation (1) fitted well for all micronutrients except Cu (**Figure 4A**). The absorption rate increased for all nutrients from 14 to 111 days. *N* is described as a decreasing power function for all micronutrients, resulting in absorption rate functions whose shape differs from the sigmoid function *D*.

The total B uptake at the R_1 growth stage (70 days) was 80 g ha^{-1} and at the end of the crop cycle, R_6 growth stage (139 days) 110 g ha^{-1}. The B absorption rate presents a maximum value λ_{max} on 112 days (R_4/R_5 growth stage) (**Figure 4A**) with 6 g ha^{-1} d^{-1}. Thereafter the rate decreases to the end of the crop cycle.

Cu had a similar behavior to Mn and Zn, but with a slightly worse fit. Nevertheless, the power model for Cu content was chosen, in the same way as for the contents of Mn. The total

TABLE 3 | Empirical parameter values (k_5, $mgkg^{-1}d^{-k_6}$ and k_6) and Pearson correlation coefficient (r) for micronutrient (m) content (*N*, mg kg^{-1}) power model, for B, Cu, Fe, Mn, and Zn in leaf, stalk, tassel, style-stigma, straw, ear, cob, grain, and above ground maize plant (hybrid DKB390VTPRO2).

m	Leaf	Stalk	Tassel	Style-stigma	Straw	Ear	Cob	Grain	Plant
EMPIRICAL PARAMETER VALUES (k_5, $mgkg^{-1}d^{-k_6}$)									
B	93.693	336.03	16.258	95.93	3,593.6	1.0E+13	336,647	4.0E+06	4,828.994
Cu	8.408	21.183	0.2974	7.2660	12.275	55.104	63,469	17.101	0.251
Fe	2,213.2	38,302	0.0535	0.0015	4.1593	2.0E+06	0.2858	0.0323	112,537.98
Mn	75,824	33,114	0.0631	0.0183	19.722	292,534	353.59	103.96	71.34
Zn	38,335	370.12	20.272	8.2008	1,083.9	2,431.2	206,959	0.0153	111.305
EMPIRICAL PARAMETER VALUES (k_6)									
B	−0.395	−0.807	0.1481	−0.165	−1.209	−6.254	−2.45	−2.959	−1.4147
Cu	0.093	−0.253	0.7613	0.1234	−0.107	−0.382	−1.877	−0.29	−0.2397
Fe	−0.489	−1.383	1.7164	2.4275	0.5916	−2.514	1.0133	1.3793	−1.5903
Mn	−0.142	0.089	1.4705	1.4864	0.0127	−2.158	−0.751	−0.65	−0.1207
Zn	−0.115	−0.666	0.7936	0.3853	−0.868	−0.913	−1.84	1.5549	−0.3687
PEARSON CORRELATION COEFFICIENT (r)									
B	0.604*	0.914*	0.203	0.075	0.602*	0.950*	0.785*	0.917*	0.869*
Cu	0.303	0.728*	0.771*	0.115	0.087	0.338	0.766*	0.149	0.741*
Fe	0.538*	0.876*	0.715*	0.629*	0.626*	0.877*	0.341	0.355	0.965*
Mn	0.582*	0.353	0.860*	0.762*	0.014	0.921*	0.633	0.360	0.742*
Zn	0.554*	0.894*	0.464	0.561*	0.808*	0.978*	0.979*	0.811*	0.945*

**Significant at level of 10% for micronutrient content power model by the critical Pearson correlation coefficient test.*

TABLE 4 | Mean values of nutrient content (*N*, $mgkg^{-1}$), absorption (*A*, gha^{-1}) and absorption rate (λ, $gha^{-1}d^{-1}$) of B, Cu, Fe, Mn and Zn in maize crop (hybrid DKB390VTPRO2) in relation to time after sowing (t, d).

t	Nutrient content (*N*)					Absorption (*A*)					Absorption rate (λ)				
	B	Cu	Fe	Mn	Zn	B	Cu	Fe	Mn	Zn	B	Cu	Fe	Mn	Zn
14	–	13	1,693	52	42	–	3	430	10	10	–	1	20	3	3
21	65	12	888	49	36	50	10	690	40	30	2	1	20	3	2
28	43	11	562	48	33	60	20	760	60	40	2	1	30	4	2
42	24	10	295	45	28	70	30	810	120	80	2	1	30	4	2
50	19	10	224	44	26	70	40	830	170	100	3	1	40	5	3
56	16	10	187	44	25	70	40	840	200	110	3	1	40	6	3
70	12	9	131	43	23	80	60	890	290	160	4	1	50	7	3
77	10	9	113	42	22	80	70	910	340	180	4	2	60	8	4
84	9	9	98	42	22	90	80	940	400	210	4	2	70	9	4
91	8	9	86	41	21	90	100	970	470	240	5	2	70	10	4
104	7	8	70	41	20	100	120	1,030	600	300	6	2	80	11	5
111	6	8	63	40	20	100	140	1,050	680	330	6	2	90	12	5
118	6	8	57	40	19	110	150	1,070	750	360	6	2	90	11	4
127	5	8	51	40	19	110	170	1,080	840	400	5	2	80	10	4
139	4	8	44	39	18	110	180	1,040	930	430	3	1	50	6	2

absorption of Cu also increased during the crop development. At day 14, the absorption of Cu was 3 g ha^{-1}, at the end of the vegetative growth stage (day 56, V_{15} growth stage), the absorption was 40 g ha^{-1} and at the end of the cycle, it increased to 180 g ha^{-1}. With respect to the absorption rate, the behavior of Cu was similar to Zn, being high at the first sampling (14 days), decreasing until 38 days (V_8 growth stage) and then increasing up to the maximum value at 109 days (between R_4 and R_4/R_5 growth stages) of 2 g ha^{-1} d^{-1} (**Figure 4A**). Thereafter it decreases continuously until the end of the crop cycle.

The total absorption of Fe increased throughout the development of the crop, at 14 days (V_2 growth stage) showing a value of 430 g ha^{-1}. At R_1 growth stage (70 days), the Fe uptake was 890 g ha^{-1} and at the end of the cycle 1,040 g ha^{-1}. With respect to the absorption rate of Fe, it was high initially, at the V_2 growth stage (14 days) with a value of 20 g ha^{-1} d^{-1}. The

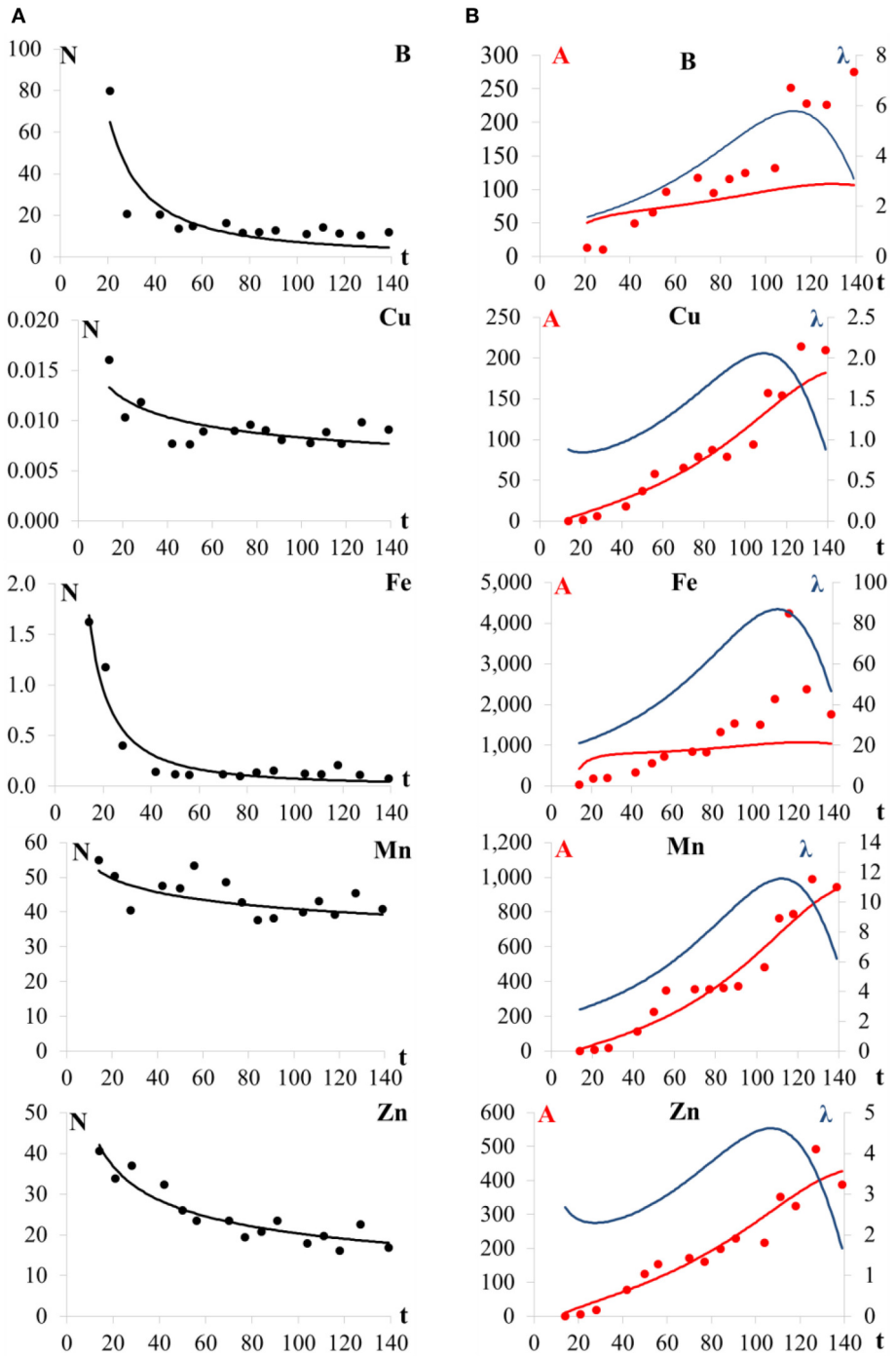


FIGURE 4 | (A) micronutrient contents of the above ground maize plant (*N*, mgkg^{-1}), and **(B)** total absorption (*A*, gha^{-1}) and absorption rate (λ, gha^{-1}d^{-1}) for B, Cu, Fe, Mn, and Zn in the maize (hybrid DKB390VTPRO2) crop (weighted average of all above ground organs) as a function of the number of days after sowing (t, d).

nutrient absorption rate of Fe increased to the maximum value of 90 g ha^{-1} d^{-1} at 112 days. The Fe rate decreases thereafter to the end of the cycle (**Figure 4B**).

Mn behaved similarly to Cu. The total absorption increased during crop development from 290 g ha^{-1} at the R_1 growth stage (70 days) to 930 g ha^{-1} at the end of the cycle (R_6 growth stage, 139 days). At 14 days, the absorption rate was low, increasing thereafter continuously until 112 days. From this period, increasing values were observed, and the maximum absorption rate (λ_{max}) occurred at 112 days (between R_4/R_5 and

R_5 growth stages) (**Figure 4B**), with a value of 12 g ha^{-1} d^{-1}. Then, the Mn absorption rate decreased until harvest.

Zn total absorption was equal to10 g ha^{-1} at 14 days and 430 g ha^{-1}at the end of the cycle. The absorption rate peak was 5 g ha^{-1} d^{-1} at 107 days (between R_4 and R_4/R_5 growth stages) (**Figure 4A**), thereafter decreasing until harvest.

The Pearson correlation coefficient values (0.869, 0.741, 0.965, 0.742, and 0.945) (**Table 4**) for micronutrient (B, Cu, Fe, Mn, and Zn) content model (**Figure 4A**) were larger than the critical value, at the 10% of significance level.

Micronutrient Content in Stalk Sap

The values for B, Cu, Fe, Mn, and Zn contents (N), total absorption (A) and the maximum absorption rates for each micronutrient (λ_{max}), and maximum concentrations (C_C) are shown in **Table 5**.

The maximum absorption rate (λ_{max}, g ha^{-1} d^{-1}) of the micronutrients occurred between 107 and 112 days, i.e., during the R_4 (50% of the plants exh0ibiting farinaceous grains) and R_5 (50% of the plants exhibiting hard farinaceous grains) growth stages (Ritchie et al., 1996). These growth stages are therefore the most important with respect to the nutritional needs of the maize crop. At these growth stages, the starch accumulation in the maize grain increases featuring a period of grain filling, resulting in greater dry mass of grain.

The ETa values in the corresponding days of λ_{max} are relatively small because the period coincides with the local winter season, due to cloudy days. Estimates of the maximum concentration (C_C, mg L^{-1}) in the gross plant sap were highest for Fe, followed by Mn (**Table 5**).

DISCUSSION

Evapotranspiration

Water absorption refers to the sum of the transpiration, the amount of water retained by the maize crop and the amount of water required to produce the total dry matter. Maize crop productivity depends fundamentally on the water absorption and the carbon dioxide assimilation, since carbon and oxygen (from CO_2), and hydrogen (from H_2O) represent about 96% of the dry matter.

The maize crop potential evapotranspiration (water requirement) during the whole cycle was 180 mm. The total actual evapotranspiration was 135 mm. There was a water deficit of 45 mm. A mild water stress occurred during 38 days of crop establishment, from April 14 to May 21.

Water Deficit

Effects of water deficit vary according to the development growth stage of the crop, and when occurring in the vegetative period, plant height was reduced, leaves become smaller, and consequently present a lower leaf area. However, the stress is mainly harmful to grain productivity when it occurs at the beginning of the tassel development up to grain filling. The demand for water increases rapidly about 2 weeks before the development of the tassel, and reaches a maximum at the flowering peak, continuing high for about two more weeks when it decreases rapidly.

Crop Cycle

The cropping period was from March to August, a season that is called "second maize harvest" in Brazil, nowadays more

TABLE 5 | Maximum concentration (C_C, mgL^{-1} or Cc*, μM, as function of the molecular weight M, gmol^{-1}) of each micronutrient (m) in crude sap (xylem) of maize crop (hybrid DKB390VTPRO2) at t days after sowing, corresponding to the growth stage (*GS*), above ground plant content (N, mgkg^{-1}), nutrient absorption (A, gha^{-1}), maximum absorption rate (λ_{max}, gha^{-1}d^{-1}), actual evapotranspiration (*ETa*, mmd^{-1}), evaporation (E, mmd^{-1}), absorbed water by the plant to form dry matter, the water retained by the plant (α, kg m^{-2} d^{-1}), total dry matter stored on the day of calculation ($\Delta D/\Delta t$, kgm^{-2}d^{-1}) and relative difference (ΔC, %) of *Cc* values estimated by equations 12 (Cc_{12}) and 14 (Cc_{14}).

m	M g mol^{-1}	t das	N mg kg^{-1}	A g ha^{-1}	λ_{max} g ha^{-1} d^{-1}	ETa mm d^{-1}	E[a] mm d^{-1}	Δ[b] kg m^{-2} d^{-1}
B	11	112	5.8	103.6	6.1	1.2	0.060	0.1717
Cu	64	109	8.1	131.5	2.1	1.5	0.075	0.1711
Fe	56	112	62	1,054.6	86.8	1.2	0.060	0.1717
Mn	55	112	40.4	686.4	11.6	1.2	0.060	0.1717
Zn	65	107	19.9	309.3	4.6	1.5	0.075	0.1699
m	***GS***	**t das**	$\Delta D/\Delta t$ kg ha^{-1} d^{-1}	Cc_{12} mg L^{-1}	Cc_{12}* μM	Cc_{14} mg L^{-1}	Cc_{14}* μM	ΔC[3] %
B	R_4/R_5	112	289	0.47	42.3	0.51	46.2	−9.31
Cu	R_4/R_5	109	288	0.13	2.1	0.14	2.2	−6.40
Fe	R_4/R_5	112	289	6.62	118.2	7.23	129.2	−9.31
Mn	R_4/R_5	112	289	0.88	16.1	0.97	17.6	−9.31
Zn	R_4/R_5	107	286	0.29	4.4	0.31	4.7	−6.33

[a] $E = 0.05ETa$.
[b] $u = 0.9$ kg kg^{-1} (plant water content) and $T_H = 0.06$ kg kg^{-1} (hydrogen content in the dry matter of maize crop).
[c] $\Delta C = 100(Cc_{12}-Cc_{14})/Cc_{14}$.

important than the "main maize harvest," from September to March. The crop cycle had a duration of 139 days. The vegetative phase ended 70 days (R_1 growth stage), accumulating 863°C days, and the reproductive phase extended from 70 to 139 days (R_6 growth stage), with a total of 733°C days (**Table 1**).

Crop Growth and Development

At the end of the vegetative phase, the maize plant gives priority to the development of the tassel and ear, which causes a remobilization of the photo-assimilates and nutrients from the leaf to the production and dispersion of pollen and to grain filling. It is also possible to observe a slight reduction of stalk dry matter during 91–104 days (R_3 and R_4 growth stages) (**Table 2**).

During the reproductive phase, the ear is a significant physiological drain. The plant redistributes photo-assimilates and nutrients to the grain, for filling. Therefore, the closer to physiologic maturity, the more leaves and stalk enter in senescence.

The accumulation and mobilization of dry matter in maize crop has a characteristic sequence over the growth cycle. For the maize crop, some authors report senescence as a process that encompasses the above ground plant, besides being caused by internal and external factors and mediated by a genetic program.

At 112 days, the values of the growth rate continue to be positive (dY/dt), but with decreasing daily gains (d^2Y/dt^2), i.e., the maize plant reduces its rate of dry matter accumulation because of the senescence process.

Micronutrient Content

The demand for micronutrients depends mainly on the crop productivity (production of total dry matter mass per unit area) (Bray, 1948), and the variation of dry matter composition of genotypes of the species of interest (least significant component).

Studies of Ciampitti et al. (2013), evaluating the contents of B, Cu, Fe, Mn, and Zn in maize plants, showed similar results as those found in this study. These micronutrient contents are larger at the beginning of the development of the maize crop, and then decrease up to physiological maturity.

During the initial growth, as there is a low production of plant dry matter, a high concentration of micronutrients was found coming from the soil or being remobilized from other parts of the plant.

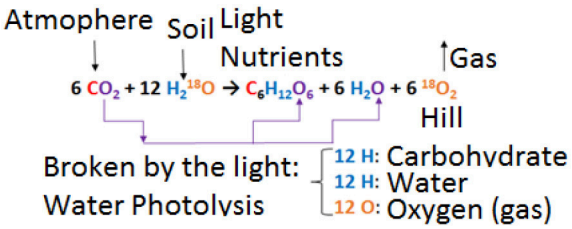
The absorption of micronutrients depends on the water absorption and the effective content of the micronutrient in the soil solution. Transpiration depends on the difference between water potential in the leaf and in the atmosphere. The absorption of water and nutrients depends on the elements of the climate (such as temperature, relative humidity and wind speed), crop (such as root system architecture, leaf area index and productivity) and soil (water and micronutrients content).

With crop growth, which usually follows a sigmoidal model, the dry matter accumulation is more expressive than the capacity of the plant to absorb and concentrate micronutrients. Therefore, we have a dilution effect due to the maize crop growth. Furthermore, it is known that higher concentrations of Fe and B, for example, are related to leaf and stalk. Over time, other structures gain greater proportion in the share of total dry matter thus contributing to part of this dilution effect.

Micronutrient Content in Stalk Sap

Due to complex factors related to the interaction between genotype and environment (climate and soil), the determination of the critical content in the xylem solution is the proposed procedure to evaluate soil fertility. Among the different factors, we highlight: phenological stage, effective depth of the root system corresponding to at least 90% of the potential evapotranspiration of the crop, distribution of the root system in the soil profile, soil water flow density, root trapping, mass flow and diffusion of micronutrients in soil, pH, temperature and content of the different micronutrients in soil solution and evapotranspiration, mainly.

During photosynthesis, the produced carbohydrate (CH_2O) molecules are composed of carbon (C) and oxygen (O) atoms from atmospheric CO_2, whereas the hydrogen (H) originates from water molecules from the soil. The produced O_2 returning to the atmosphere also originates from the soil water molecule (Taiz and Zeiger, 2006):



(11)

This reaction is endothermic, requiring energy that, in the photosynthesis process, is provided by solar radiation, where (i) 6 carbon and 6 oxygen atoms of the produced carbohydrate molecule are derived from the enzymatic breakdown of atmospheric carbon dioxide by RuBisCO (Ribulose-1,5-bisphosphate carboxylase/oxygenase); (ii) 12 atoms of hydrogen of the produced carbohydrate molecule are derived from the extracted water from the soil and broken by the light in the leaf, process known as water photolysis; (iii) the other 12 atoms of hydrogen with 6 atoms of oxygen will build 6 molecules of water; and (iv) the 12 atoms of oxygen of the molecule of oxygen (O_2), that return as gas to the atmosphere, are derived from water molecule: reaction proven by Robin Hill in the years 1960 using labeled oxygen (^{18}O). Then, the absorbed water responsible for the total dry matter produced during the day of calculation can be estimated as $9T_H \frac{dD}{dt}$ and the water retained by the plant can be estimated as $9T_H \frac{dD}{dt} / (1-u)$, so that the maximum concentration of a micronutrient in the stalk sap (Cc, mg L^{-1}) is:

$$Cc = \frac{\lambda_{\max}}{10^4 \left[ETa - E + 9 \left(\frac{2-u}{1-u} \right) T_H \frac{dD}{dt} \right]} \quad (12)$$

where ETa is the actual evapotranspiration (mm d^{-1}), here calculated using the simple Thornthwaite and Mather (1955) water balance, E the soil surface evaporation (mm d^{-1}), u the plant water content (kg kg^{-1}), T_H the hydrogen content in the

dry matter of maize crop (0.06 kg kg^{-1}) and d*D*/*dt* the total dry matter stored on the day of calculation (kg m^{-2} d^{-1}):

$$\frac{dD}{dt} = \frac{-2k_2{k_4}^2\left(t-k_3\right)}{\left[{k_4}^2+\left(t-k_3\right)^2\right]^2} \quad (13)$$

Cc represents the critical concentration λs (mg L^{-1}) in the soil solution. In this way, knowing *Cc*, it should be possible to develop a methodology for characterizing soil fertility and to recommend fertilization optimized to reach the maximum productivity as a function of the limiting nutrient.

With the aim of estimating the micronutrient maximum concentrations (Cc, mg L^{-1}) in the maize stalk sap at the time of maximum absorption, Equation (12) was simplified to:

$$Cc = \frac{\lambda_{\max}}{10^4.ETa} \quad (14)$$

The elimination of the term $[-E + 9\left(\frac{2-u}{1-u}\right)T_H\frac{dD}{dt}]$ results in an underestimation of 6.33% (Zn) to 9.31% (B, Fe, and Mn) (**Table 5**).

Experiments carried out with the aim of evaluating the absorption of nutrients by a maize crop also report that the increased absorption of nutrients B, Cu, Fe, Mn, and Zn occurred during reproductive phase, between R_4 and R_5 growth stages (Ciampitti et al., 2013).

The maximum concentration (C_C, mg L^{-1}) of each nutrient in the xylem sap is here assumed to be related to the soil solution absorbed by plant roots (**Table 4**). Based on the analysis of the results obtained in this research, it is suggested that future studies should be conducted in more than one growing season, with replicates of several years or even at different times. Such experiments may include different genotypes, as well as different regions, varying the population of plants in the experimental area, and simulate high, medium and low technology managements.

It is not possible to separate the effect of lack of water from the lack of micronutrients in the loss of maize crop productivity. The critical micronutrient content (*Cc*) is related to the transpiration and productivity (related to the maximum maize crop demand λ_{max}). Theoretically, the micronutrient with lowest content in the stalk sap (related to the lowest soil offer λ*s*) defines productivity. On the other hand, the limiting maximum micronutrient absorption rate (λ_{max}) corresponding to the lower productivity defines the attainable productivity limited by water and micronutrient (Liebig's minimum law).

Water and Nutrients Absorption by Plants

In nature, water and nutrient absorption occurs simultaneously (with and without energy expenditure), because the solute movement in the soil occurs by the combined processes of diffusion and mass flow, which in both cases are related to dynamics of water in the natural system composed by soil, plant and atmosphere phases.

From the thermodynamic point of view, the soil chemical fertility depends on the physical process related to the water movement in nature. The transpiration depends, among other factors, on the water potential difference between vapor in the atmosphere and liquid water in the leaf.

The water potential (ψ, atm) in the atmosphere phase defines the magnitude order of transpiration, which depends on air temperature (T, K) and relative humidity (Ω, kPa kPa^{-1}):

$$\Psi = 55.5.R.T.\ln(\Omega) \quad (15)$$

which depend on other climate elements such as solar radiation, rainfall and wind, for example.

Maximum Micronutrient Concentration in the Sap, Productivity and Transpiration

The water absorption by a crop is given by the sum of transpiration and constitutional water dependent on dry matter production. Absorption of nutrients is dependent on water absorption (or transpiration if we neglect constitutional water) and nutrient concentration on the absorbed solution. Therefore, high nutrient concentration in the soil solution (higher chemical fertility in the current classical model, since it does not reach salinization levels) is correlated with high concentration in the plant xylem solution, greater nutrient absorption, high dry matter production (of the different organs) and greater productivity (yield).

Therefore, this approach (preliminary studies to characterize the temporal variation of micronutrient composition of the above ground organs of maize and correlated uptake rates) allows developing a new concept to change the classic criteria of fertilization recommendation taking into account the reality of the facts (high dependence of micronutrients absorption to water absorption). The experimental data are used merely to give an example of application and to show the order of magnitude of the micronutrients contents in different organs and the micronutrients concentrations in the sap.

Figure 5 shows the micronutrient (B, Cu, Fe, Mn, and Zn) maximum concentrations in the maize stalk sap at the time t_{max} estimated by Equations (12) and (14) for the maize crop as a function of low (lower than 10,000 kg ha^{-1}) and high (higher than 10,000 kg ha^{-1}) yield and actual evapotranspiration. It is observed that the maximum concentration Cc presents high variation as a function of productivity in an environment of low evapotranspiration, which does not occur under high evapotranspiration. It was also observed that the simplified equation (Equation 14) presents smaller errors under the condition of high evapotranspiration.

Final Considerations

This preliminary study may serve as a basis for other researchers to develop an alternative methodology to the current procedure using chemical soil extractors. This alternative may vary from the measurement of the micronutrient content in the xylem in the indicator plant, which may be the species of interest at the point of maximum growth rate as described in this work, or even through the correlation of this value with the value measured in seedlings of the same species or using a native species present in the area (Cate and Nelson, 1965).

Chlorine, molybdenum and nickel were not considered in the present study due to the low amounts of these micronutrients

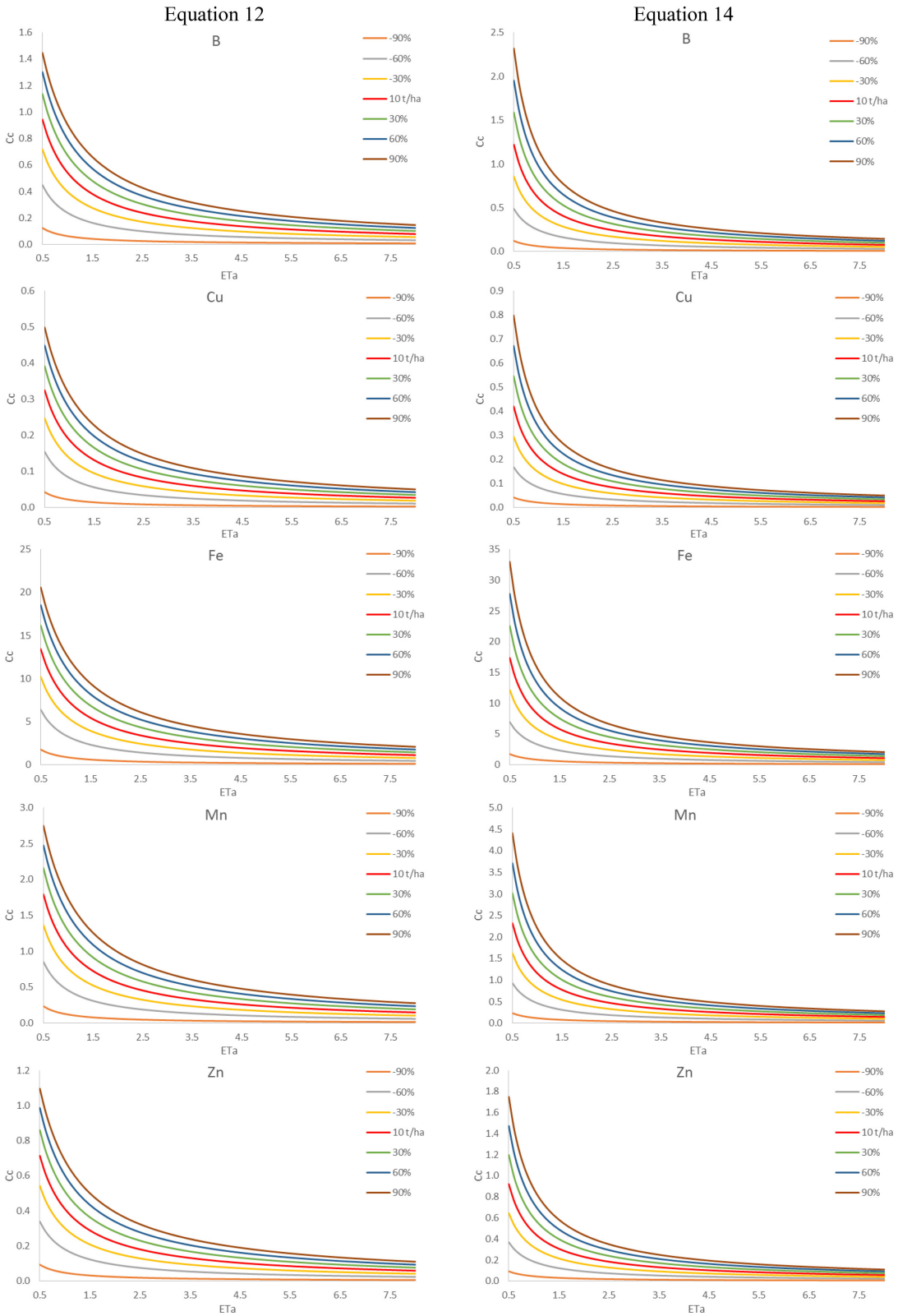


FIGURE 5 | B, Cu, Fe, Mn, and Zn maximum concentrations (Cc, mgL^{-1}) in the maize stalk sap estimated by complete (Equation 12) and simplified (Equation 14) equations for the maize crop as a function of yield (lower and higher than $10tha^{-1}$) and actual evapotranspiration (ETa, mmd^{-1}).

used in maize fertilization programs, in general, for the following reasons: (i) chlorine is supplied when potassium is applied as KCl (about 40 kg ha^{-1} of K_2O), (ii) molybdenum is most frequently used in soybean cultivation due to the symbiosis with *Bradyrhizobium* sp. and (iii) nickel presents few research results because its essentiality was found only recently.

The next steps of this preliminary study would be including roots and validating the model under different environmental conditions and with different genotypes. In this way, the model is able to take into account the effect of the environment and of the genotype that were not considered in this initial work, as well as to characterize the chlorine, molybdenum and nickel contents of different organs and their respective extractions.

CONCLUSIONS

We proposed a methodology that can be used for characterizing the micronutrient absorption rate of crops and tested it for a maize growing scenario. Results show that: (i) the micronutrient content variation in time follows a power model, with higher values for the initial growth stages of development and leveling off to minimum values at the R_6 growth stage, (ii) the maximum micronutrient absorption rates occur in the reproductive growth stages, and (iii) these evaluations allowed to predict the maximum need for micronutrients, considered equal to their concentration in the stalk sap. The proposed methodology can be used as a basis for further improvement in micronutrient fertilization of maize and other crops.

AUTHOR CONTRIBUTIONS

KM and DD designed the research; JF, FS, GF, and SM performed the research; KM, DD, KR, and Qd analyzed data; KM, DD, KR, and Qd wrote the manuscript.

ACKNOWLEDGMENTS

This research was supported by the "National Council for Scientific and Technological Development" (CNPq) of "Ministry of Science, Technology, Innovations and Communications" (Grant No. 312470/2014-2), "Commission for the Improvement of Higher Education" (CAPES) of "Ministry of Education" and "Luiz de Queiroz College of Agriculture" of University of São Paulo, Brazil.

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. Rome: FAO.

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., and Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorol Z.* 22, 711–728. doi: 10.1127/0941-2948/2013/0507

Bender, R. R., Haegele, J. W., Ruffo, M. L., and Below, F. E. (2013). Nutrient uptake, partitioning, and remobilization in modern transgenic insect-protected corn hybrids. *Agron J.* 105, 161–170. doi: 10.2134/agronj2012.0352

Berger, K. C., and Truog, E. (1940). Boron deficiencies as revealed by plant and soil tests. *J. Am. Soc. Agr.* 32, 297–301. doi: 10.2134/agronj1940.00021962003200040007x

Bray, R. H. (1948). "Correlation of soil tests whit crop response to added fertilizers and with fertilizer requirement," in *Diagnostic Techniques for Soils and Crops*, ed H. B. Kitchen (Washington, DC: The American Potash Institute), 53–86.

Cantarutti, R. B., Barros, B. F., Martinez, H. E. P., and Novais, R. F. (2007). "Avaliação da fertilidade do solo e recomendação de fertilizantes," in *Fertilidade do solo*, eds R. F. Novais, V. H. Alvarez, B. F. Barros, R. L. F. Fontes, R. B. Cantarutti, and J. C. L. Neves (Viçosa, MG: Sociedade Brasileira de Ciência do Solo), 769–850.

Cate, R. B., and Nelson, L. A. (1965). *A Rapid Method for Correlation of Soil Test Analysis with Plant Response Data*. Raleigh, NC: North Carolina State University.

Ciampitti, I. A., Camberato, J. J., Murrell, S. T., and Vyn, T. J. (2013). Corn nutrient accumulation and partitioning in response to plant density and boron rate: I. Micronutrients. *Agron J.* 105 783–795. doi: 10.2134/agronj2012.0467

CONAB - Companhia Nacional de Abastecimento (2015). *Séries históricas: de área plantada, produtividade e produção, relativas às safras 1976/77 a 2014/15 de grãos, 2001 a 2014 de café, 2005/06 a 2014/15 de cana-de-açúcar*. Available online at in: http://www.conab.gov.br/conteudos.php?a=1252&. Accessed in: 10/25/2015.

Edwards, J. (2009). *Maize Growth and Development*. Orange, NSW: NSW Department of Primary Industries, 60.

Lindsay, W. L., and Norvell, W. A. (1978). Development of a DTPA soil test for zinc, iron, manganese, and copper. *Soil Sci. Soc. Am. J.* 42, 421–428. doi: 10.2136/sssaj1978.03615995004200030009x

Oliveira, E. L. (2003). *Sugestões de adubação e calagem para culturas de interesse econômico no Estado do Paraná*. Londrina: Instituto Agronômico do Paraná.

Ribeiro, A. C., Guimarães, P. T. G., and Venegas, V. H. A. (1999). *Recomendações para o uso de corretivos e fertilizantes em Minas Gerais: 5. Aproximação*. Viçosa: Comissão de Fertilidade do Solo do Estado de Minas Gerais, 359.

Ritchie, S. W., Hanway, J. J., and Benson, G. O. (1996). *How a Corn Plant Develops*. Ames: Iowa State University of Science and Technology.

Sociedade Brasileira de Ciência do Solo (SBCS) (2004). *Manual de Adubação e de Calagem Para os Estados do Rio Grande do Sul e de Santa Catarina, 10th Edn.* Porto Alegre: Sociedade Brasileira de Ciência do Solo, Núcleo regional sul. Comissão de Química e Fertilidade do Solo, 400.

Soil Survey Staff (1975). *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*. Washington, DC: USDA Agric. Handbook 436.

Systat Software (2000). *Table Curve 2D: Automated Curve Fitting and Equation Discovery*. Chicago: SPSS Science.

Taiz, L., and Zeiger, E. (2006). *Plant Physiology*. Sunderland: Sinauer Associates.

Thornthwaite, C. W., and Mather, J. R. (1955). *The Water Balance*. New Jersey, NJ: Drexel Institute of Technology.

van Raij, B., Cantarella, H., Quaggio, J. A., and Furlani, A. M. C. (1996). *Recomendações de adubação e calagem para o estado de São Paulo*. Campinas: Instituto Agronômico; Fundação IAC.

Von Pinho, R. G., Borges, I. D., Pereira, J. L. A. R., and Reis, M. C. (2009). Marcha de absorção macronutrientes e acúmulo de matéria seca em milho. *Rev. Bras. de Milho e Sorgo, Sete Lagoas* 8, 157–173.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer YZ and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 05 February 2019
doi: 10.3389/fenvs.2019.00014

Multi-Scale Phenology of Temperate Grasslands: Improving Monitoring and Management With Near-Surface Phenocams

Christopher J. Watson[1,2,3]*, *Natalia Restrepo-Coupe*[1,4] *and Alfredo R. Huete*[1]

[1] School of Life Sciences, University of Technology Sydney, Sydney, NSW, Australia, [2] New South Wales Office of Environment and Heritage, Sydney, NSW, Australia, [3] Centre de Recherche sur les Interactions Bassins Versants-Écosystèmes Aquatiques, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada, [4] Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, United States

OPEN ACCESS

Edited by:
Urs Feller,
University of Bern, Switzerland

Reviewed by:
Michael John Hill,
University of North Dakota, United States
Nathaniel K. Newlands,
Agriculture and Agri-Food Canada (AAFC), Canada

***Correspondence:**
Christopher J. Watson
chris.watson@uqtr.ca

Specialty section:
This article was submitted to Agroecology and Ecosystem Services, a section of the journal Frontiers in Environmental Science

Received: *11 June 2018*
Accepted: *18 January 2019*
Published: *05 February 2019*

Citation:
Watson CJ, Restrepo-Coupe N and Huete AR (2019) Multi-Scale Phenology of Temperate Grasslands: Improving Monitoring and Management With Near-Surface Phenocams. Front. Environ. Sci. 7:14. doi: 10.3389/fenvs.2019.00014

Grasslands of the Australian Southern Tablelands represent a patchwork of native and exotic systems, occupying a continuum of C_3-dominated to C_4-dominated grasslands where composition depends on disturbance factors (e.g., grazing) and climate. Managing these complex landscapes is both challenging and critical for maintaining the security of Australia's pasture industries, and for protecting the biodiversity of native remnants. Differentiating C_3 from C_4 vegetation has been a prominent theme in remote sensing research due to distinct C_3/C_4 seasonal productivity patterns (phenology) and high uncertainty about how C_3/C_4 vegetation will respond to a changing climate. Phenology is used in northern hemisphere ecosystems for a range of purposes but has not been widely adopted in Australia, where dynamic climate often results in non-repetitive seasonal vegetation patterns. We employed time-lapse cameras (phenocams) to study the phenology of twelve grassland areas dominated by cool season (C_3) and warm season (C_4), native or exotic grasses near Canberra, Australia. Our aims were to assess phenological characteristics of the functional types and to determine the drivers of phenological variability. We compared the fine-scale phenocam seasonal profiles with field sampling and MODIS/Landsat satellite products to assess paddock-to-landscape functioning. We found C_3/C_4 species dominance to be the primary driver of phenological differences among grassland types, with C_3 grasslands demonstrating peak greenness in spring, and senescing rapidly in response to high summer temperatures. In contrast, C_4 grasslands showed peak activity in Austral summer and autumn (January-March). Some sites displayed primary and secondary peaks dependent on rainfall and species composition. We found that the proportion of dead vegetation is an important biophysical driver of grassland phenology, as were grazing pressures and species-dependent responses to rainfall and temperature. The satellite and field datasets were in general agreement with the phenocam results. However, the higher temporal fidelity of the cameras captured changes in vegetation not observed in the coarser satellite or field

results. Our phenocam data shows consistent periods of increasing and decreasing greenness over as little as 5 days. Applications for management of grasslands in temperate Australia include the identification of remnant native grasslands, tracking biosecurity issues, and assessing productivity responses to climate variability.

Keywords: phenology, phenocam, grassland, vegetation dynamics, remote sensing

INTRODUCTION

Grasslands represent one of the most dynamic and widespread biomes on Earth are the dominant ecosystems in a variety of climatic conditions (Scurlock and Hall, 1998). However, despite their importance in grazing systems and their acknowledged provision of ecosystem services, historical, and ongoing land management practices have degraded grasslands throughout the world (Ceballos et al., 2010). This is particularly true for temperate grasslands, which are facing many threats to their sustainable future, including modification for agriculture, habitat fragmentation, weed invasions, and changes in species composition due to a changing climate (Peart, 2008).

In Australia's temperate Southern Tablelands region, grasslands support both unique native flora and an important grazing industry. Grazing favors a community shift from tall perennial grasses to short grasses, and fertilization favors exotic annuals over native perennials (Moore and Biddiscombe, 1964; Gott et al., 2015). Historical land use of the Southern Tablelands therefore drives a patchwork of grasslands dominated by a variety of native grasses, exotic pasture grasses, invasive weeds, or a continuum of intermediate states. These grasslands differ greatly in their composition, structure, and functional attributes (Benson, 1994). Many native temperate grassland communities are only present as remnants and occupy a small fraction of their pre-European range (Groves, 1979; Benson, 1994). Their conservation and restoration is recognized as a priority, however there is an acute need to integrate conservation and agricultural values to ensure success (Wong and Dorrough, 2015). Classification of grasslands based on these attributes is the first step in being able to determine appropriate ecological and agricultural management.

Data for effective classification can be provided through field surveys, however these can be labor-intensive and impractical on a large scale. As an alternative, remote sensing has been explored as a potential approach to identify grassland types and condition. Efforts to discriminate grassland communities have had some success both worldwide (e.g., Price et al., 2002) and within Australia (Hill et al., 1999; Agrecon, 2004; Lymburner et al., 2011), though the classification groupings can be broad. Achieving finer-scale classification of temperate grasslands remains challenging due to their dynamic, heterogeneous nature (Hill, 2013), habit of retaining dead material on the plant (Tremont and McIntyre, 1994; Morgan and Lunt, 1999), unique shading issues (Shimada et al., 2012), and the continuum between disturbed and undisturbed conditions (Psomas, 2008).

While many factors can be used to classify grassland vegetation, the distinction between C_3 (cool season) or C_4 (warm season) photosynthetic types is fundamental (Epstein et al., 1997; Still et al., 2003; Adjorlolo et al., 2012). The nature of C_3 or C_4 dominance dictates patterns of growth and productivity during different times of the year; C_3 species are more productive in cooler, mesic climates, whereas C_4 species have a greater advantage in warmer and drier regions (Wand et al., 1999; Baldocchi, 2011). In the Australian Southern Tablelands, a continuum occurs from C_3-dominated to C_4-dominated grasslands without any defined spatial distribution. Much of the grassland composition depends as much on disturbance factors (e.g., grazing) as climate (Wimbush and Costin, 1979; Benson, 1994). There is a high uncertainty about how C_3 and C_4 vegetation will respond to increased CO_2 concentration and temperature and to modified moisture regimes predicted in a changing climate (Baldocchi, 2011; IPCC, 2014), in particular how this will impact agricultural productivity (Winslow et al., 2003; Howden et al., 2008; Cullen et al., 2009; Pau et al., 2013). Rising temperatures and lower available moisture are expected to favor C_4 grasses, while higher CO_2 concentrations should favor C_3 grasses (Morgan et al., 2011).

Differentiating C_3-dominant from C_4-dominant grasslands has been a prominent theme in remote sensing research due to distinct C_3/C_4 seasonal productivity patterns (Wang et al., 2013; Dye et al., 2016). Satellite data products characterize "land surface phenology" of vegetation types across landscape to global spatial scales (de Beurs and Henebry, 2004; Broich et al., 2015). These typically use a time-series of vegetation indices calculated from measured spectral reflectance, which can reliably estimate biophysical parameters such as biomass and vegetation cover for a diverse range of vegetation types (Weiser et al., 1986; Huete et al., 2002). Several satellite-based phenology studies include grasslands (Justice and Hiernaux, 1986; Fontana et al., 2008; Cui et al., 2012; Horion et al., 2013; Wang et al., 2013), though the majority of these focus on northern hemisphere grasslands where phenology is strongly driven by temperature. One notable study from southeastern Australia provided a classification of pastures types using Advanced Very High Resolution Radiometer (AVHRR) time-series data (Hill et al., 1999). This study successfully grouped broad land use types (e.g., native pastures, sown pastures, mixed pastures/cropping, and forest) based on similar time-series phenology profiles. More recent landscape-scale phenological research in Australia focuses on arid and semi-arid regions where vegetation dynamics are primarily driven by rainfall (Ma et al., 2013; Petus et al., 2013). The unique vegetation dynamics in many Australian environments (e.g., missing an annual growing season or having multiple greening periods) result in non-seasonal behavior and requires the development of different phenological approaches

than those used in typical northern hemisphere systems (Zhang X. et al., 2006; Broich et al., 2015).

Satellite remote sensing has the advantage of capturing large areas consistently, but its usefulness in phenological studies is constrained by temporal (i.e., time of satellite revisit) and spatial resolution (i.e., size of pixel) limitations. In contrast, time-lapse fixed cameras (termed "phenocams") have no such constraints and have shown great promise in capturing phenological information in a wide range of biomes (Brown et al., 2016), including northern hemisphere broadleaf forest (Ahrends et al., 2008; Richardson et al., 2009; Nagai et al., 2011; Mizunuma et al., 2013), Brazilian cerrado (Alberton et al., 2014), European alpine grasslands (Migliavacca et al., 2011; Julitta et al., 2014), Malaysian tropical forest (Nagai et al., 2016), and grasslands in Japan (Inoue et al., 2015). In Australia, Moore et al. (2016) provided an overview of phenocam imagery captured across the continent at different ecosystems including a tropical rainforest, a tropical savannah and a temperate evergreen forest. Generally, phenocams sample a smaller area than satellites and lack the spectral resolution of modern satellite sensors. However, they have the advantage of capturing high frequency (sub-daily) imagery, they can be positioned to directly monitor the vegetation of interest, atmospheric effects have less impact, and users can visually examine imagery to explain observed data patterns or anomalies. Phenocam imagery is typically converted to a vegetation index such as the Excess Green (e.g., Woebbecke et al., 1995) or the Green Chromatic Coordinate (g_{CC}) (Gillespie et al., 1987; Sonnentag et al., 2012) through manipulation of the red, green, and blue brightness values. Phenocam-based phenology has shown a good correspondence of phenophase timing when compared with eddy-covariance towers, satellite imagery, and field observations (Richardson et al., 2007; Migliavacca et al., 2011; Nagai et al., 2011; Mizunuma et al., 2013; Toomey et al., 2015; Moore et al., 2017), albeit with quantifiable time lags or restrictions to certain times of year (e.g., remotely sensed observations can be unavailable during the wet-cloudy season).

Remote sensing data is often validated through field biophysical observations (Mutanga and Skidmore, 2004; Zhang Q. et al., 2006; Shen et al., 2008; Liang et al., 2011; Psomas et al., 2011), with some agencies in Australia providing substantial investment and support to this aim (Muir et al., 2011). Some research has shown successful scaling from field measures to remote sensing (e.g., Fisher and Mustard, 2007; Studer et al., 2007). However, others have highlighted the sometimes weak relationship between *in situ* and satellite observations (Badeck et al., 2004; Ahl et al., 2006; Soudani et al., 2012). One of the more pressing challenges in phenological research is to understand the sources of variability between spatial scales (Friedl et al., 1994; Reed et al., 2009). This is particularly relevant for heterogeneous grasslands, where variability in spatial scales of field measurements can be problematic (Klimeš, 2003).

Given the importance of grasslands for food security and ecosystem preservation and the need for a better understanding of remote sensing-derived phenology over pastures and grasslands, this research aims to:

a) Assess the variability in phenology among of C_3/C_4 temperate grassland types with the use of phenocams;
b) Identify the biophysical drivers that cause changes to grassland land surface phenology;
c) Evaluate the utility of phenocams for capturing temperate grassland phenology; and
d) Compare scales of phenocam phenology data with field measurements and satellite phenology products.

MATERIALS AND METHODS

Study Sites

The study area is located in the Southern Tablelands region of New South Wales (NSW) and the Australian Capital Territory (ACT), and is part of the South Eastern Highlands bioregion (Environment Australia, 2000). The study area is approximately bounded by the towns of Bungendore (35.2500°S, 149.4500°E), Gungahlin (35.1831°S, 149.1330°E), and Bredbo (35.9420°S, 149.2009°E) (**Figure 1**). The region has distinct seasonal temperature values (mean monthly ranging from 0 to 30°C) and contains several types of native and exotic grasslands co-occurring within a similar climatic envelope. The climate is characterized by warm summers (December–February) with maximum daily temperatures frequently reaching 35°C. Winters (June–August) are cold, with frequent daily minimum temperatures below 0°C. Rainfall is relatively consistent throughout the year, with a mean of between 30 and 90 mm per month, and an annual average of 650 mm, although rainfall in the region is impacted by elevation, latitude, and aspect and may be spatially sporadic.

Modeled temperature and rainfall data generated from *MODIS* Terra Global Land Surface Temperature and Tropical Rainfall Measuring Mission version 7 (data product 3B43-v7) are presented to compare the monthly rainfall and temperature of the study period with the 1998–2017 average (**Figures 2A,B**). The study period (May 2014–April 2015) had temperatures that were generally consistent with the mean, though the period of December 2014–April 2015 was cooler than average (**Figure 2A**). Precipitation during the study period was more sporadic; a few months (June December, January and April) had unusually heavy rainfall and many months had lower than average rainfall (**Figure 2B**).

Three replicate areas of four distinct perennial grassland types common to the area were selected. Elevation of the study locations is between 550 and 750m above sea level. Sites were grouped based on the C_3/C_4 and the native/exotic status of the dominant perennial grass species (summarized in **Table 1**). Site groups were selected based on important grassland types in the region, including those that are used for conservation purposes, and those that are used for grazing agriculture. Sites within each category were not required to have the same dominant species *per se*, but rather be dominated by the same functional group (e.g., native C_4 grass). This dominance was designated by the terminology used throughout this text as: C_4 Native, C_4 Exotic, C_3 Native, and C_3 Exotic. C_4 Native sites were dominated by the common native grass *Themeda triandra*, a typical indicator species of low disturbance. C_4 Exotic sites were dominated

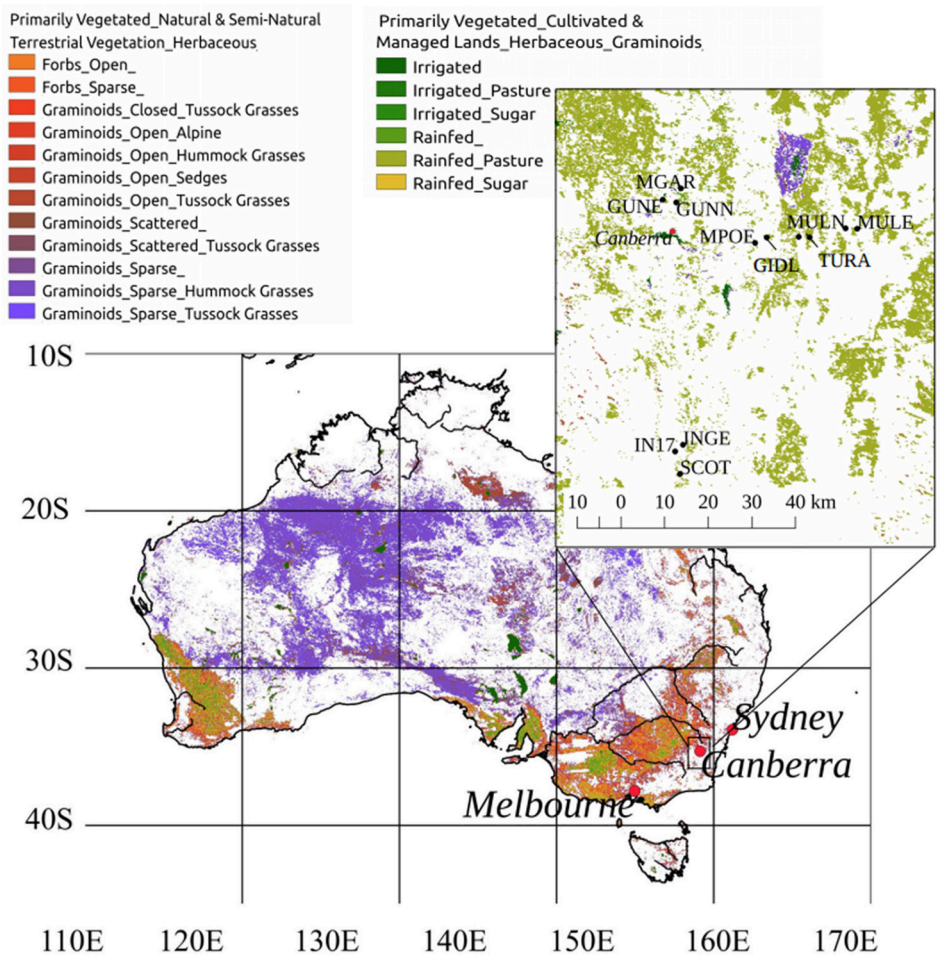


FIGURE 1 | Location of study area, south east of the continent, near the city of Canberra, Australia. Background vegetation mapping refers to the Dynamic Land Cover Dataset (Lymburner et al., 2011). Inset shows the location of individual study sites.

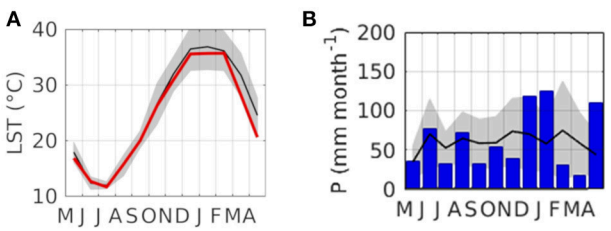


FIGURE 2 | Study site satellite derived climatic variables sampled at Mullunggari Nature Reserve (center pixel) **(A)** temperature from MODIS Terra Global Land Surface Temperature (LST; °C) (MOD11A2) mean (black line) and standard deviation 1998–2017 (gray area) and this study period (May 2014–April 2015; red line) **(B)** precipitation (P; mm month-1) from the version 7 Tropical Rainfall Measuring Mission (TRMM) data product (3B43-v7) (0.25 deg) mean (black line) and standard deviation 1998–2017 (gray area) and this study period (May 2014–April 2015; blue bars).

by the invasive agricultural weed *Eragrostis curvula*. C_3 Native sites contained a mixture of *Austrostipa* and *Rytidosperma* C_3 species typical of grazed native pastures in the region. C_3 Exotic sites were dominated by exotic pasture grasses, typically *Phalaris aquatica, Dactylis glomerata*, and *Festuca arundinaceae*. All grassland types contained secondary components of species outside the dominant functional group. For example, C_4 Native sites contained a small fraction of C_3 native grasses and exotic species. C_3 forbs occurred at most site at low vegetative cover.

All sites met the following criteria: homogenous cover of the selected grassland type; consistent land management throughout the study period; and the grassland area being >20 hectares, with adequate coverage in all dimensions to incorporate satellite pixels. None of the field sites were artificially irrigated.

Time-Lapse Digital Photography

Time-lapse RGB Wingscapes™ cameras (phenocams) in weatherproof housing were installed at each site to capture vegetation changes at a high temporal capacity. The phenocam was mounted 2.3 m above ground level and angled downward ∼15° from horizontal. The field-of-view included the horizon in the image but only a small quantity of sky, i.e., most of the scene was the target grassland. This field-of-view captured an area between 2 and 4 hectares. Each camera was positioned

TABLE 1 | Characteristics of study sites by grassland functional type.

Grassland functional type	Site name	Code	Location (Dec.Degrees)	Grazing pressure	Dominant genera	Area (ha)	Floristic composition	Tenure
C_3 Native	Mullunggari Nature Reserve	MGAR	−35.17377, 149.15075	Moderate (kangaroos)	*Rytidosperma, Austrostipa*	48	Native/exotic forbs and grasses	Nature Reserve
C_3 Native	Gungaderra Reserve (native paddock)	GUNN	−35.20961, 149.13885	Moderate (kangaroos)	*Rytidosperma, Austrostipa*	35	Native and exotic forbs (particularly *Hypochaeris*); native grasses	Nature Reserve
C_3 Native	Mulloon Creek Natural Farms (native paddock)	MULN	−35.27538, 149.57024	Light (cattle)	*Rytidosperma, Austrostipa*	24.1	Very diverse mix of native and exotic forbs and grasses	Private Land
C_4 Native	Gidleigh Traveling Stock Reserve	GIDL	−35.29711, 149.45078	Light (horses, sheep)	*Themeda*	15.5	Very diverse in native forbs; few exotics	Traveling Stock Reserve
C_4 Native	Turallo Nature Reserve	TURA	−35.2983, 149.47868	Light (sheep)	*Themeda*	23.4	Very diverse in native forbs	Nature Reserve
C_4 Native	Millpost Farm (native paddock)	MPON	−35.29873, 149.36943	Heavy (sheep)	*Themeda*	17.5	Extremely diverse in native forbs; few exotics	Private Land
C_3 Exotic	Mulloon Creek Natural Farms (exotic paddock)	MULE	−35.27643, 149.59981	Heavy (cattle)	*Phalaris, Festuca, Dactylis*	26.3	Exotic pasture grasses with some exotic legumes.	Private Land
C_3 Exotic	Gungaderra Reserve (exotic paddock)	GUNE	−35.20252, 149.10452	Light (kangaroos, wallabies)	*Phalaris*	25	Exotic pasture grasses and sedges	Nature Reserve
C_3 Exotic	Millpost Farm (exotic paddock)	MPOE	−35.3126, 149.33973	Heavy (sheep, rabbits)	*Phalaris Festuca, Dactylis,*	12.4	Mostly exotic grass/legumes with some native grass species	Private Land
C_4 Exotic	Scottsdale Bush Heritage Reserve	SCOT	−35.90046, 149.1482	Light (kangaroos)	*Eragrostis*	46	Exotic grass, scattered exotic annual forbs with some areas of native forbs	Nature Reserve
C_4 Exotic	Ingelara Farm	INGE	−35.82609, 149.15601	Heavy (cattle)	*Eragrostis*	14.3	Exotic perennial grass with annual exotic forbs and grasses	Private Land
C_4 Exotic	Ingelara Paddock 17	IN17	−35.84286, 149.13597	Moderate (cattle/kangaroos)	*Eragrostis*	26.6	Exotic perennial grass with annual exotic forbs and grasses	Private Land

to face south to minimize the impacts of sun glint on the images. We collected one image at hourly intervals between 9:00 and 15:00 Australian Eastern Standard Time (UTC +10). No standardization of color was used through the use of reference cards as can be found in similar studies (Ahrends et al., 2008; Richardson et al., 2009; Sonnentag et al., 2012). Color balance drift has been reported in a study using phenocams (Ide and Oguma, 2010) however this only became apparent after 2 years of use. As such, no significant color balance drift is assumed for this study.

The Green Chromatic Coordinate index (g_{CC}; Equation 1) was used as a surrogate of greenness as it represents the relative brightness of the green fraction compared to the sum of the green, red, and blue bands (Gillespie et al., 1987; Sonnentag et al., 2012). It is a unitless index that pilot studies suggest ranges between 0.25 (no green vegetation) and 0.5 (abundant green vegetation) in the subject grasslands. A variety of phenocam-based studies have preferentially used this index because of its dynamic response to changes in plant biophysical variables and robustness to variations in image brightness due to sky conditions, time of day, or shadowing (Ide and Oguma, 2010; Sonnentag et al., 2012; Julitta et al., 2014).

$$gCC = \frac{G}{R + G + B} \quad (1)$$

Images were manually filtered and viable images were processed with ImageJ software (Abràmoff et al., 2004) to extract RGB values and calculate g_{CC}. The mean g_{CC} across the target area of the image was used for each hourly data point and was used to establish a daily time-series, known as a phenology profile. For some phenocam profiles, we fitted a non-parametric Locally Weighted Scatterplot Smoothing (LOESS) curve to improve visualization and the identification of trends (Cleveland, 1979).

Field Measurements and Floristic Surveys

Monthly biometric measurements included pasture height and vegetation cover using non-destructive methods, and aboveground biomass using destructive harvesting.

- Percent vegetation cover was taken using a point-transect method (NSW Catchment Management Authority, 2005). One 100 m transect was established across a representative part of the study site with cover noted at 1 m intervals. Vegetation cover was classified as: green vegetation (photosynthetic vegetation, PV), dead vegetation (non-photosynthetic vegetation, NPV), and background (bare soil and substrate, BS).
- Average pasture height was measured using a falling plate method (Rayburn and Rayburn, 1998). A standard acrylic plate was mounted on a ruler and allowed to fall on the vegetation at the sample point. The mean value of 20 points was taken.
- Total biomass was harvested within six replicates of 1 m^2 quadrats to ∼1 cm above ground level. The location of each quadrat was randomly selected at each monthly visit without replication. Biomass samples were stored in a plastic bag in a cool environment (a cooler in the field and a refrigerator in the laboratory) prior to processing to prevent degradation. Vegetation was separated into "grasses," and "forbs" (including pasture legumes, native forbs, and exotic weeds), then further separated into photosynthetic ("green") and non-photosynthetic ("dead"). Samples were placed in paper bags, oven-dried at 60°C for 72 h and results converted to kg/ha.

Floristic surveys were conducted at each location (three 20 × 20 m plots) each month to monitor species composition. Monthly floristic data is not presented.

Satellite Data Sources

Satellite data were obtained from three sources that represent typical sources of land surface phenology data: the Moderate Resolution Imaging Spectrometer (MODIS) sensor aboard the Terra satellite, the Operational Land Imager (OLI) sensor aboard Landsat 8, and the Enhanced Thematic Mapper Plus (ETM+) sensor on Landsat 7.

The Terra MODIS 16-day composite NDVI (MOD13Q1) at 250 m spatial resolution was downloaded from the NASA Land Processes Distributed Active Archive Center (http://e4ftl01.cr.usgs.gov/) over the period May 1, 2014 to April 30, 2015. The NDVI provided by this product was calculated using the bands specified in **Table 2**. The data were filtered based on the quality assurance flags provided in the quality control layers of the product, with periods of clouds, high aerosol loads, and low quality were removed. The 16-day composite data reduces impact of cloud cover on long-term data sets, though at the expense of higher temporal resolution. Due to the relatively small size of most of our grassland sites, one pixel (250 × 250 m) was used for analysis.

Landsat data (OLI and ETM+) was obtained from the Climate Data Record surface reflectance from the Science Processing Architecture System of USGS Earth Resources Observation and Science Center (http://espa.cr.usgs.gov/), corrected for atmospheric effects and BRDF. We subsequently calculated NDVI from reflectance data as per Equation 2 using the red and near infra-red (NIR) bands specified in **Table 2**. Landsat 7 and Landsat 8 data were combined into one NDVI time-series as the data have been shown to be equivalent (Li et al., 2014; Ahmadian et al., 2016). The Landsat 7 and 8 data are collected at a nominal 16-day frequency; however this is subject to effects of clouds that can reduce or eliminate the usability of parts of an image. As such, temporal resolution frequently extends beyond 16 days. To capture a similar area as MODIS data sources, we used a 5 × 5 grid of 30 × 30 m Landsat pixels resulting in a total area of 150 × 150 m.

TABLE 2 | Satellite sensor spectral bands used in the calculation of NDVI.

Sensor	Satellite	Red (RED)	Near infrared (NIR)
MODIS	Terra	Band 1 (620–670 nm)	Band 2 (841–876 nm)
OLI	Landsat 8	Band 4 (640–670 nm)	Band 5: (850–880 nm)
ETM+	Landsat 7	Band 3 (630–690 nm)	Band 4 (770–900 nm)

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{RED})}{(\mathrm{NIR} + \mathrm{RED})} \quad (2)$$

On average, there were 19.2 of a possible 23 MODIS 16-day temporal data points for each site (range: 18–20) over the annual study period, and 17.3 Landsat OLI/ETM+ 16-day data points (range: 15–19) of a possible 23. Some large gaps, due primarily to cloud obstruction, were evident in the temporal satellite data which can impact phenology studies, with a maximum 126 day gap for Landsat data and a 48 day gap for MODIS data.

Data processing, including graphical analysis and statistical analysis, was conducted with the R software package (R Core Team, 2013). For quantitative comparison between methods, it was necessary to match data as closely as possible in time. As field sampling was the least frequent data type (monthly), we selected 12 data points from satellite and phenocam time-series that were closest to the 12 field sampling dates (number of days before or after). Pearson's correlations were conducted across all sites using this matched data.

RESULTS

Phenocam Imagery

The phenocam derived g_{CC} profiles (**Figure 3**) showed C_3 sites having maximum greenness as a distinct peak in October-November, whereas C_4 sites peaked later (January-February) and showed a broader peak. All sites started greening after August as temperatures warmed which is indicative of the presence of C_3 species (albeit as a secondary component in C_4-dominated sites). C_3 vegetation had a second greening period in late February in response to increased rainfall. Fine-scale vegetation dynamics are observed at all sites as small but rapid increases and decreases in g_{CC}.

C_3 Exotic grassland sites showed peak greenness in October-November (**Figure 3A**). However, these sites were variable: one

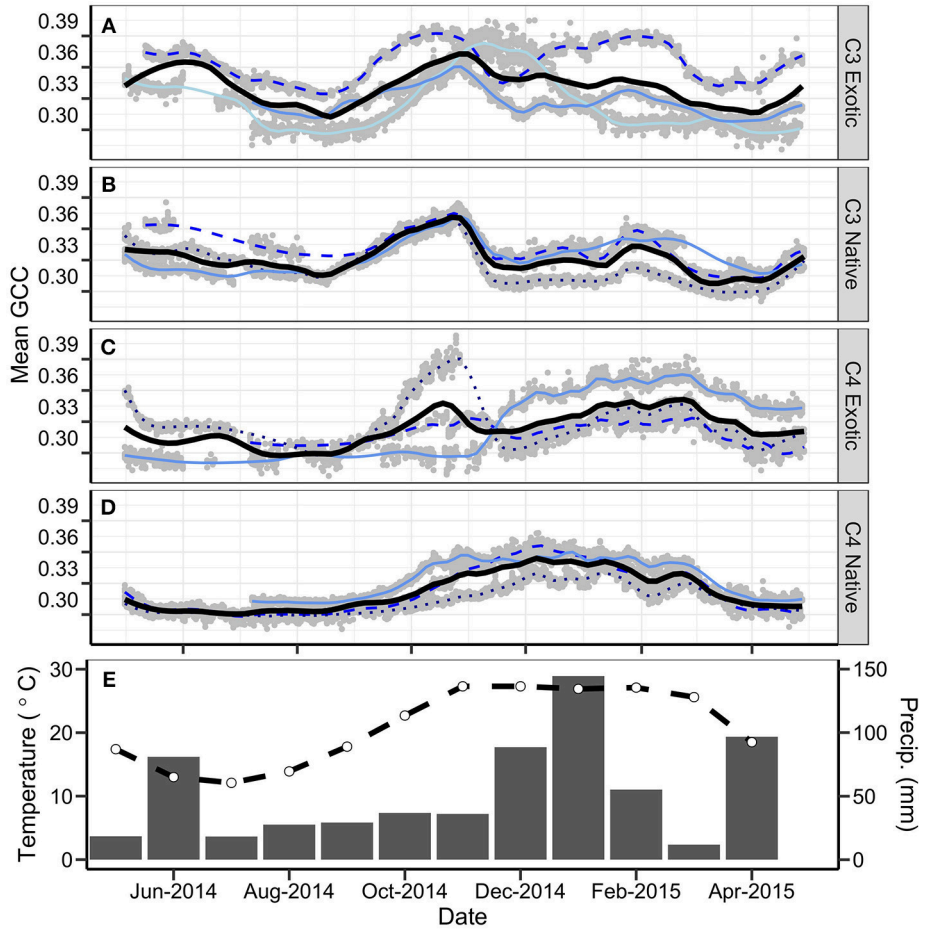


FIGURE 3 | Annual combined g_{CC} phenology profiles for all sites, grouped by functional type, **(A)** C_3 Exotic, **(B)** C_3 Native, **(C)** C4 Exotic, **(D)** C4 Native. Gray dots represent hourly data points. Blue lines represent individual sites. The thick black line of each panel is a LOESS fitted curve (span = 0.1) for each functional type. Panel **(E)** represents the monthly mean maximum temperature (°C; dashed line; Tuggeranong Bureau of Meteorology station) and the monthly rainfall (mm; solid bars; Australian National Botanic Gardens Bureau of Meteorology station) for the study period.

site showed an almost unimodal profile, increasing in g_{CC} from September, reaching a peak in November, and steadily decreasing to the baseline in January. At the second site we observed green-up steadily climbing to a peak in late October, before decreasing in December. Multiple small increases in g_{CC} occur until February, followed by a steady decline through March. The third site had two equally dominant peaks in October and February.

The C_3 Native profiles demonstrated a gradual decline of g_{CC} from May through to August and a characteristic greening pattern commencing in August and reaching a maximum in October-November (**Figure 3B**). The g_{CC} then abruptly decreased in November. Patterns changed slightly between the sites from this point; some showed periods of greening and browning through the summer, whereas others had only one small greening period. Two sites showed steady decline in g_{CC} through February to a minimum in March, whereas the other demonstrated an influence from secondary C_4 species (a low wide peak from December and reaching a minimum in late March).

The C_4 Exotic profiles show consistency in summer peak greenness (January-February) (**Figure 3C**). The g_{CC} values were all very low through the winter months, and shared identical timing of the characteristic broad summer peak. The individual sites showed different patterns during spring. One site had a distinctive peak of C_3 vegetation resulting from a flush of spring pasture grasses. Another showed a delayed but very rapid increase in g_{CC}, commencing in November and lagging biomass-measured greening by 2 months.

The C_4 Native grasslands show a consistent group of profiles (**Figure 3D**). All sites have low g_{CC} from May to August, with an indistinct inflection point in late August marking the start of a gradual greening. High g_{CC} is maintained through January-February followed by a characteristic drop in March as temperatures start to decrease.

Vegetation Cover

All groups showed distinct seasonal changes in cover throughout the year, with the mean time-series presented as **Figure 4**. The C_4

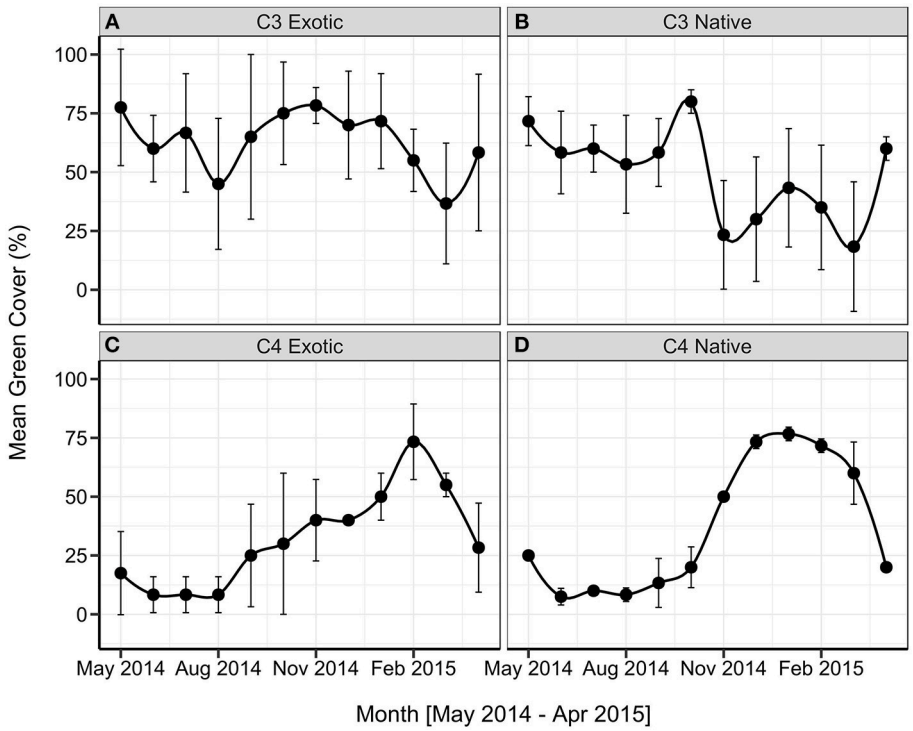


FIGURE 4 | Monthly mean green cover (% ± s.d., $n = 3$), presented by grassland functional type, **(A)** C_3 Exotic, **(B)** C_3 Native, **(C)** C_4 Exotic, **(D)** C_4 Native.

Native sites demonstrate particularly distinctive and predictable patterns, with low green cover during the winter months, rapid greening in October, and a peak of 75% green cover maintained from December to February. After March, the green cover at C_4 Native sites decreased quickly to 20%. The seasonal green cover variation between C_4 Native sites was low. The C_4 Exotic grasslands display a similar seasonal pattern, though with greater variation between sites: low green cover (∼10%) through winter, green-up starting in August with a steady increase to November, and a rapid rise to a peak of 75% in February. C_3 Native sites show a high variability throughout the year, with a peak of 80% green cover in October. Minima in November and March correspond with high temperatures and low rainfall. C_3 Exotic peak green cover is in November (76%) though high green cover is maintained from September through January. The minimum green cover occurs in March for both C_3 Exotic and C_3 Native groups.

Biomass Sampling

Green and dead biomass varied considerably between sites and seasonally at the same location (**Figure 5**). Exotic groups tended to have the greater quantity of mean green biomass, where we observed the maximum value at a single site of (2,769 kg/ha). The mean green biomass for C_4 Native sites did not exceed 1,000 kg/ha, and C_3 Native sites did not exceed 500 kg/ha. The C_4 Exotic sites had higher green biomass during the summer months than the winter months. By contrast, C_3 Exotic sites fluctuated throughout the year in response to grazing pressures and seasonal drivers, particularly periods of low rainfall. In general, sites that were dominated by C_4 species had low quantities of green biomass during the winter months, though notably still >0. The key features of the C_3 biomass pattern are an increase in September which reaches a peak in November (mean green biomass 1,536 kg/ha for exotic; 586 kg/ha for native). The C_4 sites commenced green-up a month later, in October, and reached a peak in December which was sustained through the summer (1,438 kg/ha exotic; 922 kg/ha native). At two of the three C_4 Exotic sites, we observed an additional peak in late summer (February; mean live biomass 2,026 kg/ha) and hence the associated standard deviation is large.

Dead biomass was greatest in C_4-dominated sites, with a mean maximum of 5,810 kg/ha at exotic C_4 sites and 4,689 at native C_4 sites. This accumulation of dead material was driven by minimal grazing or disturbance at these sites. C_4-dominated sites tended to have a higher biomass during the winter months but exhibited a very high variability between sites and replicates. All C_3 Native sites had a more consistent quantity of dead biomass throughout the year; however, variability was higher between C_3 Exotic sites, hence higher monthly standard deviation. At C_4-dominated sites the dead biomass was much greater than green biomass throughout the year. Forbs contributed a very minor proportion of the overall biomass.

Satellite Data

Satellite phenology profiles are presented as individual sites to allow visualization and comparison of the phenology trends

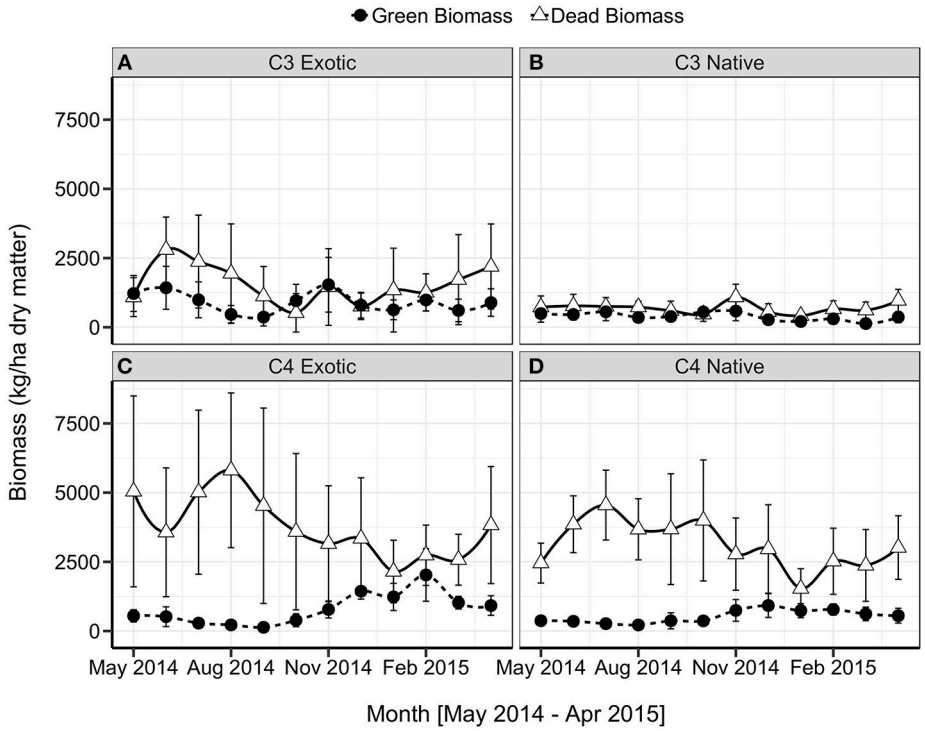


FIGURE 5 | Mean monthly green (solid line) and dead (broken line) biomass (kg/ha ± s.d.; $n = 18$) by grassland functional type, **(A)** C_3 Exotic, **(B)** C_3 Native, **(C)** C_4 Exotic, **(D)** C_4 Native.

of different satellite greenness products (**Figure 6**). While the satellite data shows substantial fluctuation, the timing of the maximum NDVI at each site corresponds with the C_3/C_4 functional group.

Landsat and MODIS data at C_4 Native sites (**Figure 6A**) show similar trends though they deviate substantially in magnitude. The paucity of Landsat data beyond January 2015 due to cloud contamination hindered meaningful comparison. The temporal resolution of C_4 Native seasonal pattern is less clear in satellite data than in other data sources. Two sites (*GIDL* and *TURA*) both show a consistently low NDVI below 0.5 throughout the autumn and winter months until green up. After a peak in October, satellite data shows a decrease in greenness toward late November that is not apparent in other data sources. The NDVI increases throughout the remainder of the summer months to a maximum of 0.62 and then slowly tails off toward the winter baseline. By contrast, at the third site (*MPON*), NDVI is more variable throughout the measurement period, with two peaks evident in October and January.

At C_4 Exotic sites (**Figure 6B**), Landsat data exhibited smooth trends whereas MODIS fluctuated. Site *IN17* showed the most consistency between the two satellite data sources. From a typically low greenness during winter, NDVI started to increase in August to an October peak of 0.6. After a small browning period, NDVI remained relatively high during the summer and decreased to 0.49 by late August. Landsat data for *INGE* showed a prominent peak in October (NDVI = 0.65), whereas the MODIS peak was apparent in February. Landsat data at *SCOT* showed a long, gradual increase in NDVI from late August to late December which contrasts with the rapid increase in g_{CC} observed in the phenocam data. MODIS data at *SCOT* showed a much more variable signal, with the only clear peak occurring in October and a high NDVI being maintained throughout summer.

The MODIS and Landsat NDVI values for C_3 Native sites (**Figure 6C**) were comparable, both in magnitude and trend. Most inconsistent data points can be attributed to data gaps. The MODIS data shows consistent temporal patterns with phenocam and biomass trends at all three sites.

Trends of Landsat and MODIS NDVI corresponded well for C_3 Exotic sites (**Figure 6D**), though deviations in magnitude were particularly evident at one site (*GUNE*). Like other functional groups, the paucity of Landsat data beyond January 2015 made comparisons difficult. In general, patterns consistent with C_3 characteristics were shown by the satellite data. All three sites showed greening commencing in July–August with a peak in October, and multiple additional peaks in February and April. Two of the C_3 Exotic sites had very high NDVI values, with peak NDVI exceeding 0.8.

Relationship Between Remotely-Sensed and Biophysical Variables

Variables were identified that were expected to be related to the quantity and quality of green vegetation: green grass biomass, green biomass (i.e., grass + forbs), total biomass, green vegetation cover, dead vegetation cover, pasture height, phenocam g_{CC}, MODIS NDVI, and Landsat NDVI (**Figure 7**).

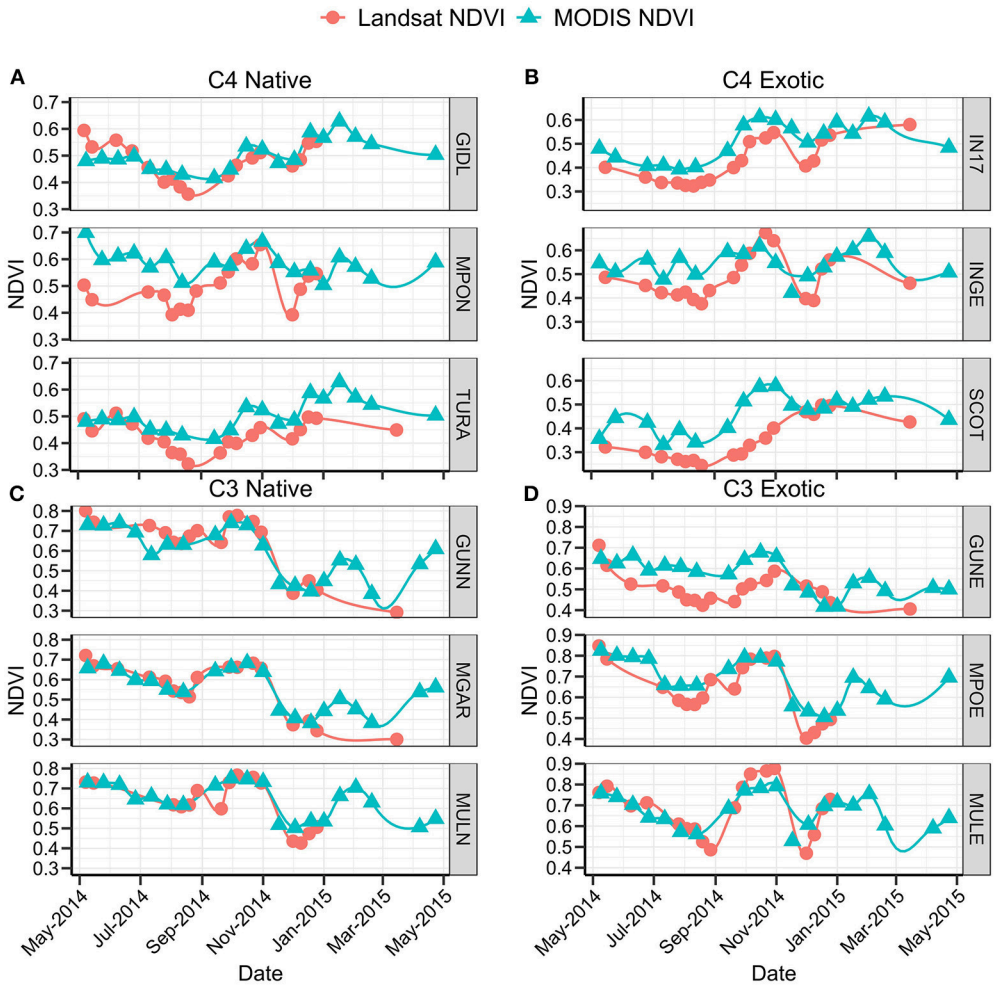


FIGURE 6 | Terra MODIS (▲) and Landsat OLI/ETM+ (•) NDVI data for 1 May 2014–30 April 2015 at **(A)** C_4 Native sites, **(B)** C_4 Exotic sites, **(C)** C_3 Native sites, **(D)** C_3 Exotic sites.

Correlations are presented across all sites as separation of grassland functional types yielded only minor differences when tested. Longitudinal graphical comparisons of all parameters are presented as **Figures 8**, **9** for sites *GIDL* (C_4 Native) and *MULN* (C_3 Native). These figures are examples of the comparison between phenology variables measured at individual sites.

Across all sites, green biomass was strongly and significantly correlated with green grass biomass ($r = 0.99$, $p < 0.001$). This demonstrates that the biomass was dominated by the influence of grass rather than forbs. Green biomass was poorly-correlated with total biomass ($r = 0.25$), indicating that the contribution of the dead biomass component has a strong influence on the overall biomass. Green biomass was only weakly correlated with satellite estimates of NDVI (MODIS NDVI $r = 0.17$; Landsat NDVI $r = 0.26$). Height was not significantly correlated with green cover, dead cover or phenocam g_{CC}, and only weakly correlated with satellite and biomass variables.

Phenocam g_{CC} has a stronger relationship to the fraction of green cover ($r = 0.7$) than its relationship with any of the biomass variables (green biomass $r = 0.41$; green grass biomass $r = 0.38$). A similar strength of correlation is present between green cover and NDVI (Landsat: $r = 0.71$, MODIS: $r = 0.73$), with satellite data also having a stronger correlation to green cover than to green biomass (**Figure 7**). MODIS and Landsat NDVI values are strongly and significantly correlated with one another ($r = 0.88$, $p < 0.001$). The correlation between phenocam g_{CC} and Landsat NDVI ($r = 0.69$) was stronger than that between g_{CC} and MODIS NDVI ($r = 0.59$).

DISCUSSION

Assessment of Phenocams for Monitoring Temperate Grasslands

This research found that the sub-daily image capture available through phenocams allowed for the detection of fine-resolution changes in greenness that were not observed in other methods. Temperate grasslands of the Australian Southern Tablelands

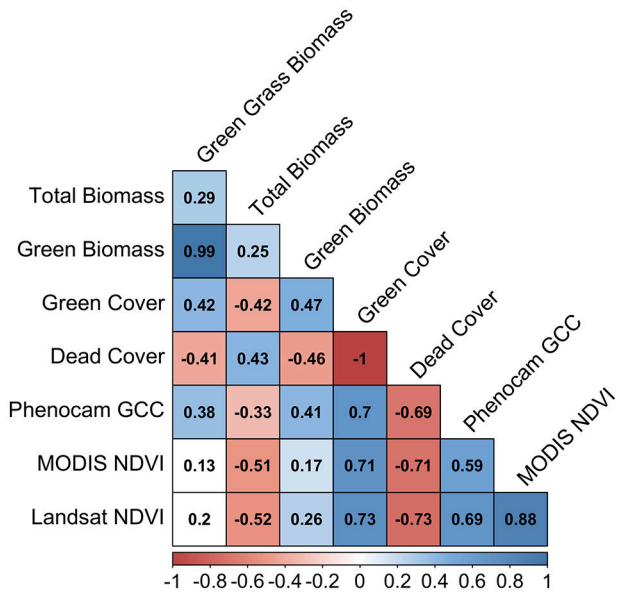


FIGURE 7 | Pearson's correlation of variables across all sites that are relative to quality and quantity of green vegetation. Shaded values represent significant correlation at $p = 0.05$. Values are Pearson's correlation coefficient (r). Negative values (red) indicate a negative correlation; positive values (blue) indicate a positive correlation.

show fluctuations in g_{CC} that represent rapid responses to climatic and environmental change, and phenocams add value for interpreting the dynamics of this vegetation. Changes in phenology due to climate trends typically report differences in scales of days per decade (Parmesan and Yohe, 2003; Badeck et al., 2004; Graham et al., 2009). In many cases, data collected coarser than daily frequency (i.e., satellite and biomass data here presented) will render changes under a certain threshold undetectable. Sampling data at daily frequency also has the capacity to resolve very subtle trends driven by community composition and environmental drivers that are not possible to resolve using other means.

We found the g_{CC} to be a consistent and repeatable vegetation index for capturing the dynamics of temperate grasslands in the subject region. Similar to other studies (Sonnentag et al., 2012) this index was found to be relatively invariant to changes in illumination. Differences in relative angle between the camera and the target have rarely been explored in the phenocam literature and have the potential to be more confounding on RGB digital numbers and greenness indices than errors associated with illumination effects. Standardization of sensor-target angle is recommended for future studies of groundcover vegetation types.

Phenocams currently lack the spectral resolution shown by many satellites and cannot match their spatial coverage. However, phenocams have enormous potential as tools to support ecological monitoring at an intermediate scale that can reliably estimate biophysical variables at sub-daily frequency. Phenocams may provide advantage to the agricultural sector in determining appropriate times for pasture stocking, mowing, or other management actions. On a larger scale, a network of phenocams could be used to track weed invasions, or report on drought impacts—items that are particularly important to Australian agriculture. As this field develops further, there is a growing need to conduct further testing and analysis in novel biomes, quantify illumination, and camera angle effects on RGB indices, and to promote the standardization of methodology to enable cross-continental phenological comparisons (Brown et al., 2016). These is also significant scope for advances in statistical methods to characterize and analyse time-series imagery (Gray and Song, 2013).

C_3/C_4 Phenological Response

We found C_3/C_4 species composition to be the primary driver of phenology patterns in temperate grasslands of the Australian Southern Tablelands, with several key phenological features identified for C_3- and C_4-dominated grasslands. Some of these features are most prominent in the higher temporal frequency methods (e.g., phenocam g_{CC}) and are partially obscured by coarser data sources (e.g., MODIS NDVI). Generally in this region C_3-dominated grasslands showed a steady decline in green signal from May to August (austral late autumn/winter). A relatively sharp greening commenced in late August, reaching a peak in late October to early November as C_3 green leaf expansion was at its maximum. Elevated temperatures and low rainfall in November caused a decline in C_3 greenness until secondary greening occurred in late January driven by higher rainfall. In contrast, C_4-dominated grasslands demonstrated a consistently low quantity of green vegetation from May to August. Green-up commenced in September (early spring) and green-up rates tended to be slower than those observed for C_3-dominated vegetation. Major C_4 vegetation greening in October resulted in either a single peak in December (early summer) or consistently high greenness from December through February with a steady return to low green signal in April. The C_4 grasslands exhibited only minor senescence during spring/summer when higher temperatures and low rainfall promoted browning in C_3-dominated systems. The observed patterns were less variable in C_4 than in C_3 grasslands, which may be attributed to a consistent dominant species in the two C_4 grasslands groups.

Comparison of Field and Remotely Sensed Methods

The relationship between phenocam, field, and satellite phenology variables was analyzed using linear correlation, a similar approach to other researchers (e.g., Zhang et al., 2003). We found that green biomass was not-well correlated with total biomass. This stems from the proportionally high contribution of standing litter in some of the subject temperate grasslands, particularly C_4 grasslands that typically retain standing litter throughout the year (Tremont and McIntyre, 1994). The very strong correlation of green biomass to green grass highlights the importance of the grass component for remote sensing: although forbs contribute most to ecosystem species richness in Australian temperate grasslands (Wimbush and Costin, 1979;

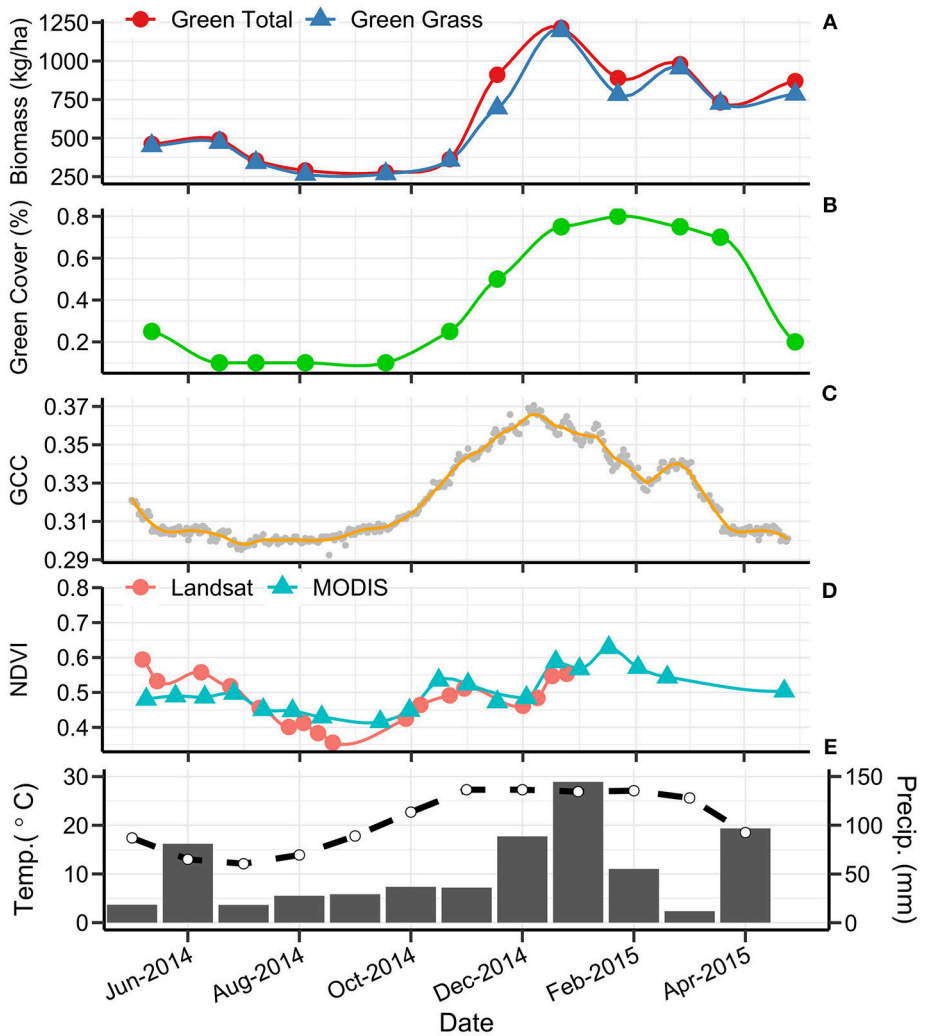


FIGURE 8 | Multi-scale phenology from 1 May 2014 to 30 April 2015 at site GIDL (C_4 Native). From top to bottom panel: **(A)** total green biomass (•) and green grass biomass (▲) in kg/ha; **(B)** green cover (%); **(C)** phenocam 13:00 daily g_{CC}; **(D)**Terra MODIS (▲) and Landsat OLI/ETM+ (•) NDVI; **(E)** monthly mean maximum temperature (°C; dashed line; Tuggeranong Bureau of Meteorology station) and the monthly rainfall (mm; solid bars; Australian National Botanic Gardens Bureau of Meteorology station).

Tremont and McIntyre, 1994), they contribute a small fraction of total biomass. Green cover was only moderately correlated with green biomass ($r = 0.43$) but was negatively correlated with total biomass, again stressing the influence of standing litter. This influence of litter may confound attempts to use remote sensing for the classification, management, and monitoring of grasslands in the study region.

Like other grassland studies (e.g., Paruelo et al., 2000; Migliavacca et al., 2011; Inoue et al., 2015), we found a statistically significant relationship between (a) phenocam indices and green biomass and (b) between phenocam indices and satellite NDVI. However, phenocam g_{CC} was more strongly correlated with green cover than with green biomass and suggests that phenocam data may be more appropriate for estimating cover rather than biomass for Southern Tableland temperate grasslands. Similar research on temperate grasslands in the USA (Vanamburg et al., 2006) suggests that digital camera-based estimates of biomass are poor. However, in European alpine grasslands, Migliavacca et al. (2011) found that their Greenness Index (equivalent to g_{CC}) was significantly correlated with green biomass and visual greenness estimates. Such inconsistencies between temperate grasslands communities suggest differences the biophysical characteristics of different grassland types and indicate that the more complex the grassland structure, the lower the likely correlation between g_{CC} and green biomass. As such, results from this study should not be assumed to apply to functionally different temperate grasslands systems, even within Australia.

Green biomass was only weakly correlated with satellite NDVI, contrary to some grassland studies in northern hemisphere biomes (Migliavacca et al., 2011). The abundant

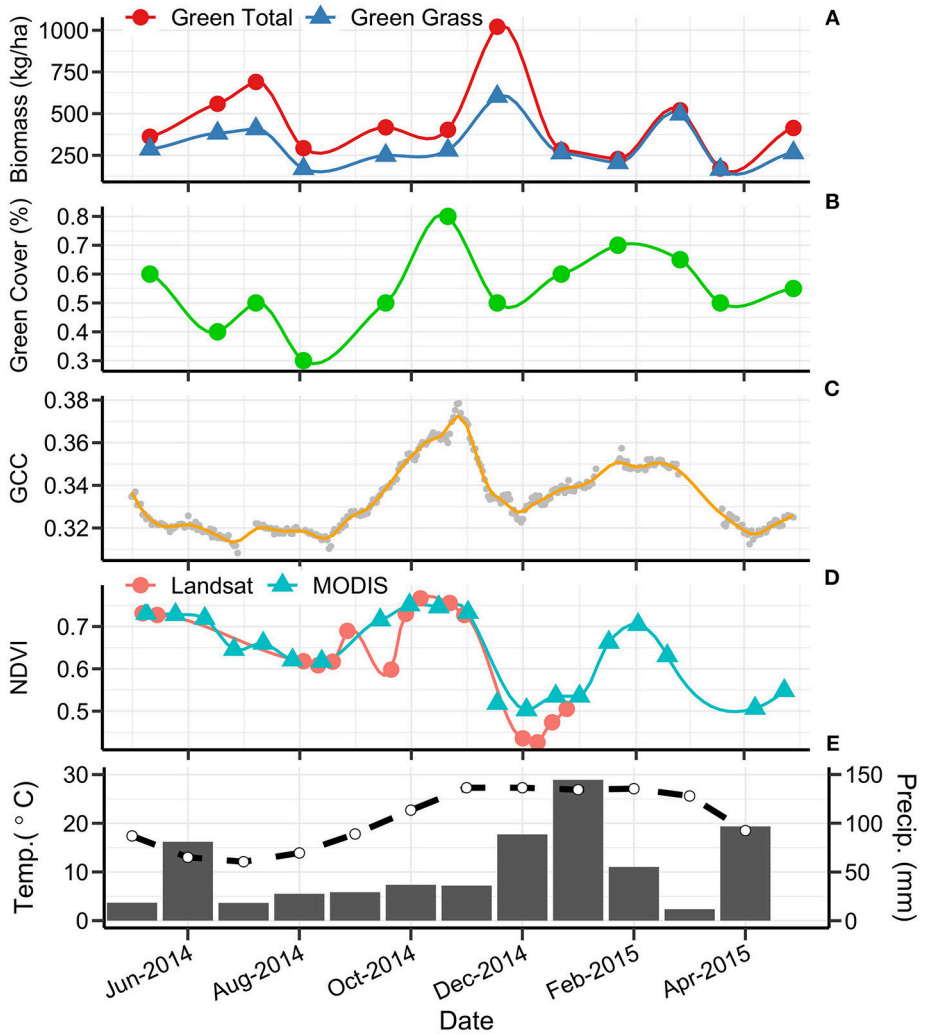


FIGURE 9 | Multi-scale phenology from 1 May 2014 to 30 April 2015 at site MULN (C_3 Native). From top to bottom panel: **(A)** total green biomass (•) and green grass biomass (▲) in kg/ha; **(B)** green cover (%); **(C)** phenocam 13:00 daily g_{CC}; **(D)** Terra MODIS (▲) and Landsat OLI/ETM+ (•) NDVI; **(E)** monthly mean maximum temperature (°C; dashed line; Tuggeranong Bureau of Meteorology station) and the monthly rainfall (mm; solid bars; Australian National Botanic Gardens Bureau of Meteorology station).

standing litter we observed not only obscures green vegetation, but it reflects more light in the red wavelengths, reducing the NDVI signal (Watson et al., 2013). The relatively weak correlation between satellite and ground variables may also be due to data gaps in the satellite time-series and temporal registration between field and satellite data. Overall, we found phenocam g_{CC} to be better correlated with Landsat NDVI than MODIS NDVI. This may be due to the smaller Landsat footprint and highlights the importance of sampling equivalent size plots for remote sensing comparisons (Rienke and Jones, 2006).

Despite the reasonable correlation between phenocam, satellite and field data, we suggest that a 16-day temporal resolution is too coarse to quantify fine changes that occur in dynamic grassland systems in our study region, particularly given that cloud cover impacts often lengthens this period (Stow et al., 2004). Our phenocam data shows consistent periods of increasing and decreasing greenness over as little as 5 days. Other studies have indicated that an effective 16-day revisit time is insufficient to detect key phenological dates (Westergaard-Nielsen et al., 2013). Daily data products from MODIS are available that can be used to generate finer-scale phenology products (Narasimhan and Stow, 2010), but the temporal resolution comes at a trade-off with data quality (e.g., cloud contamination).

Scale transferability when estimating phenology remains a major challenge (Friedl et al., 1994; Eisfelder et al., 2016). While some studies have reported success in this regard (e.g., Fisher et al., 2006; Fisher and Mustard, 2007), other authors caution that there are still major challenges in scaling biophysical measures to satellites (Huete et al., 2002; Soudani et al., 2012). Some researchers in this field (e.g., Hufkens et al., 2012) suggest that

issues of scale and representation (i.e., what is being sampled) strongly influence the relationship between near-surface and satellite remote sensing measures of phenology. This concern is particularly relevant in temperate grasslands within our study region that often have heterogeneous composition in time as well as space. Another fundamental difficulty in comparing different methods is that no method provides a single point of truth. It should be recognized that each method provides subtly different information, uncertainty and errors (Hill et al., 2006).

Sources of Variability and Divergence

The C_4 Native grasslands within this study show the most consistent pattern within the functional groups. This group has the most homogenous composition, dominated by one climax species, *Themeda triandra*, and sites have comparable levels of grazing and other external disturbances. Within other groups, variations from the typical profiles were mostly due to changes in species composition throughout the year. For example, the site *INGE* was dominated by the C_4 exotic grass *Eragrostis curvula*. However, high winter rainfall resulted in a flush of C_3 annual pasture grasses and forbs that produces a green peak in early spring, atypical of a C_4-dominated system. The C_3 Native site *MULN* was found to have a secondary composition of C_4 grasses in the summer, hence exhibited a higher and more consistent greenness through summer months than other C_3-dominated locations. This heterogeneity is difficult to control in natural dynamic systems. Exotic pasture grasses are ubiquitous in low numbers in native pastures (Moore and Perry, 1970), but can flourish if conditions are optimal. This makes grassland ecosystems a challenge for classification—species groups can be abundant 1 year, and rare the next (Vivian and Baines, 2014). Periodic assessment of species composition should be a crucial part of remotely-sensed phenology studies in dynamic systems.

Forbs had low contribution to biomass in the subject grasslands but may be under-represented because common prostrate herbs (e.g., *Trifolium* spp., *Hypochaeris radicata, Solenogyne dominii*) are not as readily collected during sampling. However, the canopy architecture and leaf morphology of planophile forbs intercept and reflect more light than erectophile grasses (Jackson and Pinter, 1986) and they have a low proportion of dead vegetation. As such forbs may contribute proportionally more to measures of vegetation cover and remotely-sensed vegetation indices which may diverge from biomass data. Furthermore, planophile forbs may not be detected by oblique-viewing phenocams, however they are more evident to nadir-viewing satellite remote sensing.

At some sites, the vegetation was trampled by domestic stock during grazing. This changes the canopy architecture, which in turn impacts the spectral reflectance properties and satellite VIs (Mutanga et al., 2005). As the phenocam g_{CC} was more responsive to changes in green cover rather than green biomass, this may represent a cause of divergence in responses between near-surface and satellite methods.

Significant attention has been given to the impact that standing litter has on vegetation indices and phenology estimates (van Leeuwen and Huete, 1996; Nagler et al., 2000; Watson et al., 2013). From the perspective of phenocams, satellites and cover estimates, the growth of new green leaves takes longer to emerge through the standing litter than at a site with no litter. This was demonstrated by our data: live biomass data was recorded even when the phenology curve was at its lowest and green cover data was nil; however, in the case of phenocams, the oblique angles of the cameras further suppress detection of emergent green leaves. The timing of greening estimated from phenocams is likely to be delayed at grasslands when high quantities of standing litter are present. This influence may be further explored in future research by utilizing the MODIS and Landsat fractional cover products that estimate the cover of green vegetation, dead vegetation and background across Australia (Guerschman et al., 2009, 2015).

Study Limitations

The effects of grazing were noted throughout the study but were not controlled. Grazing reduces biomass and has been shown to decrease vegetation index scores (Wylie et al., 2002; Yang and Guo, 2011). Our sites show a variety of grazing pressures from known grazers—notably domestic stock and conspicuous native grazing animals (kangaroos)—but also will have grazing effects from other herbivores (e.g., rabbits, invertebrates). As such, grazing is a difficult variable to control on a large scale. However, phenocam g_{CC} showed that, within the same functional group, sites with higher grazing pressures have similar phenology curves to less heavily grazed sites. Since we have shown g_{CC} to be more closely correlated with cover than biomass in these temperate grasslands, grazing may have less of an effect on g_{CC} than it does on other indices.

Unlike similar studies that have compared radiometric properties of vegetation at field vs. remote scales (Westergaard-Nielsen et al., 2013; Inoue et al., 2015), the current study aimed to use commonly measured field biophysical parameters as the basis for ground-scale comparison. However, adding a field level spectral assessment would increase the detail of inter-scale comparability (Garrity et al., 2011; Hmimina et al., 2013).

The conclusions of this research have been drawn from 1 year of monitoring, as well as historical research on phenology in this region (Hill et al., 1999). This annual window precludes any investigation of temporal climatic factors on phenology profiles. Given the inter-annual variability of Southern Tablelands climate, further years of study would be necessary to disentangle the influence of temperature/rainfall seasonal differences on temperate grassland phenology.

CONCLUSION

In the Australian Southern Tablelands, temperate grasslands represent a continuum from highly productive exotic pastures to diverse native grasslands. Given the significance of grazing agriculture in this region, there is a need to classify and manage different grassland types to integrate conservation and agricultural values under a changing climate. Remote sensing offers the capability to conduct this using land surface phenology of different grassland types, but an understanding of the biophysical and ecological principles underpinning phenology drivers is essential.

The primary driver of phenology in this study was found to be C_3/C_4 species composition. The C_3 grasslands of this study showed moderate greenness in autumn and winter, rapidly increasing to a peak greenness in mid-spring, with secondary peaks following late summer rains. They senesced rapidly when high temperatures and low rainfall coincided. The C_4 grasslands exhibited very low green levels in the winter, began steadily greening from early spring to a summer peak and maintained relatively high values until autumn. C_4 grassland phenology was influenced by the large quantities of standing litter that most sites contained. Previous work to classify vegetation in this region based on phenological profiles successfully distinguished native pastures from sown pastures and forests (Hill et al., 1999). Our distinction of C_3/C_4 dominant functional type adds an extra dimension to the classification process and—coupled with ground truthing—may guide finer-scale discrimination of grasslands communities.

Phenocams were found to be useful for monitoring temperate grassland dynamics as they capture dynamic changes in greening and browning trends over as little as 5 days. High temporal frequency allows for greater resolution than satellite data sources can provide, particularly for regions that have high cloud cover for all or part of the year. Satellite data collection is far superior over larger areas (region to continental scale) but the accuracy of phenology metrics may suffer from decreased temporal collection. Phenocams may assist with agricultural management of temperate grasslands by informing optimal timing for grazing, destocking, and other management actions.

Correlations between phenocam greenness, biomass estimates, and satellite vegetation indices were significant across the data set and fall within the range of agreement found in similar cross-scalar studies. The significance and strength of these relationships was found to differ between grassland functional types. Imperfect correlations between measured variables occur due to different structural, spatial, and spectral differences in the variables being measured.

In temperate grasslands in the Australian Southern Tablelands, phenocams were more effective at estimating green vegetation cover than green biomass. Phenocams showed similar effectiveness as commonly-used satellite products at predicting green cover, however they had the advantage of eliciting immediate greening and browning events. Scaling phenology estimates between field and satellite level is dependent on understanding the underlying biophysical variables being measured.

AUTHOR CONTRIBUTIONS

CW and AH jointly designed the project. CW conducted the fieldwork, laboratory work, analysis and drafting of the manuscript. NR-C provided initial design advice, technical expertise, and analysis.

FUNDING

This research was supported by an Australian Government Research Training Program Scholarship and an Australian Wildlife Conservancy research grant. Support was received by the Australian Terrestrial Ecosystem Research Network (TERN), through phenology workshops and remote sensing funding.

ACKNOWLEDGMENTS

Thanks are due to the custodians of the field sites that facilitated access, to Rakhesh Devadas and Zunyi Xie for providing technical assistance on MODIS/LANDSAT processing, and to Rainer Rehwinkel for assistance in sourcing field sites and tuition on temperate grassland species identification.

REFERENCES

Abràmoff, M. D., Magalhães, P. J., and Ram, S. J. (2004). Image processing with imageJ. *Biophotonics Int.* 11, 36–41. Available online at: https://dspace.library.uu.nl/handle/1874/204900

Adjorlolo, C., Mutanga, O., Cho, M., and Ismail, R. (2012). Challenges and opportunities in the use of remote sensing for C_3 and C_4 grass species discrimination and mapping. *Afr. J. Range Forage Sci.* 29, 47–61. doi: 10.2989/10220119.2012.694120

Agrecon (2004). *Remote Sensing Mapping of Grassy Ecosystems in the Monaro*. Canberra, ACT: Report to the New South Wales Department of Environment and Conservation, Agricultural Reconnaisance Technologies Inc.

Ahl, D. E., Gower, S. T., Burrows, S. N., Shabanov, N. V., Myneni, R. B., and Knyazikhin, Y. (2006). Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS. *Remote Sens. Environ.* 104, 88–95. doi: 10.1016/j.rse.2006.05.003

Ahmadian, N., Ghasemi, S., Wigneron, J. P., and Zölitz, R. (2016). Comprehensive study of the biophysical parameters of agricultural crops based on assessing landsat 8 OLI and landsat 7 ETM + vegetation indices. *GISci. Remote Sens.* 53, 337–359. doi: 10.1080/15481603.2016.1155789

Ahrends, H. E., Bräugger, R., Stöckli, R., Schenk, J., Michna, P., Jeanneret, F., et al. (2008). Quantitative phenological observations of a mixed beech forest in northern Switzerland with digital photography. *J. Geophys. Res.* 113:G04004. doi: 10.1029/2007JG000650

Alberton, B., Almeida, J., Helm, R., Torres, R., Menzel, A., and Morellato, L. P. C. (2014). Using phenological cameras to track the green up in a cerrado savanna and its on-the-ground validation. *Ecol. Inform.* 19, 62–70. doi: 10.1016/j.ecoinf.2013.12.011

Badeck, F. W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J., et al. (2004). Responses of spring phenology to climate change. *N. Phytol.* 162, 295–309. doi: 10.1111/j.1469-8137.2004.01059.x

Baldocchi, D. (2011). The grass response. *Nature* 476, 160–161. doi: 10.1038/476160a

Benson, J. S. (1994). The native grasslands of the monaro region: southern tablelands of NSW. *Cunninghamia* 3, 609–650.

Broich, M., Huete, A., Paget, M., Ma, X., Tulbure, M., Restrepo, N., et al. (2015). A spatially explicit land surface phenology data product for science, monitoring and natural resources management applications. *Environ. Model. Softw.* 64, 191–204. doi: 10.1016/j.envsoft.2014.11.017

Brown, T. B., Hultine, K. R., Steltzer, H., Denny, E. G., Denslow, M. W., Granados, J., et al. (2016). Using phenocams to monitor our changing earth: toward a global phenocam network. *Front. Ecol. Environ.* 14, 84–93. doi: 10.1002/fee.1222

Ceballos, G., Davidson, A., List, R., Pacheco, J., Manzano-Fischer, P., Santos-Barrera, G., et al. (2010). Rapid decline of a grassland system and its ecological and conservation implications. *PLoS ONE* 5:e8562. doi: 10.1371/journal.pone.0008562

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74, 829–836. doi: 10.1080/01621459.1979.10481038

Cui, X., Guo, Z. G., Liang, T. G., Shen, Y. Y., Liu, X. Y., and Liu, Y. (2012). Classification management for grassland using MODIS data: a case study in the Gannan region, China. *Int. J. Remote Sens.* 33, 3156–3175. doi: 10.1080/01431161.2011.634861

Cullen, B. R., Johnson, I. R., Eckard, R. J., Lodge, G. M., Walker, R. G., Rawnsley, R. P., et al. (2009). Climate change effects on pasture systems in south-eastern Australia. *Crop Pasture Sci.* 60, 933–942. doi: 10.1071/CP09019

de Beurs, K. M., and Henebry, G. M. (2004). Land surface phenology, climatic variation, and institutional change: analyzing agricultural land cover change in Kazakhstan. *Remote Sens. Environ.* 89, 497–509. doi: 10.1016/j.rse.2003.11.006

Dye, D. G., Middleton, B. R., Vogel, J. M., Wu, Z., and Velasco, M. (2016). Exploiting differential vegetation phenology for satellite-based mapping of semiarid grass vegetation in the Southwestern United States. *Remote Sens.* 8:33. doi: 10.3390/rs8110889

Eisfelder, C., Kuenzer, C., Dech, S., Eisfelder, C., Kuenzer, C., and Dech, S. (2016). Derivation of biomass information for semi-arid areas using remote-sensing data. *Int. J. Remote Sens.* 33, 2937–2984. doi: 10.1080/01431161.2011.620034

Environment Australia (2000). *Revision of the Interim Biogeographic Regionalisation for Australia (IBRA) and development of Version 5.1 Summary Report*. Canberra, ACT.

Epstein, H., Lauenroth, W., and Burke, I. (1997). Effects of temperature and soil texture on ANPP in the U.S. *Great Plains. Ecol.* 78, 2628–2631. doi: 10.1890/0012-9658(1997)078[2628:EOTAST]2.0.CO;2

Fisher, J. I., and Mustard, J. F. (2007). Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sens. Environ.* 109, 261–273. doi: 10.1016/j.rse.2007.01.004

Fisher, J. I., Mustard, J. F., and Vadeboncoeur, M. A. (2006). Green leaf phenology at landsat resolution: scaling from the field to the satellite. *Remote Sens. Environ.* 100, 265–279. doi: 10.1016/j.rse.2005.10.022

Fontana, F., Rixen, C., Jonas, T., Aberegg, G., and Wunderle, S. (2008). Alpine grassland phenology as seen in AVHRR, VEGETATION and MODIS NDVI time series - a comparison with *in situ* measurements. *Sensors* 8, 2833–2853. doi: 10.3390/s8042833

Friedl, M. A., Schimel, D. S., Michaelsen, J., Davis, F. W., and Walker, H. (1994). Estimating grassland biomass and leaf area index using ground and satellite data. *Int. J. Remote Sens.* 15, 1401–1420. doi: 10.1080/01431169408954174

Garrity, S. R., Bohrer, G., Maurer, K. D., Mueller, K. L., Vogel, C. S., and Curtis, P. S. (2011). A comparison of multiple phenology data sources for estimating seasonal transitions in deciduous forest carbon exchange. *Agric. For. Meteorol.* 151, 1741–1752. doi: 10.1016/j.agrformet.2011.07.008

Gillespie, A. R., Kahle, A. B., and Walker, R. E. (1987). Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques. *Remote Sens. Environ.* 22, 343–365.

Gott, B., Williams, N. S. G., and Antos, M. (2015). "Humans and grasslands - a social history," in *Land of Sweeping Plains,* eds N. S. G. Williams, A. Marshall, and J. W. Morgan (Clayton, VIC: CSIRO Publishing), 6–26.

Graham, E. A., Yuen, E. M., Robertson, G. F., Kaiser, W. J., Hamilton, M. P., and Rundel, P. W. (2009). Budburst and leaf area expansion measured with a novel mobile camera system and simple color thresholding. *Environ. Exp. Bot.* 65, 238–244. doi: 10.1016/j.envexpbot.2008.09.013

Gray, J., and Song, C. (2013). Consistent classification of image time series with automatic adaptive signature generalization. *Remote Sens. Environ.* 134, 333–341. doi: 10.1016/j.rse.2013.03.022

Groves, R. H. (1979). The status and future of Australian grasslands. *N. Z. J. Ecol.* 2, 76–81.

Guerschman, J. P., Hill, M. J., Renzullo, L. J., Barrett, D. J., Marks, A. S., and Botha, E. J. (2009). Estimating fractional cover of photosynthetic vegetation, non-photosynthetic vegetation and bare soil in the Australian tropical savanna region upscaling the EO-1 Hyperion and MODIS sensors. *Remote Sens. Environ.* 113, 928–945. doi: 10.1016/j.rse.2009.01.006

Guerschman, J. P., Scarth, P. F., McVicar, T. R., Renzullo, L. J., Malthus, T. J., Stewart, J. B., et al. (2015). Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. *Remote Sens. Environ.* 161, 12–26. doi: 10.1016/j.rse.2015.01.021

Hill, M. J. (2013). Vegetation index suites as indicators of vegetation state in grassland and savanna: an analysis with simulated SENTINEL 2 data for a North American transect. *Remote Sens. Environ.* 137, 94–111. doi: 10.1016/j.rse.2013.06.004

Hill, M. J., Furnival, P. E., Donald, G. E., and Vickery, P. J. (1999). Pasture land cover in eastern Australia from NOAA - AVHRR NDVI and classified Landsat TM. *Remote Sens. Environ.* 67, 32–50. doi: 10.1016/S0034-4257(98)00075-3

Hill, M. J., Senarath, U., Lee, A., Zeppel, M., Nightingale, J. M., Williams, R. J., et al. (2006). Assessment of the MODIS LAI product for Australian ecosystems. *Remote Sens. Environ.* 101, 495–518. doi: 10.1016/j.rse.2006.01.010

Hmimina, G., Dufrêne, E., Pontailler, J.-Y., Delpierre, N., Aubinet, M., Caquet, B., et al. (2013). Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: an investigation using ground-based NDVI measurements. *Remote Sens. Environ.* 132, 145–158. doi: 10.1016/j.rse.2013.01.010

Horion, S., Cornet, Y., Erpicum, M., and Tychon, B. (2013). Studying interactions between climate variability and vegetation dynamic using a phenology based approach. *Int. J. Appl. Earth Observ. Geoinform.* 20, 20–32. doi: 10.1016/j.jag.2011.12.010

Howden, S., Crimp, S., and Stokes, C. (2008). Climate change and Australian livestock systems: impacts, research and policy issues. *Aust. J. Exp. Agric.* 48, 780–788. doi: 10.1071/EA08033

Huete, A. R., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B. H., Milliman, T., and Richardson, A. D. (2012). Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sens. Environ.* 117, 307–321. doi: 10.1016/j.rse.2011.10.006

Ide, R., and Oguma, H. (2010). Use of digital cameras for phenological observations. *Ecol. Inform.* 5, 339–347. doi: 10.1016/j.ecoinf.2010.07.002

Inoue, T., Nagai, S., Kobayashi, H., and Koizumi, H. (2015). Utilization of ground-based digital photography for the evaluation of seasonal changes in the aboveground green biomass and foliage phenology in a grassland ecosystem. *Ecol. Inform.* 25, 1–9. doi: 10.1016/j.ecoinf.2014.09.013

IPCC (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Gobal and Sectorial Aspects.* Working Group II contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge; New York, NY: Cambridge University Press.

Jackson, R. D., and Pinter, P. J. (1986). Spectral response of architecturally different wheat canopies. *Remote Sens. Environ.* 20, 43–56. doi: 10.1016/0034-4257(86)90013-1

Julitta, T., Cremonese, E., Migliavacca, M., Colombo, R., Galvagno, M., Siniscalco, C., et al. (2014). Using digital camera images to analyse snowmelt and phenology of a subalpine grassland. *Agric. Forest Meteorol.* 198–199, 116–125. doi: 10.1016/j.agrformet.2014.08.007

Justice, C. O., and Hiernaux, P. H. Y. (1986). Monitoring the grasslands of the sahel using NOAA AVHRR data: niger 1983. *Int. J. Remote Sens.* 7, 1475–1497. doi: 10.1080/01431168608948949

Klimeš, L. (2003). Scale-dependent variation in visual estimates of grassland plant cover. *J. Veg. Sci.* 14, 815–821. doi: 10.1111/j.1654-1103.2003.tb02214.x

Li, P., Jiang, L., and Feng, Z. (2014). Cross-comparison of vegetation indices derived from landsat-7 enhanced thematic mapper plus (ETM+) and landsat-8 operational land imager (OLI) sensors. *Remote Sens.* 6, 310–329. doi: 10.3390/rs6010310

Liang, L., Schwartz, M. D., and Fei, S. (2011). Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sens. Environ.* 115, 143–157. doi: 10.1016/j.rse.2010.08.013

Lymburner, L., Tan, P., Mueller, N., Thackway, R., Lewis, A., Thankappan, M., et al. (2011). *The National Dynamic Land Cover Dataset, Record 2011/31*. Canberra, ACT: Geoscience Australia.

Ma, X., Huete, A., Yu, Q., Coupe, N. R., Davies, K., Broich, M., et al. (2013). Spatial patterns and temporal dynamics in savanna vegetation phenology across the North Australian tropical transect. *Remote Sens. Environ.* 139, 97–115. doi: 10.1016/j.rse.2013.07.030

Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., et al. (2011). Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO_2 uptake. *Agric. For. Meteorol.* 151, 1325–1337. doi: 10.1016/j.agrformet.2011.05.012

Mizunuma, T., Wilkinson, M. L., Eaton, E., Mencuccini, M., Morison, I. L., and Grace, J. (2013). The relationship between carbon dioxide uptake and canopy colour from two camera systems in a deciduous forest in southern England. *Funct. Ecol.* 27, 196–207. doi: 10.1111/1365-2435.12026

Moore, C. E., Beringer, J., Evans, B., Hutley, L. B., and Tapper, N. J. (2017). Tree-grass phenology information improves light use efficiency modelling of gross primary productivity for an Australian tropical savanna. *Biogeosci. Discuss.* 14, 111–129. doi: 10.5194/bg-14-111-2017

Moore, C. E., Brown, T., Keenan, T. F., Duursma, R. A., Van Dijk, A. I. J. M., Beringer, J., et al. (2016). Reviews and syntheses: australian vegetation phenology: new insights from satellite remote sensing and digital repeat photography. *Biogeosciences* 13, 5085–5102. doi: 10.5194/bg-13-5085-2016

Moore, R. M., and Biddiscombe, E. F. (1964). "The effects of grazing on grasslands," in: *Grasses and Grasslands*, ed C. Barnard (Melbourne, VIC: MacMillan), 221–235.

Moore, R. M., and Perry, R. A. (1970). "Vegetation," in *Australian Grasslands,* ed R. M. Moore (Canberra, ACT: Australian National University Press), 59–73.

Morgan, J. A., LeCain, D. R., Pendall, E., Blumenthal, D. M., Kimball, B. A., Carrillo, Y., et al. (2011). C_4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. *Nature* 476, 202–205. doi: 10.1038/nature10274

Morgan, J. W., and Lunt, I. D. (1999). Effects of time-since-fire on the tussock dynamics of a dominant grass (*Themeda triandra*) in a temperate Australian grassland. *Biol. Conserv.* 88, 379–386. doi: 10.1016/S0006-3207(98)00112-8

Muir, J., Schmidt, M., Tindall, D., Trevithick, R., Scarth, P., and Stewart, J. B. (2011). *Field Measurement of Fractional Ground Cover: a Technical Handbook Supporting Ground Cover Measurement for Australia*. Canberra, ACT: Australian Bureau of Agricultural and Resource Economics and Sciences.

Mutanga, O., and Skidmore, A. K. (2004). Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* 25, 3999–4014. doi: 10.1080/01431160310001654923

Mutanga, O., Skidmore, A. K., Kumar, L., and Ferwerda, J. (2005). Estimating tropical pasture quality at canopy level using band depth analysis with continuum removal in the visible domain. *Int. J. Remote Sens.* 26, 1093–1108. doi: 10.1080/01431160512331326738

Nagai, S., Ichie, T., Yoneyama, A., Kobayashi, H., Inoue, T., and Ishii, R. (2016). Usability of time-lapse digital camera images to detect characteristics of tree phenology in a tropical rainforest. *Ecol. Inform.* 32, 91–106. doi: 10.1016/j.ecoinf.2016.01.006

Nagai, S., Maeda, T., Gamo, M., Muraoka, H., Suzuki, R., and Nasahara, K. N. (2011). Using digital camera images to detect canopy condition of deciduous broad-leaved trees. *Plant Ecol. Divers.* 4, 79–89. doi: 10.1080/17550874.2011.579188

Nagler, P. L., Daughtry, C. S. T., and Goward, S. N. (2000). Plant litter and soil reflectance. *Remote Sens. Environ.* 71, 207–215. doi: 10.1016/S0034-4257(99)00082-6

Narasimhan, R., and Stow, D. (2010). Daily MODIS products for analyzing early season vegetation dynamics across the North Slope of Alaska. *Remote Sens. Environ.* 114, 1251–1262. doi: 10.1016/j.rse.2010.01.017

NSW Catchment Management Authority (2005). *Native Vegetation Regulation 2005 Clause 28 Policy - Special Provisions for Long Term Environmental Benefits*. NSW CMA.

Parmesan, C., and Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421, 37–42. doi: 10.1038/nature01286

Paruelo, J. M., Lauenroth, W. K., and Roset, P. A. (2000). Estimating aboveground plant biomass using a photographic technique. *J. Range Manage.* 53, 190–193. doi: 10.2307/4003281

Pau, S., Edwards, E. J., and Still, C. J. (2013). Improving our understanding of environmental controls on the distribution of C_3 and C_4 grasses. *Glob. Chang. Biol.* 19, 184–196. doi: 10.1111/gcb.12037

Peart, B. (2008). "Life in a working landscape: towards a conservation strategy for the world's temperate grasslands," in *A Record of the World Temperate Grasslands Conservation Initiative Workshop* (Hohhot).

Petus, C., Lewis, M., and White, D. (2013). Monitoring temporal dynamics of great artesian basin wetland vegetation, Australia, using MODIS NDVI. *Ecol. Indic.* 34, 41–52. doi: 10.1016/j.ecolind.2013.04.009

Price, K. P., Guo, X., and Stiles, J. M. (2002). Optimal landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. *Int. J. Remote Sens.* 23, 5031–5042. doi: 10.1080/01431160210121764

Psomas, A. (2008). *Hyperspectral Remote Sensing for Ecological Analyses of Grassland Ecosystems*. Unpublished PhD Thesis, Universitat Zurich.

Psomas, A., Kneubühler, M., Huber, S., and Itten, K. (2011). Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of grassland habitats. *Int. J. Remote Sens.* 32, 9007–9031. doi: 10.1080/01431161.2010.532172

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rayburn, E., and Rayburn, S. (1998). A standardized plate meter for estimating pasture mass in on-farm research trials. *Agron. J.* 90, 238–241. doi: 10.2134/agronj1998.00021962009000020022x

Reed, B. C., Schwartz, M. D., and Xiao, X. (2009). "Remote sensing phenology: status and the way forward," in *Phenology of Ecosystem Processes: Applications in Global Change Research,* ed A. Noormets (New York, NY: Springer-Verlag), 231–246. doi: 10.1007/978-1-4419-0026-5_10

Richardson, A. D., Braswell, B. H., Hollinger, D. Y., Jenkins, J. P., and Ollinger, S. V. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecol. Appl.* 19, 1417–1428. doi: 10.1890/08-2022.1

Richardson, A. D., Jenkins, J. P., Braswell, B. H., Hollinger, D. Y., Ollinger, S. V., and Smith, M.-L. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia* 152, 323–334. doi: 10.1007/s00442-006-0657-z

Rienke, K., and Jones, S. (2006). Integrating vegetation field surveys with remotely sensed data. *Ecol. Manage. Restor.* 7, S18–S23. doi: 10.1111/j.1442-8903.2006.00287.x

Scurlock, J. M. O., and Hall, D. O. (1998). The global carbon sink: a grassland perspective. *Glob. Chang. Biol.* 4, 229–233. doi: 10.1046/j.1365-2486.1998.00151.x

Shen, M., Tang, Y., Klein, J., Zhang, P., Gu, S., Shimono, A., et al. (2008). Estimation of aboveground biomass using in situ hyperspectral measurements in five major grassland ecosystems on the Tibetan Plateau. *J. Plant Ecol.* 1, 247–257. doi: 10.1093/jpe/rtn025

Shimada, S., Matsumoto, J., Sekiyama, A., Aosier, B., and Yokohana, M. (2012). A new spectral index to detect poaceae grass abundance in mongolian grasslands. *Adv. Space Res.*50, 1266–1273. doi: 10.1016/j.asr.2012.07.001

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., et al. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agric. For. Meteorol.* 152, 159–177. doi: 10.1016/j.agrformet.2011.09.009

Soudani, K., Hmimina, G., Delpierre, N., Pontailler, J.-Y., Aubinet, M., Bonal, D., et al. (2012). Ground-based network of NDVI measurements for tracking temporal dynamics of canopy structure and vegetation phenology in different biomes. *Remote Sens. Environ.* 123, 234–245. doi: 10.1016/j.rse.2012.03.012

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S. (2003). Global distribution of C_3 and C_4 vegetation: carbon cycle implications. *Glob. Biogeochem. Cycles* 17, 6–14. doi: 10.1029/2001GB001807

Stow, D. A., Hope, A., Mcguire, D., Verbyla, D., Gamon, J., Huemmrich, F., et al. (2004). Remote sensing of vegetation and land-cover change in arctic tundra ecosystems. *Remote Sens. Environ.* 89, 281–308. doi: 10.1016/j.rse.2003.10.018

Studer, S., Stöckli, R., Appenzeller, C., and Vidale, P. L. (2007). A comparative study of satellite and ground-based phenology. *Int. J. Biometeorol.* 51, 405–414. doi: 10.1007/s00484-006-0080-5

Toomey, M., Friedl, M. A., Frolking, S., Hufkens, K., Klosterman, S., Sonnentag, O., et al. (2015). Greenness indices from digital cameras predict the timing and seasonal dynamics of canopy-scale photosynthesis. *Ecol. Appl.* 25, 99–115. doi: 10.1890/14-0005.1

Tremont, R., and McIntyre, S. (1994). Natural grassy vegetation and native forbs in temperate Australia: structure, dynamics and life histories. *Aust. J. Bot.* 42, 641–658. doi: 10.1071/BT9940641

van Leeuwen, W. J. D., and Huete, A. R. (1996). Effects of standing litter on the biophysical interpretation of plant canopies with spectral indices. *Remote Sens. Environ.* 138, 123–138. doi: 10.1016/0034-4257(95)00198-0

Vanamburg, L. K., Trlica, M. J., Hoffer, R. M., Weltz, M. A., Trlica, M. J., Hoffer, R. M., et al. (2006). Ground based digital imagery for grassland biomass estimation. *Int. J. Remote Sens.* 27, 939–950. doi: 10.1080/01431160500114789

Vivian, L., and Baines, G. (2014). *Research Update 2014/4: Longitudinal Study of Groundcover Flora Condition in Select Grassy Ecosystem Sites*. Canberra, ACT: Environment and Planning Directorate.

Wand, S. J. E., Midgley, G. F., Jones, M. H., and Curtis, P. S. (1999). Elevated atmospheric CO_2 concentration: a meta-analytic test of current theories and perceptions. *Glob. Chang. Biol.* 5, 723–741. doi: 10.1046/j.1365-2486.1999.00265.x

Wang, C., Hunt, E. R., Zhang, L., and Guo, H. (2013). Phenology-assisted classification of C_3 and C_4 grasses in the U.S. Great plains and their climate dependency with MODIS time series. *Remote Sens. Environ.* 138, 90–101. doi: 10.1016/j.rse.2013.07.025

Watson, C. J., Restrepo Coupe, N., and Huete, A. R. (2013). "Hyperspectral assessments of condition and species composition of Australian grasslands," in: *Proceedings of the 2013 IEEE International Geoscience & Remote Sensing Symposium* (Melbourne, VIC), 2770–2773.

Weiser, R. L., Asrar, G., Miller, G. P., and Kanemasu, E. T. (1986). Assessing grassland biophysical characteristics from spectral measurements. *Remote Sens. Environ.* 20, 141–152. doi: 10.1016/0034-4257(86)90019-2

Westergaard-Nielsen, A., Lund, M., Ulf, B., and Peter, M. (2013). Camera-derived vegetation greenness index as proxy for gross primary production in a low Arctic wetland area. *ISPRS J. Photogr. Remote Sens.* 86, 89–99. doi: 10.1016/j.isprsjprs.2013.09.006

Wimbush, D., and Costin, A. B. (1979). Trends in vegetation at Kosciusko. II: subalpine range transects, 1959-1978. *Austr. J. Bot.* 27, 789–831. doi: 10.1071/BT9790789

Winslow, J. C., Hunt, E. R., and Piper, S. C. (2003). The influence of seasonal water availability on global C_3 versus C_4 grassland biomass and its implications for climate change research. *Ecol. Modell.* 163, 153–173. doi: 10.1016/S0304-3800(02)00415-5

Woebbecke, D. M., Meyer, G. E., Von Bargen, K., and Mortensen, D. A. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE* 38, 259–269. doi: 10.13031/2013.27838

Wong, N., and Dorrough, J. W. (2015). "Integrating grassland conservation into farming practice," in: *Land of Sweeping Plains,* eds N. S. G. Williams, A. Marshall, and J. W. Morgan (Clayton,VIC: CSIRO Publishing), 253–284.

Wylie, B. K., Meyer, D. J., Tieszen, L. L., and Mannel, S. (2002). Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands - a case study. *Remote Sens. Environ.* 79, 266–278. doi: 10.1016/S0034-4257(01)00278-4

Yang, X., and Guo, X. (2011). Investigating vegetation biophysical and spectral parameters for detecting light to moderate grazing effects: a case study in mixed grass prairie. *Central Eur. J. Geosci.* 3, 336–348. doi: 10.2478/s13533-011-0032-4

Zhang, Q., Xiao, X., Braswell, B. H., Linder, E., Ollinger, S., Smith, M., et al. (2006). Characterization of seasonal variation of forest canopy in a temperate deciduous broadleaf forest, using daily MODIS data. *Remote Sens. Environ.* 105, 189–203. doi: 10.1016/j.rse.2006.06.013

Zhang, X., Friedl, M. A., and Schaaf, C. B. (2006). Global vegetation phenology from moderate resolution imaging spectroradiometer (MODIS): evaluation of global Patterns and comparison with in situ measurements. *J. Geophys. Res.* 111, 1–14. doi: 10.1029/2006JG000217

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., et al. (2003). Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* 84, 471–475. doi: 10.1016/S0034-4257(02)00135-9

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The review editor NN is currently co-organizing a Research Topic with one of the authors AH, and confirms the absence of any other collaboration.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 October 2018
doi: 10.3389/fpls.2018.01544

Aerial Imagery Analysis – Quantifying Appearance and Number of Sorghum Heads for Applications in Breeding and Agronomy

Wei Guo[1], Bangyou Zheng[2], Andries B. Potgieter[3], Julien Diot[4], Kakeru Watanabe[5], Koji Noshita[5], David R. Jordan[6], Xuemin Wang[6], James Watson[3], Seishi Ninomiya[1] and Scott C. Chapman[2,7]*

[1] International Field Phenomics Research Laboratory, Institute for Sustainable Agro-ecosystem Services, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, [2] Agriculture and Food – Commonwealth Scientific and Industrial Research Organisation, St Lucia, QLD, Australia, [3] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Toowoomba, QLD, Australia, [4] Montpellier SupAgro, Montpellier, France, [5] Laboratory of Biometry and Bioinformatics, Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan, [6] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Warwick, QLD, Australia, [7] School of Agriculture and Food Sciences, The University of Queensland, Gatton, QLD, Australia

OPEN ACCESS

Edited by:
Yann Guédon, Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), France

Reviewed by:
Zhanguo Xin, Agricultural Research Service (USDA), United States
Thiago Teixeira Santos, Empresa Brazileira de Pesquisa Agropecuária (EMBRAPA), Brazil

***Correspondence:**
Wei Guo
guowei@isas.a.u-tokyo.ac.jp

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *23 June 2018*
Accepted: *02 October 2018*
Published: *23 October 2018*

Citation:
Guo W, Zheng B, Potgieter AB, Diot J, Watanabe K, Noshita K, Jordan DR, Wang X, Watson J, Ninomiya S and Chapman SC (2018) Aerial Imagery Analysis – Quantifying Appearance and Number of Sorghum Heads for Applications in Breeding and Agronomy. Front. Plant Sci. 9:1544. doi: 10.3389/fpls.2018.01544

Sorghum (*Sorghum bicolor* L. Moench) is a C4 tropical grass that plays an essential role in providing nutrition to humans and livestock, particularly in marginal rainfall environments. The timing of head development and the number of heads per unit area are key adaptation traits to consider in agronomy and breeding but are time consuming and labor intensive to measure. We propose a two-step machine-based image processing method to detect and count the number of heads from high-resolution images captured by unmanned aerial vehicles (UAVs) in a breeding trial. To demonstrate the performance of the proposed method, 52 images were manually labeled; the precision and recall of head detection were 0.87 and 0.98, respectively, and the coefficient of determination (R^2) between the manual and new methods of counting was 0.84. To verify the utility of the method in breeding programs, a geolocation-based plot segmentation method was applied to pre-processed ortho-mosaic images to extract >1000 plots from original RGB images. Forty of these plots were randomly selected and labeled manually; the precision and recall of detection were 0.82 and 0.98, respectively, and the coefficient of determination between manual and algorithm counting was 0.56, with the major source of error being related to the morphology of plants resulting in heads being displayed both within and outside the plot in which the plants were sown, i.e., being allocated to a neighboring plot. Finally, the potential applications in yield estimation from UAV-based imagery from agronomy experiments and scouting of production fields are also discussed.

Keywords: high-throughput phenotyping, UAV remote sensing, sorghum head detecting and counting, breeding field, image analysis

INTRODUCTION

The grain yield of cereal crops is determined by accumulated processes of resource capture (e.g., radiation, water, and nutrients) that support net photosynthesis across the growing season (i.e., the carbohydrate source) and the utilization of this source, especially in the critical period around the reproductive stage which allows establishment of a potential sink (grain number) and, later in the crop, to fill those grains. These processes and their complex interrelationships form the basis of physiological models of crop growth and the development of crops such as sorghum (*Sorghum bicolor* L. Moench) (Hammer et al., 2010). Plant breeders and agronomists work collectively to modify these processes via genetics and management to develop cropping systems that optimize adaptation to different environments, particularly those associated with drought and heat (Lobell et al., 2015; Potgieter et al., 2016).

On an area basis, the final grain yield of cereal crops in a plot can be described as the product of average values of plant population, fertile head number per plant (i.e., main stem plus tillers), seeds per head, and individual seed mass. Insights into the changes in these component traits through the season and their final values at harvest provide researchers with a better understanding of crop adaptation, and potentially allow breeders to select for different combinations of these traits in different environments. The process of tillering provides a flexible or "plastic" response to challenging environments such as drought, and the trait of fertile head number per plant is under strong genetic control in both sorghum (Lafarge et al., 2002) and wheat (*Triticum aestivum*) (Mitchell et al., 2013; Dreccer et al., 2014). Although all of these component traits can be measured through labor-intensive hand-sampling methods, plant breeders and agronomists doing large trials will typically only use measures of yield (via plot harvester) and individual seed mass (via sample of grains from each plot). Together with estimates of plant population, which can be done by counting the emerged plants using ground and aerial images (Gnädinger and Schmidhalter, 2017; Jin et al., 2017; Liu et al., 2017), rapid and precise estimates of fertile head number per unit area would allow researchers to estimate the fertile head number per plant as an indicator of "tillering propensity."

The aim of the research presented here was to develop a method that can detect and count the heads of sorghum from unmanned aerial vehicle (UAV) images, and then apply the method to specific plots to meet the needs of breeding programs. Machine-based image algorithms for detecting and counting an agriculture product with the use of harvesting robots and ground monitoring vehicles have been applied to imagery of grapes, tomato, apple, mango, and citrus fruits (Nuske et al., 2011; Payne et al., 2014; Sengupta et al., 2014; Yamamoto et al., 2014; Linker and Kelman, 2015; Gongal et al., 2016; Qureshi et al., 2016). However, these algorithms were designed to handle high-resolution images that do not include targets with large shape variations. Therefore, they are not suited for use with either the images or target object taken by UAVs in a breeding field of sorghum where different genotypes have heads that vary in color and shape, with these differences potentially changing with environment. In this paper, we propose a two-step machine-learning-based method that can detect and count sorghum heads from aerial images. To the best of our knowledge, this is the first report of research of this type.

MATERIALS AND METHODS

Field Experiments and Image Acquisition

The field experiments were part of multi-environment advanced yield-testing trials in a sorghum pre-breeding program. The trial was sown on 22 December 2015 at Hermitage, QLD, Australia (latitude: 28.21° S, longitude: 152.10° E, altitude: 459 m above sea level) during the 2015–2016 summer growing season. The target plant density was 115,000 plants/ha, with genotypes planted in plots comprising two 5-m-long rows. The plants were sown in plots within columns, and the trial used a solid row configuration, with a row spacing of 0.76 m between the two rows and a distance between two neighboring plots of 1 m as shown in **Figure 1**. In this trial, 1440 plots (laid out as 36 columns × 40 double-row plots; hereafter, we refer to double-row plots as rows, i.e., 36 columns × 40 rows) were sown, with several columns (216 plots in total) being "filler plots" to allow access for spraying. The trial comprised 22 check hybrids and 903 test-cross hybrids derived from crossing between a range of elite male parents and two female testers in the breeding program. The check hybrids were replicated at least four times, whereas 220 of the 903 test hybrids were replicated twice, with no replication at this site for the remaining 683 hybrids. The trial field was rain-fed and managed according to local management practices.

A UAV (Modified 3DR X8, Skywalker Technology Co., Ltd., China) was flown over the field with a pre-designed flight plan controlled with Mission Planner (open-source flight planning software for Pix Hawk autopilot[1]). The path included substantial overlap (i.e., 70% front-overlap and 80% side-overlap) at flight heights of 20 m and a flight speed of 3 m/s. The total flight time was approximately 50 min (five 10-min flights to cover the whole field). A commercial RGB camera (Sony Cyber-shot DSC-RX100M3, Tokyo, Japan) was mounted on the UAV in a landscape format. The resolution of the camera was set to 5472 × 3648 pixels, which resulted in an average ground sampling distance of 0.45 cm at 20 m height with a footprint of 20 m. The image sets were captured at 1-s intervals during the flights, so that about 2000 images were produced for the target field (about 35 GB+) per flight.

Data Preparation

Flight data obtained on 24 March 2016 were chosen for this study because almost all of the genotypes were heading stages at this time, so that the widest diversity of heads in terms of color and shape could be found in these images and the dataset would be large and balanced enough for later processing. We believe, however, that the image processing algorithm conducted on this dataset has the general capabilities to be used on the whole dataset, including images taken at different times of the

[1]http://planner.ardupilot.com/

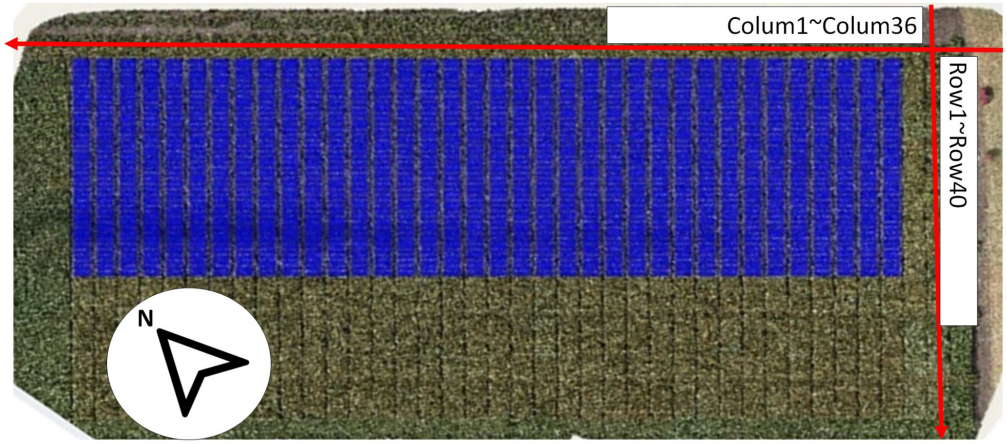


FIGURE 1 | Experimental field layout.

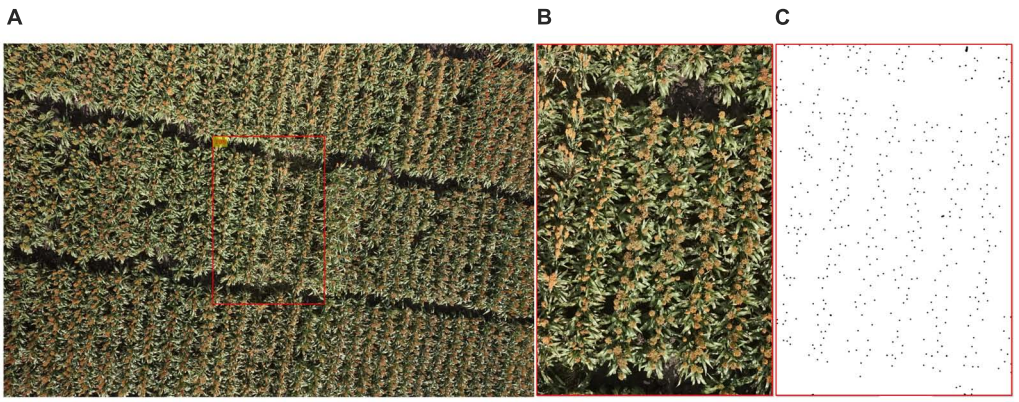


FIGURE 2 | An example of image preparation for algorithm development: **(A)** original image, **(B)** cropped image, and **(C)** manually labeled cropped image, the points represent the heads (dataset 1).

growth season. From the 2109 original images obtained on 24 March 2016 when most of the plants in the experiment had produced heads (on average, this date was about one to 2 weeks after anthesis), 52 images were randomly selected following an uniform distribution to develop and test the head detecting and counting algorithm. To minimize the influence of camera lens distortion, all of the images were cropped so that about 10% of image was used (**Figure 2**). The original image of 5472 × 3648 pixels was cropped to a central region of 1154 × 1731 pixels, which corresponded to an area of 2.3 m × 3.5 m which contained three to five plots. All of 52 cropped images were carefully hand labeled with points in Adobe Photoshop (Adobe Systems Inc., San Jose, CA, United States) as shown in **Figure 2C**. These images were grouped as Dataset 1.

Sorghum Head Detection

The main challenges of creating an image-based solution in a real breeding field are: (1) changing light conditions within a single flight (images vary in color; **Figure 3A**); (2) complex background (**Figure 3B**); and (3) head variations in color, size, and shape caused by light conditions, genotype, heading stage, source of head (main stem or tillers), angle of head stands, and overlapping of heads (**Figures 3C,D**).

To overcome the first challenge, in our previous work (Potgieter et al., 2015; Guo et al., 2016), we proposed a two-step machine-learning, voting-based method. The method uses colors (RGB, HSV [hue, saturation, and value], Lab from related color space, ExG [excess green], and ExR [excess red]) introduced by Meyer and Neto (2008); texture features (average gray level, average contrast, measure of smoothness, third moment, measure of uniformity, and entropy) from gray-scale imagery introduced by Gonzalez et al. (2010); and contrast, correlation, energy, and homogeneity from the gray-level co-occurrence matrix introduced by Haralick et al. (1973) to train several decision-tree-based pixel segmentation models (DTSM) (Guo et al., 2013). These DTSM models are then used to segment the images to sorghum and non-sorghum head regions. Based on the segmented regions created in the last step, a bag of visual words approach, modified from Guo et al. (2015), is applied to the test images again to gain a new segmentation

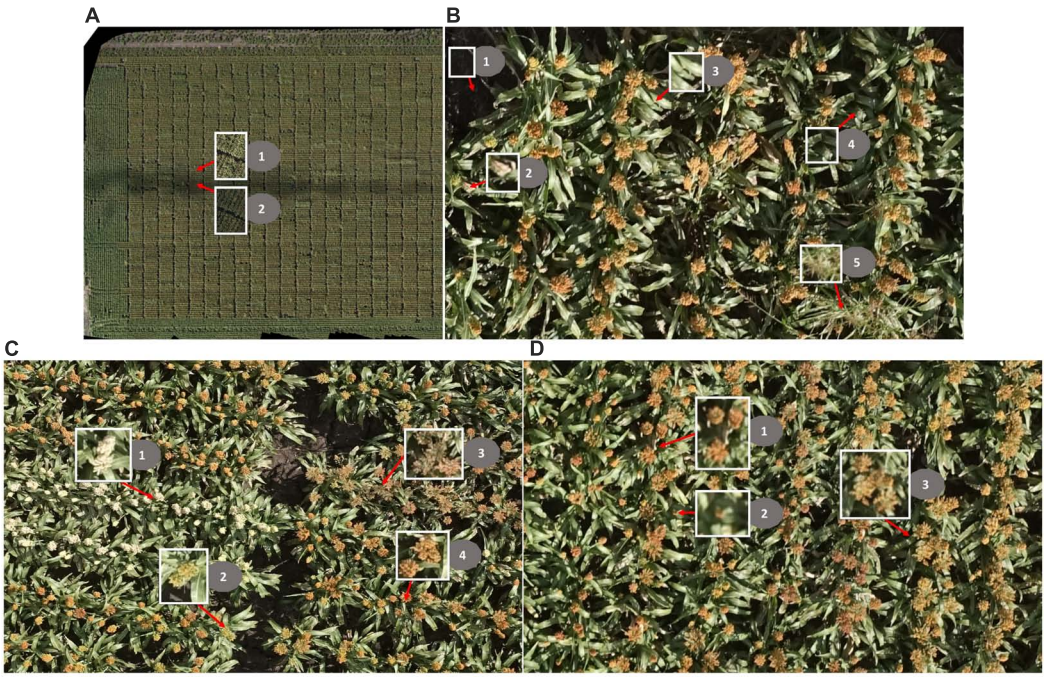


FIGURE 3 | Challenges of head detection in a real field. **(A)** Changing light conditions within one flight: (1) images taken under sunny conditions; (2) images taken under cloudy conditions. **(B)** Complex background: (1) soil/ground (shadowed partially/fully), (2) dead leaves, (3) green leaves, (4) shadowed leaves, and (5) grass. **(C)** The sorghum heads vary in color: (1) white, (2) green, (3) brown, and (4) orange. **(D)** The sorghum heads vary in size and shape: (1) heads from main stem, (2) heads from tillers, and (3) overlapping heads; note that the shape of the heads is compact in 1 and 2 but is expanded in 3.

image with misclassifications removed from the previous step. Finally, a voting process is used for all the segmented images to acquire the most reliably detected region of the sorghum heads. With only 20 test images cropped from a GoPro^{TM} Hero4 camera (GoPro, Inc., San Mateo, CA, United States), the method showed good accuracy for sorghum head detection; the precision (the proportion of correctly detected head region inside true head region) and recall (the proportion of correctly detected head region inside detected head region) were 0.95 and 0.96, respectively (Guo et al., 2016). However, since this method used texture feature and sliding window, the computation time and cost was substantial for processing of high resolution images and did not suit practical use.

Using the knowledge gained from our previous studies (Potgieter et al., 2015; Guo et al., 2016), here we only used color features to train a pixel-based segmentation model. First, seven classes – (1) background soil, (2) background shadow, (3) background dead leaves, (4) leaves, (5) green heads, (6) orange heads, and (7) white heads – were defined. For each class, a series of nine color features (r, g, b; H, S, V; L*, a*, and b*) from three standard color spaces were carefully collected from 17 images (**Figure 4**, Dataset 0) that were selected from the entire image dataset of 2109 images, considering the diversity of lighting conditions and head colors. Using these features, we trained a DTSM model and applied it to all of the test images to classify their pixels into the seven classes. DTSM is a supervised machine learning approach based on the decision tree (DT) (Guo et al., 2013, 2017). This approach generates a decision tree model using the selected color features and corresponded classes, then a constructed tree model is applied to segment test images, such that each pixel becomes assigned to one of the classes (**Figures 5A,B**). After this, the head-related pixels (green heads, orange heads, and white heads) were selected and integrated together into "head regions," as shown in **Figure 5C**.

Sorghum Head Counting

To count the number of detected regions (**Figure 5D**) from the first step with a reliable model, we randomly separated Dataset 1 into six sets, each set being eight or nine images. The images from five of the sets were used to train the model with fivefold cross validation, and the last set was used to estimate the performance of the model. In detail, hand-labeled images were used to extract the 11 morphology features of all of the candidate head regions and the corresponding head numbers (**Figures 5C,D**):

(1) Area: actual number of pixels in each candidate head region.
(2) Eccentricity: eccentricity of the ellipse that has the same second-moments as the candidate head region.
(3) Extent: ratio of pixels in the candidate head region to pixels in the total bounding box.
(4) Perimeter: total number of pixels around the boundary of the candidate head region.
(5) Major axis length: length of the major axis of the ellipse that has the same normalized second central moments as the candidate head region.

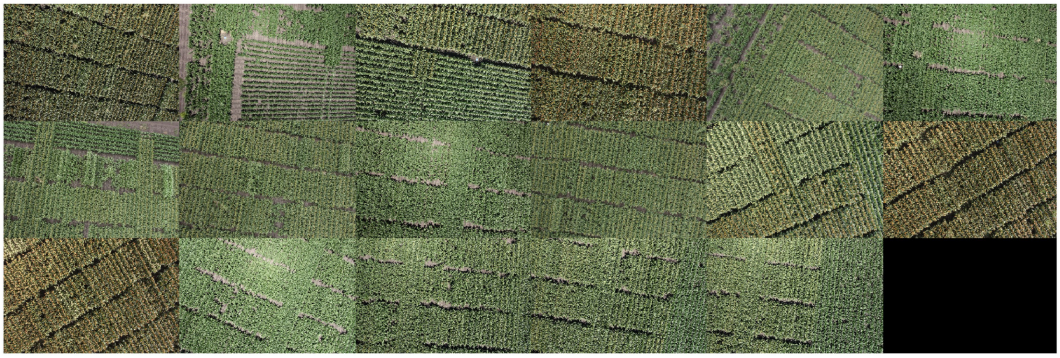


FIGURE 4 | Dataset 0 comprised 17 images for training data collection of pixel-based segmentation model. The images were selected considering light condition, head color, head shape, and background.

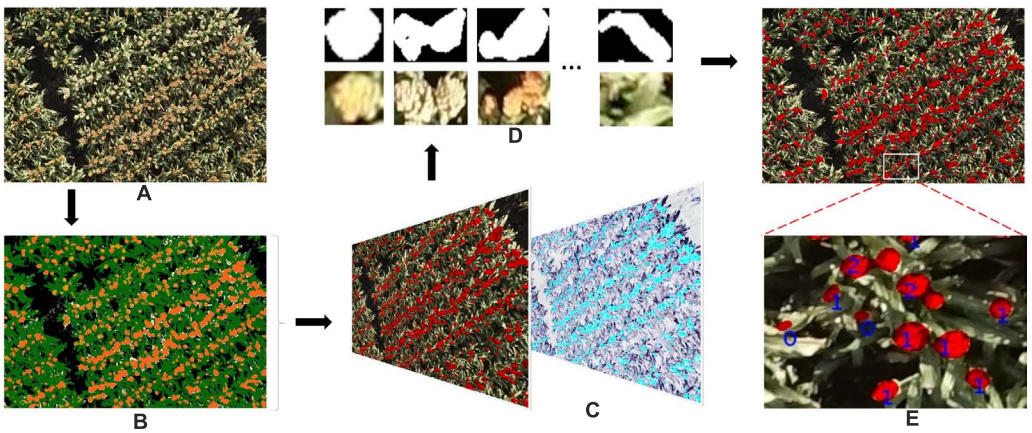


FIGURE 5 | The work flow of the proposed method of detecting and counting sorghum heads. **(A)** Original image. **(B)** Pseudo-color image demonstrating pixel classification result by DTSM: white head, yellow; soil, gray; shadows, black; dead leaves, off-white; leaves, green; orange heads, dark orange; and green heads, light orange. **(C)** Detected head regions (left) and overlapped with manually pointed head image (right). The black dots indicate heads pointed manually with Photoshop. **(D)** The head regions cropped from original images based on **(C)**. **(E)** Detected head regions and number of heads counted. The numbers shown in the image indicate the number of the sorghum heads; 0 means incorrect detection..

(6) Minor axis length, length of the minor axis of the ellipse that has the same normalized second central moments as the candidate head region.
(7) ConvexArea, number of pixels in smallest convex polygon that can contain the candidate head region.
(8) FilledArea: number of pixels in each candidate head region with all holes filled in.
(9) EquivDiameter: diameter of a circle with the same area as the candidate head region.
(10) Solidity: proportion of the pixels in the convex hull that are also in the candidate head region.
(11) Roundness: circularity of candidate head region.

These features of each candidate head region were then used as predictors with corresponded head numbers as the response, in order to train a Quadratic-SVM (Support Vector Machine) classifier with fivefold cross validation. Support Vector Machine is a supervised machine learning algorithm which has become commonly used to solve classification problems. SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes. In this paper, a quadratic kernel is used, as it is less computationally intensive but has been show to perform as well as previous work (Guo et al., 2016).

Then model was applied to all of the candidate regions from step 1 to count the numbers of heads in each image (**Figure 5E**). The training data and a guidance is also provided in **Supplementary Materials**, which can support opportunities for readers to test other classifiers (such as Decision Trees, Random Forest, SVMs with different kernel functions) using the MATLAB "*Classification Learner*" application.

Application of Method to Count Heads Within Individual Plots

In total, 2109 original images were also processed by the Pix4Dmapper software package (Pix4D, SA, Lausanne, Switzerland) to generate 3D point cloud and ortho-mosaic images of the whole experimental field. The ortho-mosaic images

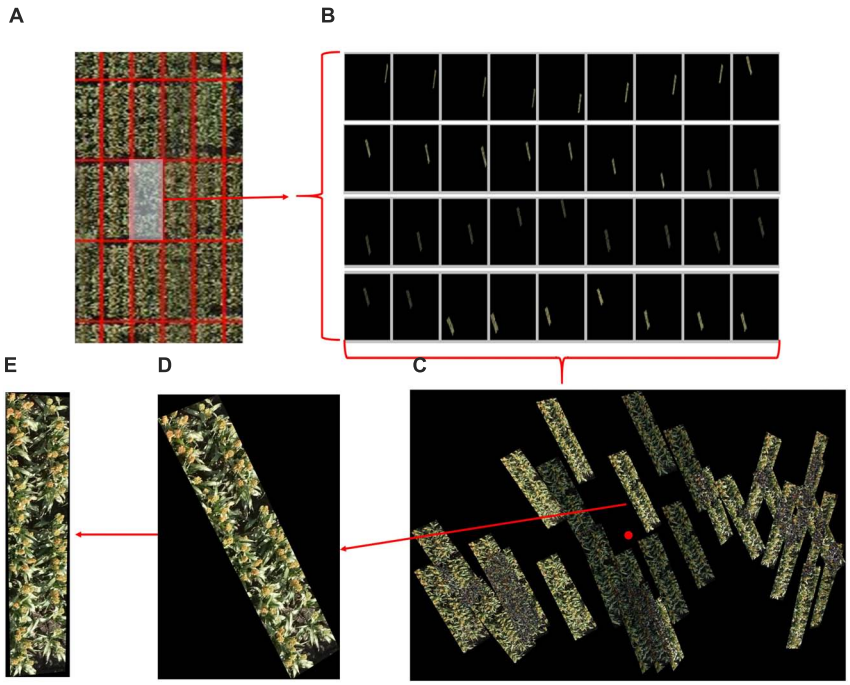


FIGURE 6 | An example of plot segmentation and identification from original images. **(A)** A plot is selected from a set of ortho-mosaic images. **(B)** The selected plot appears in several original images but in different locations. **(C)** The plots images are grouped and one is selected based on its distance from the central part of the image. **(D)** The selected plot is cropped from the corresponding original image. **(E)** The selected plot is rotated based the corner detection and orientation calculations of **(D)**.

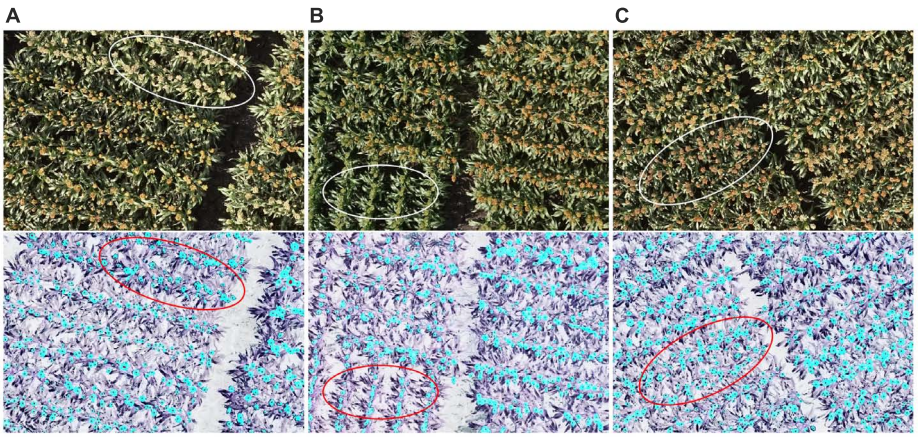


FIGURE 7 | An example of head detection. Images contain **(A)** white heads, **(B)** green heads, and **(C)** brown heads. All of the images contain orange heads. The upper panels show the original images and the lower ones show the detected head regions (blue) and hand-labeled head centers (black dots). Almost all of the heads of the different colors were detected by the proposed model.

were segmented into individual plots and projected back to the corresponding original images to segment the original pixels (cf. mosaic) following our previously reported method (Duan et al., 2016). In total, 28,825 individual plot images (i.e., many replications of each of the 1440 plots) were segmented from the dataset. The process of plot segmentation and identification from the original images is shown in **Figure 6**. Any given plot can appear in several original images but in different locations (**Figures 6A,B**); the plot with the shortest Euclidean distance to the central part of the image was selected as the candidate plot image and cropped from the original image and rotated with calculated orientation (**Figures 6C,D**). Each plot was thereby generated from 1440 images, and 40 of them were randomly selected for this study. To validate the accuracy of the proposed detection and counting method, each plot image was also carefully hand labeled by two scientists (Dataset 2).

RESULTS AND DISCUSSION

Datasets 1 and 2 were both used to evaluate the head detecting and counting capabilities of the proposed method.

Figure 7 shows head detection results from dataset 1. Almost all of the heads of different colors, shapes, and sizes were successfully detected. **Table 1** presents an evaluation of the detection results in terms of precision and recall based on the definitions of Davis and Goadrich (2006). Precision indicates that for the total number of head regions, which proportion were correctly detected (with a ratio of 1.0 being perfect) while recall indicates for all detected regions, how many are correctly detected (perfect = 1.0). The algorithm was able to accurately detect 87% of sorghum heads for dataset 1 and 82% for dataset 2, and the accuracy rates were high (recall = 0.98) for both datasets.

TABLE 1 | Evaluation of the detection results.

Dataset	TP	FP	FN	Precision	Recall	*F*-measure
1 (52 images)	15,773	2434	314	0.87	0.98	0.92
2 (40 plots)	2762	587	44	0.82	0.98	0.89

Images from datasets (1 and 2) had not been used to train the detection model.
TP (True Positive): head region correctly detected.
FP (False Positive): non-head region erroneously detected as a head.
FN (False Negative): head region undetected.
Precision = TP/(TP + FP). Recall = TP/(TP + FN).
F-measure = (2 × precision × recall)/(precision + recall).

Figure 8 shows the counting accuracy of the proposed method. First, the total number of sorghum heads in each image from both datasets was counted and double checked carefully by two researchers. Then, the test part of dataset 1, all of dataset 1, and all of dataset 2 were tested by the model, and the coefficients of determination (R^2) were 0.85, 0.88, and 0.56, respectively.

The R^2 of the head counting at the image level (Dataset 1) was relatively high, however, it was observed to be decreased substantially when applying the model to single plot images (Dataset 2). The main reason for this is likely related to the application of the plot segmentation algorithm in this experiment. In our case, there were only two rows in each plot, and hence it was common to cut out parts of the heads (which overlapped "plot" boundaries) during segmentation (**Figure 9A**). Plot segmentation accuracy could be improved either by redesigning the field experiment to enlarge plot size or by using drones with a better positioning system and a higher-resolution camera. Alternatively, we could try applying methods to the rows of heads in a multi-plot image to try to better delineate the boundaries between the plots (i.e., tracing around the heads at edges of plot).

For both of the datasets, challenges remain in dealing with the large variability among genotypes, growth stages, and growth position (main stem or tillers) of heads, all of which contribute to large differences in the morphological features of detected head regions. As shown in **Figure 9B**, detected regions may comprise multiple overlapping heads while other heads may be obscured by leaves, neither feature of which has been trained in the counting

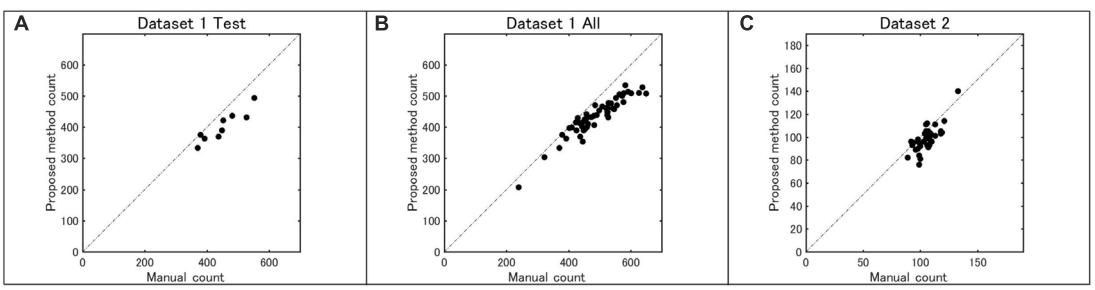


FIGURE 8 | Accuracy of head number determined by the proposed method as compared with that done by manual counting: **(A)** dataset 1 test, **(B)** dataset 1 (all), and **(C)** dataset 2 (all).

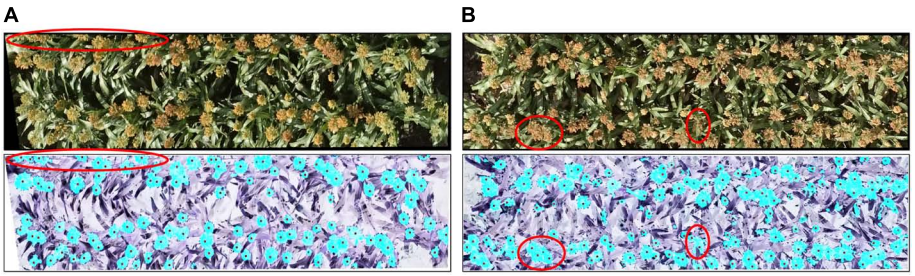


FIGURE 9 | Reasons for incorrect counting: **(A)** the plot segmentation was not perfect, so parts of some heads were cut out (upper red oval); **(B)** some regions included multiple overlapping heads (left red circle), and some heads were covered by leaves (right red circle). The upper panels show the original images and the lower ones show the detected head regions (blue) and hand-labeled head centers (black dots).

model. The model capabilities could therefore be improved with additional training data and exploring more efficient features.

With the rapid development of GPU technology in recent years, the size of electrical infrastructure has been decreased significantly without loss in performance. These types of embedded platforms allow for onboard real-time image processing, so that the proposed or other methods could be applied with real-time image input. By integrating such system with UAVs, scouting of production fields could be completed by the end of the flight.

CONCLUSION

We proposed a simple two-step machine-learning-based image processing method to detect and count the number of sorghum heads from high-resolution images captured by UAVs in a breeding field. This introduces realistic challenges given that sorghum has various genotypes with different growth stages, and the heads can have different colors, shapes, and sizes. Using carefully selected training data, the precision and recall of head detection were 0.87 and 0.98, respectively, for dataset 1 and 0.82 and 0.98 for dataset 2. The coefficients of determination (R^2) for head counting were 0.88 and 0.56 for datasets 1 and 2, respectively.

Head number per unit area is an important component of the yield of cereal crops. As well as being useful to agronomists and breeders, the method described here has utility in production agriculture, e.g., by using UAVs to survey a field to estimate head number, and then manually sampling a range of head sizes in order to estimate yield as product of weighted average head size (grain weight per head) and head number. Counting can also be used to characterize spatial variability in the field, as well as non-uniformity of development or head size over time (multiple monitoring flights). In a research context, the ability of this method to count heads of contrasting genotypes in diverse measurement conditions provides a better capability to estimate head number in plant breeding trials. The main limitation at present in being able to correctly delineate the boundaries of plots and we are investigating ways of doing this, e.g., by defining plots early in the season and tracking the head positions relative to original plant positions.

The application of machine-learning-based image analysis technologies is become increasingly important in field-based plant phenotyping tasks. By using these rapidly improving techniques, we believe the accuracy of phenotyping will increase while the computational cost will decrease, both of which will help researchers reach the goal of real-time phenotyping (Fuentes et al., 2017; Naik et al., 2017). However, key techniques such as training data preparation, model selection, and feature definition still rely on highly specialized knowledge in both plant science and computer science (Singh et al., 2016). Deep learning is a possible solution to reduce the difficulties, but generating ground-truth data (image annotation) to train the models is still very labor intensive (Sa et al., 2016, 2017; Ghosal et al., 2018). A method is needed to automatically generate reliable training data, and the detection feature of the proposed method could be used as a semi-automatic tool to provide candidate training datasets. We encourage the plant research community to share the existing annotated dataset to accelerate plant-phenotyping community growth in a similar way that ImageNet is used (Russakovsky et al., 2015). To aid in this growth, datasets 1 and 2 along with the manual labeling used in this study are available in the **Supplementary Materials**.

AUTHOR CONTRIBUTIONS

WG analyzed the data and interpreted results with the input of JD. AP, BZ, SC, KW, KN, and SN conceived the research. WG and XW labeled the images. DJ and XW conceived, designed, and coordinated the field experiments. WG wrote the paper with input from all authors.

FUNDING

This study was partially funded by the CREST Program JPMJCR1512 and the SICORP Program "Data Science-based Farming Support System for Sustainable Crop Production under Climatic Change" of the Japan Science and Technology Agency. This research was partially funded by the Australian Government through the Australian Research Council Centre of Excellence for Translational Photosynthesis and by the partners in that Centre: CSIRO, Australian National University, The University of Queensland, University of Sydney, Western Sydney University, and International Rice Research Institute. The trials from which images and head counts were collected are part of the Sorghum Core-Breeding Program supported by the Grains Research and Development Corporation and the Queensland Department of Agriculture and Fisheries.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01544/full#supplementary-material

REFERENCES

Davis, J., and Goadrich, M. (2006). "The relationship between precision-recall and roc curves," in *Proceedings of the 23rd International Conference on Machine Learning ICML '06*, (New York, NY: ACM), 233–240. doi: 10.1145/1143844.1143874

Dreccer, M. F., Wockner, K. B., Palta, J. A., McIntyre, C. L., Borgognone, M. G., Bourgault, M., et al. (2014). More fertile florets and grains per spike can be achieved at higher temperature in wheat lines with high spike biomass and sugar content at booting. *Funct. Plant Biol.* 41, 482–495. doi: 10.1071/FP13232

Duan, T., Zheng, B., Guo, W., Ninomiya, S., Guo, Y., Chapman, S. C., et al. (2016). Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by UAV. *Funct. Plant Biol.* 44, 169–183. doi: 10.1071/FP16123

Fuentes, A., Yoon, S., Kim, S. C., and Park, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases

and pests recognition. *Sensors* 17:E2022. doi: 10.3390/s17092022

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., and Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proc. Natl. Acad. Sci.* 115, 4613–4618. doi: 10.1073/pnas.1716999115

Gnädinger, F., and Schmidhalter, U. (2017). Digital counts of maize plants by unmanned aerial vehicles (UAVs). *Remote Sens.* 9:544. doi: 10.3390/rs9060544

Gongal, A., Silwal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. (2016). Apple crop-load estimation with over-the-row machine vision system. *Comput. Electron. Agric.* 120, 26–35. doi: 10.1016/j.compag.2015.10.022

Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2010). *Digital Image Processing Using MATLAB*, 2nd Edn. Upper Saddle River, NJ: Pearson/Prentice Hall.

Guo, W., Fukatsu, T., and Ninomiya, S. (2015). Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. *Plant Methods* 11:7. doi: 10.1186/s13007-015-0047-9

Guo, W., Potgieter, A. B., Jordan, D., Armstrong, R., Lawn, K., Kakeru, W., et al. (2016). "Automatic detecting and counting of sorghum heads in breeding field using RGB imagery from UAV," in *Proceedings of the Conference on CIGR-AgEng*, Aarhus.

Guo, W., Rage, U. K., and Ninomiya, S. (2013). Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Comput. Electron. Agric.* 96, 58–66. doi: 10.1016/j.compag.2013.04.010

Guo, W., Zheng, B., Duan, T., Fukatsu, T., Chapman, S., and Ninomiya, S. (2017). EasyPCC: benchmark datasets and tools for high-throughput measurement of the plant canopy coverage ratio under field conditions. *Sensors* 17, 1–13. doi: 10.3390/s17040798

Hammer, G. L., van Oosterom, E., McLean, G., Chapman, S. C., Broad, I., Harland, P., et al. (2010). Adapting APSIM to model the physiology and genetics of complex adaptive traits in field crops. *J. Exp. Bot.* 61, 2185–2202. doi: 10.1093/jxb/erq095

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* 3, 610–621. doi: 10.1109/TSMC.1973.4309314

Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Remote Sens. Environ.* 198, 105–114. doi: 10.1016/j.rse.2017.06.007

Lafarge, T. A., Broad, I. J., and Hammer, G. L. (2002). Tillering in grain sorghum over a wide range of population densities: identification of a common hierarchy for tiller emergence, leaf area development and fertility. *Ann. Bot.* 90, 87–98. doi: 10.1093/aob/mcf152

Linker, R., and Kelman, E. (2015). Apple detection in nighttime tree images using the geometry of light patches around highlights. *Comput. Electron. Agric.* 114, 154–162. doi: 10.1016/j.compag.2015.04.005

Liu, S., Baret, F., Andrieu, B., Burger, P., and Hemmerlé, M. (2017). Estimation of wheat plant density at early stages using high resolution imagery. *Front. Plant Sci.* 8:739. doi: 10.3389/fpls.2017.00739

Lobell, D. B., Hammer, G. L., Chenu, K., Zheng, B., Mclean, G., and Chapman, S. C. (2015). The shifting influence of drought and heat stress for crops in Northeast Australia. *Glob. Chang. Biol.* 21, 4115–4127. doi: 10.1111/gcb.13022

Meyer, G. E., and Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293. doi: 10.1016/j.compag.2008.03.009

Mitchell, J. H., Rebetzke, G. J., Chapman, S. C., and Fukai, S. (2013). Evaluation of reduced-tillering (tin) wheat lines in managed, terminal water deficit environments. *J. Exp. Bot.* 64, 3439–3451. doi: 10.1093/jxb/ert181

Naik, H. S., Zhang, J., Lofquist, A., Assefa, T., Sarkar, S., and Ackerman, D. (2017). A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. *Plant Methods* 13:23. doi: 10.1186/s13007-017-0173-7

Nuske, S., Achar, S., Bates, T., Narasimhan, S., and Singh, S. (2011). "Yield estimation in vineyards by visual grape detection," in *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Francisco, CA. doi: 10.1109/IROS.2011.6048830

Payne, A., Walsh, K., Subedi, P., and Jarvis, D. (2014). Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging. *Comput. Electron. Agric.* 100, 160–167. doi: 10.1016/j.compag.2013.11.011

Potgieter, A. B., Hammer, G. L., Armstrong, R. N., Mclean, G., Laws, K., Chapman, S., et al. (2015). "The use of in-situ proximal sensing technologies to determine crop characteristics in sorghum crop breeding," in *Proceedings of the 2015 Conference on Tropical Agriculture*, Brisbane, QLD.

Potgieter, A. B., Lobell, D. B., Hammer, G. L., Jordan, D. R., Davis, P., and Brider, J. (2016). Yield trends under varying environmental conditions for sorghum and wheat across Australia. *Agric. For. Meteorol.* 228, 276–285. doi: 10.1016/J.AGRFORMET.2016.07.004

Qureshi, W. S., Payne, A., Walsh, K. B., Linker, R., Cohen, O., and Dailey, M. N. (2016). Machine vision for counting fruit on mango tree canopies. *Precis. Agric.* 18, 224–244. doi: 10.1007/s11119-016-9458-5

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* 115, 211–252. doi: 10.1007/s11263-015-0816-y

Sa, I., Chen, Z., Popovic, M., Khanna, R., Liebisch, F., Nieto, J., et al. (2017). weedNet: dense semantic weed classification using multispectral images and MAV for smart farming. *IEEE Robot. Autom. Lett.* 3, 588–595. doi: 10.1109/LRA.2017.2774979

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., and McCool, C. (2016). Deepfruits: a fruit detection system using deep neural networks. *Sensors* 16:E1222. doi: 10.3390/s16081222

Sengupta, S., Lee, W. S., Aleixos, N., Blasco, J., and Moltó, E. (2014). Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions. *Biosyst. Eng.* 117, 51–61. doi: 10.1016/j.biosystemseng.2013.07.007

Singh, A., Ganapathysubramanian, B., Singh, A. K., and Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends Plant Sci.* 21, 110–124. doi: 10.1016/j.tplants.2015.10.015

Yamamoto, K., Guo, W., Yoshioka, Y., and Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors* 14, 12191–12206. doi: 10.3390/s140712191

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 September 2017
doi: 10.3389/fpls.2017.01532

Multi-Spectral Imaging from an Unmanned Aerial Vehicle Enables the Assessment of Seasonal Leaf Area Dynamics of Sorghum Breeding Lines

Andries B. Potgieter[1†], Barbara George-Jaeggli[2,3*†], Scott C. Chapman[4,5,6], Kenneth Laws[3], Luz A. Suárez Cadavid[1], Jemima Wixted[2], James Watson[1], Mark Eldridge[2], David R. Jordan[2] and Graeme L. Hammer[5]*

[1] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Toowoomba, QLD, Australia, [2] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Warwick, QLD, Australia, [3] Agri-Science Queensland, Department of Agriculture and Fisheries, Warwick, QLD, Australia, [4] School of Agriculture and Food Sciences, University of Queensland, Gatton, QLD, Australia, [5] CSIRO Agriculture and Food, St. Lucia, QLD, Australia, [6] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, QLD, Australia

Genetic improvement in sorghum breeding programs requires the assessment of adaptation traits in small-plot breeding trials across multiple environments. Many of these phenotypic assessments are made by manual measurement or visual scoring, both of which are time consuming and expensive. This limits trial size and the potential for genetic gain. In addition, these methods are typically restricted to point estimates of particular traits, such as leaf senescence or flowering and do not capture the dynamic nature of crop growth. In water-limited environments in particular, information on leaf area development over time would provide valuable insight into water use and adaptation to water scarcity during specific phenological stages of crop development. Current methods to estimate plant leaf area index (LAI) involve destructive sampling and are not practical in breeding. Unmanned aerial vehicles (UAV) and proximal-sensing technologies open new opportunities to assess these traits multiple times in large small-plot trials. We analyzed vegetation-specific crop indices obtained from a narrowband multi-spectral camera on board a UAV platform flown over a small pilot trial with 30 plots (10 genotypes randomized within 3 blocks). Due to variable emergence we were able to assess the utility of these vegetation indices to estimate canopy cover and LAI over a large range of plant densities. We found good correlations between the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) with plant number per plot, canopy cover and LAI both during the vegetative growth phase (pre-anthesis) and at maximum canopy cover shortly after anthesis. We also analyzed the utility of time-sequence data to assess the senescence pattern of sorghum genotypes known as fast (senescent) or slow senescing (stay-green) types. The Normalized Difference Red Edge (NDRE) index which estimates leaf chlorophyll content was most useful in characterizing the

OPEN ACCESS

Edited by:
Yann Guédon, Agricultural Research Centre For International Development, France

Reviewed by:
Nadine Hilgert, Institute National de la Recherche Agronomique Centre Montpellier, France
Yuhui Chen, Noble Research Institute, LLC, United States

***Correspondence:**
Andries B. Potgieter
a.potgieter@uq.edu.au
Barbara George-Jaeggli
b.georgejaeggli@uq.edu.au

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *03 May 2017*
Accepted: *21 August 2017*
Published: *08 September 2017*

Citation:
Potgieter AB, George-Jaeggli B, Chapman SC, Laws K, Suárez Cadavid LA, Wixted J, Watson J, Eldridge M, Jordan DR and Hammer GL (2017) Multi-Spectral Imaging from an Unmanned Aerial Vehicle Enables the Assessment of Seasonal Leaf Area Dynamics of Sorghum Breeding Lines. Front. Plant Sci. 8:1532. doi: 10.3389/fpls.2017.01532

leaf area dynamics/senescence patterns of contrasting genotypes. These methods to monitor dynamics of green and senesced leaf area are suitable for out-scaling to enhance phenotyping of additional crop canopy characteristics and likely crop yield responses among genotypes across large fields and multiple dates.

Keywords: crop cover, mosaics, UAV, leaf area dynamics, water use, sorghum breeding

INTRODUCTION

Sorghum is the dominant dry-land summer crop in the north-eastern Australian grain belt. The growing environments of this area are characterized by high temperatures and variable rainfall, although many of the soils have sufficient water-holding capacity to allow crops to grow on stored sub-soil moisture (Pratley, 2003). As sub-soil moisture is depleted, mild or severe drought stress frequently develops toward the end of the growing season (Chapman et al., 2002), reducing crop yield. In the next decades, this situation is expected to occur even more frequently with increasing climate variability and weather patterns becoming more extreme (Lobell et al., 2015a) as was seen during the last two decades globally as well as in Australia (IPCC, 2014).

With changes in climate, quantitative breeding for specific traits that enhance yield in water-limited environments, will become even more important. One such trait is leaf area index (LAI), as the size of the crop canopy has important consequences for water use (Borrell et al., 2014a,b). Being able to accurately characterize leaf area would greatly enhance the selection of sorghum genotypes that are well adapted to water-limited environments. For example, in environments with mild to severe terminal drought stress, crops with smaller leaf area per plant have been found to have a yield advantage, as their reduced water use before flowering conserves sub-soil moisture that can be accessed during the critical grain-filling period (He et al., 2016). The stay-green trait in sorghum, which is associated with reduced leaf senescence and yield benefits under post-anthesis drought is thought to operate via this mechanism by conferring reduced tillering and smaller plant leaf areas before flowering (Borrell et al., 2014a,b). Stay-green has been an important trait in Australia's sorghum breeding programs, which has partly contributed to significant increases in sorghum yield trends in dry environments compared to moderate and wet environments over the last three decades (Potgieter et al., 2016). Up to now, breeders have positively selected for stay-green by visually rating leaf senescence after flowering. However, this only works in trials in which the right drought conditions develop for the trait to be expressed.

Apart from these links to evapotranspiration (George-Jaeggli et al., 2017), LAI is also useful to evaluate the fraction of absorbed photosynthetically active radiation, which is required to model canopy photosynthesis (Weiss et al., 2004). Being able to measure leaf area development over time would therefore allow the estimation not only of the water use pattern of a genotype, but also its likely photosynthetic output.

While visual scores of stay-green during grain-filling can be reasonably accurate when assessed at the right time and under the right level of water limitation, it is difficult to estimate plant leaf area or leaf area index (LAI) earlier in the season, and actual measurements of leaf area are time-consuming. Measurement of leaf area on thousands of plots at one time point, let alone several time points throughout the growing season is impractical. A low-cost high-throughput method for phenotyping canopy size of sorghum genotypes is needed.

The first application of remotely-sensed multi-spectral imagery and the development of vegetation indices to monitor crops goes back to the first NASA LANDSAT series in the 1970's (Tucker, 1979). The application of remote-sensing technology, in particular, hyperspectral imaging (Goetz, 2009), in vegetation mapping and yield forecasting has been steadily developing since then, and many different indices using specific wavelengths have been developed that can be used to assess plant growth parameters (Beeri and Peled, 2006; White et al., 2012). More recently this has been extended to predicting crop and vegetation biophysical attributes like net primary production (NPP), fraction of absorbed photosynthetically active radiation and LAI. This was done through the use of spectral indices (e.g., NDVI, EVI) derived from visible and near infrared reflectance spectra at moderate to high spatial resolutions across large scales (Huete et al., 2002; Hanes, 2014).

While the use of a digital camera attached to an Unmanned Aerial Device Unmanned aerial vehicles (UAV) was first proposed as a cost-effective way to monitor small wheat plots nearly a decade ago (Lelong et al., 2008) it was not until very recently that cheap, but highly precise positioning and digital imaging technologies and unmanned aerial device technology have become mainstream so that their use has become practical for farmers and research programs alike (Haboudane et al., 2004; Chapman et al., 2014; Candiago et al., 2015). The combination of these technologies provides the potential for high-throughput phenotyping to allow plant breeding programs to undertake quantitative screens of large breeding populations.

This paper presents results from a pilot study using a multi-rotor UAV fitted with a narrow-band multispectral camera (five bands of 10–40 nm width) to capture images of sorghum breeding lines with diverse canopy attributes across seven dates. We evaluated three narrow-band vegetation indices i.e., the normalized difference vegetation index (NDVI), the enhanced vegetation index (EVI) and the normalized difference red edge index (NDRE) to estimate traits, such as canopy cover, leaf area index and leaf chlorophyll content that are of particular interest to sorghum breeders in the northern grain belt of Australia.

Previous studies have demonstrated the utility of such vegetation indices to estimate LAI in soybean and maize (Viña et al., 2011) and wheat (Haghighattalab et al., 2016), but no such studies previously existed for sorghum. The objective of our study was to assess the suitability of vegetation indices calculated from

spectral data captured with a multi-spectral camera mounted on a UAV to estimate canopy cover, leaf area and leaf chlorophyll content of a diverse set of sorghum genotypes grown in breeding plots. We also discuss the utility of such an approach to assess sorghum breeding lines for differences in canopy size and leaf chlorophyll content during critical crop stages, such as around flowering and during grain fill.

MATERIALS AND METHODS

Experiment and Genotypes

An experiment was conducted to test the ability of multi-spectral sensing technologies on-board a UAV platform to calculate various vegetation indices to estimate canopy characteristics, such as plant cover, leaf area, leaf greenness or chlorophyll content and biomass of single plots sown to different sorghum genotypes. This paper only focuses on the outcomes related to plant cover, LAI and chlorophyll content.

Ten grain sorghum genotypes known for differences in canopy traits, such as plant height, leaf angle and leaf area were selected, including 4 genotypes with contrasting senescence type (i.e., rapid senescence after flowering = senescent type, or slow senescence = stay-green type).

The 10 sorghum genotypes were arranged in a randomized complete block design with 3 blocks (10 genotypes × 3 rows), resulting in 30 plots (**Figure 1**). Plots were 4 rows wide with 0.76 m row spacing by 10 meters long (i.e., 30.4 m^2) and planted in an east-west direction.

The study site was located at the Hermitage Research Facility (28°12′ S, 152°06′ E; 480 m above sea level) in north-eastern Queensland. The soil of the trial area was conditioned 6 months prior to planting via incorporation of 3.5 t ha^{-1} of Gypsum, 350 kg ha^{-1} of NatraMin (AgSolutions, Australia) and 6 t ha^{-1} feedlot manure. One month prior to planting the trial area was fertilized with 220 kg ha^{-1} of GRAN-AM (20% Nitrogen, 24% Sulfur, Incitec Pivot, Australia) and 100 kg ha^{-1} of Urea (46% Nitrogen). The plots were sown with a precision planter on the 19th of November 2015. The trial was planted on a near-level site on a self-mulching alluvial clay with a high montmorillonite clay content (McKeown, 1978) that had a full sub-soil moisture profile at sowing. The trial was not irrigated, but regular in-crop rainfall and the sub-soil reserves prevented the development of significant water limitation. Crop establishment was variable due to surface flooding just after sowing. However, data are compared at the sample quadrat level (see details below) so that 30 sample quadrats at each harvest can essentially be considered as samples of potential leaf area for a diverse set of genotypes.

Data Capturing Missions

At sowing time, accurate ground control points (GCP) were collected using a 1 cm resolution handheld GPS (Global Positioning System) unit (Trimble XT, Trimble, Sunnyvale California). Each of these GCPs was marked with a square concrete paver painted with blue triangles so that they were easily identifiable from above.

Data capturing missions were conducted at different critical times during the crop growth period (**Table 1**). Sample quadrat cuts of evenly established areas of two central rows (1 lineal meter from each) were taken within each plot at two different stages: pre-anthesis (ca. 8 weeks after sowing, or 3 weeks prior to anthesis) and at or within 1 week of anthesis.

To reduce effects of ambient light condition, we limited data capturing missions to clear and cloudless days and conducted them around the middle of the morning.

UAV Platform

The UAV platform used was a 3D Robotics X8+ multi-rotor (Berkeley, California). The X8+ has the advantage of being able to fly at very low altitudes and at low speeds, which is critical for

TABLE 1 | Experiment details, dates and variables collected.

Experimental design	**10 genotypes × 3 replicates (blocks), randomized complete block design**
Genotypes	R55637 (senescent), MR Buster (senescent), R931945-2-2 (stay-green), R931945-2-2TM (stay-green), 84G22, 85G56, FF_B963676, A1*F_B010054/F9_R04377-31, A1*F_B02055-9/R986087-2-4-1, R974443-1-2
UAV flights (2016)	12 January, 26 January, 3 February, 10 February, 16 February, 25 February, 31 March
Crop stages	Sowing—19 November 2015; Flowering—9 February 2016; Final Harvest—11 May 2016
Leaf area index (LAI) quadrat cuts	Pre-Anthesis (13 January) and Anthesis (9 February)
Number of culms per m^2	Total number of culms (main stems and tillers together) at anthesis

MR Buster and the two A1 genotypes are hybrids while the remaining genotypes are inbred lines.

FIGURE 1 | Experimental layout of plots at the Hermitage site as shown in aerial photo mosaic taken on 3 February. Plots were arranged in three rows and ten columns per row. Numbers in figure refer to the Column-Row position of each plot. Rows were treated as blocks and genotypes were completely randomized within each block (row).

creating accurate and high-resolution mosaics (Corrigan, 2015). Flight altitudes for each flight were set at 20 m resulting in a ground sampling distance (GSD) or pixel size of ∼0.5 cm.

Multi-Spectral Camera

A RedEdge™ narrow-band multispectral camera (MicaSense, Seattle, Washington) (http://www.micasense.com/rededge/) simultaneously capturing 5 bands at specific nanometre (nm) wavelength peaks was fitted to the UAV platform. The bands captured were Blue (B: 475 nm center wavelength, 20 nm bandwidth), Green (G: 560 nm, 20 nm), Red (R: 668 nm, 10 nm), Red Edge (RE: 717 nm, 10 nm), and Near Infrared (NIR: 840 nm, 40 nm) (**Figure 2**). The camera captured the images and GPS information to a local digital card in 16-bit raw GeoTIFF files. This allowed for post geo-rectification and mosaicking. The horizontal field of view was 47.2 degrees with a 5.5 mm focal length producing an image resolution of 1,280 × 960 pixels.

The RedEdge camera includes factory calibration coefficients in each image for optics chain properties, such as lens distortion and optical vignetting. Atlas uses a CMOS sensor (Complementary metal–oxide–semiconductor) model along with extracted regions from images of a calibrated Lambertian reflectance panel to convert raw image digital number (DN) to reflectance units. These images are then linearly combined through a photogrammetry process to estimate the surface reflectance of each pixel in the final reflectance map.

Mosaicking, Ortho Rectification and Reflectance

After each flight, images for each of the five wavelengths were uploaded to the ATLAS cloud (MicaSense, Seattle, Washington) (http://www.micasense.com/atlas/). The cloud service uses the Pix4d software (PIX4d, Lausanne, Switzerland) (www.pix4d.com) and proprietary algorithms to stitch images together to create a geo-referenced multi-layer ortho-mosaic of the flight for each date. Stitched GEOTIFF format images for each band were downloaded from ATLAS and imported into ArcGIS (https://www.arcgis.com/home/index.html) for layer stacking and geo-rectification to GCP for each date.

To be able to convert DNs into reflectance, an image of a white reflectance panel was taken at the start and end of each flight and was uploaded with the images prior to the cloud processing. During the mosaicking process, the reflectance of the reference panel was used so that each of the 5 downloaded GeoTIFF files was a calibrated reflectance map for the respective band. The pixel values are proportional to % reflectance, with a pixel value of 32,768 being equal to 100% reflectance (65,535 is equal to 200% reflectance). Once the 5 bands had been layer stacked for a single flight date, data from each date was geo-rectified to high-precision GCP. Pixel values were then converted to reflectance values between 0 and 1 by dividing each pixel by the max reflectance value of 32,768. Pixels with specular reflectance (e.g., bright mirror like reflectance) and missing values from the mosaics were omitted in the analysis by masking.

After adjustment of pixel reflectance, indices per plot and per quadrat cut from each mosaic were generated in ArcGIS software and extracted and saved into an ASCII file format for comparison with measured data. The reflectance of areas of the quadrat cuts for any single flight could be determined by examining a subsequent flight (after cutting) to exactly identify where the cuts were made.

Narrowband Vegetation Indices and Percent Cover

Two spectral indices were calculated from the reflectance measured by the RedEdge™ sensor. These indices relate to canopy health and canopy architecture (i.e., leaf area and biomass). The most widely used vegetation index is the Normalized Difference Vegetation Index (NDVI). NDVI is a simple normalized ratio between the NIR and R wavebands and is therefore a comparable metric between dates (Rouse et al., 1974):

$$\mathbf{NDVI} = (\mathbf{NIR} - \mathbf{R})/(\mathbf{NIR} + \mathbf{R}) \quad (1)$$

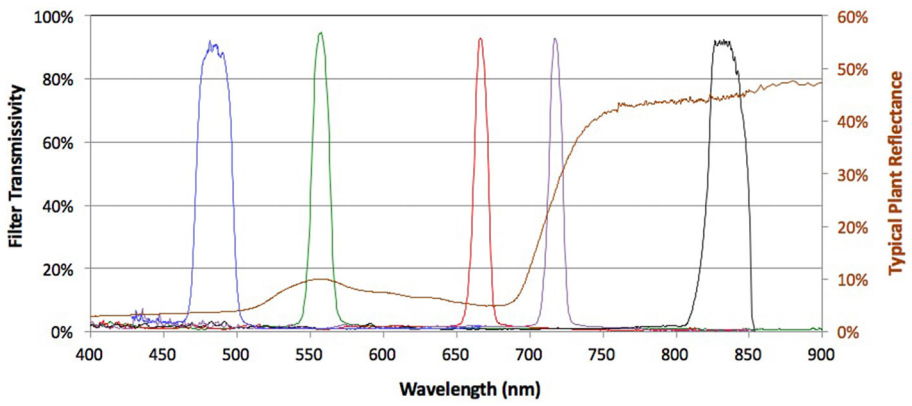


FIGURE 2 | Multispectral bands of the Micasense™ camera across the spectrum of visible and infrared light. Peaks for each band's transmissivity are shown across the electromagnetic spectrum at specific nanometer (nm) wavelengths. Blue (B: 475 nm center wavelength, 20 nm bandwidth), Green (G: 560 nm, 20 nm), Red (R: 668 nm, 10 nm), Red Edge (RE: 717 nm, 10 nm), and Near Infrared (NIR: 840 nm, 40 nm). Source: http://www.micasense.com.

We used a NDVI threshold of > 0.5 to capture reflectance from green leaves only and exclude soil background reflectance. This threshold has the greatest effect when plants are small i.e., at the time of the pre-anthesis measurement.

The enhanced vegetation index (EVI), which relates to canopy architecture was computed as follows (Huete et al., 2002):

$$\mathbf{EVI} = 2.5*[\frac{\mathbf{NIR\text{-}R}}{\mathbf{NIR} + 6*\mathbf{R} - 7.5*\mathbf{B} + 1}] \quad (2)$$

NDVI and EVI pixel values were aggregated to generate individual plot index metrics for each 30.4 m^2 sorghum plot at each flight date. Maximum ($NDVI_{max}$, EVI_{max}) and average ($NDVI_{avg}$, EVI_{avg}) values for each index were derived from this time series.

In order to assess the degree of crop establishment (i.e., number of plants visible after emergence) we calculated the crop cover (CC, %) for each plot. An RGB image was generated from the Micasense mosaics. CC was then derived as the proportion of green pixels per plot. We defined pixels as "green" if their hue was between 70 and 140 degrees.

Plant Counts and Leaf Area Index

Total number of plants per entire plot area (i.e., 30.4 m^2) were counted 26 days after sowing (DAS). Leaf area (LAI) was measured destructively (sample quadrats) on the ground during the vegetative period (pre-anthesis, 54 DAS) and 2 weeks after the last genotype started flowering (anthesis, 83 DAS). At each sampling time, all plants within a 1.52 m^2 sampling quadrat (2 × 1 m from the middle 2 rows of each 4-row plot), were cut at ground level and brought up to the laboratory for processing. Plants were separated into stems, leaves and panicles, main stems and tillers separately, and dried in a forced draft oven at 80°C until dry weight reached a minimum and then weighed. During the anthesis sample, culm numbers (main stem and tillers together) were recorded.

Crop Senescence

To analyze differences in rate of senescence between genotypes, we calculated the normalized difference red edge index (NDRE) (Gitelson and Merzlyak, 1994; Sims and Gamon, 2002):

$$\mathbf{NDRE} = (\mathbf{NIR\text{-}RE})/(\mathbf{NIR{+}RE}) \quad (3)$$

The NDRE index is highly correlated with chlorophyll content within plants and therefore is a good surrogate for photosynthetic capacity (Gitelson and Merzlyak, 1994; Sims and Gamon, 2002; Gitelson et al., 2003). The difference between NDRE at maximum (peak) canopy cover and the NDRE at maturity (final flight date) was used as a simple metric for the rate of senescence (RS NDRE). To test whether this index was useful to differentiate between genotypes that were known to be senescent (tendency to senesce rapidly after flowering) or stay-green (tendency to senesce slowly after flowering), we grouped a subset of 4 genotypes into 2 groups (Senescent and Stay-green).

Statistical Validation Metrics

All analyses and graphs were done using R (R Core Team, 2016).

We used simple linear regressions or logarithmic functions depending on best fit between vegetation indices (i.e., NDVI and EVI) and measured data at sample quadrat levels.

To test for significant genotype or group effects on individual vegetation indices we used linear mixed models in the lme4 package in R (Bates et al., 2015).

The general form of the mixed models used was:

$$Y = X\beta + Z\mu + \varepsilon \quad (4)$$

where the response (vector y) is modeled by a set of fixed effects (vector β) and random effects (vector μ) and ε is the random error term. The design matrices X and Z assign the fixed and random effects, respectively to the observations.

For the time series of NDVI or NDRE vs. days after sowing (DAS) for individual genotypes, the mixed model included vector β comprising Genotype (factor with 10 levels) and DAS (factor with 7 levels) (fixed effects) and vector μ comprising Block (3 levels) within plot and vector ε comprising error (random effects).

To test whether Group (senescent or stay-green) had a significant effect on the difference between NDRE at maximum canopy cover and NDRE at maturity, we first tested a mixed model with vector β comprising Genotype (factor with 4 levels) and Group (factor with 2 levels), μ comprising Block (3 levels) and vector ε comprising error, but as Genotype had no significant effect, we only included Group in vector β in the final model. Assumptions of normality were tested with a quantile-quantile plot and seemed to have been met. Analysis of covariance was conducted with Group as variable to test whether the slopes of the relationship of NDRE vs. days after sowing during the post-anthesis period were significantly different between the senescent and stay-green genotypes.

RESULTS

Vegetation Indices Aggregated at Entire Plot Level

Averaged across all plots, NDVI (>0.5) and EVI values aggregated over the entire 30.4 m^2 plot areas were 0.75 and 0.37, respectively, at the first flight date (**Figure 3**). Maximum values for NDVI occurred between 68 and 83 DAS and ranged from 0.72 to 0.86, depending on genotype. After this time, NDVI decreased due to crop senescence and reached values of between 0.62 and 0.67 by the end of the experiment (133 DAS) (**Figure 3**). Average EVI values remained relatively consistent for all, but the last flight date, when EVI was significantly lower (**Figure 3**).

Percent Cover at Plot Level

Due to surface flooding in the first week after sowing affecting emergence, crop cover varied from as low as 7 and 18% to as high as 57 and 77% 54 days after sowing (DAS) (12 January) and 76 DAS (3 February), respectively (**Figure 4**). Actual plant counts ranged from 36 to 204 plants per plot (30.4 m^2). This variability gave us an opportunity to test the validity of using NDVI to estimate crop cover across a broad range of plant covers. When aggregated over the entire plot area, NDVI (>0.5) was

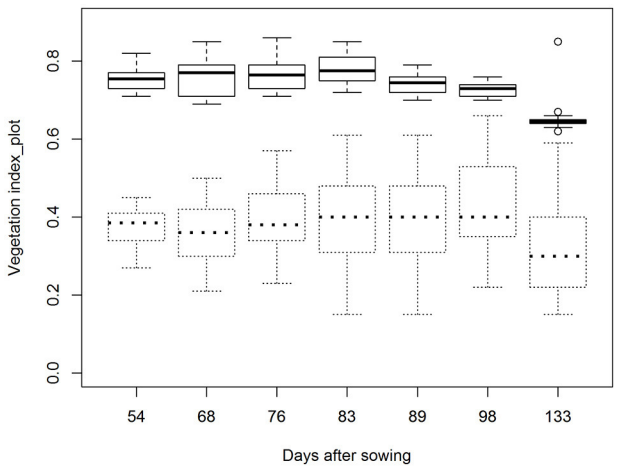


FIGURE 3 | Boxplots showing NDVI (—) and EVI (—) values aggregated over entire 30.4 m^2 plot areas and averaged across all 30 plots at seven flight dates. Flight dates were: 12 January or 54 days after sowing (DAS), 26 January (68 DAS), 3 February (76 DAS), 10 February (83 DAS), 16 February (89 DAS), 25 February (98 DAS), 31 March 2016 (133 DAS).

significantly and strongly correlated with plants per plot (R^2 = 0.58, RMSE = 0.03, **Figure 5**).

Maximum crop cover at entire plot level was mostly >30 and 50% at the pre-anthesis (January image) and anthesis stages (February image), respectively (**Figure 4**).

Correlation of Vegetation Indices with LAI Measured by Sampling Quadrats

Leaf area index (LAI) values derived from the quadrat cut sampling areas ranged from 0.71 to 4.01 (m^2/m^2) at pre-flowering and increased to between 1.31 and 4.71 2 weeks after flowering in all plots.

Vegetation indices derived from pixels aggregated over the entire plot area at both the pre-flowering and the anthesis sampling dates were strongly linearly correlated with LAI from quadrat cuts (**Table 2**). The correlations were better for NDVI than for EVI as can be seen from greater regression coefficients (R^2) and smaller root mean square errors (RMSE). $NDVI_{max}$ also correlated well with LAI at the anthesis sampling date.

When just aggregating the pixels over the actual sample quadrat areas the vegetation indices explained more of the variation in LAI as indicated by larger R^2, but the RMSE did not always improve (**Table 2**).

When leaf area index data from quadrat cuts and NDVI aggregated over the entire plots was combined for both pre-anthesis and anthesis sampling dates (60 samples in total) a logarithmic function fitted the data slightly better than a linear one (RMSE of 0.038 vs. 0.041 for logarithmic and linear, respectively) (**Figure 6**).

Temporal Dynamics of NDVI and NDRE

Normalized Difference Vegetation Index (NDVI) (aggregated over the entire plot) gradually increased and reached maximum values (>0.9; blue colored) by the anthesis sampling date (83 DAS or 10th February) (**Figure 7**). After this date, NDVI values decreased to around 0.5 (light green) due to progressive senescence of leaves as genotypes approached maturity.

NDRE values were much lower, but showed a similar pattern of slowly increasing up to about 2 weeks after flowering (83 DAS) and then decreasing as NDVI (**Figure 8**). As NDRE is related to chlorophyll, differences in NDRE values from peak canopy NDVI (i.e., $NDVI_{max}$) to maturity (last flight date) are associated with the rate of senescence. The trial included genotypes that are known to senesce quickly (senescent genotypes; MR Buster and R955637) and others that have the stay-green trait meaning they have a slower rate or senescence and retain more green leaf area during grain fill compared with senescent types, particularly when water is limited (stay-green genotypes; R931945-2-2 and R931945-2-2TM) (**Figure 8**).

When these genotypes were grouped (i.e., stay-green group: R931945-2-2 and R931945-2-2TM and senescent group: MR Buster and R955637), RS NDRE (the difference between NDRE at maximum canopy cover and NDRE at harvest maturity), was significantly greater for the senescent group (0.19 vs. 0.13, $p < 0.05$, $n = 12$), indicating that these genotypes senesced at a faster rate compared with genotypes classified as stay-green genotypes. Consistent with this, the slope of the relationship between NDRE and DAS from maximum NDRE until maturity, was significantly steeper (−0.003 units per day) for the senescent group, compared with the stay-green group (−0.002 units per day) $p < 0.01$, indicating that the senescent genotypes lost chlorophyll significantly faster than the stay-green genotypes (**Figure 9**).

DISCUSSION

Due to variable plant numbers among plots in this pilot study, we were able to test the suitability of the vegetation indices to estimate percent cover and LAI over a range of densities. Actual plant counts per 30.4 m plot ranged from 36 to 204 plants. Correlations between actual plant number and percent cover estimated using NDVI were significant and moderately strong. The goodness of fit increased when masking was applied and pixels with NDVI lower than 0.5 (i.e., soil and non-living materials) were disregarded.

Correlations between vegetation indices and quadrat-estimated LAI improved when index values were aggregated over the sample quadrat area only rather than the entire plot area. Aggregating over the quadrat area alone provided a more direct correlation between the index and LAI, avoiding heterogeneity in canopy cover across the plot, associated with variable establishment. Sample cuts were selected where plant cover was more homogenous and therefore plot-level cover and derived vegetation index values differed from values derived from just sampling cut areas. Hence it is reasonable to assume that in more uniform trials, the expected relationship should be similar to that found for the quadrat comparisons (**Table 2**). Further experiments are being undertaken to confirm these relationships.

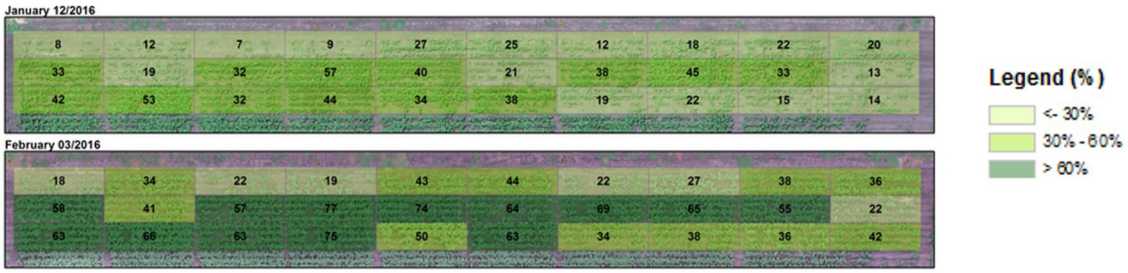


FIGURE 4 | Aggregated crop cover calculated for entire plots super-imposed on a visible (narrow band RGB) image of the sorghum breeding experiment taken pre-anthesis (upper image) or a week before anthesis (lower). PC ranged from low (light green) to high (dark green).

TABLE 2 | Relationships between NDVI (>0.5) and EVI aggregated over entire plots (plot; 30.4 m^2) or only the quadrat sampling areas (quadrat; 1.5 m^2) with leaf area index (LAI) at 54 (pre-anthesis) and 83 (anthesis) days after sowing.

Stage	Pixel aggregation level	Formula	R^2	RMSE	*P*-value
Pre-anthesis	NDVI_plot	NDVI = 0.025 * LAI + 0.703	0.55	0.019	<0.001
	NDVI_quadrat	NDVI = 0.034 * LAI + 0.690	0.85	0.011	<0.001
	EVI_plot	EVI = 0.027 * LAI + 0.316	0.19	0.045	<0.05
	EVI_quad	EVI = 0.139 * LAI + 0.138	0.81	0.056	<0.001
Anthesis	NDVI_plot	NDVI = 0.037 * LAI + 0.679	0.59	0.024	<0.001
	$NDVI_{max}$_plot	NDVI = 0.035 * LAI + 0.685	0.56	0.025	<0.001
	NDVI_quadrat	NDVI = 0.050 * LAI + 0.664	0.66	0.053	<0.001
	EVI_plot	EVI = 0.089 * LAI + 0.153	0.33	0.099	<0.001
	EVI_quadrat	EVI = 0.110* LAI + 0.360	0.70	0.247	<0.001

R^2, Regression Coefficient; RMSE, Root Mean Square Error.

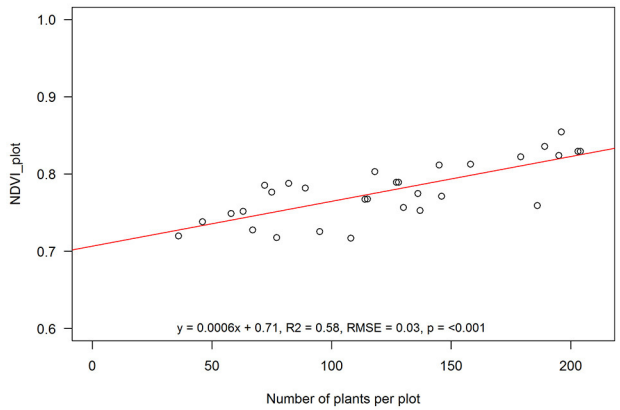


FIGURE 5 | Aggregated NDVI for entire plots at anthesis against total plant number for each 30.4 m^2 plot. Linear function: EVI = 0.0006*Plants + 0.71; R^2: regression coefficient; RMSE, root mean square error; p, statistical significance level. Solid line is the fitted linear function through the sampling points (open circles).

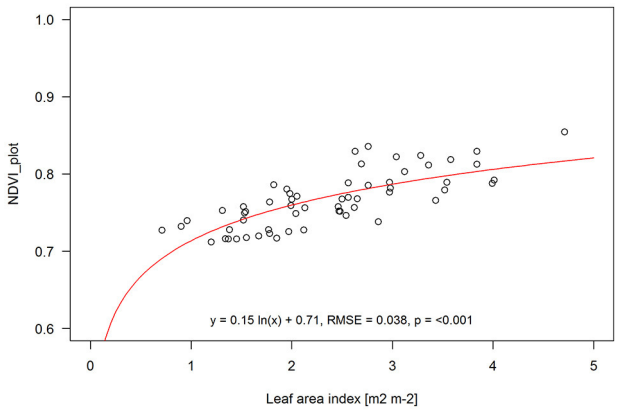


FIGURE 6 | Aggregated NDVI for entire plots vs. LAI from quadrat cuts within each plot at both the pre-anthesis and anthesis sampling dates combined. Logarithmic function NDVI = 0.15 ln (x) + 0.71; R2: regression coefficient; RMSE, root mean square error; p, statistical significance level. Solid line is the fitted function through the sampling points (open circles).

Likely limitations could exist when out-scaling this approach to other locations and crops. For example, the threshold used showed significant improvement in statistical analysis, similar to that of the EVI metric, but its utility requires further investigation. In addition, capturing data multiple times during the pilot study demonstrates the potential of these methods to study canopy dynamics. A likely constraint of comparing indices from different dates is that ambient light conditions may vary

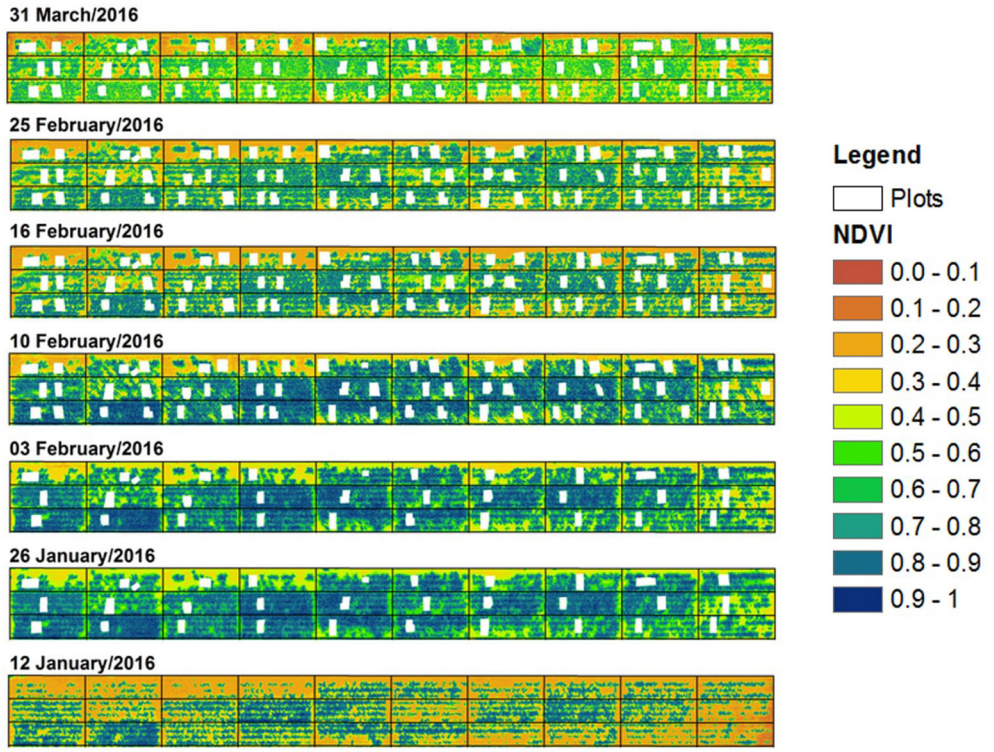


FIGURE 7 | Normalized difference vegetation index (NDVI) for each date across the study area during the main growing period. Black dividing lines indicate plot boundaries, while white mask-out areas represent the areas where sample quadrat cuts were taken.

between flights. This was limited here by flying only on clear days with no clouds and during the middle of the morning. Furthermore, NDVI is less sensitive to such changes since it is a ratio index.

Previous studies have reported a saturation of NDVI at higher LAI values (i.e., LAI > 4) and thus in dense vegetation canopies using EVI might be preferable to NDVI (Huete et al., 2002; Myneni et al., 2002). Our experimental plots were all planted at a target population density of 5 plants per square meter and LAI at anthesis ranged from 1.3 to 4.7. When combining pre-anthesis and anthesis data we also observed a slight improvement in prediction power when fitting a logarithmic instead of a linear function (**Figure 6**). To assess LAI in sorghum breeding plots with higher LAI, it might also be better to use EVI instead of NDVI.

Peak NDVI values varied from 0.72 to 0.86 and end NDVI values from 0.62 to 0.67. In this trial, the end values were not greatly lower than maximum values, given that drought stress was not substantial, with plot yields in the uniform plots being over 9 t ha^{-1}. Lines with the stay-green trait, R931945-2-2 and R931945-2-2TM, had a slower decline in NDRE after anthesis, compared with the two senescent genotypes, MR Buster and R955637. The stay-green trait has been associated with increased yield under post-anthesis drought (Borrell et al., 1999, 2000; Jordan et al., 2012) and due to the frequency of post-anthesis drought in sorghum growing areas, it has been actively selected for in Australian sorghum breeding programs. Being able to monitor senescence over time will assist breeders in selecting for stay-green under drought.

Breeding for yield under water-limitation has been the focus of sorghum breeding activities in Australia for the last three decades. This may well explain why sorghum yield advances in dry environments are currently more than double those in wet environments (Potgieter et al., 2016). However, there is potential to further improve yields in water-limited environments by improving the matching of leaf area and water-use dynamics to the temporal characteristics of drought (Chapman et al., 2000). The approach presented here offers the opportunity to monitor LAI of different genotypes throughout the crop-growing season, thus providing breeders with information on canopy dynamics. This will support the accelerated development and release of commercial hybrids that are matched to specific environments types.

In addition to plant breeders, agronomists and growers will also benefit from having access to information on crop canopy dynamics as it will allow them to estimate water use and expected yields for their sorghum crops as the season unfolds. Besides directly affecting crop water use (George-Jaeggli et al., 2017), LAI also relates to the fraction of absorbed photosynthetically active radiation (PAR) and therefore is one of the most important canopy attributes (Weiss et al., 2004; Sadras and McDonald, 2012; Sibley et al., 2014; Sadras and Calderini, 2015). LAI is

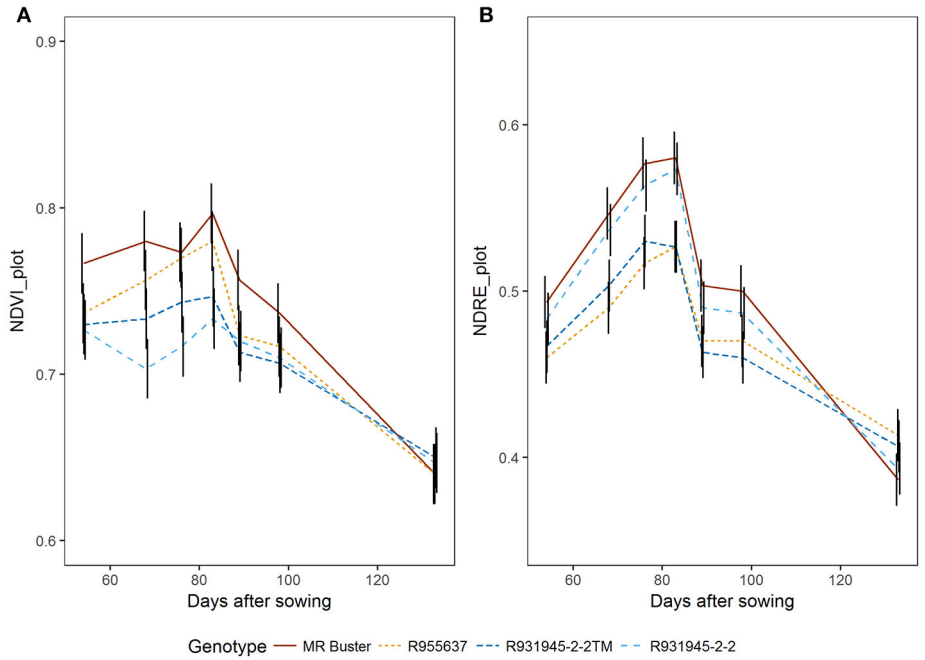


FIGURE 8 | $NDVI_{avg}$ **(A)** and $NDRE_{avg}$ **(B)** aggregated over entire plot area from flowering to maturity for four sorghum genotypes contrasting in stay-green characteristics. MR Buster (brown) and R955637 (yellow) are both senescent types, while R931945-2-2 (light blue) and R931945-2-2TM (dark blue) are lines with the stay-green trait. Points are least squares means for NDVI and NDRE, respectively, predicted by the linear mixed model. Black vertical bars represent standard errors for three replicates at each time point.

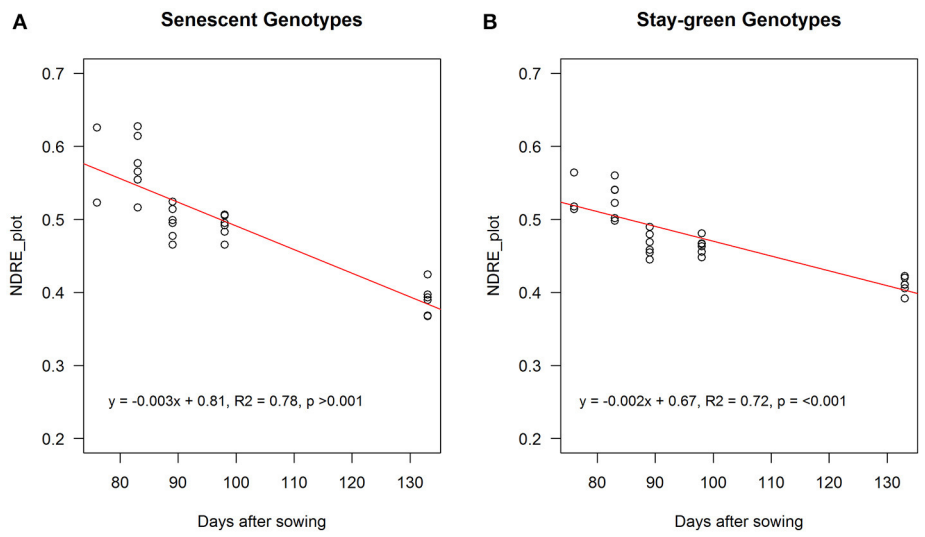


FIGURE 9 | $NDRE_{avg}$ aggregated over entire plot from maximum canopy cover to maturity (final flight) for senescent (MR Buster and R955637; **A**) and stay-green genotypes (R931945-2-2 and R931945-2-2TM; **B**). Points are values for individual plots. Solid lines are the fitted functions through the sampling points (open circles).

an important input variable for crop models, such as APSIM (Keating et al., 2003) that are used for yield predictions at field and regional scales (Lobell et al., 2015b). An improvement in this methodology would be to be able to monitor the LAI as it increases toward a maximum value, and by accurately accounting for heads, soil and senescing leaves, to estimate the LAI as it changes during grain filling. A full-season measurement of LAI would allow use of these crop models in the estimation of seasonal crop growth and potential water use.

Apart from the capacity to scale phenotyping up from a few to thousands of breeders' plots, the approach presented here will facilitate the scaling-out of phenotyping from plant to plot to field

scales and thus enabling industry to maximize yield potential at both the genetic and the agronomic level.

AUTHOR CONTRIBUTIONS

AP and BG contributed equally to this manuscript. GH and DJ obtained funding for the study. AP, BG, SC, DJ, and GH planned and designed the experiments. BG organized the data collection and JWi, ME, and KL collected the on-ground and aerial data, respectively. AP, LS, JWa, and KL processed the images. BG and AP performed the statistical analyses, interpreted the results and drafted the manuscript with contributions from SC. All authors have read and approved the final manuscript.

FUNDING

This study was funded by the Center of Excellence for Translational Photosynthesis, Australian Research Council (grant CE140100015), the Bill & Melinda Gates Foundation (grant OPPGD1197 iMashilla "A targeted approach to sorghum improvement in moisture stress areas of Ethiopia") and a Major Equipment and Infrastructure Grant "Phenotype Sensing Platform to Enhance Plant Breeding" by the University of Queensland.

ACKNOWLEDGMENTS

We thank the sorghum breeding and farm staff of the Queensland Department of Agriculture and Fisheries Hermitage Research Facility in Warwick for the preparation of seed and planting of the trial, Marie Bouteillé-Pallas and casual staff for assistance with on-ground data collection, Colleen Hunt for trial design and advice on statistical analysis, DuPont Pioneer and PacificSeeds for use of their hybrid seed and Gabriel Torres and Justin McAllister from MicaSense for the use of **Figure 2** and useful comments on the manuscript. We would also like to acknowledge the contributions of the Australian Grains Research and Development Corporation and the Queensland Government to the development of germplasm and on-going funding of the Queensland sorghum breeding program.

REFERENCES

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 48. doi: 10.18637/jss.v067.i01

Beeri, O., and Peled, A. (2006). Spectral indices for precise agriculture monitoring. *Int. J. Remote Sens.* 27, 2039–2047. doi: 10.1080/01431160612331392950

Borrell, A. K., Bidinger, F. R., and Sunitha, K. (1999). Stay-green associated with yield in recombinant inbred sorghum lines varying in rate of leaf senescence. *Int. Sorghum Millets Newsl.* 40, 31–34.

Borrell, A. K., Hammer, G. L., and Henzell, R. G. (2000). Does maintaining green leaf area in sorghum improve yield under drought? II. Dry matter production and yield. *Crop Sci.* 40, 1037–1047. doi: 10.2135/cropsci2000.4041037x

Borrell, A. K., Mullet, J. E., George-Jaeggli, B., van Oosterom, E. J., Hammer, G. L., Klein, P. E., et al. (2014a). Drought adaptation of stay-green sorghum is associated with canopy development, leaf anatomy, root growth, and water uptake. *J. Exp. Bot.* 65, 6251–6263. doi: 10.1093/jxb/eru232

Borrell, A. K., van Oosterom, E. J., Mullet, J. E., George-Jaeggli, B., Jordan, D. R., Klein, P. E., et al. (2014b). Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. *N. Phytol.* 203, 817–830. doi: 10.1111/nph.12869

Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., and Gattelli, M. (2015). Evaluating multispectral images and vegetation indices for precision farming applications from UAV images. *Remote Sens.* 7, 4026–4047. doi: 10.3390/rs70404026

Chapman, C. S., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, F. M., et al. (2014). Pheno-Copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Chapman, S. C., Cooper, M., and Hammer, G. L. (2002). Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. *Aust. J. Agric. Res.* 53, 379–389. doi: 10.1071/AR01070

Chapman, S. C., Cooper, M., Hammer, G. L., and Butler, D. G. (2000). Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. *Aust. J. Agric. Res.* 51, 209–221. doi: 10.1071/AR99021

Corrigan, F. (2015). *Top 3DR X8 Quadcopter For Mapping And Aerial Filming* [Online]. Available online at: http://www.dronezon.com/drone-reviews/3dr-x8-quadcopter-for-mapping-and-aerial-filming/ (Accessed November 9, 2016).

George-Jaeggli, B., Mortlock, M. Y., and Borrell, A. K. (2017). Bigger is not always better: reducing leaf area helps stay-green sorghum use soil water more slowly. *Environ. Exp. Bot.* 138, 119–129. doi: 10.1016/j.envexpbot.2017.03.002

Gitelson, A. A., Gritz, Y., and Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282. doi: 10.1078/0176-1617-00887

Gitelson, A., and Merzlyak, M. N. (1994). Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. spectral features and relation to chlorophyll estimation. *J. Plant Physiol.* 143, 286–292.

Goetz, A. F. H. (2009). Three decades of hyperspectral remote sensing of the Earth: a personal view. *Remote Sens. Environ.* 113 (Suppl. 1), S5–S16. doi: 10.1016/j.rse.2007.12.014

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., and Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* 90, 337–352. doi: 10.1016/j.rse.2003.12.013

Haghighattalab, A., González Pérez, L., Mondal, S., Singh, D., Schinstock, D., Rutkoski, J., et al. (2016). Application of unmanned aerial systems for high throughput phenotyping of large wheat breeding nurseries. *Plant Methods* 12, 35. doi: 10.1186/s13007-016-0134-6

Hanes, J. M. (2014). *Remote Sensing and Photogrammetry*. Berlin; Heidelberg: Springer-Verlag.

He, J., Du, Y.-L., Wang, T., Turner, N. C., Yang, R.-P., Jin, Y., et al. (2016). Conserved water use improves the yield performance of soybean (*Glycine max* (L.) Merr.) under drought. *Agric. Water Manage.* 179, 236–245. doi: 10.1016/j.agwat.2016.07.008

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

IPCC (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. eds R. K. Pachauri and L. A. Meyer. Geneva: IPCC; Core Writing Team.

Jordan, D. R., Hunt, C. H., Cruickshank, A. W., Borrell, A. K., and Henzell, R. G. (2012). The relationship between the stay-green trait and grain yield in elite sorghum hybrids grown in a range of environments. *Crop Sci.* 52, 1153–1161. doi: 10.2135/cropsci2011.06.0326

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., et al. (2003). An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18, 267–288. doi: 10.1016/S1161-0301(02)00108-9

Lelong, C. C. D., Burger, P., Jubelin, G., Roux, B., Labbe, S., and Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors* 8, 3557–3585. doi: 10.3390/s8053557

Lobell, D. B., Hammer, G. L., Chenu, K., Zheng, B., McLean, G., and Chapman, S. C. (2015a). The shifting influence of drought and heat stress for crops in northeast Australia. *Glob. Chang. Biol.* 21, 4115–4127. doi: 10.1111/gcb.13022

Lobell, D. B., Thau, D., Seifert, C., Engle, E., and Little, B. (2015b). A scalable satellite-based crop yield mapper. *Remote Sens. Environ.* 164, 324–333. doi: 10.1016/j.rse.2015.04.021

McKeown, F. R. (1978). *A Land Classification of the Hermitage Research Station*. Division of Land Utilisation: Queensland Department of Primary Industries, Brisbane, Australia.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* 83, 214–231. doi: 10.1016/S0034-4257(02)00074-3

Potgieter, A. B., Lobell, D. B., Hammer, G. L., Jordan, D. R., Davis, P., and Brider, J. (2016). Yield trends under varying environmental conditions for sorghum and wheat across Australia. *Agric. Forest Meteorol.* 228–229, 276–285. doi:10.1016/j.agrformet.2016.07.004

Pratley, J. (2003). *Principles of Field Crop Production*. Melbourne, VIC: Oxford University Press.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rouse, J., Jr, Haas, R., Schell, J., and Deering, D. (1974). "Monitoring vegetation systems in the Great Plains with ERTS," in *3d ERTS-1 Symp*., Vol. 1 Sect A (Goddard Space Flight Center: NASA), 309–317.

Sadras, V. O., and Calderini, D. F. (2015). *Crop Physiology*. London; Waltham, MA; San Diego, CA: Academic Press.

Sadras, V. O., and McDonald, G. (2012). *Water Use Efficiency of Grain Crops in Australia: Principles, Benchmarks and Management*. Kingston, VIC: Grains Research and Development Corporation.

Sibley, A. M., Grassini, P., Thomas, N. E., Cassman, K. G., and Lobell, D. B. (2014). Testing remote sensing approaches for assessing yield variability among maize fields. *Agron. Soils Environ. Qual.* 106, 24–32. doi: 10.2134/agronj2013.0314

Sims, D. A., and Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* 81, 337–354. doi: 10.1016/S0034-4257(02)00010-X

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0

Viña, A., Gitelson, A. A., Nguy-Robertson, A. L., and Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sens. Environ.* 115, 3468–3478. doi: 10.1016/j.rse.2011.08.010

Weiss, M., Baret, F., Smith, G. J., Jonckheere, I., and Coppin, P. (2004). Review of methods for *in situ* leaf area index (LAI) determination: part, I. I. Estimation of LAI, errors and sampling. *Agric. Forest Meteorol.* 121, 37–53. doi: 10.1016/j.agrformet.2003.08.001

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 March 2017
doi: 10.3389/fpls.2017.00421

High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling

*Kakeru Watanabe[1], Wei Guo[2], Keigo Arai[3], Hideki Takanashi[4], Hiromi Kajiya-Kanegae[1], Masaaki Kobayashi[5], Kentaro Yano[5], Tsuyoshi Tokunaga[6], Toru Fujiwara[7], Nobuhiro Tsutsumi[4] and Hiroyoshi Iwata[1]**

[1] *Laboratory of Biometry and Bioinformatics, Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan,* [2] *Institute for Sustainable Agro-ecosystem Services, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan,* [3] *Air4D Co., Ltd., Tokyo, Japan,* [4] *Laboratory of Plant Molecular Genetics, Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan,* [5] *Bioinformatics Laboratory, Department of Life Sciences, School of Agriculture, Meiji University, Kanagawa, Japan,* [6] *Earthnote Co., Ltd., Okinawa, Japan,* [7] *Laboratory of Plant Nutrition and Fertilizers, Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan*

OPEN ACCESS

Edited by:
Julie Dickerson, Iowa State University, USA

Reviewed by:
David Mayerich, University of Houston, USA
Jagadish Rane, Indian Council of Agricultural Research, India

****Correspondence:***
Hiroyoshi Iwata
aiwata@mail.ecc.u-tokyo.ac.jp

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *12 January 2017*
Accepted: *13 March 2017*
Published: *28 March 2017*

Citation:
Watanabe K, Guo W, Arai K, Takanashi H, Kajiya-Kanegae H, Kobayashi M, Yano K, Tokunaga T, Fujiwara T, Tsutsumi N and Iwata H (2017) High-Throughput Phenotyping of Sorghum Plant Height Using an Unmanned Aerial Vehicle and Its Application to Genomic Prediction Modeling. Front. Plant Sci. 8:421. doi: 10.3389/fpls.2017.00421

Genomics-assisted breeding methods have been rapidly developed with novel technologies such as next-generation sequencing, genomic selection and genome-wide association study. However, phenotyping is still time consuming and is a serious bottleneck in genomics-assisted breeding. In this study, we established a high-throughput phenotyping system for sorghum plant height and its response to nitrogen availability; this system relies on the use of unmanned aerial vehicle (UAV) remote sensing with either an RGB or near-infrared, green and blue (NIR-GB) camera. We evaluated the potential of remote sensing to provide phenotype training data in a genomic prediction model. UAV remote sensing with the NIR-GB camera and the 50th percentile of digital surface model, which is an indicator of height, performed well. The correlation coefficient between plant height measured by UAV remote sensing (PH_{UAV}) and plant height measured with a ruler (PH_R) was 0.523. Because PH_{UAV} was overestimated (probably because of the presence of taller plants on adjacent plots), the correlation coefficient between PH_{UAV} and PH_R was increased to 0.678 by using one of the two replications (that with the lower PH_{UAV} value). Genomic prediction modeling performed well under the low-fertilization condition, probably because PH_{UAV} overestimation was smaller under this condition due to a lower plant height. The predicted values of PH_{UAV} and PH_R were highly correlated with each other ($r = 0.842$). This result suggests that the genomic prediction models generated with PH_{UAV} were almost identical and that the performance of UAV remote sensing was similar to that of traditional measurements in genomic prediction modeling. UAV remote sensing has a high potential to increase the throughput of phenotyping and decrease its cost. UAV remote sensing will be an important and indispensable tool for high-throughput genomics-assisted plant breeding.

Keywords: genomic prediction, high-throughput phenotyping, near-infrared (NIR), sorghum plant height, UAV remote sensing

INTRODUCTION

Improving the throughput of phenotyping in the field is a big challenge in plant genetics, physiology, and breeding. The emergence of the next-generation sequencing technologies enables us to obtain genome-wide DNA polymorphism data for a large number of samples easily and rapidly (Mardis, 2007; Davey et al., 2011). Statistical methods, such as genome-wide association study (GWAS; Brachi et al., 2011; Korte and Farlow, 2013; Huang and Han, 2014) and genomic selection (GS; Meuwissen et al., 2001; Jannink et al., 2010) allow us to associate DNA polymorphism data, which is extremely high-dimensional, to phenotypic variations in agronomic traits. Boosted by these technological developments, the efficiency of plant breeding is expected to improve rapidly (Huang and Han, 2014). However, phenotyping is still time consuming and labor intensive, and may be more costly than genotyping. Thus, phenotyping has become a serious bottleneck in the acceleration of plant breeding (Furbank and Tester, 2011). Field experiments at multiple plant breeding stations over a large geographic area are indispensable to evaluate the adaptability of new candidate genotypes and to examine the pattern of genotype-environment interaction (Chapman et al., 2014). At each breeding station, a large number of genotypes are tested in the field. Most of the measurements conducted in the field are destructive and labor- and time-intensive, and thus cannot be repeated frequently in the course of plant growth. Because phenotypic data is necessary for genomics-assisted breeding, it is the first priority to develop a high-throughput phenotyping method.

Remote sensing using a low-altitude unmanned aerial vehicle (UAV), such as radio-controlled multicopter, can solve the problem described above. Besides low-altitude UAVs, measurements using satellites (Inoue, 1997) and ground-based vehicles (Lee and Searcy, 1999) have been applied for remote sensing of growth conditions of crop plants (Sugiura et al., 2005). However, satellites have low resolution, poor sensitivity under cloudy conditions, and slow data transmission (Sugiura et al., 2005; Zhang and Kovacs, 2012), and ground-based vehicles cannot enter fields with tall crops or muddy soil (Sugiura et al., 2005; Chapman et al., 2014). Low-altitude UAVs have no such disadvantages and can be used without expert skills (Merz and Chapman, 2011). Most low-altitude UAVs have an autopilot function to fly automatically along a route designed by mission planning software (Berni et al., 2009; Chapman et al., 2014; Zarco-Tejada et al., 2014; Díaz-Varela et al., 2015). Another widespread remote sensing technology is light detection and ranging (LiDAR). However, this technology has some shortcomings, e.g., high cost of data acquisition and processing (Díaz-Varela et al., 2015). The emergence of computer-vision technologies, such as the structure-from-motion and multi-view-stereo algorithms, enables reconstruction of accurate 3D-models from a series of images with a considerable overlap between adjacent images. These technologies are attractive alternatives to LiDAR, due to their high performance, flexibility, and relatively low cost (Díaz-Varela et al., 2015). For remote sensing of plants, near-infrared (NIR) cameras have been used in many studies (Lee and Searcy, 1999; Sugiura et al., 2005; Berger et al., 2010; Cabrera-Bosquet et al., 2011; van Maarschalkerweerd et al., 2013; Colomina and Molina, 2014; Díaz-Varela et al., 2015; Torres-Sánchez et al., 2015), because plant leaves (or chlorophylls) strongly reflect NIR light (Knipling, 1970; Tucker, 1979; Fahlgren et al., 2015) and some indices based on NIR reflectance rate, such as normalized difference vegetation index (NDVI; Rouse et al., 1974), are useful for identifying plants and assessing their growing conditions via remote sensing. Some studies have indicated that NIR sensors have advantages over standard RGB sensors in plant monitoring (Nijland et al., 2014; Zhang et al., 2016). Nevertheless, NIR cameras are less common, and often more expensive than RGB cameras (or extra cost is needed to modify RGB cameras into NIR cameras). Because remote sensing is a promising tool for phenotyping, we compared the advantages of RGB and NIR cameras in phenotyping and genomic prediction modeling.

Genomic selection is a novel breeding method that allows selection of complex traits with genome-wide markers. Because the selection is performed on the basis of the genetic potential predicted from these markers, GS requires building an accurate prediction model based on a dataset of individuals or lines that have been genotyped and phenotyped (Meuwissen et al., 2001; Jannink et al., 2010). A large dataset is needed to build an accurate prediction model. As mentioned above, however, phenotyping is time consuming and labor intensive, and is a serious bottleneck in building an accurate model. If UAV remote sensing can streamline the collection of phenotypic data, it will greatly enhance the potential of GS.

Using image-processing software for photogrammetry, we can obtain ortho-mosaic and a digital surface model (DSM) from UAV images (Gini et al., 2013). Ortho-mosaic is a distortion-corrected image. DSM provides information on the altitude. In plant science, DSM information has been applied to estimate biomass and plant height of barley (Bendig et al., 2014), and plant height, volume, and canopy size of olive trees (Zarco-Tejada et al., 2014; Díaz-Varela et al., 2015; Torres-Sánchez et al., 2015). Currently, we are using genomics-assisted breeding to develop a sorghum [*Sorghum bicolor* (L.) Moench] variety that can be used for high bioethanol production for biofuel. Plant height is one of the most important traits affecting bioethanol yield. Because some sorghum accessions may be taller than 4 m, they are usually cut for measurements, which are labor intensive, whereas GS requires phenotypic data and marker genotype data for a large number of accessions.

The objectives of this study are the validation of the usefulness of UAV remote sensing for measurement of sorghum plant height and for genomic prediction modeling. First, we confirmed the accuracy of plant height estimates from UAV images under the conditions of small plot size. Next, we examined the accuracy of genomic prediction of plant height trained by UAV remote sensing data and data manually measured with a ruler. To evaluate the robustness of this method to plant height variation related to environmental differences such as nutrition level, sorghum plants were grown at two levels of nitrogen availability. We also compared the measurement and prediction accuracy of different cameras and different procedures of remote sensing data analysis.

MATERIALS AND METHODS

Field Experiment

In this study, we used 172 accessions from sorghum germplasm collections (Supplementary Table 1). Of these, 78 accessions were from the world core collection of NIAS (National Institute of Agrobiological Science, Ibaraki, Japan; integrated into the National Agriculture and Food Research Organization from April 1, 2016), 91 were from the sorghum mini core collection of ICRISAT (International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India) and 3 were original cultivars developed by EARTHNOTE Co., Ltd. (Okinawa, Japan). Seeds were sown on 200-cell plug trays on May 8, 2014. Seedlings were transplanted to a field of the Institute for Sustainable Agro-ecosystem Services, the University of Tokyo (Tokyo, Japan; 35°44′09.1″N, 139°32′23.7″E, 60 m above the sea level) on June 9, 2014. An outline of the field design is shown in **Figure 1**. To investigate the effect of fertilization on plant growth, we used normal (N-P-K: 1.2-1.8-1.6 kg/a) and low (N-P-K: 0.6-1.8-1.6 kg/a) fertilization in two replications per treatment per accession ($172 \times 4 = 688$ plots in total). Five plants were grown in each plot (inter-plant spacing, 0.3 m; inter-row spacing, 0.72 m). On October 2 and 3, two plants per plot were harvested and their height was measured with a ruler. In total, $688 \times 2 = 1{,}376$ plants were measured.

Remote Sensing Experiment

The radio-controlled quadcopter USM-S1 (Air4D Co., Ltd., Tokyo, Japan; **Figure 2A**), was used as a UAV for remote sensing. Two digital cameras, Canon PowerShot ELPH 110HS (Canon Inc., Tokyo, Japan), were installed on the UAV (resolution, ca. 16.1 million pixels; sensor size, 6.2 mm × 4.7 mm; focal length, 4.3–21.5 mm). One was a normal RGB camera, and the other one was modified to capture NIR, green and blue (NIR-GB). The NIR-GB camera was purchased at MaxMax Inc. (Carlstadt, NJ, USA). The focal length was set at 4.3 mm. The focus was adjusted by the camera auto-focus function. On the ground at each corner of the field, we installed a white acrylic disk (27 cm in diameter) as a ground control point (GCP) (**Figure 2C**). The positions of GCPs in the World Geodetic System were measured by using GPS (Geo7X, Trimble Inc., Sunnyvale, CA, USA) and used in image processing.

To compare plant height measured with the UAV (PH_{UAV}) and plant height measured with rulers (PH_R), we performed a remote-sensing experiment on the first day of harvest (October 2). The weather on that day was cloudy. ISO sensitivity, which is an indicator of light sensitivity provided by the International Organization for Standardization (Vernier, Geneva, Switzerland), was set at 320 and shutter speed at 1/1,250 s for the RGB camera; ISO sensitivity was 800 and shutter speed was 1/800 s for the NIR-GB camera. The UAV was controlled by an autopilot system with GPS to fly along a pre-defined course designed by the PC Ground Station software (DJI Co., Ltd., Shenzheng, China). The outline of the flight course is shown in **Figure 2B**. The course was designed so that the vertical overlap of images was 70% and horizontal overlap was 30%. Photographs were taken at an altitude of 40 m, total flight time was about 10 min, and 30 photographs were taken. From an altitude of 40 m, the resolution was ca. 13 mm per pixel. A preliminary remote sensing experiment was performed on July 23 with the RGB camera NEX-7 (Sony Corporation, Tokyo, Japan; resolution, ca. 24 million pixels; sensor size, 23.5 mm × 15.6 mm) with a lens of focal length 20 mm. The following parameters were used in the preliminary experiment: ISO sensitivity, 100; shutter speed, 1/800 s; altitude, 50 m; 78 photographs were taken. From an altitude of 50 m, the resolution was ca. 10 mm per pixel. Because plants were still small on June 23, we used the data collected on July 23 to obtain DSM data on the ground surface of the field. Although NEX-7 has higher resolution than PowerShot ELPH 110HS, it is heavier and thus we could not mount two types of cameras (RGB and NIR) simultaneously on the UAV. At an altitude of 40 m, the resolution of PowerShot ELPH 110HS was similar to that of

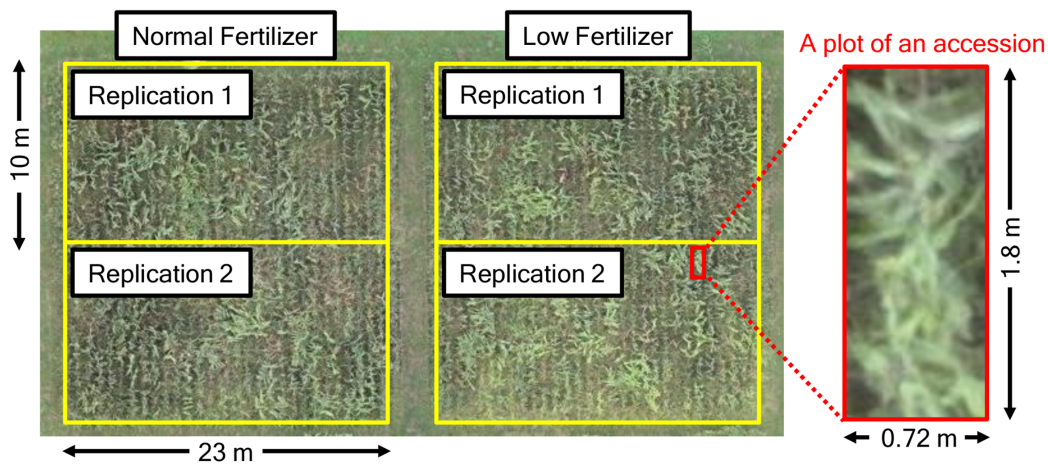


FIGURE 1 | Field design in this study.

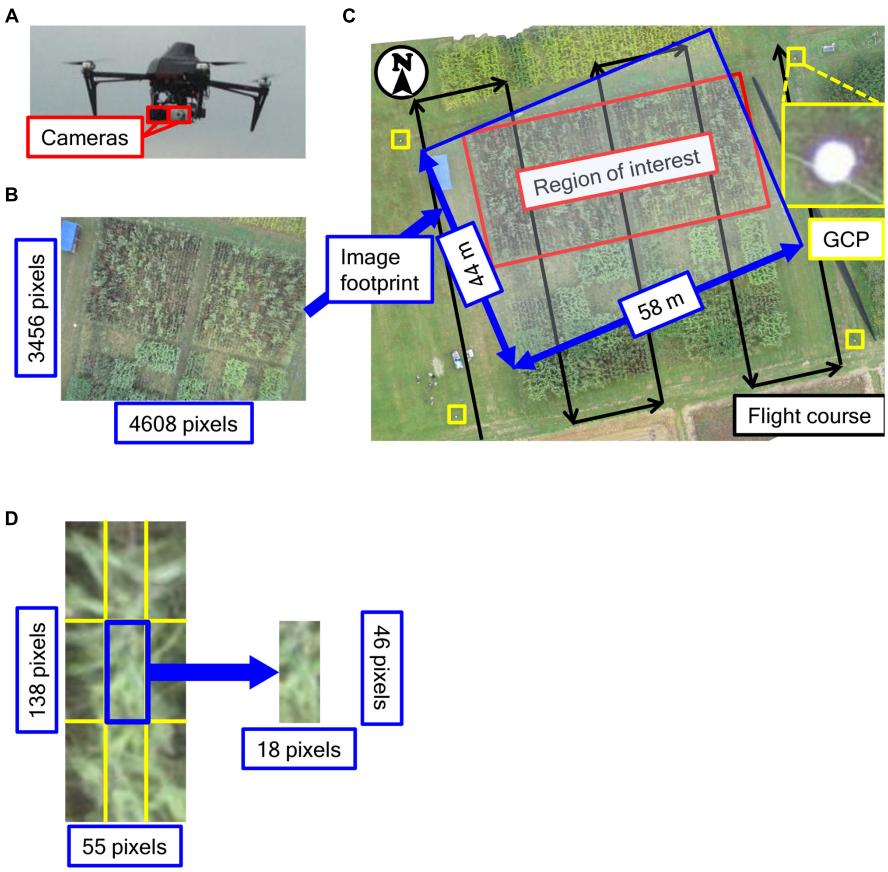


FIGURE 2 | Unmanned aerial vehicle used. (A) Images of UAV used in the study, **(B)** An image taken in this study, **(C)** Outline of the flight course and **(D)** The size of one plot and one divided block. Positions of GCPs in the World Geodetic System were measured by using GPS.

NEX-7. Therefore, we used PowerShot ELPH 110HS for remote sensing.

Image Processing

Both RGB and NIR-GB images were analyzed in the same way. By using the in-house structure-from-motion software Nadir-metry (Air4D Co., Ltd., Tokyo, Japan), ortho-mosaic images and DSM data were constructed from images taken by the UAV with the geographic coordinates of GCPs. Although the structure-from-motion algorithm of Nadir-metry is similar to that in other software, it has some advantages in feature point matching and generating point clouds. In the algorithm of Nadir-metry, feature point matching is performed by taking into account the correspondence between overlapping images estimated from their geographic coordinates. As a result, spatial skew hardly occurs. In the generation of point clouds, all pixels were analyzed to detect matching points. This decreased the number of missing matches and prevented point clouds from being sparse. Because DSM values were calculated based on the World Geodetic System 1984 and they did not directly reflect the ground height of objects, we estimated the height of sorghum plants by subtracting the DSM values of the ground surface of the field on July 23 from the DSM values on October 2, as in Bendig et al. (2013). From the location of each plot determined on the ortho-mosaic image, we obtained PH_{UAV} for each plot from the DSM data. Because adjacent plots were close to each other, DSMs of plot boundaries were contaminated with data originated from adjacent plots, and might have higher error than those inside a plot. To exclude marginal areas, we divided each plot into 9 blocks (3 × 3) and analyzed only the central block. That is, the plot size was 0.72 m × 1.8 m and corresponded to ca. 55 × 138 pixels (**Figure 2D**) at a resolution of 13 mm per pixel in DSM. The size of the central block was 0.24 m × 0.6 m, corresponding to ca. 18 × 46 pixels of DSM (**Figure 2D**). Each pixel had a DSM value that was construed as the altitude of the location. We calculated the 50th (median), 75th, 90th, and 99th percentiles of DSM values of the central block as the representative values of PH_{UAV} for the plot. We evaluated the accuracy of PH_{UAV} from its correlation coefficients with PH_R and also from root mean square difference (RMSD):

$$\mathrm{RMSD} = \sqrt{\frac{1}{688}\sum_{i=1}^{688}(\mathrm{PH}_{\mathrm{UAV},i} - \mathrm{PH}_{\mathrm{R},i})^2},$$

where $PH_{UAV,\ i}$ and $PH_{R,\ i}$ are the PH_{UAV} value and the PH_R value of the *ith* plot, respectively.

Treatment of Unexpected Values

We assumed that if a plot of a tall accession and that of a small accession were adjacent in the field, the plants of the small accession might be overlapped by those of the tall accession on images, and DSMs of the small accession might be overestimated because they were strongly contaminated by data originated from the tall accession. To confirm this, we compared the correlation between PH_{UAV} and PH_R under three conditions: (1) using data of the replication with higher PH_{UAV} value for each accession; (2) using data of the replication with lower PH_{UAV} value for each accession; (3) using average PH_{UAV} value of two replications for each accession. As described later, the result using the lower PH_{UAV} was better than others, and the replication with lower PH_{UAV} value for each accession was used in the following analysis.

Genomic Prediction Modeling

We built genomic prediction models for both PH_{UAV} and PH_R, and compared the predicted values. To obtain DNA polymorphism data, we used restriction site-associated DNA sequencing (RAD-Seq; Baird et al., 2008); which is cheaper than whole genome sequencing especially for analysis of DNA polymorphism of large number of accessions. We obtained the data for 66,132 SNPs in 151 accessions (Supplementary Table 1). Genomic best linear unbiased prediction (G-BLUP) using rrBLUP (Endelman, 2011) was used for the modeling. To calculate predicted values for the 151 accessions and compare them for PH_{UAV} and PH_R, we performed leave-one-out cross-validation. In cross-validation, a model built with the PH_{UAV} or PH_R data of 150 of the 151 accessions was used to predict the PH_{UAV} or PH_R values of the remaining accession from its DNA polymorphism data. By comparing the observed and predicted values of both PH_{UAV} and PH_R, we evaluated whether manual measurements (PH_R) can be replaced with the measurements using UAV remote sensing (PH_{UAV}) in the collection of data for building a model.

RESULTS

Image Processing and Measurement of Plant Height via UAV Remote Sensing

For both RGB and NIR-GB cameras, ortho-mosaic images and DSM heat maps of the experimental field before harvest were constructed by using 30 remote-sensing images taken on October 2 (**Figure 3**).

To assess the accuracy of plant height measurements, we compared the correlation between PH_{UAV} and PH_R obtained with the two cameras at the 50th, 75th, 90th, and 99th percentiles of DSM values (**Figure 4**). Although there was no significance between two correlation coefficients (RGB vs. NIR-GB) at any combinations, correlation coefficients were higher for the NIR-GB than for RGB camera: 0.523 vs. 0.518 (50th percentile), 0.507 vs. 0.504 (75th), and 0.496 vs. 0.491 (90th). However, correlation coefficients were higher for RGB (0.475) than for NIR-GB (0.473) at the 99th percentile. RMSD between PH_{UAV} and PH_R was lower for the NIR-GB than RGB camera: (0.649 vs. 0.883 (50th percentile), 0.626 vs. 0.827 (75th), 0.628 vs. 0.792 (90th), and 0.665 vs. 0.759 (99th). PH_{UAV} obtained from the RGB camera underestimated PH_R because the points were distributed below the $y = x$ line in **Figure 4**. PH_{UAV} obtained from the NIR-GB camera estimated PH_R more accurately, especially at the 90th and 99th percentiles, as evidenced by smaller RMSD in RGB than in NIR-GB.

The relationship between PH_{UAV} and PH_R suggested that PH_{UAV} was overestimated at PH_R values of <2.0 m (**Figure 4**). To assess whether the presence of taller plants on adjacent plots resulted in overestimation, we calculated two types of ratio for each plot as follows:

$$r_1 = \frac{\max\{PH_{R,i,k} | k = 1, 2, ..., N\}}{PH_{R,i}}, r_2 = \frac{PH_{UAV,i}}{PH_{R,i}}$$

where $PH_{R,\ i}$ and $PH_{UAV,\ i}$ are the PH_R and PH_{UAV} values of the *ith* plot, respectively, and $N = 8$ in this case. $PH_{R,\ i,\ k}$ is the PH_R value of the *kth* (1–8) plot adjacent to the *ith* plot. The ratio r_1 represents the degree of height difference among the plants on adjacent plots. The ratio r_2 represents the degree of over- or under- estimation of PH_{UAV} against PH_R. A scatter plot of these ratios for one combination (NIR-GB camera and the 50th percentile of DSM values) is shown in Supplementary Figure 1A; the results for other combinations were similar. Theoretically, the PH_{UAV}/PH_R ratio is expected to be approximately constant if UAV measurements are accurate, because both PH_{UAV} and PH_R are expected to be the true values of plant height (with measurement errors). If we regarded PH_R as the true plant height, the PH_{UAV}/PH_R ratio became large, i.e., the PH_{UAV} was overestimated against PH_R (Supplementary Figure 1A) when PH_R on adjacent plots was 1.5 times that on the central plot. This result suggests that the presence of taller plants on an adjacent plot prevents accurate construction of the DSM of the target plot because of the overlapping effect.

We analyzed the relationships between PH_{UAV} and PH_R for all the combinations of the two cameras and four percentile values separately for each of two replications (one with the lower PH_{UAV} value, the other one with the higher PH_{UAV} value) and the average (**Figure 5**). For all combinations, the correlation coefficients were highest (around 0.65) for the replications with lower PH_{UAV} values and lowest (around 0.40) for the replications with higher PH_{UAV} values. Correlation coefficients with lower PH_{UAV} were significantly higher than those with mean PH_{UAV} at significant level of 10% at all combinations. This result suggested that the lower PH_{UAV} values were more reliable.

Genomic Prediction Modeling

To evaluate the accuracy of genomic prediction for PH_{UAV} and PH_R and the agreement between their predicted values, we obtained these values via leave-one-out cross- validation for all sorghum accessions cultivated under normal and low fertilization conditions. The relationships between the observed and predicted PH_{UAV} and PH_R values are shown in **Figure 6**. If the combination of two cameras and four representative values is different, the combination

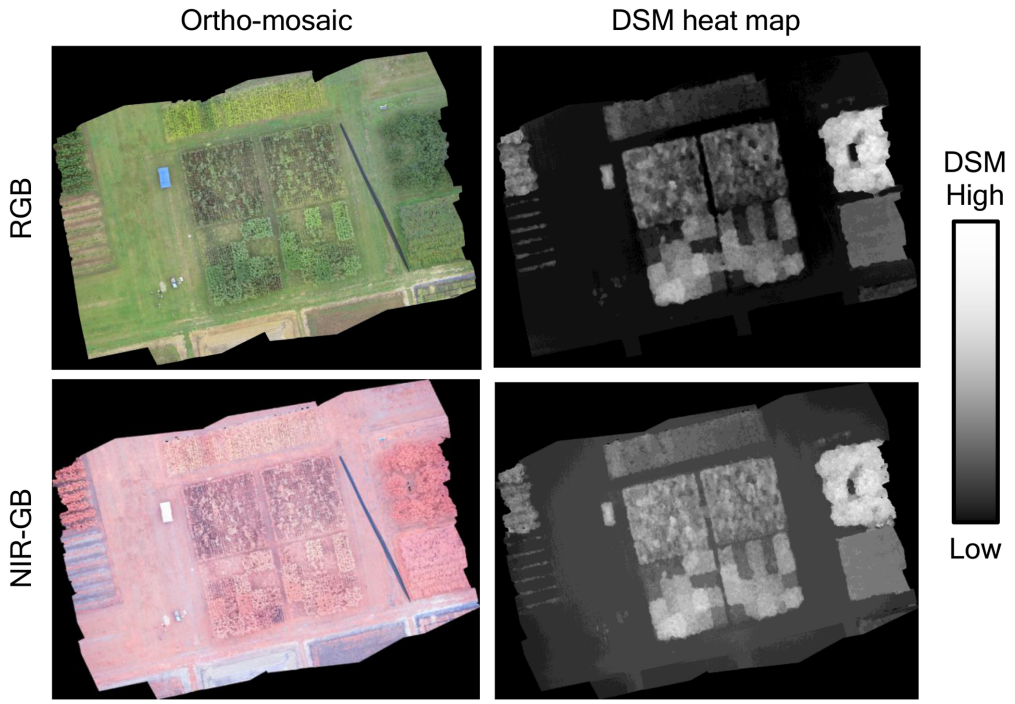


FIGURE 3 | Ortho-mosaic and DSM heat map of October 2. Upper images were originated from the standard RGB cameras and lower images were originated from the NIR-GB (near-infrared, green, blue) camera. DSM provides information of the altitude. The resolution was ca. 13 mm per pixel. The sizes of both the ortho-mosaic and DSM were 10,701 × 8,061 pixels in RGB and 10,749 × 8,041 in NIR-GB; the ortho-mosaic and DSM were built from 30 images (4,608 × 3,456 pixels each).

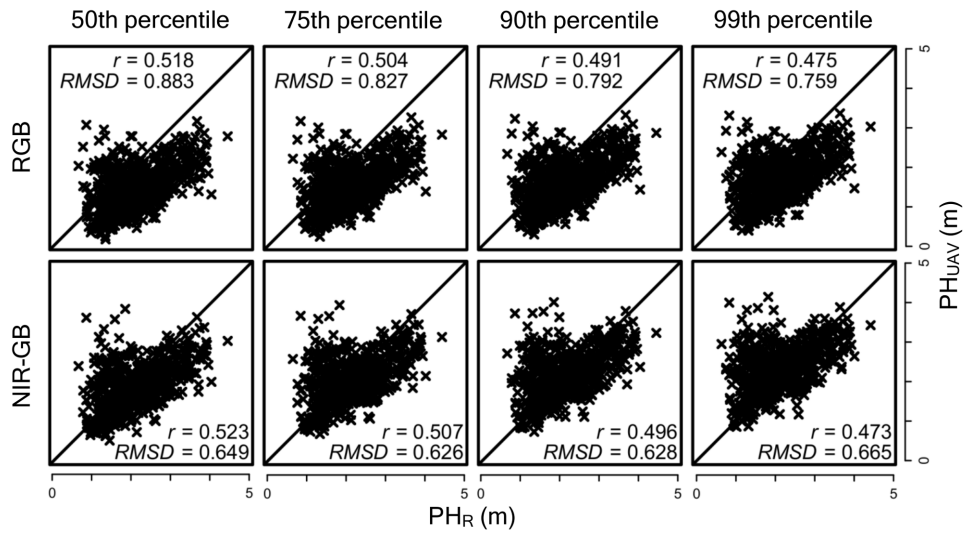


FIGURE 4 | Accuracy under different cameras and different percentiles of DSM. The diagonal line indicates $y = x$; r, correlation coefficient; RMSD, root mean square difference. PH_{UAV} is the plant height measured with the UAV and PH_R is that with rulers. PH_{UAV} values were calculated as 50th, 75th, 90th, and 99th percentile of DSM values in each plot.

of selected replications for each fertilization condition is also different. For example, we consider two combinations: (i) (PH_{UAV} of replication 1) < (PH_{UAV} of replication 2) for accession A and (PH_{UAV} of replication 1) < (PH_{UAV} of replication 2) for accession B in a combination of two cameras and four percentiles of DSM values (combination 1), and (ii) (PH_{UAV} of replication 1) < (PH_{UAV} of replication 2) for accession A and (PH_{UAV} of replication 1) > (PH_{UAV}

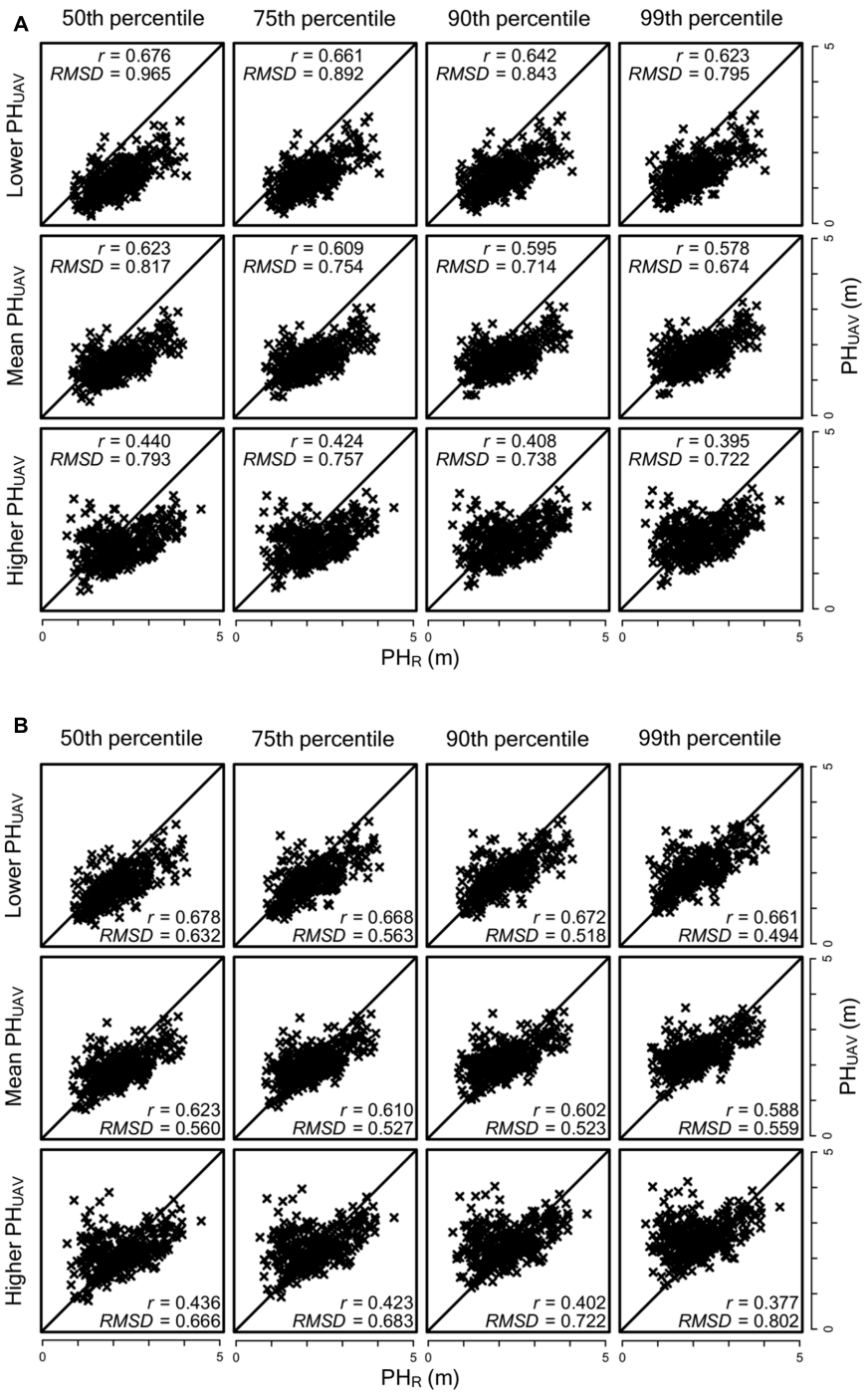


FIGURE 5 | Comparison of lower, high and average PH_{UAV} from two replications. (A) RGB camera and **(B)** NIR-GB camera. The diagonal line indicates $y = x$; r, correlation coefficient; RMSD, root mean square difference.

of replication 2) for accession B in another combination (combination 2). Then, for accession A, replication 1 is selected as the plot with the lower PH_{UAV} value in both combinations, whereas for accession B replication 1 is selected in combination 1 and replication 2 is selected in combination 2 as the plot with lower PH_{UAV} value. The same combinations of replications were also used for PH_R. However, the combinations of replications were almost the same between different combinations of two cameras and four percentiles. Because of this, the results of PH_R prediction with different cameras and different percentiles were similar to each other. The combinations with the highest and the lowest correlation coefficients are shown in **Figure 6**

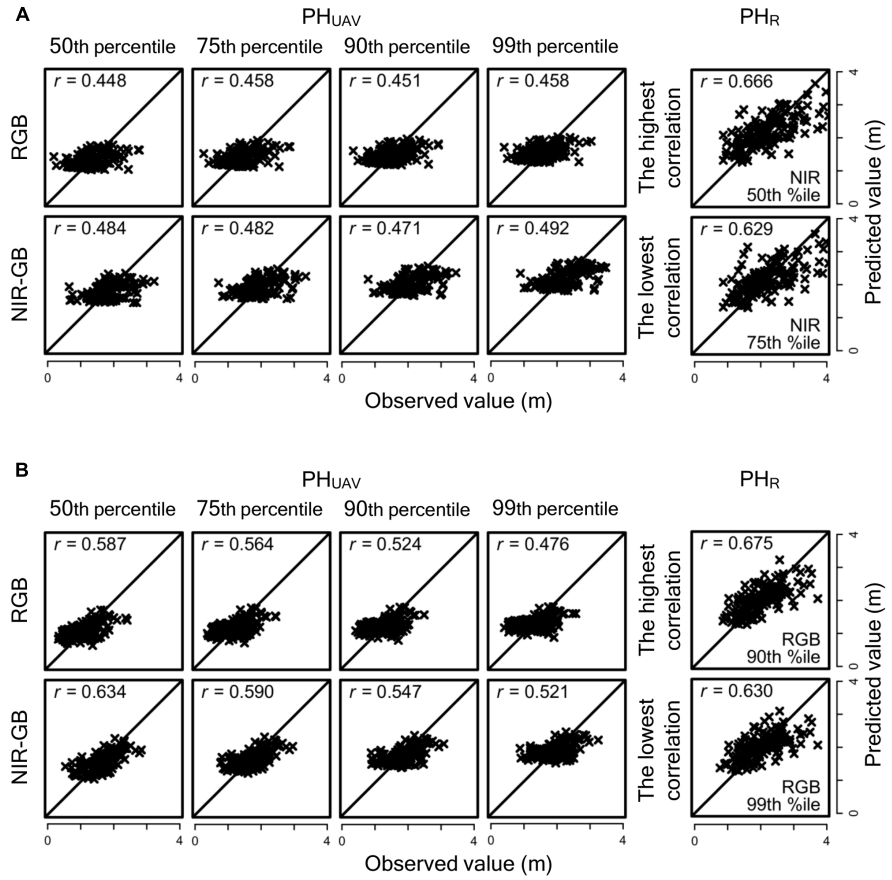


FIGURE 6 | Relationships between observed and predicted PH_{UAV} and PH_R values in genomic prediction under normal and low fertilization conditions. (A) Normal fertilization and **(B)** Low fertilization. *r*, correlation coefficient. For PH_R, only the combinations the highest and lowest correlation coefficients are shown.

for PH_R prediction. Under normal fertilization, correlation coefficients between observed and predicted PH_{UAV} were less than 0.5 (range, 0.448–0.492) for all combinations of the type of cameras and the percentiles of DSM, whereas those for PH_R ranged from 0.629 to 0.675. Correlation coefficients for PH_{UAV} were higher under low fertilization than under normal fertilization in all combinations. Although almost all combinations had lower correlation coefficients for PH_{UAV} than for PH_R, the correlation coefficient for PH_{UAV} in the combination of NIR-GB and 50th percentile was as high as that for PH_R (**Figure 6**). Predicted PH_{UAV} and predicted PH_R highly correlated with each other in all combination, (correlation coefficients $\geq$ 0.66 under low fertilization and $\geq$0.78 under normal fertilization except for one combination; **Figure 7**).

DISCUSSION

We introduced UAV remote sensing for high-throughput measurement of sorghum plant height, and applied it to genomic prediction modeling. The results of this study suggest the potential of UAV remote sensing for the high-throughput phenotyping of plant height in sorghum. Some sorghum genotypes are too tall to measure the height of the plants without harvesting them. Our approach would not only decrease labor cost but would also allow observation of plant growth over time. Traditionally, all accessions are measured only once, although their growth stages at the measurement time may differ. Our remote sensing approach would allow comparison of accessions at the same growth stage.

Using UAV remote sensing, we could not measure sorghum plant height as accurately as that of barley (which is smaller than sorghum) measured by Bendig et al. (2014). One reason may be that tall sorghum plants overlap (Supplementary Figure 1). Another reason may be low plant density: 3.9 plants/m^2 vs. 300 plants/m^2 in Bendig et al. (2014); we can see the sparseness of plants in a plot in **Figure 1**. If plant density is high enough to form a continuous canopy, most of the matched points are captured from the canopy and DSM reflects plant height precisely. However, because of the sparseness, matched points included not only the top of canopy but also the ground or lower parts of plants. This could cause errors on

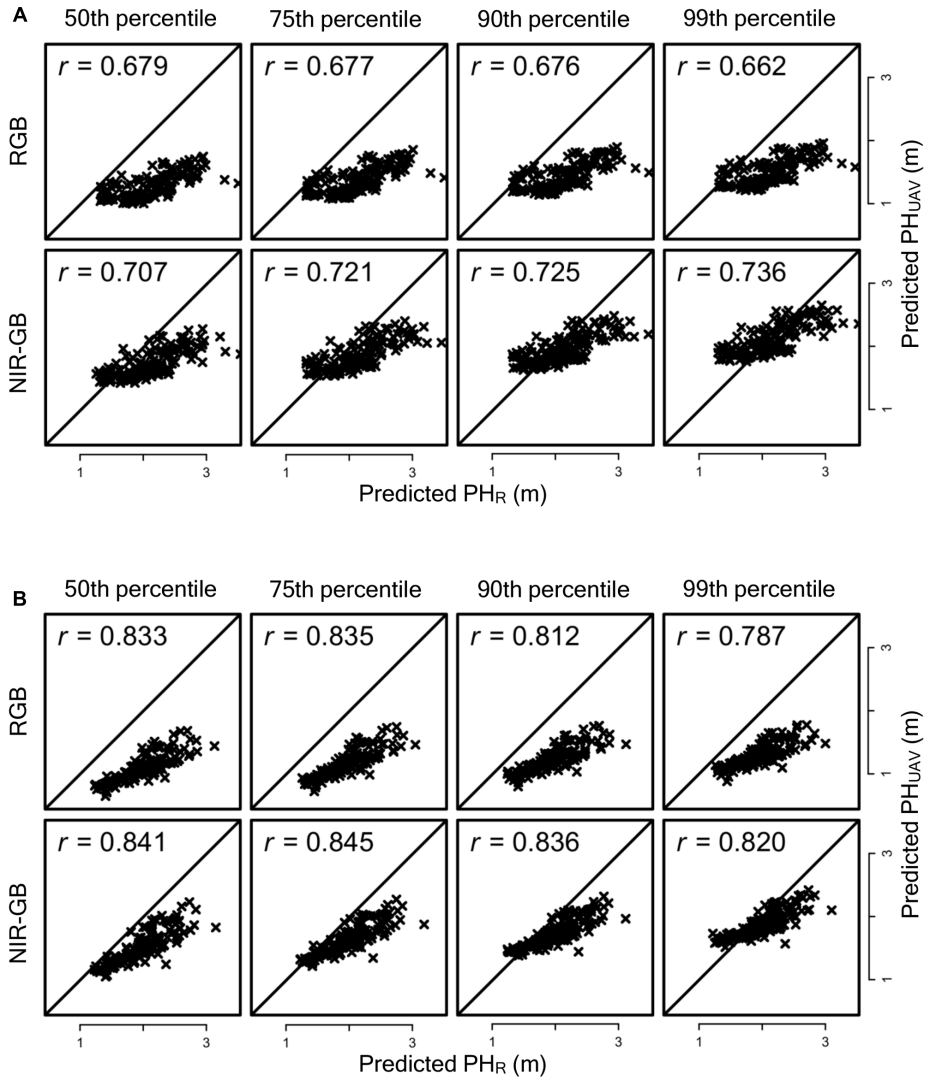


FIGURE 7 | Relationships between predicted PH_R and predicted PH_{UAV} under normal fertilization and low fertilization condition. (A) Normal fertilization and **(B)** Low fertilization. *r*, correlation coefficient.

DSM in plant height measurement. Increasing plot size and plant density may improve the measurement accuracy. We can easily measure plant height multiple times with UAV remote sensing, which would probably also reduce measurement error. An important point is that the predicted PH_{UAV} values were highly correlated with the predicted values of PH_R, even though the correlation between the observed values of PH_{UAV} and PH_R was not high. This is because the observed values of PH_{UAV} and PH_R had different types of errors (e.g., errors from manual measurements and UAV measurements). Higher correlation between the predicted values suggests that they were less affected by errors than observed values. If we can reduce the measurement errors of UAV remote sensing by improving technologies or by repeating measurements, UAV remote sensing will perform better than manual measurements for genomics-assisted breeding.

We also found that using the lower value of PH_{UAV} of the two replications for each association seemed to reduce the overlapping effect (**Figure 5**). For the replications with the lower value PH_{UAV} values, the correlation coefficients of PH_{UAV} and PH_R were not largely different for the RGB and NIR-GB cameras. The RGB camera underestimated PH_{UAV} and the RMSD value of this camera was higher than that of the NIR-GB camera. For genomic prediction modeling, the correlation coefficients between observed and predicted values were higher for NIR-GB than RGB (**Figures 6**, **7**), indicating that NIR-GB was slightly superior for this purpose. Only the combination of NIR-GB and 50th percentile under low fertilization resulted in a correlation coefficient between observed and predicted values for PH_{UAV} close to that for PH_R (**Figure 6**). Prediction of plant height was less accurate under normal than under low fertilization. This may suggest that the overlapping effect was stronger when plants

were larger. NIR-GB was better than RGB under both fertilization conditions. Like in crop identification (Zhang et al., 2016) and in monitoring of plant condition (Nijland et al., 2014), a NIR sensor may perform better than a standard RGB sensor for remote sensing in plant breeding.

In the case of the NIR-GB camera, the correlation coefficient was the highest at the 50th percentile of DSM values, whereas the RMSD value was the lowest at the 99th percentile. At the 99th percentile, PH_{UAV} was almost the same as PH_R. RMSD became higher, but the number of outliers also increased, which reduced the correlation coefficient. At the 50th percentile, the number of outliers was low in PH_{UAV} measurements and the correlation coefficient became higher, but the difference between the PH_{UAV} and PH_R values increased and RMSD became lower. There is probably a trade-off between the correlation coefficient and RMSD. If accurate plant height measurements are required, PH_{UAV} at the 99th percentile of DSM will perform well. In this study, PH_{UAV} at the 50th percentile of DSM was better than at other percentile values regarding the accuracy of both plant height measurements and genomic prediction modelings.

The collection of PH_{UAV} data required 3 people × 10 min, while collection of PH_R data required over 10 people × 2 days. Because GS and GWAS require phenotyping of a large number of accessions or plants, UAV remote sensing will be an important and indispensable tool for high-throughput genomic-assisted plant breeding.

Using digital cameras, we can measure canopy cover from images taken right above the plants (Purcell, 2000) and relate it to plant density, early vigor, leaf size, and radiation interception (Liebisch et al., 2015). Using NIR sensors, we can measure NDVI of the canopy and relate it to canopy biomass and nitrogen status (Hansen and Schjoerring, 2003). Using thermal sensors, we can measure canopy temperature (Berni et al., 2009) and relate it to water stress (Jackson et al., 1981).

Not all of the studies used UAV remote sensing. However, by attaching appropriate sensors to an UAV, we can obtain various types of information from plants grown in the field and measure important target trait-related characteristics. Combination of machine learning and image analysis enables the evaluation of complex traits, such as flowering date (Guo et al., 2015). In the future, various kinds of plant phenotyping data will be measured in parallel and in a high-throughput manner by UAV remote sensing.

AUTHOR CONTRIBUTIONS

KW, TF, NT, and HI planned and designed the experiments. TT and NT contributed to the preparation of materials. KW, WG, and KA performed remote sensing experiments. HT performed the RAD-Seq experiment, and MK, HK-K, and KY analyzed the RAD-Seq data. KW performed other analyses, interpreted results, and drafted the manuscript with the contributions from WG, KA, and HI. NT obtained funding for the study. HI supervised the study. All authors read and approved the final manuscript.

FUNDING

The research was funded by CREST, Japan Science and Technology Agency, Japan (research subject: "Novel techniques of tailor-made breeding for energy crop improvement using high-throughput genotyping").

ACKNOWLEDGMENTS

The authors thank the technical support staff at the Institute of Sustainable Agro-ecosystem Services, The University of Tokyo for their help in field experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00421/full#supplementary-material

REFERENCES

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, A. L., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376. doi: 10.1371/journal.pone.0003376

Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfus, S., and Bareth, G. (2014). Estimation biomass of barley using crop surface models (CMSs) derived from UAV-based RGB imaging. *Remote Sens.* 6, 10395–10412. doi: 10.3390/rs61110395

Bendig, J., Bolten, A., and Bareth, G. (2013). UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability. *PFG.* 6, 551–562.

Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought response. *J. Exp. Bot.* 261, 3519–3528. doi: 10.1093/jxb/erq201

Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., and Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sens.* 47, 722–738. doi: 10.1109/TGRS.2008.2010457

Brachi, B., Morris, G. P., and Borevitz, J. O. (2011). Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biol.* 12:232. doi: 10.1186/gb-2011-12-10-232

Cabrera-Bosquet, L., Moloero, G., Stellacci, A. M., Bort, J., Nogués, S., and Araus, J. L. (2011). NDVI as a potential tool for predicting biomass, plant nitrogen content and growth in wheat genotypes subjected to different water and nitrogen conditions. *Cereal Res. Commun.* 39, 147–159. doi: 10.1556/CRC.39.2011.1.15

Chapman, S., Merz, T., Chan, A., Jackway, P., Hrabar, S., Dreccer, M. F., et al. (2014). Pheno-copter: a low-altitude, autonomous remote-sensing robotic helicopter for high-throughput field-based phenotyping. *Agronomy* 4, 279–301. doi: 10.3390/agronomy4020279

Colomina, I., and Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: a review. *ISPRS J. Photogramm. Remote Sens.* 92, 79–97. doi: 10.1016/j.isprsjprs.2014.02.013

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

Díaz-Varela, R. A., de la Rosa, R., León, L., and Zarco-Tejata, P. J. (2015). High-resolution airborne UAV imagery to assess olive tree crown parameters using 3D photo reconstruction: application in breeding trials. *Remote Sens.* 7, 4213–4232. doi: 10.3390/rs70404213

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Fahlgren, N., Gehan, M. A., and Baxter, I. (2015). Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Curr. Opin. Plant Biol.* 24, 93–99. doi: 10.1016/j.pbi.2015.02.006

Furbank, R. T., and Tester, M. (2011). Phenomics-technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Gini, R., Pagliari, D., Passoni, D., Pinto, L., Sona, G., and Dosso, P. (2013). UAV photogrammetry: block triangulation comparisons. *Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci.* Volume XL-1/W2, 157–162. doi: 10.5194/isprsarchives-XL-1-W2-157-2013

Guo, W., Fukatsu, T., and Ninomiya, S. (2015). Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images. *Plant Methods* 11:7. doi: 10.1186/s13007-015-0047-9

Hansen, P. M., and Schjoerring, J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least square regression. *Remote Sens. Environ.* 86, 542–553. doi: 10.1016/S0034-4257(03)00131-7

Huang, X., and Han, B. (2014). Natural variation and genome-wide association studies in crop plants. *Ann. Rev. Plant Biol.* 65, 531–551. doi: 10.1146/annurev-arplant-050213-035715

Inoue, Y. (1997). Remote sensing of crop and vegetative environment (Part 2). *J. Japan Remote Sens. Soc.* 17, 57–67.

Jackson, R. D., Idso, S. B., Reginato, R. J., and Pinter, P. J. Jr. (1981). Canopy temperature as a crop water stress indicator. *Water Resour. Res.* 17, 1133–1138. doi: 10.1029/WR017i004p01133

Jannink, J. L., Lorenz, A. J., and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genomics* 9, 166–177. doi: 10.1093/bfgp/elq001

Knipling, E. B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sens. Environ.* 1, 155–159. doi: 10.1016/S0034-4257(70)80021-9

Korte, A., and Farlow, A. (2013). The advantage and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29. doi: 10.1186/1746-4811-9-29

Lee, W., and Searcy, S. (1999). "Multispectral sensor for detecting nitrogen in corn plants," in *Proceedings of the 2000 ASAE Annual International Meeting. Paper No. 001010*, Milwaukee, WI.

Liebisch, F., Kirchgessner, N., Schneider, D., Walter, A., and Hund, A. (2015). Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. *Plant Methods* 11:9. doi: 10.1186/s13007-015-0048-8

Mardis, E. R. (2007). The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24, 133–141. doi: 10.1016/j.tig.2007.12.007

Merz, T., and Chapman, S. (2011). "Autonomous unmanned helicopter system for remote sensing missions in 655 unknown environments," in *Proceedings of the International Conference on Unmanned Aerial Vehicle 656 in Geomatics*, (Zurich: ISPRS).

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Nijland, W., de Jong, R., de Jong, S. M., Wulder, M. A., Bater, C. W., and Coops, N. C. (2014). Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. *Agric. For. Meteorol.* 184, 98–106. doi: 10.1016/j.agrformet.2013.09.007

Purcell, L. C. (2000). Soybean canopy coverage and light interception measurements using digital imagery. *Crop Sci.* 40, 834–837. doi: 10.2135/cropsci2000.403834x

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1974). "Monitoring vegetation systems in the Great Plains with ERTS," in *Proceedings of the Third ERTS-1 Symposium, NASA SP-351*, (Washington, DC: NASA), 309–317.

Sugiura, R., Noguchi, N., and Ishii, K. (2005). Remote-sensing technology for vegetation monitoring using an unmanned helicopter. *Biosyst. Eng.* 90, 369–379. doi: 10.1016/j.biosystemseng.2004.12.011

Torres-Sánchez, J., López-Granados, F., Serrano, N., Arquero, O., and Peña, J. M. (2015). High-throughput 3-D monitoring of agricultural-tree plantations with unmanned aerial vehicle (UAV) technology. *PLoS ONE* 10:e0130479. doi: 10.1371/journal.pone.0130479

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0

van Maarschalkerweerd, M., Bro, R., Egebo, M., and Husted, S. (2013). Diagnosing latent copper deficiency in intact barley leaves (*Hordeum vulgare*, L.) using near infrared spectroscopy. *J. Agric. Food Chem.* 61, 10901–10910. doi: 10.1021/jf402166g

Zarco-Tejada, P. J., Diaz-Varela, R., Angileri, V., and Loudjani, P. (2014). Tree height quantification using very high resolution and imagery acquired from an unmanned aerial vehicle (UAV) and 3D photo-reconstruction methods. *Eur. J. Agronomy* 55, 89–99. doi: 10.1016/j.eja.2014.01.004

Zhang, C., and Kovacs, J. M. (2012). The application of small unmanned aerial system for precision agriculture: a review. *Precis. Agric.* 13, 693–712. doi: 10.1007/s11119-012-9274-5

Zhang, J., Yang, C., Song, H., Hoffmann, W. C., Zhang, D., and Zhang, G. (2016). Evaluation of an airborne remote sensing platform consisting of two consumer-grade cameras for crop identification. *Remote Sens.* 8:257. doi: 10.3390/rs8030257

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 02 March 2017
doi: 10.3389/fpls.2017.00282

Differentiating Wheat Genotypes by Bayesian Hierarchical Nonlinear Mixed Modeling of Wheat Root Density

Anton P. Wasson[1], Grace S. Chiu[2]*, Alexander B. Zwart[3] and Timothy R. Binns[4]*

[1] Commonwealth Scientific and Industrial Research Organisation (CSIRO) Agriculture & Food, Canberra, ACT, Australia, [2] Research School of Finance, Actuarial Studies and Statistics, College of Business and Economics, Australian National University, Canberra, ACT, Australia, [3] Commonwealth Scientific and Industrial Research Organisation (CSIRO) Data61, Canberra, ACT, Australia, [4] Australian Taxation Office, Sydney, NSW, Australia

Ensuring future food security for a growing population while climate change and urban sprawl put pressure on agricultural land will require sustainable intensification of current farming practices. For the crop breeder this means producing higher crop yields with less resources due to greater environmental stresses. While easy gains in crop yield have been made mostly "above ground," little progress has been made "below ground"; and yet it is these root system traits that can improve productivity and resistance to drought stress. Wheat pre-breeders use soil coring and core-break counts to phenotype root architecture traits, with data collected on rooting density for hundreds of genotypes in small increments of depth. The measured densities are both large datasets and highly variable even within the same genotype, hence, any rigorous, comprehensive statistical analysis of such complex field data would be technically challenging. Traditionally, most attributes of the field data are therefore discarded in favor of simple numerical summary descriptors which retain much of the high variability exhibited by the raw data. This poses practical challenges: although plant scientists have established that root traits do drive resource capture in crops, traits that are more randomly (rather than genetically) determined are difficult to breed for. In this paper we develop a hierarchical nonlinear mixed modeling approach that utilizes the complete field data for wheat genotypes to fit, under the Bayesian paradigm, an "idealized" relative intensity function for the root distribution over depth. Our approach was used to determine *heritability*: how much of the variation between field samples was purely random vs. being mechanistically driven by the plant genetics? Based on the genotypic intensity functions, the overall heritability estimate was 0.62 (95% Bayesian confidence interval was 0.52 to 0.71). Despite root count profiles that were statistically very noisy, our approach led to denoised profiles which exhibited rigorously discernible phenotypic traits. Profile-specific traits could be representative of a genotype, and thus, used as a quantitative tool to associate phenotypic traits with specific genotypes. This would allow breeders to select for whole root system distributions appropriate for sustainable intensification, and inform policy for mitigating crop yield risk and food insecurity.

Keywords: generalized linear mixed models, heritability, hierarchical modeling, root architecture, wheat phenotyping

OPEN ACCESS

Edited by:
Nathaniel Newlands, Agriculture and Agri-Food Canada (AAFC), Canada

Reviewed by:
David Zamar, University of British Columbia, Canada
Yann Guédon, Agricultural Research Centre For International Development, France

****Correspondence:***
Anton P. Wasson
anton.wasson@csiro.au
Grace S. Chiu
grace.chiu@anu.edu.au

Specialty section:
This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

Received: *04 December 2016*
Accepted: *15 February 2017*
Published: *02 March 2017*

Citation:
Wasson AP, Chiu GS, Zwart AB and Binns TR (2017) Differentiating Wheat Genotypes by Bayesian Hierarchical Nonlinear Mixed Modeling of Wheat Root Density. Front. Plant Sci. 8:282. doi: 10.3389/fpls.2017.00282

1. INTRODUCTION

Meeting the food production requirements of a growing human population who are encroaching on arable land and generating a changing climate will require an intensification of agriculture, where greater yields are obtained from crops on existing farms with sustainable inputs of water and fertilizer (Gregory et al., 2013). This will involve identifying the constraints on yield in agricultural systems, many of which are to be found below ground in the root systems of crops. There are calls for a "second Green Revolution" (Lynch, 2007) focused on breeding crops with "ideotypic" (Donald, 1968) root systems (i.e., possessing desirable root system traits) that can overcome these constraints. This approach, called physiological breeding, is to be contrasted with breeding for increased yield alone, an approach which is no longer keeping pace with growing demands (Fischer and Edmeades, 2010; Richards et al., 2010; Hall and Richards, 2013).

However, identifying ideotypic root systems for crops is fraught with difficulty. Root traits which can be identified in the laboratory are often difficult to translate to the field (Watt et al., 2013) because they are devoid of the developmental context of the soil. The soil environment is complex, and has a dominant effect on root system development (Rich and Watt, 2013). Furthermore, crop physiological models—which are used to formulate strategies for plant breeding and crop yield risk mitigation, and even to develop government policy—are often inadequate in addressing the spatial heterogeneity of root systems and soil properties (Holzworth et al., 2014). It is also difficult to sample roots in soil in the field, and the data obtained are complex to interpret. Nevertheless, it is in the field where the effects of soil, climate, and agronomy are integrated with the developmental genetics of the plants growing together as a crop, and hence it is also in the field where measuring root traits, identifying crop ideotypes, and modeling root development are most valuable. Selecting for root ideotypes in the field may speed up the identification of the best germplasm for breeding programs (Wasson et al., 2012; Rich and Watt, 2013).

Therefore, integrating improved measures of root distribution/development into crop physiological models will improve farm management decision making and crop yield risk mitigation. Yet, indirect measurements of crop root systems are problematic, and most direct measurements are destructive, time-consuming and/or labor intensive (e.g., root washing, minirhizotrons) (Wasson et al., 2012). Hence the core-break method was developed as a method of rapidly observing and quantifying the presence of roots as a function of depth (Drew and Saker, 1980; van Noordwijk et al., 2001); a soil core sample is taken from the crop and broken at regular intervals (corresponding to increasing depth) and the exposed roots are counted. The counts correlate with the root length in the corresponding volume of soil. This technique has been used to phenotype root count distributions in 43 genotypes (Wasson et al., 2014) and efforts have been made to automate the root count process (Wasson et al., 2016) to reduce the labor requirements. However, root counts from the core-break method are subject to a high degree of variation between samples (van Noordwijk et al., 2001), which makes it challenging to identify genotype-specific traits from root counts or to associate genotypes with discernible properties of root count profiles.

Similar types of experimental field data may have been analyzed by statistical linear models (Faraway, 2014) under an analysis-of-variance (ANOVA) framework (Wasson et al., 2014). However, a major limitation of linear models is their assumption of Gaussian (normally distributed) response data, whereas root counts are discrete, bounded below by zero, and with a distribution whose substantial right-skewness may not be easily removed by variable transformation. Indeed, root count data are more appropriately modeled as Poisson distributed, although a phenomenon known as *overdispersion* (McCulloch, 2008), commonly encountered in count data from field experiments, must be handled with care. More specifically, the Poisson distribution is characterized by a single parameter that represents the distribution mean as well as its variance. However, in practice, the count variable of concern often has a recognizably larger variance than its mean (hence, "overdispersion"), although the overall distribution still resembles Poisson in other respects.

Therefore, linear models applied to field data thus far have focused on analyzing core-level summary metrics, such as maximum rooting depth, which, after variable transformation if necessary, can approximately behave as Gaussian (Wasson et al., 2014). However, such summary metrics by definition cannot reflect root structure over depth, discarding valuable information contained at the level of individual core segments, and consequently resulting in an undesirable loss of statistical power.

To better facilitate our scientific objective of associating genotypes with discernible properties of root count profiles, in this paper we scrutinize the many facets of the inherent variability of the root count profile produced based on a field trial (**Figure 1**) that involved twenty genotypes ($n_G = 20$), each generating four replicated soil cores ($n_C = 4$) extracted *in situ* from each of four replicated plots or blocks ($n_B = 4$). Growing in a plot, as they would in a farmer's field, the plants' root systems interact and respond to each other. Their development is driven by the exploration of cracks and pores (White and Kirkegaard, 2010), which are randomly distributed. Likewise, variation in soil chemistry and nutrients, which can be patchy and vary with depth, drives the branching of roots. In contrast, impenetrable material and compaction can inhibit growth. As each soil core only captures a comparatively small piece of variation due to the various sources, results found in adjacent replicated cores can differ substantially.

Moreover, we note that many of the profiles of the average root count depicted in Figure 6 of Wasson et al. (2014) are consistent with random observations whose mean follows a functional form that roughly resembles the density function of the gamma probability distribution. Based on this observation, in this paper we develop a statistical modeling approach that can rigorously handle the non-standard nature of our root count data. Specifically, root counts at the observed depths (denoted by t) within a core are formally related through a nonlinear parametric expression $\theta(t)$ to reflect the one-dimensional spatial nature of individual soil cores. The parametric expression (with a small number of unknown parameters) is the common

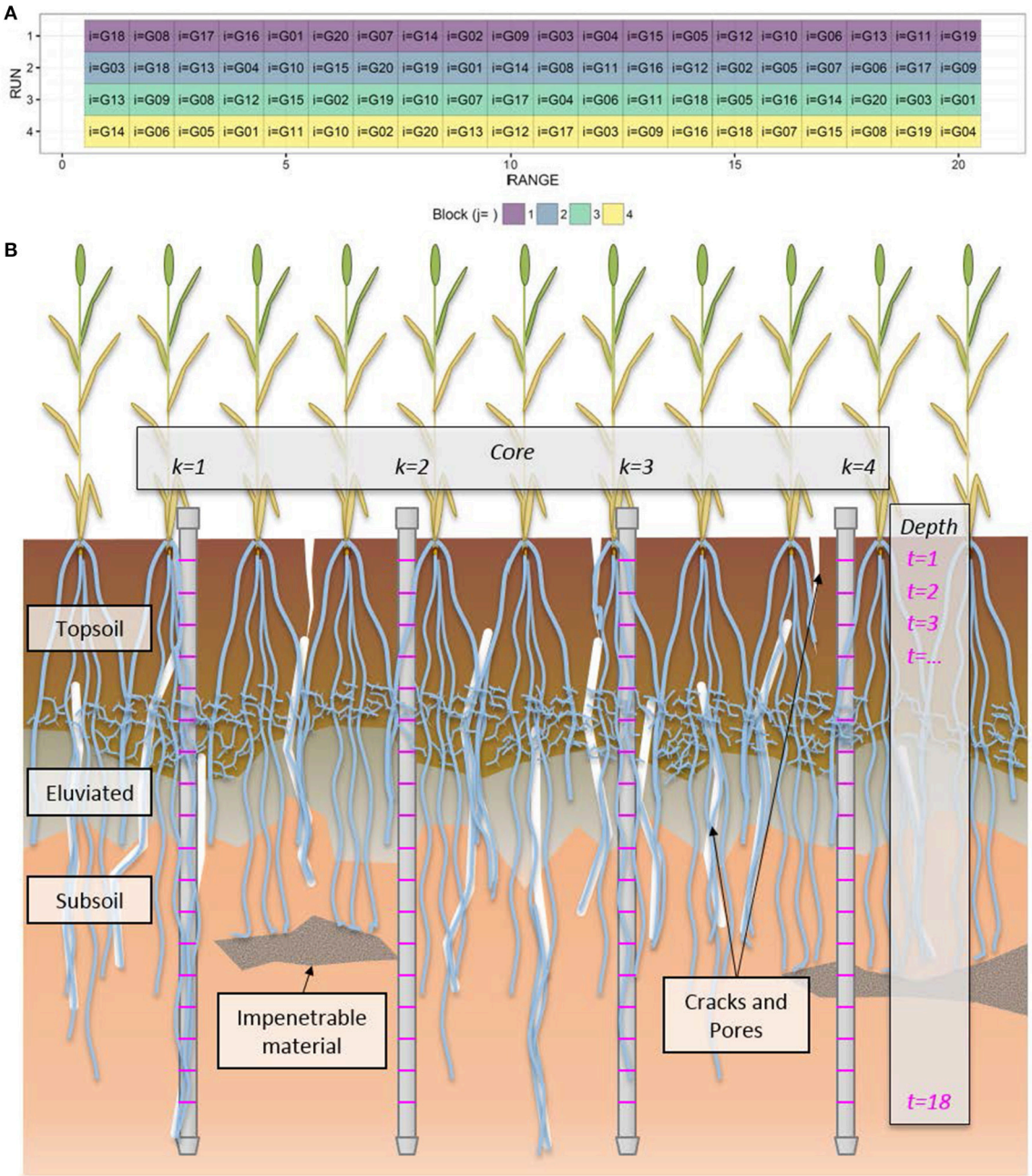


FIGURE 1 | Schematics of the field experiment. (A) Surface layout of the field experiment involving twenty genotypes (indexed by *i*) randomized within four blocks (indexed by *j*) of twenty ranges of plots. **(B)** Cartoon depicting the sampling in each plot. Four soil cores (indexed by *k*) were sampled from each plot in a steel tube. Each core was broken into 10 cm increments. The root count (*y*) at each 10 cm depth increment (indexed by *t*) is the sum of the counts on the lower face of the upper fragment and the upper face of the lower fragment. Thus, each root count, y_{ijkt}, has four unique index values. The cartoon further depicts the variability that might be encountered by sampling soil cores in a single plot, e.g., contrast cores $k = 1$ and $k = 4$.

denominator that unifies this spatial behavior among all cores. Obviously, an appropriate parametric structure imposed on the root counts within a core would lead to much greater statistical power when compared to, say, an oversimplified ANOVA approach that regards depth as a mere design feature in a factorial experiment.

We also note that our field experimental setup was such that the randomness in our data exhibits a hierarchical structure (Gelman and Hill, 2006) that comprises layers of mean and variance functions. In particular, the complex, non-standard experimental design features inherent in our data require hierarchical nonlinear mixed modeling (HNLMM), an approach

which addresses our need to model overdispersed, Poisson distributed data (McCulloch, 2008) via a hierarchy of nonlinear mean functions and associated variance components due to the formulation of $\theta(t)$.

Therefore, our approach in this paper is distinguished from existing studies particularly because of (a) our scrutiny of the root count profiles themselves, rather than the relationship between the counts and the root length density, and (b) our hierarchical modeling approach that integrates all identified facets of variability among all observed root count profiles in a comprehensive and collective manner. Additionally, our modeling framework gives rise to new heritability metrics that describe spatial and overall root architectural traits, the latter at the overall genotypic level.

The remainder of our paper is structured as follows. Under the section Materials and Methods, we provide some details of the field sampling procedure that gave rise to our root count profile data, some data visualizations, a primer on the specification of our HNLMM under a Bayesian framework (Gelman et al., 2013), and biological interpretations of model parameters and their use in defining novel multiresolution heritability measures. Statistical inference results and corresponding key biological insights appear in the Results section, followed by the section Model Validation which briefly discusses the rigor and adequacy of our approach (Technical details that supplement these sections appear in Appendices A–F and the online Supplementary Material). Our paper concludes with an in-depth Discussion section on the biological and practical implications of our integrative modeling approach in the general context of facilitating effective wheat breeding programs via root phenotyping.

2. MATERIALS AND METHODS

2.1. Data and Modeling Framework

Each soil core sampled was partitioned in the field into five-centimeter segments from which the number of roots, y, was determined every 10 cm up to 180 cm using a fluorescence imaging system (Wasson et al., 2016). Each value of y at Depth $t(= 1, ..., n_D$ where $n_D = 18)$ is the sum of the count imaged from the bottom face of the segment above t and that from the top face of the segment below t (See Appendix A for details on data collection). Let y_{ijkt} denote the total number of imaged live roots of Core k at Depth t for Genotype i in Block j. Thus, each ith genotype is associated with 288 $(= n_B n_C n_D)$ observations of y in total. Equivalently, each tth depth is associated with 320 $(= n_B n_C n_G)$ observed counts.

Data visualizations for Genotype G18 (**Figure 2**) and other genotypes (not shown) suggest that our observed root counts, y, perceivably follow a smooth nonlinear trend over core depth, but subject to substantive noise from the effects of soil physical and chemical properties described above, plus sampling and handling errors. These sources of noise culminate in the profile plots (**Figures 2A,B**) and associated boxplots (**Figure 2A**) for y. Therefore, a modeling framework comprising the following main model statements was developed to capture the complex noise structure around an idealized smooth trend:

$$\begin{aligned} y_{ijkt} &\sim \text{Poisson}(\theta_{ij}(t)), \\ \theta_{ij}(t) &= \psi_{ij} \bullet \gamma_{\alpha_i\beta_i}(t) \bullet e^{\phi_{ijt}}, \\ \psi_{ij} &= e^{\psi_0 + \kappa_j} e^{\tau_i} \end{aligned}$$

where $\theta_{ij}(t)$ denotes the underlying plot-specific Poisson *intensity curve* over depth, i.e., the modeled mean root count at Depth t $(= 1, 2, ..., 18)$ from the $\{i, j\}$th plot (for Genotype i $(= 1, 2, ..., 20)$ observed in Block j $(= 1, 2, 3, 4)$).

Intensity $\theta_{ij}(t)$ itself is decomposed into fixed and random effects (shaded nodes in **Figure 3**). Specifically, $\theta_{ij}(t)$ comprises a smooth genotype-specific "kernel function," $\gamma_{\alpha_i\beta_i}(t)$, and two sources of multiplicative Gaussian errors: genotype-specific deviation τ_i and core segment-specific deviation ϕ_{ijt}. The intensity function's proportionality multiplier ψ_{ij}, on the logarithmic scale, represents the plot-specific intercept of the $\{i, j\}$th intensity function. The intercept can be regarded as the modeled mean count (log scale) of the root system just below the soil surface. Therefore, τ_i corresponds to the genotypic random effect on this near-surface mean count. As such, ψ_{ij} itself is random. It is modeled as log-linear, where its mean can be expressed as a study-wide constant ψ_0 plus a non-random block-specific shift κ_j (both taken to be fixed effects) (see Appendix B).

2.2. The Root System's *Bulk* and *Exploration* Parameters

The idealized function $\gamma_i(t) = \gamma_{\alpha_i\beta_i}(t) = t^{\alpha_i - 1} e^{-\beta_i t}$ has two genotype-specific parameters, α_i and β_i, respectively representing the non-negative *shape* and *rate* of the gamma probability density function. Holding β_i constant and increasing α_i causes the ith kernel function to (a) peak at a lower depth and (b) exhibit more spread around the peak (**Figure 4A**). Thus, α_i corresponds to both the depth at which the root system is most dense and its tendency to explore spatially around this depth. Henceforth, we refer to α_i as the "bulk parameter."

Similarly, holding α_i constant and increasing β_i causes the ith kernel function's tail to taper off more quickly, i.e., to exhibit a more slender tail (**Figure 4B**). Thus, β_i roughly corresponds to the decline rate of the root system's downward exploration. In other words, the less slender (i.e., fatter) the kernel function's tail, the slower the decline of the root system's downward exploration (or, the bigger the tendency for the root system to explore downwards). Henceforth, we refer to β_i as the "exploration parameter."

For each ith genotype, parameters α_i and β_i are modeled as bivariate log-normal random variables (i.e., they are bivariate Gaussian on the logarithmic scale with unknown correlation ρ). These parameters and the noise terms τ_i and ϕ_{ijt} are each modeled to have a mean that is constant across the study (i.e., not indexed by i, j, k, or t), and similarly for all the (co)variance parameters in the model (See Appendix B). A visualization of the

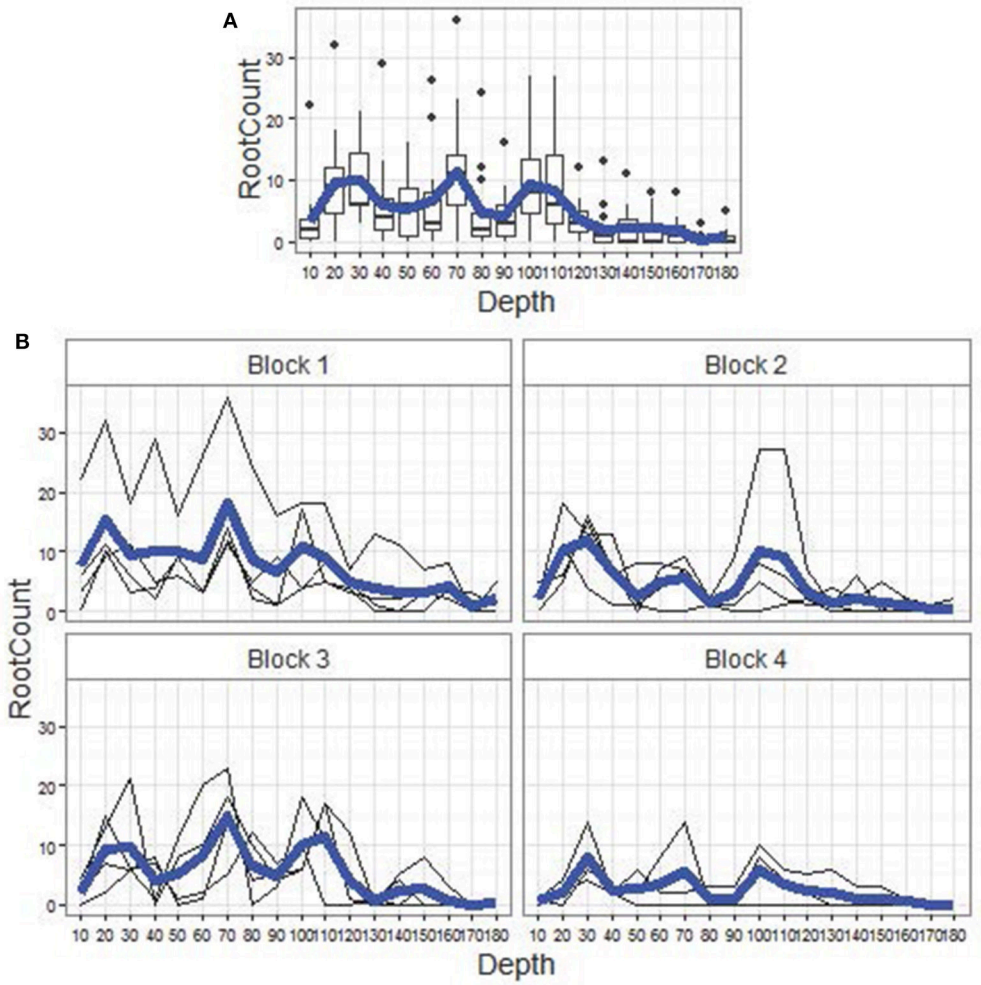


FIGURE 2 | Data visualizations. (A) Boxplots of root counts, by depth for genotype G18, pooled across replicate plots (4) and depth-specific core segments (4 per plot). The horizontal axis is depth from 10 cm to 180 cm, at 10 cm intervals. The blue line is the empirical mean root count profile over depth, which, along with the corresponding mean profiles for other genotypes, resembles those in Figure 6 of Wasson et al. (2014). **(B)** Root count profiles (in thin black) over depth, by block (replicate plot, shown as panel label), for genotype G18. Superimposed in bold blue within each block is the within-block empirical mean root count profile.

overall hierarchical structure of our modeling approach appears in **Figure 3**.

Finally, under the Bayesian inference framework, we specify reasonably non-informative prior distributions to reflect our lack of knowledge, in the absence of data, about the model parameters (see Appendix B). Collectively, the HNLMM and prior distributions as specified above are referred to as *Model 1*. Details on the implementation of *Model 1* appear in Appendix C.

2.3. Novel Heritability Measures

The general notion of heritability is the proportion of phenotypic variation that can be attributed to genetics. Loosely, we have

$$\begin{aligned} &\text{Phenotype} = \text{Genotype} + \text{Environment},\\ &\text{heritability} = \text{Var}(G)/\text{Var}(P). \end{aligned}$$

This definition of heritability assumes that genotypic and environmental variables are independent, linear components of the phenotypic response variable of interest. In practice, the biological notions of phenotype, genotype, and environment are abstract, and their quantifications that can be measured in an experiment may exhibit a complex co-dependence in a nonlinear fashion. Indeed, Moran (1973) pointed out that a quantification of heritability that purely stems from a linear decomposition of the phenotypic response can be nonsensical in practical settings.

In the case that the measurable quantities and experimental design can be reasonably described using Poisson regression, Foulley et al. (1987) adapts the linear (Gaussian) model-based definition of heritability to the scale of the linear predictor in a Poisson regression model, rather than the scale of the phenotypic response. Recently, this definition was extended to a longitudinal Poisson mixed model (Mair et al., 2015). We further extend

these ideas to define heritability measures based on segment-level count data.

Our adaptations below emphasize the challenge of detecting trends in root architecture from root count data that are both highly noisy and highly non-Gaussian, and that deviate substantially from a simple Poisson model; while considering data at a reasonably high spatial resolution may mitigate the challenge due to noise, it necessarily requires additional model complexity to address the non-standard statistical behavior of the data, and consequently, a novel quantification of heritability based on our new modeling paradigm.

The formulation of *Model 1* as presented in Appendix B gives rise to a mean number of roots that is nonlinear in its parameters, even on the logarithmic scale. Hence, this mean is not a linear predictor in the usual context of generalized linear models. Nevertheless, at each tth depth, we decompose the variability of $\log\theta(t)$ into $\sigma^2_{\log\theta}(t)$ and $\sigma^2_{\text{genes}}(t)$ both of which are spatial in nature.

Here, we must address various aspects of complexity that are non-standard in heritability studies: (1) our analog of Var(G), namely, $\sigma^2_{\text{genes}}(t)$, is attributable to the variability of the trio of genotypic parameters τ_i, α_i, and β_i; and (2) it is a spatial function. Thus, it is reasonable to further decompose this Var(G) analog into τ-, α-, and β-specific components, as each of the trio pertains to different root architectural features; and the α- and β-specific components are also functions of t and are co-dependent except in the naïve case. In Appendix D, we present the four mathematical definitions of heritability (corresponding to σ^2_{genes}, τ, α, and β) to handle such complexity.

Finally, we pool depth-specific values by taking the harmonic mean across depths, thus defining a quantity at the genotypic level that summarizes the particular architectural feature across all depths (see Appendix D). The pooling of spatial elements to form an overall heritability measure gives rise to the multiresolution nature of our approach.

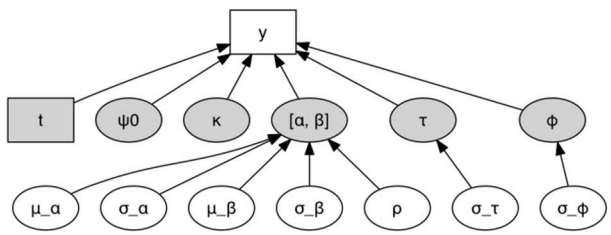


FIGURE 3 | Hierarchical structure of our modeling framework. Boxes denote data, and ovals denote model parameters (unobservable). Shaded nodes collectively determine the modeled Poisson intensity, θ.

3. RESULTS

We discuss three major biological insights that arise from the Bayesian inference, i.e., the joint *posterior distribution* among the parameters of *Model 1*.

3.1. Root Intensity Profiles Are Statistically Distinguishable among Genotypes

Posterior inference allows us to examine the intensity profiles $\theta_{ij}(t)$ and their idealized (denoised) counterparts $\psi_{ij}\gamma_i(t)$ for any

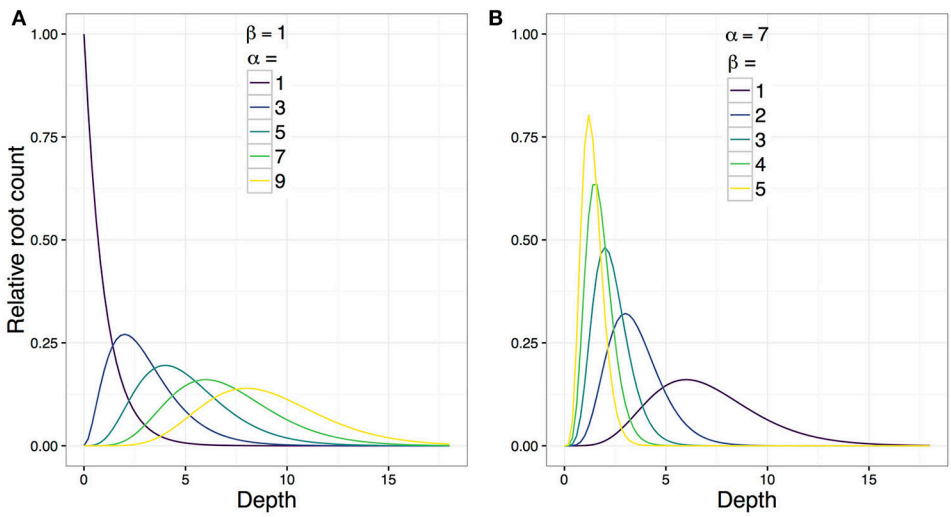


FIGURE 4 | An illustration of the effect of changing the parameters α_i and β_i on the shape of the genotype-specific kernel function ($\gamma_{\alpha_i\beta_i}(t)$) which is proportional to the probability density function of the gamma distribution. The vertical axis is the idealized relative root count (a dimensionless value). **(A)** Increasing α_i while fixing $\beta_i(= 1)$ causes the idealized function's peak to be located deeper under the soil surface and to be less concentrated. Thus, α_i is a "bulk parameter" that reflects the depth and density of the "bulk" of the root system. **(B)** The effect of increasing β_i while fixing $\alpha_i(= 7)$ causes an increased skewness in the tail of the idealized function, and consequently decreases the depth of the function's peak from the soil surface and increases its concentration. Hence, β_i is an "exploration parameter."

given replicate block j. A $\psi_{ij}\gamma_i(t)$ profile is effectively the intensity profile $\theta_{ij}(t)$ but ignoring the random genotype-block interaction ϕ_{ijt}. Note that ψ_{ij} is log-linear in τ_i and κ_j without an interaction term. Thus, the behavior among the 20 idealized profiles within any jth block is necessarily consistent across all four blocks but for an intercept shift κ_j. Hence, **Figure 5** focuses on $j = 1$ to represent the study-wide behavior of the idealized profiles.

All posterior means (Bayesian estimates) of the 20 genotypic idealized profiles $\psi_{ij}\gamma_i(t)$ are visually distinguishable (**Figure 5A**); and the study-wide statistical power is very high in determining that the genotypes do not all exhibit the same idealized profile (**Figure 5B**): 95% credible bands (Bayesian confidence bands) around the maximum and minimum idealized profiles (G12 and G18, respectively) are clearly non-overlapping. (It is analogous to rejecting the null hypothesis in a classical ANOVA at a very low significance level.) This lack of overlap at such a high credible level indicates that, among the 20 genotypes, at least G12 and G18 are highly statistically discernible with respect to their idealized intensity profiles.

While $\psi_{ij}\gamma_i(t)$ necessarily behaves similarly across all j, the genotype-block interaction intrinsic in the plot-specific intensity profile $\theta_{ij}(t)$ induces variability in the 20 profiles' collective behavior across j, as is evident in **Figure 6**: in each block, this variability reduces the statistical distinguishability among the 20 genotypes, although in each of Blocks 2, 3, and 4, at least two intensity profiles are highly discernible. Specifically, despite the noisy nature of $\theta_{ij}(t)$, **Figure 6** shows that in each of Blocks 2–4, at least two intensity profiles $\theta_{ij}(t)$ (respectively, $(i =)\{G2, G17\}$ in Block $(j =)2$, {G6, G13} in Block 3, and {G6, G15} in Block 4) are highly statistically discernible due to the general lack of overlap between the pair of block-specific 95% credible bands around $\theta_{ij}(t)$.

3.2. Root Intensity Profiles Are Substantially Heritable

Each of our four genotypic heritability measures is a model parameter that exhibits a posterior distribution, shown in black in **Figure 7**; three of these are pooled measures, each comprising 18 depth-specific components (Appendix D), shown in color.

Focusing on the genotypic level (**Figure 7** in black; **Figure 8A**), the Bayesian estimate and 95% credible interval for heritability of the intensity function are, respectively, 0.65 and (0.52, 0.75); for that of the near-surface mean count they are 0.14 and (0.03, 0.26); the "bulk" parameter, 0.62 and (0.35, 0.82); and the "exploration" parameter, 0.19 and (0.05, 0.37).

Note in **Figure 7** that (i) the depth-specific components of each of $h^2_h, h^2_{\alpha(-\beta)}$, and $h^2_{\beta(-\alpha)}$ tend to increase as depth increases, and (ii) the near-surface intensity of root count has low heritability (h^2_τ). These features of our results indicate that root count features at deeper depths are more heritable than those at shallower depths. In other words, our results provide quantitative rigor for three ideas: the heritability of root architectural traits varies substantively across depth; traits that are associated with a deeper spatial location tend to be more informative about plant genetics; and the depth at which the root system develops its bulk is negatively associated with its tendency to explore deeper. The third notion is further

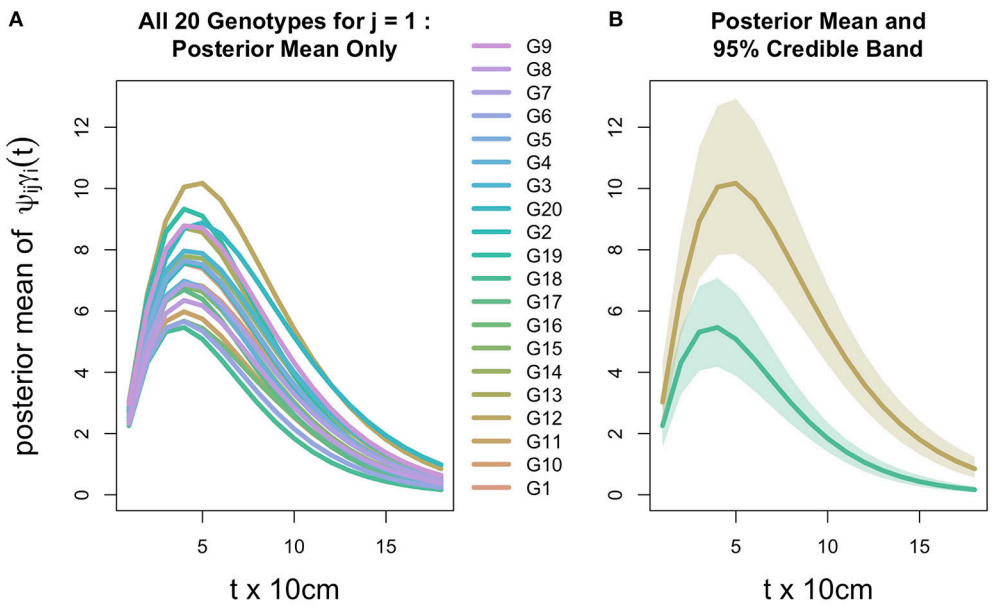


FIGURE 5 | Inference for root distribution. (A) Posterior mean (Bayesian estimate) of idealized intensity profile $\psi\,\gamma(t)$ for replicate block $j = 1$ for all 20 genotypes. Other blocks appear similarly, differing only in the intercept due to the block-specific fixed effect κ in which $\log\psi$ is linear. **(B)** Posterior means for Genotypes G12 and G18 (from panel **A**) which are respectively the maximum and minimum curves, each surrounded by a 95% credible band (Bayesian confidence band). Note that credible bands are constructed from depth-wise 95% credible intervals of $\psi_{ij}\gamma_i(t)$; thus, the lower band limit is constructed by connecting, across the 18 values of t, the 2.5th percentiles of the $\psi_{ij}\gamma_i(t)$ posterior distribution; similarly, the upper band is constructed by connecting the corresponding 97.5th posterior percentiles.

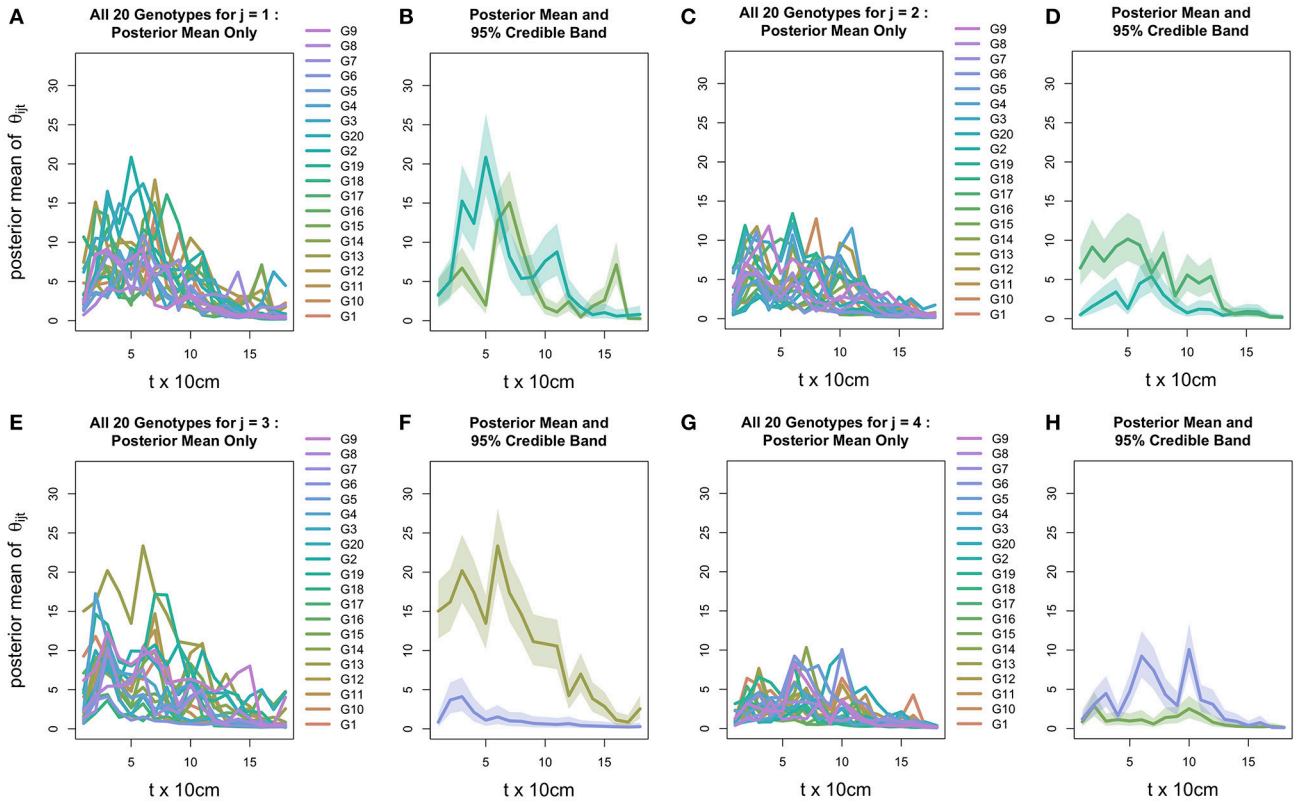


FIGURE 6 | Posterior mean of intensity profile $\theta(t)$ for all 20 genotypes, coupled with those genotypes with the maximum and minimum curves and their 95% credible intervals for $j = 1$ (A,B), $j = 2$ **(C,D)**, $j = 3$ **(E,F)**, and $j = 4$ **(G,H)**.

evidenced by another feature of *Model 1*, which is discussed in the next subsection. It is also interesting to note that, although overall heritability h_h^2 is constituted from h_h^2, $h_{\alpha(-\beta)}^2$, and $h_{\beta(-\alpha)}^2$, **Figure 8A** (which summarizes the genotypic aspects of **Figure 7**) suggests that each of the latter three tends to be less than h_h^2 itself, thus root architecture on the whole tends to be more heritable than any of these standalone features of the root system.

3.3. Linkage Exists between Near-Surface Root Density Development and Downward Exploration

The modeled correlation, ρ, between the bulk and exploration parameters (both on the log scale), is estimated to be 0.64, with 95% credible interval = (−0.85, 0.90) (see **Figure 8B**). Due to skewness, the posterior probability for ρ to be positive is 0.88, substantiating that the root system's bulk and downward exploration are generally positively associated architectural features. Specifically, a shallower and more concentrated bulk (small α) is associated with a larger tendency for the root system to explore deeper (small β). This phenomenon may be regarded as "a small β canceling out a large α," or, the tendency of exploring downwards to exhibit the effect of negating the tendency to develop root density further away from the surface. We elaborate on this discovery under Discussion.

4. MODEL VALIDATION

Details of model validation procedures appear in Appendix E. In summary, residual analyses suggest only minor statistical inadequacies of *Model 1*. With respective to the model's predictive performance as measured by the Watanabe-Akaike information criterion (WAIC) (Gelman et al., 2013), its hierarchical structure is essential. Specifically, ignoring the hierarchical structure between the root count intensity function and its various random components that are specific to genotypes, plots, and depths leads to a naïve model that agrees poorly with the empirical behavior of our root count data. Employing the hierarchical structure, the model's predictive performance remains effectively unaffected whether *a priori* dependence between the bulk and exploration parameters is considered; however, we regard this extra dependency as a key biological feature because (a) it improves the interpretability of the model by providing an explicit assessment of the interplay between the bulk and exploration parameters, and (b) this interplay is shown to be substantive based on our field data (as indicated by a smaller *effective number of parameters* for *Model 1* despite its extra complexity due to the *a priori* dependency).

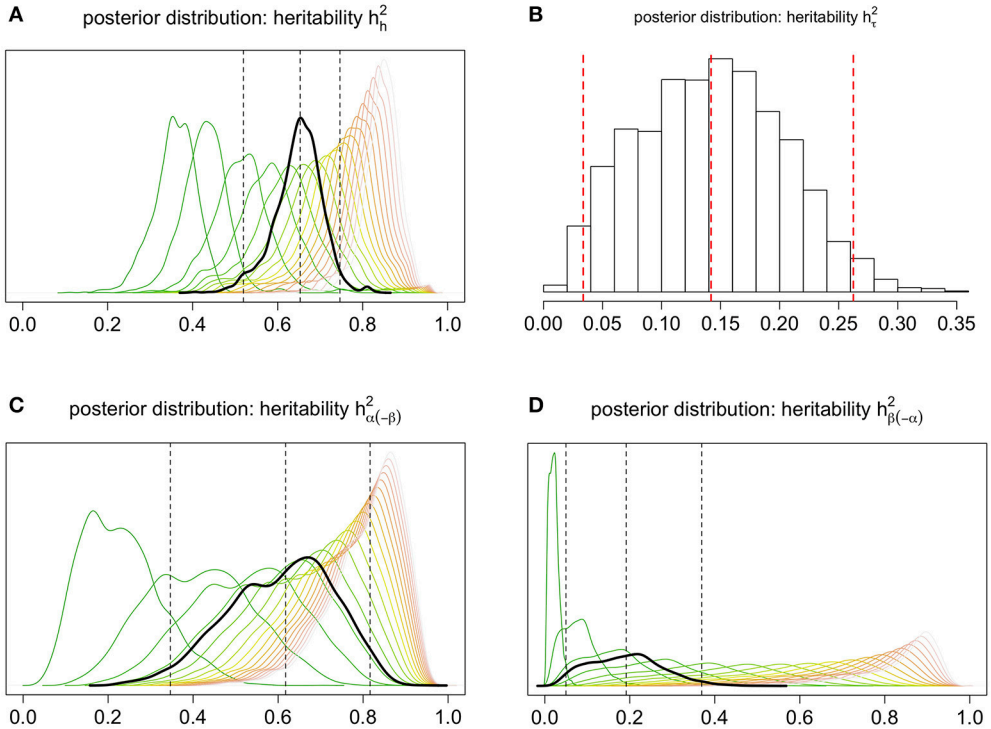


FIGURE 7 | Posterior distributions of pooled measures of heritability (black), pertaining to (A) overall root architecture, **(B)** the near-surface intensity, **(C)** the root bulk's location (and size), and **(D)** the root system's decline of penetration; the middle vertical line marks the posterior median (a Bayesian estimate), and the outer lines delimit the 95% credible interval. In **(A,C,D)**, pooling corresponds to integrating depth-dependent heritability over all 18 depths via the harmonic mean of the 18 depth-specific heritability values; the posterior distribution of the unpooled heritability at a given depth is shown in shades of "burnt grass," where more burnt corresponds to greater depth.

5. DISCUSSION

The development of roots in response to the extreme heterogeneity of the soil results in a lack of discernible root system characteristics that can be measured in the field and integrated into crop breeding programs. The inability to breed on the root development of crops and the inadequacy of conventional crop physiological models in addressing the spatial heterogeneity of root systems and soil properties are barriers to the sustainable intensification of agriculture, where root traits are known to be critical to resource-use efficiency and resistance to climatic extremes. Consequently, they are barriers to effective crop yield risk mitigation and food security. Crop physiological models with better predictive ability are much sought after, and novel statistical models can facilitate this pursuit by effectively teasing apart root system physiology from severe heterogeneity. In this paper, we have addressed this knowledge gap by scrutinizing root counts observed using a core-break count method, and by developing a novel modeling approach that accounts for all root count data holistically. Our approach gave rise to new multiresolution heritability metrics, each describing a specific feature of the root count distribution spatially and at the overall genotypic level, which we showed to be substantially heritable. Our integrative approach can allow selective pre-breeding programs for root distribution and may facilitate the identification of genetic markers from field data.

The holistic nature of our approach is an inherent advantage of hierarchical modeling. For model inference, we employed the Bayesian paradigm, which is intrinsically hierarchical in structure. It also has the potential of being greatly flexible: as long as the model is mathematically sound and sufficient computational resources and algorithms are used to implement the model, rigorous statistical inference can be straightforward even for a model with highly complex nonlinear parameters and random quantities that follow non-standard probability distributions. In contrast, classical statistical inference can be too impractical when models or data structures deviate from well-studied scenarios. In our case, the experimental setup and the notion of root architecture together led to a highly non-standard scenario that, under a classical paradigm, would have been much less straightforward to model and subsequently draw inference from. Not only was y (the response variable of interest) strictly non-Gaussian, the data were also three-dimensionally spatial in nature, where replicate plots were arranged in a certain 2-dimensional structure (indexed by $\{i, j\}$), and in turn each plot generating numerous 1-dimensional spatial observations of y (indexed by t).

Irrespective of the inference paradigm, a caveat of hierarchical modeling is that model complexity in the

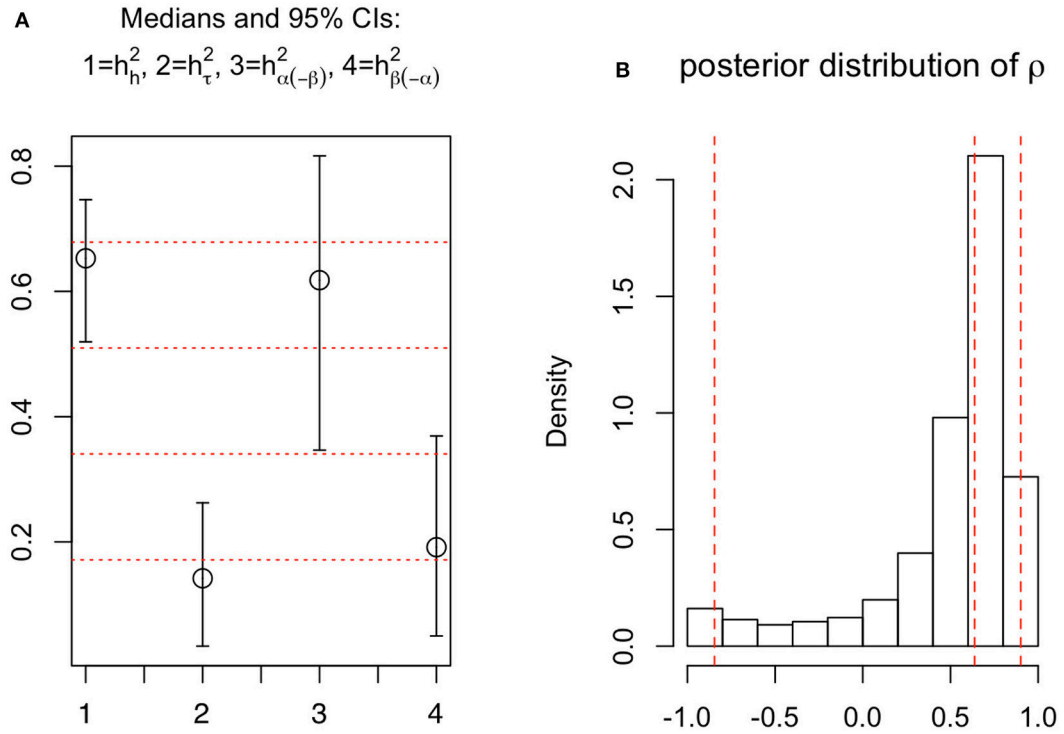


FIGURE 8 | (A) 95% credible intervals and posterior medians for the four heritability measures (1 = root architecture; 2 = near-surface mean count, log scale; 3 = bulk parameter; 4 = exploration parameter). **(B)** Posterior distribution of ρ, with vertical lines indicating the 2.5, 50, and 97.5% quantiles.

form of highly nonlinear functional forms and/or intricate hierarchical dependence structures can render the inference so computationally challenging that determining the posterior distribution (for Bayesian inference) or the sampling distributions of estimators and test statistics (for classical inference) would require novel numerical algorithms that are yet to be developed. However, for Bayesian inference in our case, model implementation and model diagnosis/validation were reasonably straightforward to conduct. The satisfactory predictive performance of our *Model 1* (Appendix E) suggests that *Model 1* is scientifically sensible and has yielded biological insights that are superior to what could have been drawn from previous linear models applied to core-level metrics (from collapsing segment-level data).

Although *Model 1* does not account for potential within-core spatial dependency among segment-level root count data (see Appendix F), the biological implications of this model nonetheless will help to define root traits for breeding. The canonical model of root distribution with depth is that of a negative exponential function (Gerwitz and Page, 1974). Gale and Grigal (1987) describe a nonlinear function $Y = 1 - \beta^d$ where Y is the cumulative root fraction from the surface to the depth (d cm), and the coefficient β is genotypically determined. This model was later employed by Jackson et al. (1996) to model root distributions across a range of terrestrial biomes.

However, this 1-dimensional model takes no account of the horizontal distance from the base of the plant. It has been observed that root distribution is 1-dimensional with depth in grassland, 2-dimensional in crops planted densely in rows, and 3-dimensional where plants are widely spaced (Bengough et al., 2000). The simulation studies by Grabarnik et al. (1998) showed that root length density—typically the length of root per volume of soil (cm/cm^3)—for maize decreased nonlinearly with horizontal distance from the stem in the top 40 cm, but below that depth they were homogeneously distributed with horizontal distribution from the plant. Grabarnik et al. (1998) also showed that the roots were subject to clustering at all depths, and that whilst there was no preferential growth in a horizontal plane, the orientation of root growth deviated from the uniform distribution with increasing depth. Similar findings were generated in the simulation study by Bengough et al. (1992), and both studies drew attention to the likely effect of soil structure to further perturb the uniform directional distribution of root development parameters.

The similarity between the model by Gale and Grigal (1987) and the special case of our gamma kernel function where $\alpha = 1$ should be noted (**Figure 4A**). Rather than root length density, our model accounts for root counts that are random with respect to sampling position by row in a crop. The model is designed to explain the distribution of root counts with depth at the crop level (and not the plant level). However, the sampling position is

likely to have a strong influence on the surface root counts, which explains the low heritability of τ in our model.

Interpreting the biological meaning of the "bulk" and "exploration" parameters (α and β, respectively) is also interesting. In the gamma kernel function, β also affects the depth and intensity of the peak otherwise defined by α (**Figure 4**). Indeed, our data analysis implicated that α and β were positively correlated (**Figure 8B**). For our HNLMMs, predictive performance remained largely unaffected whether *a priori* dependence between α and β is considered; however, including this extra dependency improved the interpretability of the model by providing an explicit assessment of the interplay between the root system's tendencies to branch beneath the surface and to explore vertically, deep below the surface.

An explanation for this effect may be found in the structure of the soil; root growth in deeper layers is perceivably constrained to networks of cracks and pores (Gao et al., 2016). White and Kirkegaard (2010) show that in a dense, structured subsoil 85–100% of roots below 60 cm were clumped in pores and cracks in the soil (compared to 30–40% above 60 cm), and 44% of the roots were clumped in pores with more than three other roots. Exploration of the soil for cracks and pores may define the exploitation of the soil by a root system. It has been suggested that plants have evolved randomness and instability in their root system development (Forde, 2009), which may facilitate exploration. The exploration of the shallow layers for cracks and pores may be what determines the eventual depth; our model implies that more branching near the surface gives better access to the subsoil.

The primary purpose of our modeling approach was to distinguish genotypes from root count data that are statistically noisy. The inference for heritability based on the intensity functions suggests that our approach can be used to identify genetic markers of root system distribution in field data; identified markers then could be integrated into breeding programs. The high heritability of the "bulk" parameter also suggests that a breeding program could successfully alter the depth at which a root system proliferates.

Notwithstanding, residual plots (Appendix E and Supplementary Material: Supplementary Figures) suggest some minor statistical inadequacies of *Model 1*. Therefore, it may be advantageous to (1) explicitly model gene-environment interactions (which are implicitly modeled by our current HNLMM due to the marginal dependence among genotypic terms indexed by i and environmental terms indexed by j and/or t); (2) formally model the within-core spatial dependence (possibly at a higher spatial resolution of core depths than the current 10 cm intervals); and (3) also incorporate an additional two-dimensional spatial correlation structure among field plots. In Appendix F, we suggest a possible decomposition at Level 1 of the model hierarchy to address (1), and discuss practical implications of modeling the 3-dimensional spatial dependence to address (2) and (3).

Finally, it may also be of benefit to develop a new quantitative framework to predict root length density from the posterior mean root count profiles while accounting for trials in different soil and climate conditions, under which the response of the intensity functions and their underlying parameters to subsoil constraints could be rigorously exploited.

AUTHOR CONTRIBUTIONS

AW and GC designed research; AW, GC, AZ, and TB performed research; AW and GC contributed new data processing and analytic tools; GC, AZ, and TB analyzed data; and AW, GC, and AZ wrote the paper.

FUNDING

Bayer CropScience (a) funded CSIRO for this research project, part of which was subcontracted to the Australian National University to sponsor GC's involvement after June 2015; and (b) sponsored the CSIRO Agriculture Vacation Scholarship 2014 that was held by TB, under which he contributed to the basic model that was later extended to the approach in this paper.

ACKNOWLEDGMENTS

We thank the Associate Editor, both reviewers, and Dr. A. H. Westveld (ANU) for their valuable comments and suggestions; and Prof. M. Watt (Institute of Bio- and Geosciences 2, Forschungszentrum Jülich) for support and guidance. GC thanks the University of Washington and University of Waterloo for online resources made available to her as an affiliate/adjunct faculty member.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00282/full#supplementary-material

REFERENCES

Adler, D. (2005). *vioplot: Violin plot*. Available online at: http://cran.r-project.org/package=vioplot

Bengough, A. G., Castrignano, A., Pagès, L., and van Noordwijk, M. (2000). "Sampling strategies, scaling, and statistics," in *Root Methods*, eds D. A. L. Smit, D. A. G. Bengough, P. D. C. Engels, D. M. V. Noordwijk, D. S. Pellerin, and D. S. C. V. D. Geijn (Berlin; Heidelberg: Springer), 147–173.

Bengough, A. G., Mackenzie, C. J., and Diggle, A. J. (1992). Relations between root length densities and root intersections with horizontal and vertical planes using root growth modelling in 3-dimensions. *Plant Soil* 145, 245–252. doi: 10.1007/BF00010353

Chiu, G. S., and Lockhart, R. A. (2010). Bent-cable regression with autoregressive noise. *Can. J. Stat.* 38, 386-407. doi: 10.1002/cjs.10070

Donald, C. M. (1968). The breeding of crop ideotypes. *Euphytica* 17, 385–403. doi: 10.1007/BF00056241

Drew, M. C., and Saker, L. R. (1980). Assessment of a rapid method, using soil cores, for estimating the amount and distribution of crop roots in the field. *Plant Soil* 55, 297–305. doi: 10.1007/BF02181809

Faraway, J. J. (2014). *Linear Models with R, 2nd Edn.* Boca Raton, FL: CRC Press.

Fischer, R. A., and Edmeades, G. O. (2010). Breeding and cereal yield progress. *Crop Sci.* 50(Suppl. 1), S85–S98. doi: 10.2135/cropsci2009.10.0564

Forde, B. G. (2009). Is it good noise? The role of developmental instability in the shaping of a root system. *J. Exp. Bot.* 60, 3989–4002. doi: 10.1093/jxb/erp265

Foulley, J. L., Gianola, D., and Im, S. (1987). Genetic evaluation of traits distributed as Poisson-binomial with reference to reproductive characters. *Theor. Appl. Genet.* 73, 870–877. doi: 10.1007/BF00289392

Gale, M., and Grigal, D. (1987). Vertical root distributions of northern tree species in relation to successional status. *Can. J. For. Res.* 17, 829–834. doi: 10.1139/x87-131

Gao, W., Hodgkinson, L., Jin, K., Watts, C. M., Ashton, R. M., Shen, J., et al. (2016). Deep roots and soil structure. *Plant Cell Environ.* 39, 1662–1668. doi: 10.1111/pce.12684

Gelfand, A., Diggle, P., Fuentes, M., and Guttorp, P. (eds.). (2010). *Handbook of Spatial Statistics*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. CRC Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis, 3rd Edn.* Boca Raton, FL: CRC Press.

Gelman, A., and Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

Gerwitz, A., and Page, E. R. (1974). An empirical mathematical model to describe plant root systems. *J. Appl. Ecol.* 11, 773–781. doi: 10.2307/2402227

Grabarnik, P., Pagès, L., and Bengough, A. G. (1998). Geometrical properties of simulated maize root systems: consequences for length density and intersection density. *Plant Soil* 200, 157–167. doi: 10.1023/A:1004382531671

Gregory, P. J., Atkinson, C. J., Bengough, A. G., Else, M. A., Fernández-Fernández, F., Harrison, R. J., et al. (2013). Contributions of roots and rootstocks to sustainable, intensified crop production. *J. Exp. Bot.* 64, 1209–1222. doi: 10.1093/jxb/ers385

Hall, A. J., and Richards, R. A. (2013). Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. *Field Crops Res.* 143, 18–33. doi: 10.1016/j.fcr.2012.05.014

Holzworth, D. P., Huth, N. I., DeVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., et al. (2014). APSIM – Evolution towards a new generation of agricultural systems simulation. *Environ. Model. Soft.* 62, 327–350. doi: 10.1016/j.envsoft.2014.07.009

Jackson, R., Canadell, J., and Ehleringer, J. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia* 108, 389–411. doi: 10.1007/BF00333714

Lynch, J. P. (2007). Roots of the second green revolution. *Aust. J. Bot.* 55:493. doi: 10.1071/BT06118

Mair, C., Stear, M., Johnson, P., Denwood, M., Jimenez de Cisneros, J. P., Stefan, T., et al. (2015). A Bayesian generalized random regression model for estimating heritability using overdispersed count data. *Genet. Select. Evol.* 47:51. doi: 10.1186/s12711-015-0125-5

McCulloch, C. E. (2008). *Generalized, Linear, and Mixed Models, 2nd Edn*. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley.

Moran, P. A. (1973). A note on heritability and the correlation between relatives. *Ann. Hum. Genet.* 37:217. doi: 10.1111/j.1469-1809.1973.tb01828.x

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rich, S. M., and Watt, M. (2013). Soil conditions and cereal root system architecture: review and considerations for linking Darwin and Weaver. *J. Exp. Bot.* 64, 1193–1208. doi: 10.1093/jxb/ert043

Richards, R., Rebetzke, G., Watt, M., Condon, A., Spielmeyer, W., and Dolferus, R. (2010). Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Funct. Plant Biol.* 37, 85–97. doi: 10.1071/FP09219

Stan Development Team (2016). *RStan: The R Interface to Stan*. Available online at: http://mc-stan.org/citations/

van Noordwijk, M., Brouwer, G., Meijboom, F., Oliveira, M. D. R. G., and Bengough, A. G. (2001). "Trench profile techniques and core break methods," in *Root Methods*, eds D. A. L. Smit, D. A. G. Bengough, P. D. C. Engels, D. M. V. Noordwijk, D. S. Pellerin, and D. S. C. V. D. Geijn (Berlin; Heidelberg: Springer), 211–233.

Vehtari, A., and Gelman, A. (2014). *WAIC and Cross-Validation in Stan*. Available online at: http://www.stat.columbia.edu/~gelman/research/unpublished/waicstan.pdf

Wasson, A., Bischof, L., Zwart, A., and Watt, M. (2016). A portable fluorescence spectroscopy imaging system for automated root phenotyping in soil cores in the field. *J. Exp. Bot.* 67, 1033–1043. doi: 10.1093/jxb/erv570

Wasson, A. P., Rebetzke, G. J., Kirkegaard, J. A., Christopher, J., Richards, R. A., and Watt, M. (2014). Soil coring at multiple field environments can directly quantify variation in deep root traits to select wheat genotypes for breeding. *J. Exp. Bot.* 65, 6231–6249. doi: 10.1093/jxb/eru250

Wasson, A. P., Richards, R. A., Chatrath, R., Misra, S. C., Prasad, S. V., Rebetzke, G. J., et al. (2012). Traits and selection strategies to improve root systems and water uptake in water-limited wheat crops. *J. Exp. Bot.* 63, 3485–3498. doi: 10.1093/jxb/ers111

Watt, M., Moosavi, S., Cunningham, S. C., Kirkegaard, J. A., Rebetzke, G. J., and Richards, R. A. (2013). A rapid, controlled-environment seedling root screen for wheat correlates well with rooting depths at vegetative, but not reproductive, stages at two field sites. *Ann. Bot.* 112, 447–455. doi: 10.1093/aob/mct122

Weaver, J. E. (1926). *Root Development of Field Crops*. New York, NY: McGraw-Hill.

White, R. G., and Kirkegaard, J. A. (2010). The distribution and abundance of wheat roots in a dense, structured subsoil implications for water uptake. *Plant Cell Environ.* 33, 133–148. doi: 10.1111/j.1365-3040.2009.02059.x

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.

Wickham, H., and Francois, R. (2015). *dplyr: A Grammar of Data Manipulation*. Available online at: http://cran.r-project.org/package=dplyr

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX A

Data Collection

The field trial was conducted at Ginninderra Experiment Station in Canberra, Australia (35°12′29.0″S 149°04′59.0″E), from late May to late December 2013 (the typical wheat growing season for the region) in alluvial cracking clay plots that were 1.3 m long. Twenty spring wheat genotypes (anonymized in this paper) were drawn from a collection of standard cultivars and from a multigenic mapping population on the basis of prior experimentation on root distributions in the field; each was sown with a tractor-drawn plot seeder in ten rows spaced 18 cm apart in a randomized block design with four replicated blocks of plots (**Figure 1A**). A seed was sown roughly 3 cm apart in each row; the final sowing density was ∼150 plants/m^2. A fertilizer (N:P:K:S = 14:12.7:0:11) was applied at 120 kg/ha at sowing, with urea added for additional N during the growing season. Prophylactic fungicide and herbicide treatments were applied to the trial to suppress weeds and prevent disease. In early January 2014 after the trial had matured and been harvested, four soil cores of ∼1.8 m in length were collected from each plot using 2 m long, 42 mm diameter stainless steel coring tubes driven into the soil vertically with a tractor-mounted hydraulic push press (Wasson et al., 2014). Our field sampling technique ensured that within each plot the cores were reasonably independent of each other (**Figure 1B**). Each core was *broken* into segments rather than *sliced*, so that the roots traversing the plane of the break would emerge intact from one of the two broken faces; the same root could not be visibly intact on both faces simultaneously. (Slicing the roots would have left only the cross sectional area on the face: 50–150 microns in diameter and difficult to detect.) Hence, the root counts on the adjoining faces can be regarded as independent values which, when combined to form y, represent the number of roots traversing the break plane at that depth. The fluorescence imaging system generates root counts (Wasson et al., 2016) which necessarily differ from an observer's manual counts, although both are subject to measurement error. The raw imaging data were processed (available from Supplementary Material: Dataset) and visualizations produced with the statistical programming language R (R Core Team, 2015) using the packages "dplyr" (Wickham and Francois, 2015) and "ggplot2" (Wickham, 2009).

APPENDIX B

Formulation of *Model 1* and Variants

Recall that y_{ijkt} denotes an observed root count, where $\{i, j, k, t\}$ indexes {genotype, block, core, depth}, for $i \in \{1, 2, ..., 20\}, j \in \{1, 2, 3, 4\}, k \in \{1, 2, 3, 4\}$, and $t \in \{1, 2, ..., 18\}$. Taking $\gamma_i(t) = t^{\alpha_i - 1} e^{-\beta_i t}$ which is the kernel of the gamma probability density function and letting "N" and "BVN," respectively denote "normal" and "bivariate normal," our model statements can be rewritten as follows:

$$\text{Level 1}: \begin{cases} \left[y_{ijkt} \,\middle|\, \theta_{ijt}\right] \sim \text{Poisson}(\text{mean} = \theta_{ijt}), \\ \log \theta_{ijt} = \psi_0 + \kappa_j + (\alpha_i - 1)\log t - \beta_i t + (\tau_i + \phi_{ijt}); \end{cases}$$

$$\text{Level 2}: \begin{cases} [\log \alpha_i, \log \beta_i]' \mid \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim \text{BVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \\ \left[\phi_{ijt} \,\middle|\, \sigma_\phi^2\right] \sim N(\text{mean} = 0, \text{var} = \sigma_\phi^2), \quad \left[\tau_i \,\middle|\, \sigma_\tau^2\right] \\ \qquad\qquad\qquad\qquad\qquad\qquad \sim N(0, \sigma_\tau^2) \end{cases}$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_\alpha \\ \mu_\beta \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_\alpha^2 & \rho\sigma_\alpha\sigma_\beta \\ \rho\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{bmatrix};$$

and κ_js are fixed effects that require a linear constraint to ensure model identifiability: we take $\kappa_4 = 0$. (See **Figure 3** for a schematic of the two levels of our HNLMM.)

For Bayesian inference, prior distributions are required for all fixed-effects parameters κ_js, $\psi_0, \mu_\alpha, \mu_\beta$, and (co)variance parameters $\rho, \sigma_\alpha^2, \sigma_\beta^2, \sigma_\tau^2$, and σ_ϕ^2. To reflect our lack of *a priori* insight (in the absence of data) into the likely values of these parameters, each was given a standard diffuse prior: the Fisher-transformation $arctanh(\rho)$ and fixed effects were all assumed to be independent zero-mean Gaussians with a variance of 10^4, and the variance parameters were assumed to follow independent and identical inverse-gamma distributions with values 1 and 0.1, respectively, for the shape and rate parameters. The resulting Markov chains of posterior draws exhibited very poor mixing for *Model 1* (as well as *Model 2*, obtained by prespecifying $\rho = 0$) when a smaller rate parameter value, namely, 0.01 or 10^{-4}, was used. As smaller rate parameter values correspond to more diffuse inverse-gamma priors, the poor mixing suggests mildly weak identifiability (even for the smaller *Model 2*). This also suggests that to improve inferential power for and the identifiability of *Model 1*, one could conduct a future field study that consists of a larger number of plots and/or depths, and/or employ stronger priors based on the inference we have presented in this current article.

For model validation (Appendix E), we also considered smaller models: *Model 2* by prespecifying $\rho = 0$ in *Model 1*; and *Model 3* by taking all *Model 2* parameters (including those that are genotype-specific) to be fixed effects.

Remarks. Note that alternative parametrizations of depth are possible. Under Supplementary Material: Model Parametrization, we discuss the so-called *canonical scale* for depth, on which the statistical inference is invariant to certain reparametrizations including the conventional t scale (i.e. $1, ..., n_D$) in this paper. This invariance is similar to that of Chiu and Lockhart (2010), where the rigor of the statistical inference is developed on the canonical scale rather than the conventional scale.

APPENDIX C

Implementation of *Model 1* and Variants

Bayesian inference requires the derivation of the joint posterior distribution of all model parameters. In our case, this distribution is intractable and Markov chain Monte Carlo (MCMC) was used to approximate it. For this, we used the RStan MCMC software (R Core Team, 2015; Stan Development Team, 2016) to fit *Model 1* and its variants for model refinement purposes (see Supplementary Material: Computer Code). Of the 320 cores

sampled, nine were ignored as they failed to yield root count data. While missing data can be imputed under extra model assumptions, the Stan framework did not yet readily allow simulation of discrete parameters, and thus imputation of the missing root counts was not performed.

APPENDIX D

Multiresolution Heritability under *Model 1*

To handle the non-linearity—even on the logarithmic scale—of the mean number of roots θ in *Model 1*, we decompose the variability of $\log\theta(t)$ into

$$\sigma^2_{\log\theta(t)} = \sigma^2_{\text{genes}}(t) + \sigma^2_\phi$$

where

$$\sigma^2_{\text{genes}}(t) = \sigma^2_\tau + (\log t)^2(e^{\sigma^2_\alpha}-1)e^{2\mu_\alpha+\sigma^2_\alpha} + t^2(e^{\sigma^2_\beta}-1)e^{2\mu_\beta+\sigma^2_\beta} - (t\log t)(e^{\rho\sigma_\alpha\sigma_\beta}-1)e^{\mu_\alpha+\mu_\beta+(\sigma^2_\alpha+\sigma^2_\beta)/2}$$

is attributable to the variability of the trio of genotypic parameters τ_i, α_i, and β_i, while the study-wide parameter σ^2_ϕ is attributable to the pure noise term ϕ_{ijt}. Note that the parameters $\sigma^2_\tau, \sigma^2_\alpha, \sigma^2_\beta, \mu_\alpha, \mu_\beta$, and ρ stipulate the collective statistical behavior, *a priori*, of τ, α, and β.

As such, we define four different measures of heritability, namely, $h^2_h, h^2_{\alpha(-\beta)}, h^2_{\beta(-\alpha)}$, and h^2_τ, each at the genotypic level, by letting

$$\begin{aligned}
h^2_h(t) &= \frac{\sigma^2_{\text{genes}}(t)}{\sigma^2_{\text{genes}}(t)+\sigma^2_\phi} \\
&= \text{depth-specific heritability of intensity function at } t, \\
h^2_h &= \text{heritability of overall architecture} \\
&= \text{harmonic mean of } h^2_h(t) = T\Big/\sum_{t=1}^{T}\left(1+\frac{\sigma^2_\phi}{\sigma^2_{\text{genes}}(t)}\right); \\
h^2_{\alpha(-\beta)} &= \text{heritability of root bulk's location (and size) on log scale, ignoring its relation with penetration rate} \\
&= \text{harmonic mean of } \left\{\frac{(\log t)^2(e^{\sigma^2_\alpha}-1)e^{2\mu_\alpha+\sigma^2_\alpha}}{(\log t)^2(e^{\sigma^2_\alpha}-1)e^{2\mu_\alpha+\sigma^2_\alpha}+\sigma^2_\phi} \text{ for } t>1\right\}; \\
h^2_{\beta(-\alpha)} &= \text{heritability of root's decline rate of penetration on log scale, ignoring its relation with bulk location} \\
&= \text{harmonic mean of } \left\{\frac{t^2(e^{\sigma^2_\beta}-1)e^{2\mu_\beta+\sigma^2_\beta}}{t^2(e^{\sigma^2_\beta}-1)e^{2\mu_\beta+\sigma^2_\beta}+\sigma^2_\phi}\right\}; \\
h^2_\tau &= \frac{\sigma^2_\tau}{\sigma^2_\tau+\sigma^2_\phi} = \text{heritability of intensity function's intercept on log scale.}
\end{aligned}$$

Note that each of $h^2_h, h^2_{\alpha(-\beta)}$, and $h^2_{\beta(-\alpha)}$ comprises depth-specific heritability components, but h^2_τ does not (and thus, its definition does not require the use of the harmonic mean).

APPENDIX E

Validating *Model 1*

(a) Predictive Performance

Although more complex models typically follow the data more closely, they may have poorer predictive performance due to potential overfitting. We consider the predictive performance of *Model 1* by comparing its value of the WAIC to those for the simpler *Models 2* and *3*, both nested within *Model 1*. The WAIC is a measure of a model's predictive accuracy, and it is asymptotically equivalent to the leave-one-out cross-validation method, the latter of which addresses the notion of the mean squared error but requires substantive computational effort for a dataset as large as ours (Vehtari and Gelman, 2014). In contrast, the WAIC can be easily computed as part of the MCMC implementation of the Bayesian inference (see Supplementary Material: Computer Code).

As mentioned in Appendix B, the modeled correlation between the bulk and the exploration parameters on the log scale was prespecified as $\rho = 0$ in *Model 2*; *Model 3* considers all model parameters (including those that are genotype-specific) as fixed effects by naïvely prespecifying

$$\begin{aligned}
\sigma^2_\phi = \tau_{20} = \kappa_4 = \mu_\alpha = \mu_\beta = \rho = 0, \\
\text{prior Var}(\psi_0) = \sigma^2_\tau = \sigma^2_\kappa = 25, \\
\sigma^2_\alpha = \sigma^2_\beta = 3.
\end{aligned}$$

Note, that the values 3 and 25 for the prior variance of $\psi_0, \tau_i, \kappa_j, \alpha_i$, or β_j constitute informative prior distributions for these fixed effects. Before considering 3 or 25, we had specified 10^3 or 10^4 for diffuseness. However, in either case, *Model 3* failed to converge due to a weakly identifiable α_8. Consequently, we decided to employ the more restrictive (but defensibly so) prior variances of 3 and 25 according to the following argument.

Based on Weaver (1926), we deduce that for a wheat plant, the total number of roots at any given depth has a magnitude that is $o(100)$, and thus a generous approximation for the standard deviation (SD) of θ_{ijt} is 100, or for $\text{SD}(\log\theta_{ijt})$ is $\log 100$. For *Model 3*, note that

$$\text{E}(\log\theta_{ijt}) = \psi_0 + \tau_i + \kappa_j + (\alpha_i - 1)\log t - \beta_i t$$

where the sets $\{\tau_i\}, \{\alpha_i\}, \{\beta_i\}$, and $\{\kappa_j\}$ each follows a linear constraint. Thus, for the priors of the fixed effects, heuristically we let

$$\begin{aligned}
(\sqrt{25}>) \qquad & \log 100 = \text{SD}(\psi_0) = \text{SD}(\tau_i) = \text{SD}(\kappa_j) \\
(\sqrt{3}>) \qquad & \log\log 100 = \text{SD}(\log\alpha_i) = \text{SD}(\log\beta_i)
\end{aligned}$$

for all $i \neq 20$ and $j \neq 4$.

For each of *Models 1–3*, we computed the WAIC based on Vehtari and Gelman (2014) (see Supplementary Material: Computer Code); they appear in **Table A1**. There, one can see that predictive performance improved drastically (WAIC decreased from > 46000 to < 35000) from the naïve fixed-effects *Model 3* to the mixed-effects *Models 1* and *2*, both of which are much more complex. Between the two complex models, although *Model 1* is slightly larger than *Model 2* (by a single parameter that represents *a priori* dependence between the bulk and exploration parameters), effectively they perform equally well in predictive power, as suggested by a merely nominal difference (=3) in WAIC values.

In addition to a mere nominal difference in WAIC values between *Models 1* and *2*, we also observed that the approximate values for the effective number of parameters, p_{WAIC} (used in the computation of WAIC), increased from 4011 for *Model 1* to 4014 for *Model 2*. Importantly, although model complexity was increased from *Model 2* to *Model 1* by correlating the genotypic bulk and exploration parameters through a study-wide parameter, the additional parameter reduced the models overall amount of unknownness. Thus, these values of p_{WAIC}, along with those of the WAIC and the substantively large posterior probability that $\rho > 0$, suggest the merit of retaining ρ as an unknown parameter in the HNLMM.

(b) Residual Plots

Next, we inspect violin plots produced by the R package "vioplot" (Adler, 2005) for residuals that correspond to the Level 1 noise terms τ_i and ϕ_{ijt} in *Model 1*. Non-noise-like patterns in these residual plots would suggest the statistical inadequacy of *Model 1* for our data.

Figure S1 is based on the posterior distribution of ϕ_{ijt} (posterior median shown as black dot), plotted against the posterior mean of $\log\theta_{ijt}$ rather than the log-transformed $\bar{y}_{ij+t} = \sum_k y_{ijkt}/4$. This is because $\bar{y}_{ij+t} = 0$ for 181 out of all 1440 combinations of $\{i, j, t\}$ (see Section Discussion for possible implications). Figure S1 shows that ϕ_{ijt} has a (a) slight tendency to increase with $\log\theta_{ijt}$, and (b) a distinctive non-random relationship with $\log\theta_{ijt}$ especially when the latter is small (which is typically at lower depths). We break down this relationship by the residual violin plots in Figures S2–S6.

Overall, we observe the following minor anomalies:

- τ_i vs. $\bar{y}_{i++t}$ (Figure S3): at many depths t, the random effect τ_i has a slight tendency to be negative for small observed values of $\bar{y}_{i++t}$, and positive for large observed values of $\bar{y}_{i++t}$;
- ϕ_{ijt} vs. i (Figure S4): at $t = 160, 170$, or 180 cm, the residual ϕ_{ijt} for a small number of plots ($\{i, j\}$ combinations) has a tendency to be highly positive;
- ϕ_{ijt} vs. t (Figure S5): for the 3rd replicate block ($j = 3$), many genotypes (e.g., $i = 2, 12, 14$, etc.) at the deepest six depths are associated with ϕ_{ijt} that increases with depth systematically; the same applies to $j = 4$ and $i = 8, 9, 11$, etc.; additionally, the plot $\{i, j\} = \{5, 3\}$ shows that ϕ_{ijt} has a tendency to be all positive;
- ϕ_{ijt} vs. j (Figure S6): the same conclusion as for Figure S4 for $t = 160, 170$, or 180 cm; additionally, ϕ_{ijt} for a small number of plots has a slight tendency to be positive at $t = 30$ or 100 cm.

Altogether, the residual violin plots suggest that the statistical inadequacy of *Model 1* lies in the modeled behavior of θ_{ijt} across the deeper depths for specific combinations of $\{i, j\}$. Under Section Discussion, we provide an overview of possible directions that may be taken to improve the adequacy of our HNLMM.

APPENDIX F

Some Limitations and Possible Extensions

Note, that while 5 cm segments were produced in the field, the first, third, fifth, etc. depths were ignored in the statistical modeling; only imaged counts at depths in 10 cm increments from the surface were considered. This implies that within the same core, the resulting counts y were less spatially autocorrelated due to a lower spatial resolution of the data from omitting alternate segments from consideration. With segment-level counts y thus produced, we considered for statistical modeling 18 ($= n_D$) depth values from each core, from depth 10 cm to depth 180 cm.

We now discuss possible extensions to *Model 1*. Although the hierarchical structure of our HNLMM already addresses some over- or under-dispersion in the raw root counts, an abundance ($181/1440 > 12\%$) of observed zero-mean root counts (averaged over four replicate cores) suggests a potential need to include a formal zero-inflation component in a future improved model (e.g., via a mixture model). Furthermore, at present our HNLMM does not include formal (a) gene-environment interactions or (b) spatial statistical modeling (Gelfand et al., 2010) of within-core spatial dependency among segment-level root count data at a 10 cm spatial resolution. For (a), it is possible to further decompose $\phi_{ijt} = \phi^{*}_{ij} + \phi^{**}_{it} + \phi^{***}_{ijt}$, whereby ϕ^{*}_{ij} and ϕ^{**}_{it} are explicit gene-environment interaction terms at

TABLE A1 | Values of the Watanabe-Akaike information criterion (WAIC) as a measure of predictive performance by our Bayesian HNLMMs.

Model	Description	WAIC
1	Most sophisticated among our models, as presented in this paper	34198
2	Same as *Model 1*, but with a prespecified $\rho = 0$	34195
3	Naïve preliminary model: same as *Model 2*, except with ϕ_{ijt} term missing and all of $\psi_0, \tau_i, \alpha_j, \beta_i$, and κ_j taken as fixed effects	46627

A smaller WAIC value suggests better model performance.

Level 1 of the model hierarchy, and ϕ_{ijt}^{***} is the pure noise term. For (b), spatial statistical modeling would impose substantial complexity to the statistical inference and computational burden. However, the residual plots perhaps suggest that formal spatial modeling could be a valuable additional component for our HNLMM, especially with data at the 5 cm resolution. A possible spatial structure could be an autoregressive dependence over depth, and/or a nearest-neighbor dependence (among field plots) that constitutes a Markov random field (Gelfand et al., 2010).

frontiers in Environmental Science

REVIEW
published: 19 June 2018
doi: 10.3389/fenvs.2018.00053

Perspectives and Challenges for Sustainable Management of Fungal Diseases of Mungbean [*Vigna radiata* (L.) R. Wilczek var. *radiata*]: A Review

Abhay K. Pandey[1], *Rishi R. Burlakoti*[2*†], *Lawrence Kenyon*[2] *and Ramakrishnan M. Nair*[1]

[1] World Vegetable Center South Asia, ICRISAT Campus, Hyderabad, India, [2] World Vegetable Center, Tainan, Taiwan

Mungbean (*Vigna radiata* var. *radiata*) is a key legume crop grown predominantly in South and Southeast Asia. Biotic and abiotic stresses cause significant yield reduction in mungbean, and among these, fungal diseases are particularly important. Although disease management practices, including physical, chemical, and biological methods have been researched and described in the literature, few of these are available or have been used by growers. Here we review the economic impact, and sustainable management options for the soil-borne and foliar fungal diseases of mungbean as well as major challenges to manage these diseases. Potential use of all possible components of integrated management practices including host resistance, fungicides, biocontrol agents, natural plant products, and cultural practices etc. are discussed. Major diseases include powdery mildew, anthracnose, Cercospora leaf spot, Fusarium wilt, Rhizoctonia root rot and web blight, Macrophomina charcoal rot/dry root rot and blight. Review of the literature indicated an absence of resistance to Rhizoctonia root rot, little sources of resistance for dry root rot and anthracnose. Major resistant genes (R genes) and quantitative trait loci (QTL) were identified for powdery mildew and Cercospora leaf spot, which may be potentially used in Marker assisted selection (MAS). Although the mechanisms of induced systemic resistance (ISR) by biocontrol agents have been studied with Macrophomina blight, there is little information on the mechanisms and use of systemic acquired resistance (SAR) in managing fungal diseases of mungbean. Several studies targeted exploiting biological control for soil-borne root rot diseases. Botanical products, such as plant extracts, are also found effective to manage root and foliar diseases. However, many of these studies were limited to laboratory and/or green house experiments. Thus, long-term field studies are required for further exploitation of biological methods and commercial applications.

Keywords: mungbean, fungal diseases, quantitative trait loci, host resistance, disease management

OPEN ACCESS

Edited by:
Andries B. Potgieter, The University of Queensland, Australia

Reviewed by:
Rao Chenchu Rachaputi, The University of Queensland, Australia
Prakit Somta, Kasetsart University, Thailand
Raju Ghosh, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

****Correspondence:***
Rishi R. Burlakoti
rrburlakoti@gmail.com

†Present Address:
Rishi R. Burlakoti, Agriculture and Agri-food Canada, Agassiz, BC, Canada

Specialty section:
This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Environmental Science

Received: *20 December 2017*
Accepted: *25 May 2018*
Published: *19 June 2018*

Citation:
Pandey AK, Burlakoti RR, Kenyon L and Nair RM (2018) Perspectives and Challenges for Sustainable Management of Fungal Diseases of Mungbean [Vigna radiata (L.) R. Wilczek var. radiata]: A Review. Front. Environ. Sci. 6:53. doi: 10.3389/fenvs.2018.00053

INTRODUCTION

Mungbean (*Vigna radiata* (L.) R. Wilczek var. *radiata*) is one of the important pulse crops in South and Southeast Asia. About 90% of global production is in South Asia, where India is the largest producer (Nair et al., 2012). India produces about 1.5–2.0 million tons of mungbean annually from about 3–4 million hectares (2014–2015), with an average productivity of 0.5 t ha^{-1} (Jadhav et al., 2016). Mungbean is also grown in China (Zhang et al., 2011), Australia (Clarry, 2016), and United States of America (Fery, 2002). In Australia, acreage of mungbean has increased substantially with 125,000 ha planted in 2015–2016 compared to only 1,000 ha in 1970s (Clarry, 2016). Average yield of mungbean is 0.4 t ha^{-1} in Asia but yields up to 2.5 t ha^{-1} may be attained with selected varieties and good management (AVRDC, 2012). Mungbean seeds is a good source of dietary protein for humans including marginal people, and people who live in areas with less access to meat or where people are mostly vegetarian (AVRDC, 2012). Mungbean sprouts and green pods contain high level of vitamins and minerals (Keatinge et al., 2011; Nair et al., 2015).

Abiotic and biotic stresses caused significant decline in legume yield in South Asia and South East Asia. Among biotic stresses, fungal diseases are responsible for reducing yield up to 40–60% in mungbean (Kaur et al., 2011). Fungal pathogens can infect mungbean plants at different stages, such as during emergence, seedling, vegetative and reproductive stages and cause substantial damage leading to yield loss or complete failure of production. Species of the genera *Fusarium* (wilt), *Rhizoctonia* (wet root rot), and *Macrophomina* (dry root rot) infect mungbean plants during seed/seedlings stages (seed-borne or soil borne), while species of the genera *Colletotrichum* (anthracnose), *Alternaria* and *Cercospora* (leaf spot), *Erysiphe/Podospheara* (*Sphaerotheca*) (powdery mildew) affect plants during vegetative and reproductive stages (**Figure 1**; Ryley et al., 2010). Singh et al. (2013a) reviewed the status of web blight in mungbean and recently, Naimuddin and Singh (2016) published a review on yellow mosaic in mungbean and urdbean from India. However, reviews on fungal diseases of mungbean, their economic impact and major management practices have not been compiled. This manuscript reviews the economic impact of major fungal diseases in the mungbean growing areas of South and Southeast Asia and other areas in world as well as options available for sustainable management of these diseases. The review will also cover efforts in resistant breeding or pre-breeding activities including disease evaluation techniques.

ECONOMIC IMPACT OF MUNGBEAN FUNGAL DISEASES

Mungbean mainly is grown in rain-fed climates and variability in climate such as elevated temperature and CO_2 within the rain-fed ecologies leads to varying intensities of biotic stress (Chakraborty et al., 2000; Sharma et al., 2007) which cause significant loss in production. Foliar and root rot fungal diseases are major production constraints in South Asia and South East Asia. Charcoal rot/dry root rot (**Figures 1a–c**) (*Macrophomina phaseolina)* and Rhizoctonia root rot (**Figure 1d**) (*Rhizoctonia solani)* are economically important soil-borne vascular diseases, causing wilt and root rot complex (Alam et al., 1985; Iqbal and Mukhtar, 2014). Among major soil-borne vascular diseases, dry root rot and wilt is a major concern since the pathogen affects the plant during all growth stages and subsequently causes significant yield loss. Yield loss due to dry root rot was reported to be 11% in Northern India (Kaushik et al., 1987) and up to 44% in Pakistan (Bashir and Malik, 1988). Dry root rot was also reported first time in Shanxi province of China in 2010 (Zhang et al., 2011). Mungbean plants with wilt and root rot symptoms, with incidence of 80–90% in susceptible genotypes, was also reported in 1979 in southwest Ontario (Anderson, 1985).

Anthracnose (*Colletotrichum lindemuthianum* or *C. truncatum* or *C. gloeosporioides)* (**Figure 1f**; Shen et al., 2010), Cercospora leaf spot (**Figure 1g**) (*Cercospora cruenta* or *C. canescens* or *C. kikuchii* or *C. caracallae)* (Joshi et al., 2006), and powdery mildew (**Figure 1h**) (*Erysiphe polygoni* or *Podosphaera fusca)* (Ryley et al., 2010), Macrophomina blight (*M. phaseolina)* and web blight (**Figure 1e**) (*R. solani*) (Alam et al., 1985; Iqbal and Mukhtar, 2014) are the major foliar diseases of mungbean as causing yield loss ranging 20–60% in different continents. A wide range of yield losses (23–96%) due to Cercospora leaf spot was reported from field trials conducted at different states of India (Kaur, 2007; Chand et al., 2012; Bhat et al., 2014) and up to 61% from Pakistan (Iqbal et al., 1995). The impact of powdery mildew on mungbean also was reported from different countries. Yield loss due to powdery mildew was reported up to 21% in the Philippines (Quebral and Cowell, 1978), up to 40% in Australia (Kelly et al., 2017), and from 20 to 100% in different regions of India. Yield losses from powdery mildew was reported 35% from Gujarat, western India (Khunti et al., 2002), 20–40% from Chhattisgarh, central-eastern India (Khare et al., 1998) and 20–40% in Maharashtra, western-central India (Mandhare and Suryawanshi, 2008), and from 9 to 50% in Uttarakhand and Uttar Pradesh of Northern India (Pandey et al., 2009). Reddy et al. (1994b) also reported 100% loss from Maharashtra State, India due to powdery mildew diseases at seedling stage. A wide range of yield losses (24–67%) due to anthracnose disease was estimated from several mungbean growing areas in India (Deeksha and Tripathi, 2002; Kulkarni, 2009; Shukla et al., 2014). Alternaria leaf spot (*Alternaria alternata*) is also reported in South Asia, but economic impact is minor i.e., only 10% loss reported from Jammu and Kashmir, India (Maheshwari and Krishna, 2013). Web blight has been a problem for several decades in Pakistan (Alam et al., 1985) and in India, where it was reported from diverse geographical areas including Kanpur and Uttar Pradesh, northern India (Dwivedi and Saksena, 1974), Punjab, northwest India (Bains et al., 1988), Madhya Pradesh, central India (Tiwari and Khare, 1998), Rajasthan (south India), Bihar, Haryana, and Himanchal Pradesh (northern India) (Anonymous, 2014). About 30–40% of yield loss due to web blight was reported from Rajasthan (Anonymous, 2000) and 20–40% seedling mortality due to *Rhizoctonia* infection was reported from Jabalpur, central India (Tiwari, 1993).

FIGURE 1 | Symptoms of fungal diseases of mungbean at World Vegetable Center, South Asia Hyderabad field, **(a)**. Wilted plant with dry root rot symptom (*Macrophomina phaseolina*), **(b,c)**. Root & stem infected with *M. phaseolina*, **(d)**. Roots showing wet root rot (*Rhizoctonia solani*), **(e)**. Leaf showing web blight (*R. solani*), **(f)**. Leaf showing anthracnose (*Colletotrichum* spp.), **(g)**. Leaf showing *Cercospora* leaf spot (*Cercospora* spp.), **(h)**. Leaves showing powdery mildew.

There is a little information available regarding the dynamics in the prevalence and incidence of diseases in mungbean at temporal and spatial scale. Fusarium wilt caused by *Fusarium oxysporum* and/or *F. solani* was a minor disease of mungbean in Australia. However, the incidence and severity of the disease has increased substantially in recent years and yield losses of up to 80% were reported in the susceptible mungbean cultivars (Kelly, 2017). Outbreaks and spread of diseases were reported in other legume crops, such as soybean (Sconyers et al., 2006) and chickpea (Sharma and Ghosh, 2017). For example, Asian soybean rust (*Phakospora pachyrhizi*) was a problem in Asia and South America, but then spread rapidly across eight states of southeastern United States within a few years of first detection in Louisiana in 2004 (Sconyers et al., 2006). It was speculated that an extreme weather event (hurricane) was responsible for the introduction and spread of the Asian soybean rust. Sharma et al. (2015) also reviewed dry root rot (*Macrophomina phaseolina*) as an emerging disease of chickpea in semiarid tropic region and disease intensity has been increased in a past decade. They further speculated that changes in weather pattern, such as high temperature and drought stress during reproductive stages of the chickpea increased the dry root rot intensity. Climate change could have positive or negative or neutral impact on the dynamics of crop diseases depending on the types of crops, diseases or geographical regions (Luck et al., 2011). Climate change would increase average global temperature, CO_2 level, and cause more extreme rain/drought events (Meehl et al., 2005). Severity of some diseases, such as brown spot of soybean (*Septoria glycines*) and sheath blight of rice (*Rhizoctonia solani*) increased with elevated levels of CO_2 (Kobayashi et al., 2006; Eastburn et al., 2010), whereas variable results were reported for powdery mildew of wheat and barley (Thompson et al., 1993; Hibberd et al., 1996). In South Asia, spot blotch in wheat (*Cochliobolus sativus*) has increased substantially in recent years and it is speculated that elevated night temperatures due to climate change has contributed to this (Sharma et al., 2007). In Australia, root and crown rot of wheat (*Fusarium pseudograminearum*) is expected to increase due to climate change as the disease was high with elevated CO_2, temperature, and drought (Melloy et al., 2010). There is no information available on the impact of climate change on the dynamics of mungbean diseases. However, based on knowledge of similar pathogens/diseases in other crops, we can speculate that the pressure of soil-borne diseases caused by *Fusarium*, *Macrophomina,* and *Rhizoctonia* in mungbean may increase due to climate change. It is difficult to predict the effect of climate change on foliar diseases since they are influenced by the combination of temperature, rainfall and relative humidity, which can't be precisely predicted in most situations of climate change.

PERSPECTIVES FOR SUSTAINABLE DISEASE MANAGEMENT

The options for sustainable management of fungal diseases of mungbean include cultural and physical methods, exploitation of host resistance, use of synthetic fungicides, use of natural products such as botanical extracts, bio-fungicides, and use of

bio-stimulants or defense activators, and these are discussed below.

Cultural and Physical Practices

Use of different cultural practices and physical methods to eliminate seed-borne pathogens were found effective to reduce the foliar and root rot diseases of mungbean in fields. Field sanitation, crop rotation, removal of crop debris and weed hosts in the vicinity of the crop reduced the Cercospora foliar blight in mungbean (as reviewed in Sharma et al., 2011). Removing root rot infected mungbean plants reduced sclerotia loads in the field and delayed sowing and maintaining wider spacing between the plants reduced powdery mildew incidence (as reviewed in Satyagopal et al., 2014). Plastic mulching increased sclerotial mortality of *M. phaseolina* and reduced pathogen infection (Yaqub and Shahzad, 2009). Mungbean seed treatment with gamma rays ($_{60}$Cobalt) for 0–4 min and 90 days of storage had a suppressive effect on root rot fungi (Ikram et al., 2010). Hot water emersion treatments (55–65°C) were effective to eliminate seed-borne infection with *Colletotrichum acutatum* and *C. gloeosporioides* of mungbean (Lee et al., 2007). In South Asia, mungbean is commonly rotated with rice and wheat. It is reported anecdotally that root rot diseases in mungbean have been increased in South Asia and other Asian countries due to continuous rotation with rice. Several soil-borne pathogens, such as *Rhizoctonia*, *Fusarium* etc. are common problem in rice and wheat (Kobayashi et al., 2006; Melloy et al., 2010). These fungal genera also infect mungbean, but more studies are required to determine if the same species and strains also infect mungbean. If it is practical, adding diversity in the crop rotations would help for the sustainable management of these soil-borne diseases in mungbean. Crop diversification and use of diverse cultural practices, such as crop rotation, plant residue management, adjusting the planting dates etc., are recommended as effective strategies for managing crop diseases in conditions of climate change (Juroszek and von Tiedemann, 2011).

Exploitation of Host Resistance

Use of host–resistance is an effective, economical, and eco-friendly method for managing mungbean fungal diseases. In this section, we synthesize the available information regarding the identification of sources of resistance, available methods for efficient and reliable disease reaction phenotyping, identification of molecular markers associated with disease resistance genes and their potential use to improve disease resistance traits in mungbean.

Resistant Sources for Major Fungal Diseases of Mungbean

Reliable and efficient methods are available to screen for reaction to foliar diseases of mungbean including Cercospora leaf spot, powdery mildew, and anthracnose. Screening of these foliar diseases can be successful in natural field conditions where disease pressure is high, or if artificial inoculation with pathogen spores is available (Iqbal et al., 2004; Yadav et al., 2014a,b). For other foliar diseases caused by hemibiotrophic (*Cercospora* spp.) or necrotrophic pathogens, disease can also be evaluated in the greenhouse with artificial inoculation. Several disease rating scales/systems were developed to assess foliar diseases of mungbean (Wongpiyasatid et al., 1999; Khunti et al., 2005; Suryawanshi et al., 2009). Mungbean germplasm accessions can be screened by inoculating with the pathogen in the controlled environments. Reliable and efficient methods were developed for screening mungbean seedlings against powdery mildew in the greenhouse (Wongpiyasatid et al., 1999; Kasettranan et al., 2010); and in the laboratory using a detached leaf assay (Reddy et al., 1987). For the assessment of foliar diseases, both qualitative and quantitative rating scales were used (Reddy et al., 1994b; Wongpiyasatid et al., 1999; Khunti et al., 2005; Marappa, 2008; Suryawanshi et al., 2009). Root rot and wilt diseases are sporadic and highly variables due to genotypes × environment (G × E) interaction, therefore it is very difficult to get consistent results while screening in natural fields. Therefore, host genotypes are usually screened by inoculation at seedling stages in controlled environment for soil-borne diseases (*M. phaseolina*, *R. solani*, and *F. solani*). Different methods such as paper towel (Khan and Shuaib, 2007) and sick pot/field inoculation methods by inoculating the fungus grown in sorghum or maize grains (Dubey et al., 2009; Choudhary et al., 2011) were used for the evaluation of root rot disease.

Sources of resistance against powdery mildew, Cercospora leaf spot, anthracnose, Macrophomina blight and dry root rot have been identified (**Table 1**). The majority of studies targeted resistance to Cercospora leaf spot and powdery mildews and were conducted in the field. There have been fewer studies to identify root rot and anthracnose resistance sources, and these were conducted in both lab/glasshouse and field experiments. Most of the identified resistant materials were derived from cultivars/recombinant lines /breeding lines/land races; however, some were from wild relatives (Marappa, 2008) and mutant lines (Wongpiyasatid et al., 1999). Since screening trials for resistance against Alternaria leaf spot, anthracnose, and root rot diseases are limited, more attention is required on these. These resistant lines from difference sources can be utilized as donors for developing resistant varieties.

Identification of Major Genes and Quantitative Trait Loci (QTL) Linked to Major Diseases

The success of developing varieties resistant to biotic stresses depends on the availability of good sources of resistance materials as well as identification of markers associated with disease resistance major genes or QTL, which can also be used in marker assisted selection (MAS) breeding program to accelerate the resistant screening for large population. In mungbean, exploitation of host resistance and identification of molecular markers associated with major genes or QTL were mainly targeted for powdery mildew and Cercospora leaf spot (Kasettranan et al., 2010), however, no QTL or associated molecular markers were reported for other major fungal diseases including dry root rot and anthracnose. The commercial breeding for powdery mildew and Cercospora leaf spot disease resistance in mungbean mostly utilized major Resistant (R) genes based on the classical gene-for-gene system (Kasettranan et al., 2009). To our knowledge, use of MAS has not been used

TABLE 1 | Resistant genotypes of mungbean against fungal diseases[†].

Diseases	Country where screening conducted	No. of genotypes[‡] evaluated	Resistant genotypes (R, Resistant; HR, highly resistant)
Powdery mildew	Taiwan	4000	R: V2159, V4189, V4207, V4574, V4668, V4990 R/HR: V3912, V4186 HR: V1104, V4631, V4658, V4662, V4717, V4883 (Hartman et al., 1993)
	Thailand	27	R:M5-10 and M5-25 (Wongpiyasatid et al., 1999)
	India	82	R:BPMR-145, Vaibhav, TARM-18, Phule M-2002-13, Phule M-2001-3, Phule M-2003-3, Phule M-2002-17, and Phule M-2001-5 (Mandhare and Suryawanshi, 2008)
	India	12	R:TARM-18 (Sujatha et al., 2011)
	India	60	R:LGG-460 (Yadav et al., 2014a)
	India	374	R:116 resistant lines; HR: BL 849, LM1668, BL 865, AKM 8803, PBM, PMB 63 (Ramakrishnan and Savithramma, 2014)
	India	-	HR: KGS 83, MH 96-1, Pusa 572, GS 33-5, AKM 99-4, GS 21-5, COGG 936, ML 1299, TMB 47, HUM 1, MH 429, MH 429 and MH 530 (Akhtar et al., 2014)
	India	63	HR: KMP-36,39 and 41 R: KMP−2,3,5,19,20,24,30,34,38,42,47,52 and MLGG-8 (Bhaskar, 2017)
	India	146	HR: F4: C1-34-23, F5: C1-15-10, C1-15A-11, C1-21A-17, C1-25-19, C1-28-20, C1-32-22, C1-37-23, C1-38-27, C1-41-28, C1-44-31, C1-175-111, C1-236-152, C1-246-159, C1-275-177 (Kumar et al., 2017)
Cercospora leaf spot (CLS)	Taiwan	4000	R: V1471, V2757, V2773, V4718, V5036 (Hartman et al., 1993)
	Thailand.	27	R: M5-22 and M5-25 (Wongpiyasatid et al., 1999)
	Pakistan	58	R: NCM 255-2, NCM 257-6, ML-267, NCM 251-1, NCM 259-2, NCM 251-13, NCM 257-2, NM-92, NCM 251-12, VC-3960-A88 NCM 257-10, NCM-209, Mung-6 C1/94-4-19, VC 3960-A89 HR: BRM-188, NM-98, C2/94-4-42, 98-cmg-003, NM-2, NM-1, 98cmg-018, Basanti, CO-3, PDM-11, VC3960-88, BARIMung-2 (Iqbal et al., 2004)
	India	696	R: ML5, 443, 453, 515, 610, 611, 613, 682, 688, 713, 728, 735,746, 759 and 769 (Singh et al., 2004)
	India	170	No infection: *Vigna aconitifolia, V. glabrascence, V. sublobata, V. umbellata* and a mutant PBM. R: 90 genotypes including PANT M103, PANT M3, PUSA 105, ML 613, PANT M2, ML 173, ML 347, ML 561, PANT M4, PDM 11 (Marappa, 2008)
	India	65	R: GM-02-08, GM-02-13, GM-03-03 HR: LGG-460 (Yadav et al., 2014b)
	India	113	R: ML-5, ML-4, HUM-9, HUM-4, HUM-1, SM-9-124, LGG-450, and SM-9-107 (Singh and Singh, 2014)
	India	136	R: 52 genotypes HR: 1224-52 and 12404 (Zhimo et al., 2013)
	India	-	R: AKM 9910, IPM 02-5, ML 1299 and SML 668 (Akhtar et al., 2014)
	India	63	MR: KMP-13 (Bhaskar, 2017)
CLS, anthracnose, Macrophomina blight	India	56	R: ML1486, ML1464, ML1194 and ML1349 (Kaur et al., 2011)
Dry root rot	India	25	R: MSJ 118, KM 4-44 and KM 4-59 (Choudhary et al., 2011)
	Pakistan	29	R: 40504, NCM 257-5, 40457, NCM 251-4, 6368-64-72 HR: NCM 252-10 and 40536 (Khan and Shuaib, 2007)

[†] *All the trials were conducted in the field except dry root rot screening work by Khan and Shuaib (2007), which was conducted in the greenhouse.*
[‡] *Genotypes include cultivars, landraces, wild relatives, breeding lines, mutant lines, and germplasms.*

for mungbean breeding programs targeted for fungal disease resistance in developing countries. However, identification of molecular markers associated with disease resistant major genes and QTL shows the potential application of MAS in mungbean disease resistance. Genetic studies using different sources of resistance revealed both monogenic (qualitative) and quantitative modes of inheritance in mungbean for powdery mildew resistance (Reddy et al., 1994a; Kasettranan et al., 2009). Gawande and Patil (2003) reported that both additive and dominant gene actions were important in inheritance of powdery mildew resistance including non-allelic interactions. Several earlier studies reported monogenic inheritance of powdery mildew resistance controlled by single dominant genes and studies were conducted using mungbean varieties ML3 and ML5 (AVRDC, 1979), and breeding lines VC 1560A (AVRDC, 1981), ATF 3640 (Humphry et al., 2003) and RUM (Reddy et al., 1994a). Using restriction fragment length polymorphism (RFLP) markers, Chaitieng et al. (2002) and Humphry et al. (2003) revealed single major locus conferring the resistance against powdery mildew with 65 and 80% R^2, respectively. Khajudparn et al. (2007) found non-allelic dominant gene for powdery mildew in F_2 populations developed from resistant lines V4718, V4758, and V4785 (obtained from World Vegetable Center) and susceptible line CN72. Reddy (2009) studied the inheritance of Pm3 gene (different from earlier identified resistant genes, Pm1 and Pm2), a new gene responsible for powdery mildew by using local mungbean cultivar Mulmarada from Maharashtra (India). He found that F_1, F_2, and F_3 families exhibited complete resistance to powdery mildew is controlled by single dominant gene.

Several researchers reported quantitative mode of inheritance for powdery mildew resistance (Young et al., 1993; Sorajjapinun et al., 2005). Sorajjapinun et al. (2005) reported that additive gene action was found to play a major role in controlling powdery mildew (*E. polygoni*) resistance in the population of crosses developed between moderately resistant KPS 2 and resistant VC 6468-11-1A (sourced from the World Vegetable Center). Using mapping population developed from advanced mungbean breeding line VC3890 (from World Vegetable Center) as a resistance parent, Young et al. (1993) identified three QTL associated with powdery mildew (*E. polygoni*) resistance. These QTL explained 17 to 28 and 58% of phenotypic variation (R^2) individually and together, respectively. Using SSR markers, Kasettranan et al. (2010) identified two major QTL (qPMR-1 and qPMR-2) associated with powdery mildew resistance, which explained R^2 of 20 and 58%, respectively. They used 190 F_7 recombinant inbred line (RIL) population developed from the crosses between a susceptible cultivar, Kamphaeng Saen 1 and a resistant line, VC6468-11-1A (sourced from World Vegetable Center). SSR markers flanking and closely associated with qPMR-1 (CEDG282 and CEDG191) and qPMR-2 (MB-SSR238 and CEDG166) can be useful for MAS powdery mildew resistant breeding program of mungbean. Chankaew et al. (2013) also identified a major QTL associated powdery mildew resistance on linkage group (LG) 9 and two minor QTL on LG4 in V4718 (sourced from World Vegetable Center). They also detected two major QTL on LG6 and LG9 and one minor QTL on LG4 in the mapping populations developed using mungbean genotype RUM5 (Chankaew et al., 2013).

In Cercospora leaf spot, genetic inheritance studies using different resistant sources revealed that the resistance is controlled by either a single dominant gene (Lee, 1980), a single recessive gene (Mishra et al., 1988) or quantitative genes (AVRDC, 1980; Chankaew et al., 2011). Although the above information is useful for breeders in developing the resistant varieties, progress in selecting CLS-resistant genotypes in large breeding programs is still limited. First QTL mapping for resistance to Cercospora leaf spot in mungbean was carried out in Thailand (Chankaew et al., 2011). Using F_2 (CLS susceptible cultivar Kamphaeng Saen1, KPS1 × CLS-resistance mungbean line, V4718) and BC_1F_1 [(KPS1 × V4718) × KPS1] populations, they identified one major QTL (qCLS) on LG3 located between the markers CEDG 117 and VR 393, which explained 66–81% phenotypic variation. Their study further confirmed that SSR markers flanking qCLS will facilitate transfer of CLS resistance allele from V4718 into elite mungbean cultivars.

Protection With Synthetic Fungicides

Applications of fungicides are the most common approach of managing fungal diseases of crops. The traditional ways of disease management in mungbean include use of broad spectrum fungicides as a seed treatment chemicals and foliar spray. Efficacies of different mode of fungicides evaluated to reduce the major fungal diseases of mungbean are summarized in **Table 2**. Fungicides were evaluated mostly in the field trials as seed treatment and/or foliar sprays. Majority of trials were targeted for Cercospora leaf spot, anthracnose and powdery mildew and few trials were on Macrophomina blight, web blight and dry root rot (**Table 2**). Most of these studies assessed fungicide efficacies in reducing disease incidence and/or severity and yield benefit; however missed the economic analyses of the fungicide applications, which is critical component to recommend for farmers. The major group of effective fungicides to control foliar diseases including powdery mildew, Cercospora leaf spot, web blight, and Macrophomina blight were DMI, and MBC. Application of mancozeb (dithocarbmate) was not effective for powdery mildew; however, was effective for Cercospora leaf spot and Macrophomina blight. Dinocap (QiL) and tridemorph (amines groups) were effective for powdery mildew. Carbendazim and benomyl (MBC) were effective for anthracnose. Most of the foliar spray was applied immediately after the appearance of disease symptoms followed by 2nd and 3rd spray after 15–20 days of first spray for anthracnose, powdery mildew and Cercospora leaf spot as given in the **Table 2**. Seed treatment is applied mainly against wet and dry root rot, anthracnose and Alternaria leaf spot diseases before sowing. For dry and wet root rot disease, carbendazim was found to be most effective fungicides (Rathore, 2006). The other effective fungicides for wet root rot were flutolanil and tolclofos-methyl (SDHI), carbendazim (MBC) and pencycllron (Phenylureas) (Kumari et al., 2012).

To our knowledge, disease outbreak due to break down of fungicides has not yet been reported in mungbean. However,

TABLE 2 | Efficacy of fungicides for the control of fungal diseases in mungbean.

Diseases	FRAC code and Fungicide groups†	Effective fungicides	Method and frequency of application	Efficacy Impact (Disease reduction and yield)
FOLIAR DISEASES				
Powdery mildew	***M02-**(inorganic)	Wettable Sulfur (0.4%)	Twice foliar spray	Highest cost benefit ratio (3.3) was noticed (Das and Narain, 1990)
	3-DMI	Hexaconazole 5 EC (0.005%),	First foliar spray when disease appeared, repeated after 15 days	59% with 779 kg/ha yield in treatment, while 395 kg/ha in check (Khunti et al., 2005)
	29-QiI	Dinocap 48 EC (Dinitrophenyl crotonate)		73% with 1425 kg/ha (Suryawanshi et al., 2009)
	5-Amines	Tridemorph (0.05%)	First foliar spray when disease appeared, repeated after 7 days	69% and 532 kg/ha yield, while 326 kg/ha in check (Rakhonde et al., 2011)
	3-DMI	Propiconazole (0.10%),	Single foliar spray after first disease appearance	100% with 908 kg/ha yield, while 746 kg/ha in check (Akhtar et al., 2014)
Cercospora leaf spot	**3-**DMI	Hexaconazole 5 EC (0.005%),	First foliar spray when disease appeared, repeated after 15 days	59 and 779 kg/ha yield, while 395 kg/ha in check (Khunti et al., 2005)
	3-DMI	Difenconazole (25% EC) (0.0125 %)	Foliar spray after disease initiation, repeated twice at 15 DAS	61% (Kapadiya and Dhruj, 1999)
	1-MBC	Carbendazim (0.10%)	First foliar spray when disease appeared, repeated after 15 days	61% and 690 kg/ha yield at 70 DAS (Khan et al., 2005)
	3-DMI	Hexaconazole (0.1%)	Single foliar spray when disease appeared	81% with 752 kg/ha yield, while 525 kg/ha in check (Veena et al., 2013)
	1-MBC	Carbendazim (0.1 %),	Single foliar spray when disease appeared	77% (Singh et al., 2013b)
	4-PA	Metalaxyl (1.2 kg ha^{-1})	Foliar spray after 50 days of sowing before disease appearance	55% (Shahbaz et al., 2014)
	3-DMI	Propiconazole (0.10%),	Foliar spray after first disease appearance	86% with 908 kg/ha yield, while 746 kg/ha in check (Akhtar et al., 2014)
	1-MBC + **3-**DMI	Carbendazim (0.1%) + Difenconazole (0.02 %),	First foliar spray when disease appeared, repeated after 15 DAS	82 and 72% leaf infection and 76 and 96% pod infection with 825 and 808 g/9 m^2 yield during 2009 and 2010, respectively, while in check yields were 691 and 680 g 9 m^{-2} (Bhat et al., 2015)
	1-MBC + **M03**-dithiocarbamates and relatives	Carbendazim (12%) + Mancozeb (63%) 75% WP	First foliar spray when disease appeared, repeated after 15 DAS	70 and 990 kg/ha yield, while decreased in check (570 kg/ha) (Yadav et al., 2014b)
Anthracnose	**1-**MBC	Carbendazim (0.10%),	First foliar spray when disease appeared, repeated after 15 DAS	38% with 690 kg/ha yield at 70 DAS (Khan et al., 2005)
	1-MBC	Carbendazim (0.1%)	First foliar spray when disease appeared, repeated after 15 DAS	65% with increase in grain (1090 kg/ha) and stalk yield (1470 kg/ha) than untreated plots of resistant cultivar (UPM-98) (Shukla et al., 2014)
	1-MBC	Benomyl 50% (WP),	Single foliar spray @ 1.13 kg (a.i.)/ha per 1136 L of water at 10 days interval of disease	79 and 32 % in 6,601 and M-19-19 varieties with 587 and 669 kg/ha yield, respectively, while in untreated plots yields were 327 and 90 kg/ha (Bashir et al., 1985)
Web blight	**3-**DMI	Propiconazole (0.10%),	Foliar spray after first disease appearence	78% with 908 kg/ha yield, while 746 kg/ha in check (Akhtar et al., 2014)
	1-MBC	Carbendazim 50% WP (0.1%)	First foliar spray when disease appeared, repeated after 15 DAS	59% and 620 kg/ha yield, while it was reduced to 360 kg/ha in check (Jhamaria and Sharma, 2002)
Macrophomina blight	**M03**-Dithiocarbamates and relatives	Mancozeb (0.2%)	Single foliar spray after 7 days of pathogen inoculation	80% and 15 g/plant yield (Murugapriya et al., 2011)

(Continued)

TABLE 2 | Continued

Diseases	FRAC code and Fungicide groups†	Effective fungicides	Method and frequency of application	Efficacy Impact (Disease reduction and yield)
	1-MBC, **M03-**dithiocarbamates and relatives	Carbendazim (0.1%), mancozeb (0.2%)	Foliar spray after appearance of disease	94 and 88% due to mancozeb and carbendazim with 14 and 13.5 g yield/plant, while 5 g/plant yield in check (Rana et al., 2014)
Alternaria leaf spot	**3-**DMI	Hexaconazole (0.03)	First spray immediately after disease appearance and 2nd and 3rd spray were done at 10 days of interval	85% and yield 868 kg/ha, while yield decreased in control to 432 kg/ha (Maheshwari and Krishna, 2013)
ROOT ROT DISEASES				
Dry root rot	**1-**MBC	Carbendazim	Seed treatment @ 2 g kg seeds^{-1}	Reduced 54% disease incidence in pre-emergence and 66% at post-emergence (Kumari et al., 2012)
Damping off/wet root rot	**7-**SDHI, **14–**AH, **1** MBC, **20-**Phenylureas	Flutolanil (1 um a.i. ml^{-1}), tolclofos-methyl and carbendazim (5 urn a.i. ml^{-1}), pencycuron (50 urn a.i. ml^{-1})	Seed dressing (2×3 g ai kg^{-1} seed) or as soil drench (200 and 300 p.g ml^{-1}) of all the fungicides	Flutolanil, tolclofos-methyl, carbendazim and pencyclIron were most effective completely (100%) inhibited growth of *R. solani* and also reduced disease incidence (Reddy et al., 1992)

**The bold values indicate the FRAC (Fungicide Resistance Action Committee) code designated to the fungicide group. †DMI, De Methylation Inhibitors; QiI, Quinone inside Inhibitors; MBC, Methyl Benzimidazole Carbamates; PA, Phenyl Amides; SDHI, Succinatedehydrogenase inhibitors; AH, Aromatic Hydrocarbons (Chlorophenyls, nitroanilines); DAS, Day after spray.*

disease management failures in legume crops associated with fungicide resistance have been reported from several countries (Chang et al., 2007; Lonergan et al., 2015; Price et al., 2015). For example, Price et al. (2015) reported that isolates of *C. kikuchii* (Cercospora leaf spot) from soybean fields in Louisiana State, USA were insensitive to thiophanate methyl. Isolates of *Ascochyta rabiei* (ascochyta blight of chickpea) from Canada and USA showed insensitivity with fungicides pyraclostrobin, chlorothalonil, fluxapyroxad, and prothioconazole (Chang et al., 2007; Lonergan et al., 2015). More than 90% mungbean are produced in developing countries where strict regulations for fungicides are lacking and poor extension services to educate farmers to apply fungicides properly. This may lead in future the disease outbreak due to fungicide resistance problems. Therefore, fungicide resistance management strategies, such as rotation of fungicides with different mode of actions, tank mix of broad spectrum and selective fungicides, and integrate the fungicide spray programs with other components of disease management practices, should be implemented at regional and national level as recommended by Fungicide Resistance Action Committee (FRAC). Use of next generation fungicides derived from active constituents of natural products, which are ecologically safe and effective at lower doses, would also be beneficial (Sierotzki and Scalliet, 2013).

Biological Methods

Biological Control Agents

Very limited information is available on the biological methods to manage mungbean foliar diseases including powdery mildews, Cercospora leaf spot, and anthracnose. However, more information is available for the management of root rot pathogens. Most of these studies were conducted in the laboratories to evaluate the effects of bio-control agents (*Trichoderma* species, *Pseudomonas*, *Bacillus* etc.) to inhibit growth of root rot pathogens, *Rhizoctonia* and *Macrophomina*. Few studies were also conducted in the greenhouse to study the impact of seed or soil applications of the biocontrol agents to reduce the root rot; however, only very few studies were conducted in fields.

Sharma et al. (2017) recommended that application of biocontrol agents is more effective to suppress the soil-borne diseases as effective chemical protectants are either not available or not economical. Integrated applications of biocontrol agent with organic amendments were recommended to reduce root of mungbean in fields (Raghuchander et al., 1993; Ehteshamul-Haque et al., 1995). Dubey and Patel (2002) reported that soil application of *T. viride* (8 g/kg) multiplied in pulse bran and saw dust in the greenhouse experiment showed 75% reduction in root rot disease caused by *R. solani* and also promoted plant growth. A 76% reduction in Rhizoctonia root rot was reported when *Gliocladium virens* (*Trichoderma virens*) applied as seed treatment @ 10^6 spores/ml/10 g seeds (Dubey, 2003). Bioproducts Pusa 5SD (*T. viride*) showed 72% root rot reduction and 978 kg ha^{-1}, Pusa 5SD (*T. harzianum*) showed 71% disease reduction and 940 kg ha^{-1} yield in sick field (Dubey et al., 2011). Similarly, *T. harzianum*, and *T. viride* reduced about 54–73% Rhizoctonia root rot incidences (Singh et al., 2008; Maheshwari and Krishna, 2013) in green house and field experiments, respectively.

Seed dressing and soil drenching with bacterial strains of *Pseudomonas aeruginosa* and *Bacillus subtilis* significantly reduced 42% Macrophomina root rot, 39% Fusarium root rot and 70% Rhizocotnia root rot incidences in mungbean (Siddiqui et al., 2001). Bacterial strain TNAU-1 (*Burkholderia* spp.) inhibited mycelial growth of *M. phaseolina* in *in vitro* dual culture and also reduced root rot incidence up to three-fold when applied as a seed treatment and soil application with talc

based formulations (Satya et al., 2011). *Trichoderma viride* and *T. harzianum* were found to be reduced *M. phaseolina* growth (respective 42–33 and 42–25 mm) in dual culture (Ebenezar and Yesuraja, 2000). In the field study, Kumari et al. (2012) found that mixed application of vermicompost (10%) + bavistin (0.1%) + *T. harzianum* (4%) exhibited 100% reduction of Macrophomina root rot. *Bacillus subtilis* and *T. longibrachyatum* against *M. phaseolina* exhibited 64 and 63% antagonistic activity, respectively (Tandel et al., 2014). In greenhouse study, application of 4 g kg^{-1} seeds of *T. harzianum* with 25 g kg^{-1} of phosphate solubilizing bacteria as seed dresser reduced 26% incidence of Macrophomina root rot (Deshmukh et al., 2016).

The compatibility of different bioagents against root rot pathogens has also been studied. Application of plant growth promoting rhizobacteria, *Pseudomonas aeruginosa*, with a medicinal plant *Launaea nudicaulis* @ 0.5% as soil amendment reduced 51% of Macrophomina root rot, while combined application of *L. nudicaulis* (0.1% W/W) + *P. aeruginosa* and *L. nudicaulis* (1.0% w/w) + *P. lilacinus* gave 0% infection reduced of Rhizoctonia and Fusarium root rot, respectively (Mansoor et al., 2007). In the green house and field trials (Thilagavathi et al., 2007), soil application of *Pseudomonas fluorescens* strain Pf1+ *Trichoderma viride* strain Tv1 controlled 86% Macrophomina root rot in pot culture and 59% in field conditions with 833 kg ha^{-1} yield. The authors also found that *T. viride* strain is not compatible with *B. subtilis* (Bs16), but *P. fluorescens* strain is compatible with *B. subtilis* and *T. viride* in the management of dry root rot. Yadav et al. (2017) reported that *T. viride*, *T. harzianum*, and *Pseudomonas fluorescens* were effective to reduce powdery mildew of mungbean (∼80–84% reduction).

Botanical Fungicides and Bio-Stimulants

The plant products, particularly plant extract and essential oils, showed prominent toxicity to the diverse genera of plant pathogenic fungi, bacteria, insects and nematodes (Pandey and Tripathi, 2011). Plants synthesize aromatic secondary metabolites in the form of terpenes, like phenols (carvacrol, eugenol, and thymol), phenolic acids, quinones, flavones, flavonoids, flavonols, tannins and coumarins (Cowan, 1999), these groups of compounds show fungicidal effect and serves as plant defense mechanisms against fungal pathogens (Slusarenko et al., 2008; Das et al., 2010). For mungbean diseases, most of studies were preliminary and different kinds of plant extracts or their products have been evaluated against mungbean fungal pathogens (Javaid and Amin, 2009; Murugapriya et al., 2011). Foliar spray of neem extract (1:4 w/v) was reduced 65% of Cercospora leaf spot and increased 25% yield in mungbean (Uddin et al., 2013). Leaves extracts behada (*Terminali belerica*), tapioca (*Manihot utilissimum*), and sadafuli (*Vinca rosea*) reduced 60–66% of conidial germination of powdery mildew fungus *E. polygoni* (Rakhonde et al., 2011). Similarly, *in vitro* evaluation of leaf extracts of *Adenocalymma alliaceum* and *Allium* spp. reduced about 75–77% mycelia growth of *M. phaseolina*, Macrophomina blight incidences in greenhouse experiments (Murugapriya et al., 2011; Rana et al., 2014). In the greenhouse experiments, combined applications of 10% extract of *Allium* spp., mancozeb (0.2%), and 10% extract of *Allium* spp. with zinc sulfate (0.5%) reduced about 88–94% incidence of Macrophomina root rot (Sundaramoorthy et al., 2013). Javaid and Saddique (2011) found that amendment of dry leaf manure of *Datura metel* (1.5% w/w) in the soil reduced 80% plant mortality caused by *M. phaseolina*. Similarly, soil application of *L. nudicaulis* (1% w/w) extract also reduced dry root rot (62%), wet root rot (75%) and Fusarium wilt (100%) incidences (Mansoor et al., 2007) in glasshouse. Mungbean seeds dressing with 2% concentration of palmarosa (*Cymbopogon martinii*) oil gave complete inhibition of *M. phaseolina* mycelial growth (100%) in poison food testing and also caused 72.33% reduction in dry root rot in the greenhouse trials (Kumari et al., 2012).

Induced systemic resistance (ISR) and systemic acquired resistance (SAR) are both important phenomenon in the interactions of plant-pathosystems. Both ISR and SAR increased productions of proteins (defense enzymes) like peroxidase (PO), pathogenesis related (PR), phenylalanine ammonia lyase (PAL), polyphenol oxidase (PPO), phenols etc. (Jones and Dangl, 2006; Walters et al., 2009), which showed positive associations with resistance for several fungal diseases of vegetable and legume crops (Vallad and Goodman, 2004; Abdel-Kader et al., 2013). In mungbean, limited studies have been conducted to understand the mechanism of ISR and SAR. Similar to other crops, increase production of plant defense enzymes were reported when mungbean plants were treated with bioagents and plant products and also challenged with plant pathogens. Application of 10% aqueous leaf extracts of *Allium alliaceum* and other *Allium* sp. exhibited increase in PO, PPO, PAL and total phenol contents in mungbean plants inoculated with *Macrophomina phaseolina* (Sundaramoorthy et al., 2013). Treatment of *M. phaseolina* pre-inoculated mungbean plants with Pf1 (*Pseudomonas fluorescens*) formulation amended with chitin increased the accumulation of PAL, PO, PPO, chitinase, β-1,3-glucanse and phenolics indicating that the PGPR strains amended with chitin bioformulation induced defense-related enzymes and pathogenesis related (PR) proteins (Saravanakumar et al., 2007). Higher levels of PO and PPO activity was observed in *M. phaseolina* infection treated with the bioformulation combination of plant growth promoting bacteria (*P. fluorescens*) and biocontrol agents (*Trichoderma* or *Bacillus*) than the plants treated with single biocontrol agent (Thilagavathi et al., 2007). Mechanism of SAR were studied for bacterial (Dutta et al., 2005; Farahani and Taghavi, 2016) and viral diseases (Rashid et al., 2004) of mungbean, but study regarding fungal diseases are still meager. Thus, more investigations are required to understand the SAR and ISR mechanisms in the interactions between mungbean and fungal diseases.

CHALLENGES FOR THE SUSTAINBLE DISEASE MANAGEMENT

More than 90% of mungbean are cultivated in the developing countries, where small farmers do not have proper knowledge on integrated pest management and several challenges exist in the implementation of integrated management options. For example, use of gamma rays for seed treatment is a good option to eliminate seed-borne pathogen from seed, but it is

not viable for the small holder farmers since they produce seeds in their farm in a small scale. Several developing countries in south Asia do not have strong national breeding programs in mungbean to exploit host resistance for multiple diseases. Disease resistant genotypes identified in several studies were evaluated in few locations or seasons. Variability in pathogen populations exists among diverse geography; therefore, screening trials should be conducted multi-locations and years while developing breeding lines for disease resistance. Instability and breakdown of disease resistance in mungbean cultivars is a major challenge in breeding programs due to monogenic host resistance and high pathogenic variability (Nair et al., 2017). Integration of disease resistance traits without compromising valuable agronomic traits is a key challenge for mungbean breeders as linkage drags inhibit the proper use of genetic diversity from wild germplasm into the commercial cultivars (Acosta-Gallegos et al., 2008; Keneni et al., 2011). Further, undesired and desired traits co-inheritance may affect on seed quality.

In developing countries including India, fungicides are registered by CIBRC (Central Insecticide Board and Registration Committee) with their effective dose and label claim which provides guideline to the growers. However, where fungicides are not registered, agriculture officers or fungicide retailers provide fungicides spray guidelines to the growers. Several growers do not apply fungicides with appropriate doses and timing, although majority of fungicides used are preventative (broad-spectrum), which require applying prior to pathogen infection or prior to first symptoms appearance. In addition, farmers do not commonly rotate fungicides with different mode of actions due to poor knowledge and extension on IPM. As fungicides resistance is a big concern for legume industry in several countries (Chang et al., 2007; Lonergan et al., 2015; Price et al., 2015), the problem may arise in mungbean industry. Fungicide resistance could be significant challenge for the mungbean farmers in future to manage diseases effectively. Fungicide resistance management strategies recommended by FRAC, which we have described in the section "Protection With Synthetic Fungicide," have not been deployed at regional or national levels in several developing countries.

Additionally, attempts have been made to produce and apply biopesticide commercially in the developing countries. Challenges also exist for the commercial use of biopesticides in mungbean. Most of biopesticides only suppress the diseases and are not effective as chemical fungicides, therefore growers are reluctant to use the products (Flexner and Belnavis, 2000; Felde et al., 2006). Due to poor extension, growers do not apply the biopesticides as a component of integrated approach. In addition, several abiotic and biotic factors make the biopesticides less effective in field (Meyer and Roberts, 2002; Sharma et al., 2017). Sharma et al. (2017) speculated that there could be risk of developing biocontrol agents as crop pests and therefore, careful attentions are required while developing/evaluating biocontrol agents. However, we did not find any reports in the literature showing the evidences of biocontrol agents shifted to crop pathogens. Few researchers suggested that application of biocontrol agent is effective when mixed with other biocontrol agents; however, other investigators reported that such combinations may not be always advantageous as antagonism can occur among biocontrol agents (Viaene and Abawi, 2000). Most of studies to exploit botanicals and other bio-based products were evaluated in laboratory or controlled environments and their efficacy has not been evaluated in fields. This shows future potentiality of these products for the sustainable management of diseases, however, growers do not have current access of these products.

Mungbean farmers in the developing countries are not well educated about the impact of global climate change in the disease management. Global climate change would influence the emergence of new diseases, biology of the plant pathogens, disease development and their management practices in different geographical regions (Chakraborty et al., 2000; Juroszek and von Tiedemann, 2011; Luck et al., 2011). Global rise in temperature and CO_2 due to climate change may modify aggressiveness and fecundity of the plant pathogens, increase host susceptibility, and change host architecture and host-pathogen interaction (Chakraborty et al., 2000; Luck et al., 2011). The mungbean breeding programs in the developing countries does not have enough resources and strategies for developing resistant varieties for biotic and abiotic stresses associated with climate change elevation in temperature, CO_2, and moisture stress due to climate change may also affect the efficacies and the durability of plant protection chemicals and biocontrol agents (reviewed in Juroszek and von Tiedemann, 2011), which could be also key challenge to manage mungbean diseases in future.

CONCLUSIONS AND FUTURE PROSPECTS

The present review identified that root rot complex and wilt caused by soil-borne pathogens and foliar diseases are major fungal diseases impacting mungbean production in South Asia and South East Asia. Fusarium wilt and root rot and powdery mildew are problematic in Australia. For the management of these diseases potential options such as chemical and non-chemical (cultural, physical, host-plant resistance, biological) have been investigated by the researchers. Although several field trials were conducted to evaluate fungicides and other non-chemical management options by researchers from universities and governments, very little information has been transferred to the mungbean growers in South and Southeast Asia due to the poor linkage between research and extension activities. Deployment of Integrated Disease Management (IDM) to manage mungbean diseases in a coordinated approach requires good collaborations among academia, national and international research institutes, national extension agencies and growers.

Use of resistant varieties (if available) in combination with other components of management is a most effective option to combat with these fungal diseases. Described literature

revealed that sources of resistant genotypes have been identified for Cercospora leaf spot, powdery mildew, and anthracnose diseases and few for dry root rot by screening mungbean germplasms in natural field/artificial conditions in few specific locations. The identified sources for resistance in these diseases could be region specific as they were tested in a few specific locations. For example, V4718 mungbean accession from the WorldVeg gene bank has been used as a source of resistance to powdery mildew in Thailand and India. Breeding lines developed from the above source through cross breeding, and selections have been shared to the partners in the ACIAR funded International Mungbean Improvement Network project. Therefore, evaluation of resistant genotypes for these diseases at multi-locations in a coordinated approach would help in deploying host resistance at a larger scale. Compared to foliar diseases, few resistant genotypes of mungbean are available for root rot diseases. This may be due to the less priority given to the screening of these diseases in the past. However, the incidence of dry root rot in mungbean grown as part of rice based farming system in eastern part of India (Odisha state) and in Myanmar has triggered the need for identification of sources of resistance. Sharma and Ghosh (2017) reported that chickpea genotype which showed a good level of resistance to Fusarium wilt at 24°C were susceptible at 27°C. Therefore, breeding programs should consider potential impact of climate change in the new and existing biotic stresses. Attention should be given to develop climate resilient cultivars (such as cultivars can show a good level of resistance at higher temperature) with greater diversity and incorporating traits for multiple disease resistance. Literature evidenced that molecular markers are available for powdery mildew and Cercospora leaf spot, however, there is need to validate them in breeding programs. More attention is required to develop the molecular markers for root rot and anthracnose diseases. Currently, as a part of the network, we are screening 296 mini-core accessions of mungbean (Schafleitner et al., 2015) for resistance to anthracnose, dry root rot, powdery mildew and Cercospora leaf spot diseases. The resistant accessions identified will be shared among the project partners for cultivar development.

Application of synthetic fungicides is a common practice to control fungal diseases of mungbean, and growers also integrate other cultural methods with chemical sprays. Efficacies of several fungicides (**Table 2**) were evaluated in fields and controlled environments at universities and research institutions, however, there is a knowledge gap regarding how much of these evaluated fungicides are currently used by mungbean growers. In addition, additional research are required for fungicide efficacy trials including rotating and tank mixing with different modes of actions, different rates as well as volume of water for spray coverage. Attentions should also be given to develop and evaluate new generation fungicides. Fungicide resistance problem has not yet been reported in mungbean growing areas, which could also be due to research gap to investigate fungicide sensitivity against mungbean pathogens. In literature, baseline sensitivity data are not available for any fungicides and pathogens. Therefore, future research is recommended for *in vitro* fungicide sensitivity test using large numbers of pathogen isolates from diverse areas. Fungicide resistance management strategies (such as integrating chemical fungicides with other management practices, judicial use of fungicides, rotation and tank mix of different groups of fungicides) should be deployed at regional and national level to reduce the risk of developing fungicide resistance fungal population. Future impact of climate change on diseases of other crops such as wheat, soybean, and potatoes etc. were studied (reviewed in Luck et al., 2011). Climate change could make crop disease management more challenging in the developing countries. To our knowledge, no studies have been conducted to understand the effect(s) of climate change on mungbean diseases, and thus future research should address this. Induced resistance due to bio-stimulants has been explored for a few diseases including Macrophomina blight and therefore, additional research is required to exploit induced resistance to manage anthracnose, powdery mildew, Cercospora leaf spot and root rot diseases. Regarding biological control, investigation has been focused for root rot pathogens using strains of *Trichoderma, Pseudomonas*, and *Bacillus* as seed dresser and soil application. Biocontrol agents were more effective in reducing diseases in controlled environments than in fields. Plant-based products as described understory have been extensively researched for the control of seed/soil borne and foliar pathogens, but few have yet reached the market due to lack of their large scale trials at field level. Use of genomics tools has opened avenue to understand the mode of actions of biocontrol agents and genes associated with it (Sharma et al., 2017), however, more research is required in this area. Coordinated approaches from researchers from the universities, private sectors, national and international research centers are required to evaluate promising biocontrol agents, biostimulants, and botanical products in fields at multilocations and commercialize these products. Compatibility between different products including fungicides and these bioagents should be also evaluated. Persistent efforts are required for refinement, validation, transfer and adoption of the integrated disease management modules by the mungbean growers.

AUTHOR CONTRIBUTIONS

AP lead author in reviewing the literature, compiling the information preparing the reveiw draft and revising the manuscript. RB substantial contribution in writing the manuscript from the begining of the manuscript draft. Guided lead author to outline the sections and compile the manuscript. Critically reviewed and revised the manuscript, resturctured the entire manuscript with signifcant contribution to shape the manuscript for the final version. LK crtically reviewed the manuscript and contributed to rewrite and restructure the manuscript. Substanital contribution to revise the all sections of manuscript including abstract, prospects and conclusion. RN guided lead author to compile the manuscript, Significant contribution to revise the manuscript and contributed to write the Host Resistance section of the manuscript.

ACKNOWLEDGMENTS

Funding for this review was provided by core donors to the World Vegetable Center: Republic of China (Taiwan), UK aid, United States Agency for International Development (USAID), Australian Centre for International Agricultural Research (ACIAR) through ACIAR Project on International Mungbean Improvement Network (CIM-2014-079), Germany, Thailand, Philippines, Korea, and Japan.

REFERENCES

Abdel-Kader, M. M., El-Mougy, N. S., and Lashin, S. M. (2013). Biological and chemical resistance inducers approaches for controlling foliar diseases of some vegetables under protected cultivation system. *J. Plant Pathol. Microb.* 4:200. doi: 10.4172/2157-7471.1000200

Acosta-Gallegos, J. A., Kelly, J. D., and Gepts, P. (2008). Prebreeding in common bean and use of genetic diversity from wild germplasm. *Crop Sci.* 48, 3–16. doi: 10.2135/cropsci2007.04.0008IPBS

Akhtar, J., Lal, H. C., Kumar, Y., Singh, P. K., Ghosh, J., Khan, Z., et al. (2014). Multiple disease resistant in greengram and blackgram germplasm and management through chemicals under rain-fed conditions. *Legume Res.* 37, 101–109. doi: 10.5958/j.0976-0571.37.1.016

Alam, S. S., Qureshi, S. H., and Bashir, M. (1985). A report on web blight of mungbean in Pakistan. *Pakistan J. Bot.* 17:165.

Anderson, T. R. (1985). Root rot and wilt of mungbean in Ontario. *Can. Plant Dis. Surv.* 65, 3–6.

Anonymous (2014). *Project Coordinator's Report (mungbean and urdbean), All India Coordinated Research Project on MULLaRP*. Indian Institute of Pulses Research. Kanpur, 2013–2014.

Anonymous (2000). *Annual Report-AICRP on Pulses*. Jaipur: Rajasthan Agricultural Research Institute (Sri Karan Narendra Agriculture University, Jobner), 1–23.

AVRDC (1979). *Asian Vegetable Research and Development Centre.* Mungbean report for 1975, (Tianan). 18.

AVRDC (1980). *Asian Vegetable Research and Development Center*. AVRDC Progress Report 1980, (Shanhua).

AVRDC (1981). *Asian Vegetable Research and Development Center.* AVRDC Progress Report 1981 (Shanhua).

AVRDC (2012). Asian Vegetable Research and Development Center, Mungbean, AVRDC Progress Report 2012, (Shanhua).

Bains, S. S., Dhaliwal, H. S., and Basandrai, A. K. (1988). A new blight of mung and mash in Punjab. *Ann. Biol. Ludhiana* 4, 113–114.

Bashir, M., Alam, S. S., Qureshi, S. H., and Malik, B. A. (1985). Control of mungbean anthracnose by foliar fungicides. *Pakistan J. Agric. Res.* 6, 173–175.

Bashir, M., and Malik, B. A. (1988). Diseases of major pulse crops in Pakistan—a review. *Trop. Pest Manage.* 34, 309–314. doi: 10.1080/09670878809371262

Bhaskar, A. V. (2017). Genotypes against major diseases in green gram and black gram under natural field conditions, A. Vijaya Bhaskar. *Int. J. Curr. Microbiol. App. Sci.* 6, 832–843. doi: 10.20546/ijcmas.2017.606.098

Bhat, F. A., Bhat, G. N., Anwar, A., and Mohiddin, F. A. (2015). Cost effective strategies for the management of fungal leaf spot of greengram caused by *Cercospora canescens* Ell. & Mart. under temperate condition of Jammu and Kashmir. *Legume Res.* 38, 109–114. doi: 10.5958/0976-0571.2015.00018.1

Bhat, F. A., Mohiddin, F. A., and Bhat, H. A. (2014). Reaction of green gram (*Vigna radiata*) to *Cercospora canascens* (ELL.) and Mart. *Indian J. Agric. Res.* 48, 140–144. doi: 10.5958/j.0976-058X.48.2.023

Chaitieng, B., Kaga, A., Han, O. K., Wang, X., Wongkaew, S., Laosuwan, P., et al. (2002). Mapping a new source of resistance to powdery mildew in mungbean. *Plant Breed.* 121, 521–525. doi: 10.1046/j.1439-0523.2002.00751.x

Chakraborty, S., Tiedemann, A. V., and Teng, P. S. (2000). Climate change: potential impact on plant diseases. *Environ. Pollut.* 108, 317–326. doi: 10.1016/S0269-7491(99)00210-9

Chand, R., Singh, V., Pal, C., Kumar, P., and Kumar, M. (2012). First report of a new pathogenic variant of *Cercospora canescens* on mungbean (*Vigna radiata*) from India. *New Dis. Rep.* 26:6. doi: 10.5197/j.2044-0588.2012.026.006

Chang, K. F., Ahmed, H. U., Hwang, S. F., Gossen, B. D., Strelkov, S. E., Blade, S. F., et al. (2007). Sensitivity of field populations of *Ascochyta rabiei* to chlorothalonil, mancozeb, and pyraclostrobin fungicides, and effects of strobilurin fungicides on the progress of ascochyta blight of chickpea. *Can. J. Plant Sci.* 87, 937–944. doi: 10.4141/CJPS07019

Chankaew, S., Somta, P., Isemura, T., Tomooka, N., Kaga, A., Vaughan, D. A., et al. (2013). Quantitative trait locus mapping reveals conservation of major and minor loci for powdery mildew resistance in four sources of resistance in mungbean [*Vigna radiata*(L.) Wilczek]. *Mol. Breed.* 32, 121–130. doi: 10.1007/s11032-013-9856-6

Chankaew, S., Somta, P., Sorajjapinun, W., and Srinives, P. (2011). Quantitative trait loci mapping of Cercospora leaf spot resistance in mungbean, *Vigna radiata* (L.) Wilczek. *Mol. Breed.* 28, 255–264 doi: 10.1007/s11032-010-9478-1

Choudhary, S., Choudhary, A. K., and Sharma, O. P. (2011). Screening of mungbean (*Vigna radiata*) genotypes to identify source of resistant to dry root rot. *J. Food Leg.* 24, 117–119.

Clarry, S. (2016). *The Rise and Rise of Mungbeans.* In GroundCover™ *Supplement Issue 125 November-December.* Grains Research and Development Corporation, 15. Available online at: https://grdc.com.au/resources-and-publications/groundcover/ground-cover-supplements/ground-cover-issue-125-pulse-breeding-advances/the-rise-and-rise-of-mungbeans. (Accessed Apr 4, 2018).

Cowan, M. M. (1999). Plant products as antimicrobial agents. *Clin. Microbiol. Rev.* 12, 564–582.

Das, K., Tiwari, R. K. S., and Shrivastava, D. K. (2010). Techniques for evaluation of medicinal plant products as antimicrobial agent: current methods and future trends. *J. Med. Plants Res.* 4, 104–111. doi: 10.5897/JMPR09.030

Das, S. R., and Narain, A. (1990). Management of powdery mildew of mungbean with fungicides. *Indian Phytopath.* 43, 100–101.

Deeksha, J., and Tripathi, H. S. (2002). Cultural, biological and chemical control of anthracnose of urdbean. *J. Mycol. Plant Pathol.* 32, 52–55.

Deshmukh, M. A., Gade, R. M., Belkar, Y. K., and Koche, M. D. (2016). Efficacy of bioagents, biofertilizers and soil amendments to manage root rot in greengram. *Legume Res.* 39, 140–144. doi: 10.18805/lr.v0iOF.6772

Dubey, S. C. (2003). Integrated management of web blight of urd/mung bean by bio-seed treatment *Indian Phytopath.*56, 34–38.

Dubey, S. C., Bhavani, R., and Singh, B. (2009). Development of Pusa 5SD for seed dressing and Pusa Biopellet 10G for soil application formulations of *Trichoderma harzianum* and their evaluation for integrated management of dry root rot of mungbean (*Vigna radiata*). *Bio. Contr.* 50, 231–242. doi: 10.1016/j.biocontrol.2009.04.008

Dubey, S. C., Bhavani, R., and Singh, B. (2011). Integration of soil application and seed treatment formulations of *Trichoderma* species for management of wet root rot of mungbean caused by *Rhizoctonia solani*. *Pest Manag. Sci.* 67, 1163–1168. doi: 10.1002/ps.2168

Dubey, S. C., and Patel, B. (2002). Mass multiplication of antagonists and standardization of effective dose for management of web blight of urd and mung bean. *Indian Phytopath.* 55, 338–341.

Dutta, S., Singh, R. P., and Jindal, J. K. (2005). Effect of antagonistic bacteria and plant defence activators on management of bacterial leaf spot of mungbean. *Indian Phytopath.* 58, 269–275.

Dwivedi, R. P., and Saksena, H. K. (1974). Occurrence of web blight caused by *Thanatephorus cucumeris* on mungbean. *Int. J. Farm Sci.* 2:100.

Eastburn, D., Degennaro, M., Delucia, E., Dermody, O., and McElrone, A. (2010). Elevated atmospheric carbon dioxide and ozone alter soybean diseases at SoyFACE. *Global Change Biol.* 16, 320–330. doi: 10.1111/j.1365-2486.2009.01978.x

Ebenezar, E. G., and Yesuraja, I. (2000). Effect of fungal and bacterial antagonist on *Macrophomina phaseolina* (Tassi) Goid causing root rot in green gram. *J. Trop. Agric.* 38, 73–76.

Ehteshamul-Haque, S., Abid, M., and Ghaffar, A. (1995). Efficacy of *Bradyrhizobium* sp., and *Paecilomyces lilacinus* with oil cakes in the control of root rot of mungbean. *Trop. Sci.* 35, 294–299.

Farahani, A. S., and Taghavi, M. (2016). Changes of antioxidant enzymes of mung bean [*Vigna radiata* (L.) R. *Wilczek*] in response to host and non-host bacterial pathogens. *J. Plant Prot. Res.* 56, 95–99. doi: 10.1515/jppr-2016-0016

Felde, Z. A., Pocasangre, L. E., Carnizares Monteros, C. A., Sikora, R. A., Rosales, F. E., and Riveros, A. S. (2006). Effect of combined inoculations of endophytic fungi on the biocontrol of *Radopholus similis*. *Info-Musa* 15, 12–18.

Fery, F. L. (2002). "New opportunities in *Vigna*," in *Trends in New Crops and New Uses,* eds J. Janick and A. Whipkey (Alexandria: ASHS Press), 424–428.

Flexner, J. L., and Belnavis, D. L. (2000). "Microbial Insectcides," in *Biological and Biotechnological Control of Insect Pests*, eds J. E. Rechcigl and N. A. Rechcigl (Boca Raton, FL: CRC Press; Lewis Publishers), 35–56.

Gawande, V. L., and Patil, J. V. (2003). Genetics of powdery mildew (*Erysiphe polygoni* D.C.) resistance in Mungbean (*Vigna radiata* (L.) Wilczek). *Crop Prot.* 22, 567–571. doi: 10.1016/S0261-2194(02)00202-8

Hartman, G. L., Wang, T. C., and Kim, D. (1993). Field evaluation of mungbeans for resistance to Cercospora leaf spot and powdery mildew. *Int. J. Pest Manag.* 39, 418–421. doi: 10.1080/09670879309371833

Hibberd, J.,Whitbread, R., and Farrar, J. (1996). Effect of 700 μmol mol^{-1} CO_2 and infection by powdery mildew on the growth and carbon partitioning of barley. *New Phytol.* 134, 309–315. doi: 10.1111/j.1469-8137.1996.tb04635.x

Humphry, S. M. E, Magner, T., McIntyre, C. L., Aitken, E. A., and Liu, C. L. (2003). Identification of major locus conferring resistance to powdery mildew (*Erysiphe polygoni* D.C.) in mungbean (*Vigna radiata* L. *Wiczek)* by QTL analysis. *Genome* 46, 738–744. doi: 10.1139/g03-057

Ikram, N., Dawar, S., Abbas, Z., and Zaki, M. J. (2010). Effect of ($_{60}$Co) gamma rays on growth and root rot diseases in mungbean (*Vigna radiata* L.). *Pakistan J. Bot.* 42, 2165–2170.

Iqbal, S. M., Ghafoor, A., Bashir, M., and Malik, B. A. (1995). Estimation of losses in yield components of mugbean due to *Cercospora* leaf spot. *Pakistan J. Phytopath.* 7, 80–81.

Iqbal, S. M., and Zubair, M., and Haqqani, A. M. (2004). Resistant in Mungbean to *Cercospora* leaf spot disease. *Int. J. Agric. Biol.* 6, 792–793.

Iqbal, U., and Mukhtar, T. (2014). Morphological and pathogenic variability among *Macrophomina phaseolina* isolates associated with mungbean (*Vigna radiata* L.) Wilczek from Pakistan. *Sci. World J.* 2014:950175. doi: 10.1155/2014/950175

Jadhav, M. L., Taur, N., Sapkal, S., Tathe, S., and Quadri, F. (2016). Study on effect of caffeine on growth of *Vigna radiata* L. *Int. J. Adv. Res.* 4, 596–602.

Javaid, A., and Amin, M. (2009). Antifungal activity of methanol and *n*-hexane extracts of three *Chenopodium* species against *Macrophomina phaseolina*. *Nat. Prod. Res.* 23, 1120–1127. doi: 10.1080/14786410802617433

Javaid, A., and Saddique, A. (2011). Management of *Macrophomina* root rot of mungbean using dry leaves manure of *Datura metel* as soil amendment. *Spanish J. Agric. Res.* 9, 901–905. doi: 10.5424/sjar/20110903-394-10

Jhamaria, S. L., and Sharma, O. P. (2002). Management of web blight of mungbean through chemicals and plant product. *Indian Phytopath.* 55, 526.

Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Joshi, A., Souframanien, J., Chand, R., and Pawar, S. E. (2006). Genetic diversity study of *Cercospora canescens* (Ellis & Martin) isolates, the pathogen of Cercospora leaf spot in legumes. *Curr. Sci.* 90, 564–568.

Juroszek, P., and von Tiedemann, A. (2011). Potential strategies and future requirements for plant disease management under a changing climate. *Plant Pathol.* 60, 100–112. doi: 10.1111/j.1365-3059.2010.02410.x

Kapadiya, H. J., and Dhruj, L. U. (1999). Management of mungbean *Cercospora* leaf spot through fungicides. *Indian Phytopath.* 52, 96–97.

Kasettranan, W., Somta, P., and Srinives, P. (2009). Genetics of the resistance to powdery mildew disease in mungbean (*Vigna radiata* (L.) Wilczek). *J. Crop Sci. Biotechnol.* 12, 37–42. doi: 10.1007/s12892-008-0074-4

Kasettranan, W., Somta, P., and Srinives, P. (2010). Mapping of quantitative trait loci controlling powdery mildew resistance in mungbean (*Vigna radiata* (L.) Wilczek). *J. Crop Sci. Biotechnol.* 13, 155–161 doi: 10.1007/s12892-010-0052-z

Kaur, L. (2007). Multiple disease resistant sources of mungbean. *Act Hort.* 752, 423–426. doi: 10.17660/ActaHortic.2007.752.76

Kaur, L., Singh, P., and Sirari, A. (2011). Biplot analysis for locating multiple disease resistant diversity in mungbean germplasm. *Disease Res.* 26, 55–60.

Kaushik, C. D., Chand, J. N., and Saryavir (1987). Seedborne nature of *Rhizoctonia bataticola* causing leaf blight of mungbean. *J. Mycol. Plant Pathol.* 17, 154–157.

Keatinge, J. D. H., Easdown, W. J., Yang, R. Y., Chadha, M. L., and Shanmugasundaram, S. (2011). Overcoming chronic malnutrition in a future warming world: the key importance of mungbean and vegetable soybean. *Euphytica* 180, 129–141. doi: 10.1007/s10681-011-0401-6

Kelly, L. (2017). *Fusarium Species Associated With Grain Sorghum and Mungbean in Queensland*. MS thesis, The University of Queensland.

Kelly, L, White, J., Sharman, M., Brier, H., Williams, L., Grams, R., et al. (2017). *Mungbean and Sorghum Disease Update*. Grains Research and Development Corporation (GRDC) update paper. Available online at: https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2017/07/mungbean-and-sorghum-disease-update. (Assessed Apr 4, 2018).

Keneni, G., Bekele, E., Getu, E., Imtiaz, M., Damte, T., and Mulatu, B. (2011). Breeding food legumes for resistance to storage insect pests: potential and limitations. *Sustainability* 3, 1399–1415. doi: 10.3390/su3091399

Khajudparn, P., Wongkaew, S., and Thipyapong, P. (2007). Mungbean powdery resistant identification of genes for resistant to powdery mildew in mungbean. *Afr. Crop Sci. Conf. Proc.* 8, 743–745.

Khan, A. A., Khan, R. U., and Singh, R. (2005). Management of *Cercospora* leaf spot and anthracnose diseases of mungbean by fungicides. *Ann. Plant Prot. Sci.* 13, 465–529.

Khan, K. S. H., and Shuaib, M. (2007). Identification of sources of resistant in mungbean (*Vigna radiata* L.) against charcoal rot *Macrophomina phaseolina* (Tassi) Goid. *Afr. Crop Sci. Conf. Proc.* 8, 2101–2102

Khare, N., Lankpale, N., and Agarwal, K. C. (1998). Epidemiology of powdery mildew of mungbean in Chhattisgarh region of Madhya Pradesh. *J. Mycol. Plant Pathol.* 28, 5–10.

Khunti, J. P., Bhoraniya, M. E., and Vora, V. D. (2005). Management of powdery mildew and *Cercospora* leaf spot of mungbean by some systemic fungicides. *Legume Res.* 28, 65–67.

Khunti, J. P., Bhoraniya, M. F., and Vora, V. D. (2002). Management of Powdery mildew and *Cercospora* leaf spot of mungbean by some Systemic Fungicides. *J. Mycol. Pl. Pathol.* 32:103.

Kobayashi, T., Ishiguro, K., Nakajima, T., Kim, H. Y., Okada, M., and Kobayashi, K. (2006). Effects of elevated atmospheric CO_2 concentration on the infection of rice blast and sheath blight. *Phytopathology* 96, 425–431. doi: 10.1094/PHYTO-96-0425

Kulkarni, S. A. (2009). *Epidemiology and Integrated Management of Anthracnose of Green Gram*. M.Sc. thesis, University of Agricultural Sciences, Dharwad.

Kumar, A., Adarsha, H. S., Shanthala, J., and Savithramma, D. L. (2017). Differential response of F4 and F5 green gram [*Vigna radiata* (L.) Wilczek] recombinant inbred lines (RILs) to powdery mildew infection. *J. Pharmacognosy Phytochem.* 6, 1147–1153.

Kumari, R., Shekhawati, K. S., Gupta, R., and Khokhar, M. K. (2012). Integrated management against root rot of mungbean (*Vigna radiata* L. *Wilczek)* incited by *Macrophomina phaseolina*. *J. Plant Pathol. Microbiol.* 3, 1–5. doi: 10.4172/2157-7471.1000136

Lee, J. H., Han, K. S., Kim, T. H., Bae, D. W., Kim, D. K., Kang, J. H., et al. (2007). Effective heat treatment techniques for control of mungbean sprout rot, incorporable into commercial mass production. *Plant Pathol. J.* 23, 174–179. doi: 10.5423/PPJ.2007.23.3.174

Lee, Y. B. (1980). *Inheritance Study on Resistance to Cercospora Leaf Spot in Mungbean*. Shanhua: Asian Vegetable Research and Development Center.

Lonergan, E., Pasche, J., Skoglund, L., and Burrows, M. (2015). Sensitivity of *Ascochyta* Species Infecting Pea, Lentil, and Chickpea to Boscalid, Fluxapyroxad, and Prothioconazole. *Plant Dis.* 99, 1254–1260. doi: 10.1094/PDIS-06-14-0620-RE

Luck, J., Spackman, M., Freeman, A., Tre bicki, P., Griffiths, W., Finlay, K., et al. (2011). Climate change and diseases of food crops. *Plant Pathol.* 60, 113–121. doi: 10.1111/j.1365-3059.2010.02414.x

Maheshwari, S. K., and Krishna, H. (2013). Field efficacy of fungicides and bio-agents against *Alternaria* leaf spot of mungbean. *Ann. Plant Prot. Sci.* 21, 364–367.

Mandhare, V. K., and Suryawanshi, A. V. (2008). Dual resistant against powdery mildew and yellow mosaic virus in greengram. *Agric. Sci. Digest* 28, 39–41.

Mansoor, F., Sultana, V., and Ehteshamul-Haque, S. (2007). Enhancement of biocontrol potential of *Pseudomonas aeruginosa* and *Paecilomyces lilacinus* against root rot of mungbean by a medicinal plant *Launaea nudicaulis* L. *Pakistan J. Bot.* 39, 2113–2119.

Marappa, N. (2008). Screening of mungbean genotypes and its wild relatives for resistant sources to Cercospora leaf spot disease. *Asian J. BioScience* 3, 324–326.

Meehl, G. A., Washington, W. M., Collins, W. D., Arblaster, J. M., Hu, A., Buja, L. E., et al. (2005). How much more global warming and sea level rise? *Science* 307, 1769–1772. doi: 10.1126/science.1106663

Melloy, P., Hollaway, G., Luck, J., Norton, R., and Aitken, E. S. C. (2010). Production and fitness of *Fusarium pseudograminearum* inoculum at elevated carbon dioxide in FACE. *Global Change Biol.* 16, 3363–3373. doi: 10.1111/j.1365-2486.2010.02178.x

Meyer, S. L. F., and Roberts, D. P. (2002). Combinations of biocontrol agents for management of plant-parasitic nematodes and soil borne plant-pathogenic fungi. *J. Nematol.* 34, 1–8.

Mishra, S. P., Asthana, A. N., and Yadav, L. (1988). Inheritance of Cercospora leaf spot resistance in mungbean, *Vigna radiata* (L.) Wilczek. *Plant Breed.* 100, 228–229. doi: 10.1111/j.1439-0523.1988.tb00245.x

Murugapriya, E., Alice, D., and Jayamani, P. (2011). Antifungal activity of botanicals and micro-nutrients against *Macrophomina* leaf blight in mungbean. *J. Food Leg.* 24, 113–116.

Naimuddin, A. M., and Singh, N. P. (2016). Yellow mosaic of mungbean and urdbean: current status and future strategies. *J. Food Leg.* 29, 77–93.

Nair, R. M., Götz, M., Winter, S., Giri, R. R., Boddepalli, V. N., Sirari, A., et al. (2017). Identification of mungbean lines with tolerance or resistance to yellow mosaic in fields in India where different begomovirus species and different *Bemisia tabaci* cryptic species predominate. *Eur. J. Plant Pathol.* 149, 349–365. doi: 10.1007/s10658-017-1187-8

Nair, R. M., Schalfleitner, R., Kenyon, L., Srinivasan, R., Easdown, W., Ebert, A., et al. (2012). Genetic improvement of mungbean. *SABRAO J. Breed. Gen.* 44, 177–190.

Nair, R. M., Thavarajah, D., Thavarajah, P., Giri, R. R., Ledesma, D., Yang, R. Y., et al. (2015). Mineral and phenolic concentrations of mungbean [*Vigna radiata* (L.) R. Wilczek var. radiata] grown in semi-arid tropical India. *J. Food Compost. Anal.* 39, 23–32. doi: 10.1016/j.jfca.2014.10.009

Pandey, A. K., and Tripathi, N. N. (2011). Aromatic plants of gorakhpur division: their anti-mycotic properties and medicinal value. *Int. J. Pharm. Sci. Rev. Res.* 7, 142–147.

Pandey, S., Sharma, M., Kumari, S., Gaur, P. M., Chen, W., Kaur, L., et al. (2009). "Integrated foliar diseases management of legumes," in *International Conference on Grain Legumes: Quality Improvement, Value Addition and Trade* (Kanpur: Indian Society of Pulses Research and Development, Indian Institute of Pulses Research), 143–161.

Price, P. P., Purvis, M. A., Cai, G., Padgett, G. B., Robertson, C. L., Schneider, R. W., et al. (2015). Fungicide resistance in *Cercospora kikuchii*, a soybean pathogen. *Plant Dis.* 99, 1596–1603. doi: 10.1094/PDIS-07-14-0782-RE

Quebral, F. C., and Cowell, R. (1978). "Powdery mildew and *Cercospora* leaf spot of green gram in philippines," in *First International Greengram Symposium*, ed R. A. Cowell (Taiwan: The office of the Information Services), 147–148.

Raghuchander, T., Samiyappan, R., and Arjunan, G. (1993). Biocontrol of Macrophomina root rot of mungbean. *Indian Phytopath.* 46, 379–382.

Rakhonde, P. N., Koche, M. D., and Harne, A. D. (2011). Management of powdery mildew of green gram. *J. Food Leg.* 24, 120–122.

Ramakrishnan, C. K. D., and Savithramma, D. L. (2014). Screening of mungbean germplasm for powdery mildew disease Resistant. *Int. J. Agronomy Agric. Res.* 4, 16–21

Rana, M., Ashiq, R., and Virender, A. (2014). Systemic opposition of green gram against leaf blight caused by *Macrophomina phaseolina*. *Int. J. Mycol. Plant Pathol.* 1, 72–77.

Rashid, A., Harris, D., Hollington, P., and Ali, S. (2004). On-farm seed priming reduces yield losses of mungbean (*Vigna radiata*) associated with mungbean yellow mosaic virus in the North West Frontier Province of Pakistan. *Crop Prot.* 23, 1119–1124. doi: 10.1016/j.cropro.2004.04.002

Rathore, B. S. (2006). Management of diseases of greengram with fungicides. *Indian J. Mycol. Plant Pathol.* 36, 138–141.

Reddy, K. S. (2009). Identification and inheritance of a new gene for powdery mildew resistance in mungbean (*Vigna radiata* L. Wilczek). *Plant Breed.* 128, 521–523. doi: 10.1111/j.1439-0523.2008.01609.x

Reddy, K. S., Kao, K. C., and Reddy, M. S. (1992). Evaluation of some new fungicides against Rhizoctonia solani Kühn, the incident of damping-off in mungbean. *Indian J. Plant Protect.* 20, 37–42.

Reddy, K. S., Pawar, S. E., and Bhatia, C. R. (1987). Screening for powdery mildew (*Erysiphe polygoni* DC) resistance in mungbean (*Vigna radiata* L. Wilzek) using excised leaves. *Proc. Indian Acad. Sci. U.S.A.* 97, 365–369.

Reddy, K. S., Pawar, S. E., and Bhatia, C. R. (1994a). Inheritance of powdery mildew (Erysiphe polygoni D.C.) resistance in mungbean (Vigna radiata L. *Wilczek). Theor. Appl. Gent*. 88, 945–948.

Reddy, K. S., Pawar, S. E., Wanjari, K. B., and Bhatia, C. R. (1994b). "Development of powdery mildews resistant and high yielding varieties of mungbean," in: *International Symposium on Pulses Research* (New Delhi), 146–147.

Ryley, M., Toowoomba, and Tatnell, J. (2010). *Management of the Major Foliar Diseases of Mungbeans and Peanuts in Australia*. Kingaroy: Agri-Science Qld © The State of Queensland, Department of Employment, Economic Development and Innovation.

Saravanakumar, D., Harish, S., Loganathan, M., Vivekananthan, R., Rajendran, L., Raguchander, T., et al. (2007). Rhizobacterial bioformulation for the effective management of *Macrophomina* root rot in mungbean. *Arch. Phytopathol. Plant Prot.* 40, 323–337. doi: 10.1080/03235400600587326

Satya, V. K., Vijayasamundeeswari, A., Paranidharan, V., and Velazhahan, R. (2011). *Burkholderia* sp. Strain tnau-1 for biological control of root rot in mungbean (*Vigna radiata* l.) caused by *Macrophomina phaseolina*. *J. Plant Prot. Res.* 51, 273–278. doi: 10.2478/v10045-011-0045-5

Satyagopal, K., Sushil, S. N., Jeyakumar, P., Shankar, G., Sharma, O. P., et al. (2014). *AESA based IPM package for Blackgram and Greengram*. Hyderabad. 43.

Schafleitner, R., Nair, R. M., Rathore, A., Wang, Y., Lin, C., Chu, S., et al. (2015). The AVRDC – The World Vegetable Center mungbean (*Vigna radiata*) core and minicore collections. *BMC Genomics* 16:344. doi: 10.1186/s12864-015-1556-7

Sconyers, L. E., Kemerait, R. C., , Brock, J., Phillips, D. V., Jost, P. H., Sikora, E. J., et al. (2006). *Asian Soybean Rust Development in 2005: A Perspective from the Southeastern United States*. APSnet Features.

Shahbaz, M. U., Iqbal, M. A., Rafiq, M., Batool, A., and Kamran, M. (2014). Efficacy of different protective fungicides against *Cercospora* leaf spot of mungbean (*Vigna radiata* l. Wilczek). *Pakistan J. Phytopathol.* 26, 187–191.

Sharma, M., and Ghosh, R. (2017). "Heat and soil moisture stress differentially impact chickpea plant infection with fungal pathogens," in *Plant Tolerance to Individual and Concurrent Stresses*, ed M. Senthil-Kumar (New Delhi: Springer), 47–57.

Sharma, M., Ghosh, R., and Pande, S. (2015). Dry root rot (*Rhizoctonia bataticola* (Taub.) Butler): an emerging disease of chickpea – where do we stand? *Arch. Phytopathol. Plant Prot.*48, 797–812. doi: 10.1080/03235408.2016.1140564

Sharma, M., Tarafdar, A., Ghosh, R., and Gopalakrishanan, S. (2017). "Biological control as a tool for eco-friendly management of plant pathogens," in *Advances in Soil Microbiology: Recent Trends and Future Prospects, Microorganisms for Sustainability*, eds T. K. Adhya, B. B. Mishra, K. Annapurna, D. K. Verma, and U. Kumar (Springer), 153–186.

Sharma, O. P., Bambawale, O. M., Gopali, J. B., Bhagat, S., Yelshetty, S., Singh, S. K., et al. (2011). *Field Guide Mungbean and Urdbean*. New Delhi: National Centre for Integrated Pest Management. M/s Royal Offset Printers A-89/1, Naraina Industrial Area, Phase-I. 40.

Sharma, R., Duveiller, E., and Ortiz-Ferrara, G. (2007). Progress and challenge towards reducing wheat spot blotch threat in the Eastern Gangetic Plains of South Asia: is climate change already taking its toll? *Field Crops Res.* 103, 109–118. doi: 10.1016/j.fcr.2007.05.004

Shen, Y. M., Liu, H. L., Chang, S. T., and Chao, C. H. (2010). First report of anthracnose caused by *Colletotrichum acutatum* on mungbean sprouts in Taiwan. *Plant Dis.* 94:131. doi: 10.1094/PDIS-94-1-0131C

Shukla, V., Baghel, S., Maravi, K., and Singh, S. K. (2014). Yield loss assessment in mungbean [*Vigna radiata* (L.) Wilczek] caused by anthracnose [*Colletotrichum truncatum* (schw.) Andrus and moore]. *Bioscan* 9, 1233–1235.

Siddiqui, A., Ehetshamul-Haque, E., and Shaukat, S. S. (2001). Use of *Rhizobacteria* in the control of root rot–root knot disease complex of mungbean. *J. Phytopathol.* 149, 337–346. doi: 10.1046/j.1439-0434.2001.00630.x

Sierotzki, H., and Scalliet, G. (2013). A review of current knowledge of resistance aspects for the next-generation succinate dehydrogenase inhibitor fungicides. *Phytopathology* 103, 880–887. doi: 10.1094/PHYTO-01-13-0009-RVW

Singh, G., Sharma, Y. R., Shanmugasundaram, S., Shih, S. L., and Green, S. K. (2004). "Improving income and nutrition by incorporating mungbean in cereal fallows in the Indo-Gangetic Plains of South Asia DFID Mungbean Project for 2002-2004," in *Proceedings of the Final Workshop and Planning Meeting, Status of Mung Bean Yellow Mosaic Virus Resistance Breeding.* Ludhiana: Punjab Agricultural University, 27-31 May 2004, 204–213.

Singh, J., Mishra, K. K., and Singh, A. K. (2013a). Current status of web blight of mungbean. *Asian J. Soil Sci.* 8, 495–504.

Singh, S., Chand, H., and Varma, P. K. (2008). Screening of bioagents against root rot of mungbean caused by *Rhizoctonia solani*. *Legume Res.* 31, 75–76.

Singh, S. P., and Singh, S. K. (2014). Sources of resistant in mungbean for *Cercospora* leaf spot diseases. *Ann. Agric. Biosci. Res.* 2, 280–281.

Singh, S. P., Singh, S. K., and Shukla, V. (2013b). Fungicidal management of Cercospora leaf spot of mungbean (*Vigna radiata*). *Trends Biosciences* 6, 861–863.

Slusarenko, A. J., Patel, A., and Portz, D. (2008). Control of plant diseases by natural products: allicin from garlic as a case study. *Eur. J. Plant Pathol.* 121, 313–322. doi: 10.1007/s10658-007-9232-7

Sorajjapinun, W., Rewthongchum, S., Koizumi, M., and Srinives, P. (2005). Quantitative inheritance of resistance to powdery mildew disease in mungbean (*Vigna radiata* (L.) Wilczek). *SABRAO J. Breed. Gent.* 37, 91–96.

Sujatha, K., Kajjidoni, S. T., Patil, P. V., and Somashekhar, G. (2011). Heterosis for productivity related traits involving diverse parents for powdery mildew reaction in mungbean. *J. Food Legumes.* 24, 101–105.

Sundaramoorthy, S., Murugapriya, E., Maharaja, L. G. J., and Alice, D. (2013). Induction of systemic resistant in green gram against leaf blight caused by *Macrophomina phaseolina* (Tassi.) Goid. *Afr. J. Microbiol. Res.* 7, 3976–3982. doi: 10.5897/AJMR2013.5719

Suryawanshi, A. P., Wadje, A. G., Gawade, D. B., Kadam, T. S., and Pawar, A. K. (2009). Field evaluation of fungicides and botanicals against powdery mildew of mungbean. *Agric. Sci. Digest* 29, 209–211.

Tandel, D. H., Sabalpara, A. N., Patel, R. C., Prajapati, V. R., and Patel, V. R. (2014). Antagonistic effect on growth and sclerotial formation of *Macrophomina phaseolina* (Tassi) goid. causing green gram leaf blight. *J Life Sci.* 11, 89–91

Thilagavathi, R., Saravanakumar, D., Ragupathi, N., and Samiyappan, R. (2007). A combination of biocontrol agents improves the management of dry root rot (*Macrophomina phaseolina*) in greengram. *Phytopathol. Mediterranea* 46, 157–167. doi: 10.14601/Phytopathol_Mediterr-2147

Thompson, G. B., Brown, J. K. M, and Woodward, F. I. (1993). The effects of host carbon dioxide, nitrogen and water supply on the infection of wheat by powdery mildew and aphids. *Plant Cell Environ.* 16, 687–694. doi: 10.1111/j.1365-3040.1993.tb00487.x

Tiwari, A. (1993). *Studies on the Diseases Caused by Rhizoctonia Solani Kuhn in Green Gram (Phaseolus aureus Roxb).* Ph.D. thesis, Rani Durgavati Vishwavidyalaya (Jabalpur), 338.

Tiwari, A., and Khare, M. N. (1998). Variability among isolates of *Rhizoctonia solani* infecting mungbean. *Indian Phytopath.* 51, 334–337.

Uddin, M. N., Bakr, M. A., Islam, M. R., Hossain, M. I., and Hossain, A. (2013). Bioefficacy of plant extracts to control Cercospora leaf spot of mungbean (*Vigna radiata*). *Int. J. Agric. Res. Innov. Technol.* 3, 60–65. doi: 10.3329/ijarit.v3i1.16094

Vallad, G. E., and Goodman, R. M. (2004). Systemic acquired resistance and induced systemic resistance in conventional agriculture: review and interpretation. *Crop Sci.* 44, 1920–1934. doi: 10.2135/cropsci2004.1920

Veena, Y. R. H., Math, G., and Kumar, A. G. V. (2013). Bioefficacy of fungicides against *Cercospora canescens* causing leaf spot of greengram. *Crop Res.* 46, 74–78.

Viaene, N. M., and Abawi, G. S. (2000). *Hirsutella rhossiliensis* and *Verticillium chlamydospdoiorium* as biocontrol agents of the root-knot nematode *Meloidogyne hapla* on lettuce. *J. Nematol.* 32, 85–100.

Walters, D. R., Paterson, L., Walsh, D. J., and Havis, N. D. (2009). Priming for plant defense in barley provides benefits only under high disease pressure. *Physiol. Mol. Plant Pathol.* 73, 95–100. doi: 10.1016/j.pmpp.2009.03.002

Wongpiyasatid, A., Chotechuen, S., Hormchan, P., and Srihuttagum, M. (1999). Evaluation of yield and resistance to powdery mildew, Cercospora leaf spot and cowpea weevil in mungbean mutant lines. *Kasetsart J.* 33, 204–215.

Yadav, D. L., Jaisani, P., and Pandey, R. N. (2014a). Identification of sources of resistant in mungbean genotypes and influence of fungicidal application to powdery mildew epidemics. *Int. J. Current Microbiol. Appl. Sci.* 3, 513–519.

Yadav, D. L., Pandey, R. N., Jaisani, P., and Gohel, N. M. (2014b). Sources of resistant in mungbean genotypes to Cercospora leaf spot disease and its management. *African, J. Agric. Res.* 9, 3111–3114.

Yadav, R. J., Kakraliya, S. S., Bajiya, M. R., and Abrol, S. (2017). Eco-Friendly management of powdery mildew of green gram (*Vigna radiata* L.). *Int. J. Curr. Microbiol. App. Sci.* 6, 435–439 doi: 10.20546/ijcmas.2017.607.052

Yaqub, F., and Shahzad, S. (2009). Effect of solar heating by polyethylene mulching on sclerotial viability and pathogenicity of *Sclerotium rolfsii* on mungbean and sunflower. *Pakistan J. Bot.* 41, 3199–3205.

Young, N. D., Danesh, D., Menancio-Hautea, D., and Kumar, L. (1993). Mapping oligogenic resistance to powdery mildew in mungbean with RFLPs. *Theor. Appl. Gent.* 87, 243–249. doi: 10.1007/BF00223772

Zhang, J. Q., Zhu, Z. D., Duan, C. X., Wang, X. M., and Li, H. J. (2011). First report of charcoal rot caused by *Macrophomina phaseolina* on mungbean in China. *Plant Dis.* 95:872. doi: 10.1094/PDIS-01-11-0010

Zhimo, V. Y., Panja, B. N., Saha, J., and Nath, R. (2013). Evaluation of mungbean genotypes for resistance against Cercospora leaf spot and Yellow Mosaic disease under field condition. *J. Mycopathol. Res.* 51, 273–278.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer RR and handling editor declared their shared affiliation

frontiers in Environmental Science

ORIGINAL RESEARCH
published: 27 June 2018
doi: 10.3389/fenvs.2018.00063

Model-Based Forecasting of Agricultural Crop Disease Risk at the Regional Scale, Integrating Airborne Inoculum, Environmental, and Satellite-Based Monitoring Data

Nathaniel K. Newlands [1,2]*

[1] Agriculture and Agri-Food Canada, Science and Technology Branch, Summerland Research and Development Centre, Summerland, BC, Canada, [2] Department of Geography, University of Victoria, Victoria, BC, Canada

Crop diseases have the potential to cause devastating epidemics that threaten the world's food supply and vary widely in their dispersal pattern, prevalence, and severity. It remains unclear what the impact disease will have on sustainable crop yields in the future. Agricultural stakeholders are increasingly under pressure to adapt their decision-making to make more informed and efficient use of irrigation water, fertilizers, and pesticides. They also face increasing uncertainty in how best to respond to competing health, environment, and (sustainable) development impacts and risks. Disease dynamics involves a complex interaction between a host, a pathogen, and their environment, representing one of the largest risks facing the long-term sustainability of agriculture. New airborne inoculum, weather, and satellite-based technology provide new opportunities for combining disease monitoring data and predictive models—but this requires a robust analytical framework. Integrated model-based forecasting frameworks have the potential to improve the timeliness, effectiveness, and foresight for controlling crop diseases, while minimizing economic costs and environmental impacts, and yield losses. The feasibility of this approach is investigated involving model and data selection. It is tested against available disease data collected for wheat stripe (yellow) rust (*Puccinia striiformis* f.sp. *tritici*) (Pst) fungal disease within southern Alberta, Canada. Two candidate, stochastic models are evaluated; a simpler, site-specific model, and a more complex, spatially-explicit transmission model. The ability of these models to reproduce an observed infection pattern is tested using two climate datasets with different spatial resolution—a reanalysis dataset (~55 km) and weather station network township-aggregated data (~10 km). The complex spatially-explicit model using weather station network data had the highest forecast accuracy. A multi-scale airborne surveillance design that provides data would further improve disease risk forecast accuracy under heterogeneous modeling assumptions. In the future, a model-based forecasting approach, if supported with an airborne surveillance monitoring plan, could be made operational to provide agricultural stakeholders with reliable, cost-effective, and near-real-time information for protecting and sustaining crop production against multiple disease threats.

Keywords: agriculture, Canada, disease, forecasting, modeling, risk, wheat stripe rust

OPEN ACCESS

Edited by:
Rob Swart, Wageningen Environmental Research, Netherlands

Reviewed by:
Xander Wang, University of Prince Edward Island, Canada
Robert Faggian, School of Life and Environmental Sciences, Deakin University, Australia

***Correspondence:**
Nathaniel K. Newlands
nathaniel.newlands@canada.ca

Specialty section:
This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Environmental Science

Received: *28 January 2018*
Accepted: *07 June 2018*
Published: *27 June 2018*

Citation:
Newlands NK (2018) Model-Based Forecasting of Agricultural Crop Disease Risk at the Regional Scale, Integrating Airborne Inoculum, Environmental, and Satellite-Based Monitoring Data. Front. Environ. Sci. 6:63. doi: 10.3389/fenvs.2018.00063

INTRODUCTION

Crop diseases have the potential to cause devastating epidemics that threaten the world's food supply and vary widely in their dispersal pattern, prevalence, and severity (Chakraborty and Newton, 2011). Diseases like stripe rust (*Puccinia striiformis* f. sp. *tritici)* (Pst) and fusarium head blight (*Fusarium graminearum)* (FHB) on wheat, and powdery mildew (*Erysiphe necator*) on grapes, to highlight just a few, cause major crop losses globally (Hovmøller, 2001; Carisse et al., 2009; Haran et al., 2010; Newberry et al., 2016). Plant breeding to increase host resistance remains the primary approach for managing diseases and to help sustainable agricultural yields, as crop breeding networks that deploy resistance genes decrease the likelihood that pathogens will overcome resistance (Ojiambo et al., 2017). Nonetheless, despite the introduction of crop cultivars/varieties with higher resistance, new disease races, with increased virulence, continue to emerge**.** Environmental conditions affect resistance gene performance, but the basis for this is poorly understood (Bryant et al., 2014). Moreover, environmental drivers and pressures are increasing in their influence over agroecosystems; climate change and variability is raising temperatures and lengthening growing seasons, especially in northern climates (Canada) or temperate zones (Australia), producing more days without frost, and more intense heatwave and rainfall events. Disease dynamics itself involves a complex interaction between a host, a pathogen, and their environment, representing one of the largest integrated risks facing the long-term sustainability of agriculture. Genetic factors (e.g., emergence of new diseases and of new races), environmental-driven influence (e.g., global climate change impacts on disease spread), and management-intervention driven agroecosystem interactions (e.g., crop breeding and monitoring technologies) are all important considerations in disease risk mitigation.

The cost of pesticides (e.g., fungicides, insecticides, herbicides) is a substantial burden for growers—with substantial uncertainty involved in deciding when and how much to apply to commercial fields, especially as multiple diseases often affect crops at the same time. Currently, when monitoring their fields for disease, growers often rely on simple, visual identification, assessing severity using standard area diagrams (SADs), disease progress (AUDPC) curves, and weather/forecast conditions (Contreras-Medina et al., 2009; Nopsa and Pfender, 2014; Ojiambo et al., 2017). Pesticides are then applied either preventatively, or even if no disease is detected, on a calendar-based schedule, based on perceived risk (Carisse et al., 2009). This approach is, however, limited in its ability to detect and control disease. Pesticide application must generally occur during the early stages of epidemics, and at sufficient rates. Over-application is costly and creates added selection pressure for more pesticide tolerant strains, while under-application may also be cost-prohibitive in regions where expected yield is lower (Chen, 2007). Moreover, high pesticide concentrations are not only costly, but are also associated with detrimental environmental and human health impacts (Newlands, 2016). Reducing pesticide use is a major focus of global agricultural sustainability efforts (Nicolopoulou-Stamati et al., 2016). The effectiveness of applications are also highly dependent on timing, stage of disease progression, and the strength and directionality of micro- and meso-scale wind currents (Meyer M. et al., 2017).

Integrated Pest Management (IPM) is the deployment of a variety of methods of pest control designed to complement, reduce, or replace the application of synthetic pesticides. It involves regular monitoring, use of decision thresholds, combining approaches for targeted pesticide management and substitution to broader agroecosystem considerations (Pretty and Bharucha, 2015). For a comprehensive global overview of the history, programs, and adoption of IPM programs around the world, readers are referred to Peshin et al. (2009). They highlight problems with assessing the adoption and success of IPM programs and how pesticide use has not consistently decreased in the majority of programs, despite reduction of pesticide use being one of their primary goals. Here predictive models may not only enable better program assessment and adoption, but also help to identify how to optimize changes in the timing and application of pesticides (e.g., fungicides) in time and space to reduce pesticide use where predicted disease risk is sufficiently low. Disease prediction models using advanced statistical methods (e.g., artificial neural networks) integrating weather and aerobiological monitoring data have been successfully developed and validated for Ganoderma spp. and white blister on Brassica crops (Brassica spot^{TM}) (Minchinton et al., 2013; Sadyś et al., 2016). Such prediction models need to be adapted and extended for other crop diseases and then integrated into operational IPM programs. For many IPM programs, there is a crucial need to develop and involve a more reliable and effective approach (i.e., analytical framework) for managing disease risk that establish relationships between the amount of airborne inoculum and disease development, combined use of models that integrate theoretical knowledge on crop (host) growth, disease (pathogen) development, and environmental influences; alongside data from disease monitoring, climate/weather, and other explanatory variables for assessing and predicting disease (Juroszek and von Tiedemann, 2013; Ojiambo et al., 2017). Past efforts have been hindered by sparse spatial data, limited use of field monitoring technology, and a need for greater integration and quantification. Past efforts have also concentrated mainly on understanding the physical and biological mechanisms of plant (crop) pathogen spore dispersal linked disease development, outbreaks and spatial epidemic patterning and spread. Improved detection of new airborne inoculum, weather, and satellite-based technology, however, provide new opportunities for combining disease monitoring data with predictive models. This has the potential to improve the timeliness, effectiveness and foresight for controlling crop diseases, while minimizing crop loss (Isard et al., 2011; Devadas et al., 2015; West and Kimber, 2015; Mahlein, 2016).

In a recent review of modeling the impact of climate on crop disease, improvements in measuring the uncertainty of climate change projected impacts using multi-model ensembles are highlighted. This synthesis identifies the need to explore other sources of uncertainty inherent in disease models that still remain unexplored and unreported (Newberry et al., 2016). They articulate the crucial need to investigate crop disease

dynamics at the landscape spatial scale, across a broader range of crops and pathogens. This study also identifies the need for new frameworks and models to improve the ability of models to predict impacts of climate change on crop diseases for guiding the planning of climate change adaptation strategies to ensure future food security. Because disease patterns can change in unintended ways when interventions take place across spatially heterogeneous landscapes, Newberry et al. (2016) articulate five key challenges for advancing models: (1) flexibility and accuracy, (2) interaction and contact assumptions and statistical representation, (3) how to define and estimate a critical threshold of transmissibility, while recognizing a mixture of infection types under changing susceptibility, (4) how to model disease dynamics that depends on long-distance interaction (e.g., air transmission via wind trajectories with deposition via rainfall events), and, (5) identifying the natural scale (i.e., operational resolution) for modeling (and forecasting) transmission and interaction, and how this relates to the scale at which intervention is most effective, recognizing that different sources of data are typically available at different scales (Riley et al., 2015).

Model uncertainty and reliability remain two major issues challenging the development of more robust and effective quantitative approaches for disease management. To address these aspects requires an integrated model-based framework and statistical approach that can integrate new types of data, model spatial and temporal dependence and interaction, quantify uncertainties, evaluate multiple scenarios, and bridge empirical-theory knowledge gaps (Held et al., 2004; Contreras-Medina et al., 2009; Haran et al., 2010; Savage and Renton, 2014; Kouadio and Newlands, 2015; Riley et al., 2015; Newlands, 2016; Höhle et al.,, 2017; Ojiambo et al., 2017). Dennis (1987) derived a simple, multivariate regression-based model of Pst disease infection based on air temperature and surface wetness period. Incorporating Monte-Carlo simulation to model uncertainty, El Jarroudi et al. (2017) further incorporated the Dennis disease infection model into a threshold-based weather model to guide fungicide applications for Pst. They show that an optimal combination of high humidity (>92%), temperature (4–16^{o}C) for at least 4 consecutive hours was sufficient to cause an epidemic. Audsley et al. (2005) developed a simulation model integrating it within the *Decision Support System* for Arable Crops (DESSAC) system that uses a genetic algorithm for the selection of fungicide spray plans (Parsons and Te Beest, 2004; Audsley et al., 2005), integrating major risk variables for pathogen (i.e., inoculum source and transfer), host (cultivar-specific resistance, leaf age, nitrogen uptake), and weather conditions (temperature, rain, humidity, wind), Using the simple, multiple regression modeling approach, Kuang et al. (2013) demonstrated a model-based, operational prediction system for wheat stripe rust that integrates geospatial and internet/networking technology to enable multiple users to interact, share data, automate, and update model design, combining regional predictions and testing statistical significance. The inter-comparison of assumptions (models), spatial resolution and uncertainty (climate datasets), spatial correlation/dependence of host-pathogen-environmental interaction, and its effect on model performance and forecast accuracy is lacking.

In this paper, an integrated modeling framework to forecast disease risk is proposed. The feasibility of model-based, operational disease risk forecasting is investigated, using data available for wheat stripe (yellow) rust (*Puccinia striiformis* f.sp. *tritici*) (hereafter, Pst) fungal disease within southern Alberta, Canada. Two candidate, stochastic models are evaluated; a simpler, site-specific model, and a more complex, transmission model. These models are calibrated using airborne inoculum data by a Burkard cyclone spore collector, for the first time. In addition, satellite measurements of major disease risk variables (i.e., canopy temperature and liquid water on the canopy surface) are integrated. The ability of these models to reproduce an observed infection pattern is tested using two climate datasets with different spatial resolution—a reanalysis dataset (∼55 km) and a weather station network township-aggregated data (∼10 km).

MATERIALS AND METHODS

Integrated Modeling Approach

The integrated framework was designed to take into account major aspects and considerations involved in operational model-based forecasting of crop disease at the regional-scale (**Figure 1**). This approach combines data on host, pathogen and environment, and models to capture different aspects of disease dynamics under different assumptions. In this figure, starting and end points are shown with respect to seasonal disease progression. Boxes represent model components—dashed boxes are disease aspects not considered in the current modeling. Full boxes are those that are currently considered. This design is a prototype, and while not exhaustive of all general and pathogen-specific aspects, is based on published scientific studies, evidence, and in consultation with several expert AAFC/Canadian pathologists. This design supports feasibility testing, involving the evaluation of different models, datasets, and forecast metrics. The framework integrates threshold-based infection and multivariate spatial assumptions, extending previous approaches to include a broader, more representative set of disease dynamic parameters, climate covariates, and assumptions on disease-climate interaction and spatial dependence. Its component-wide structure permits scaling-up from homogeneous to heterogeneous assumptions, whereby a region is divided into smaller subregions, with transmission assumed to occur between them and a calibrated model used to capture their specific disease dynamics, susceptibility, incidence, and risk. Further details of the model input and output parameters and variables are provided for a non-spatial/site-specific (CLR) and spatial model (hhh4) formulation in this section.

Wheat Stripe Rust Disease

Stripe (yellow) rust (Pst) is a prevalent fungal disease in all wheat (*Triticum aestivum*) growing regions around the world, occurring in most production zones having cool and moist weather conditions during the growing season (Chen et al., 2014). This disease has the potential to cause devastating

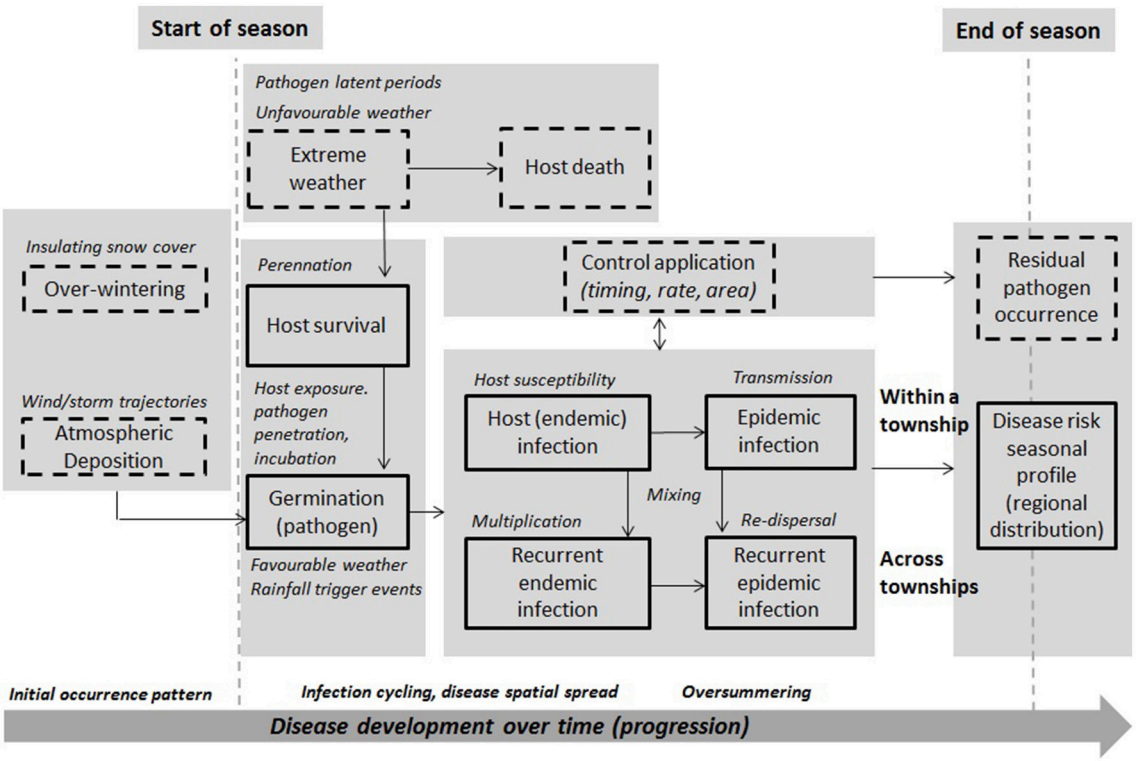


FIGURE 1 | A prototype integrated, model-based framework for forecasting disease risk. Risk components for the major disease progression stages (i.e., immigration, deposition, germination, infection, incubation/residency, multiplication, re-dispersal, perennation) are included, alongside major risk variables, e.g., weather trigger events and conditions, host exposure and susceptibility, pathogen survival, residency, latency, endemic infection, cycling, epidemic transmission and dispersal, and evolved virulence associated with the risk of resistance breakdown in the field (i.e., finer subregions compared to the ROI). Full boxes are components of the framework that are considered in the current disease risk modeling, while dashed boxes will be considered in future extended modeling.

outbreaks/severe epidemics that threaten the world's wheat supply and, in turn, global food security, as 88% of the world's wheat production of 760 million metric tons (∼$185 USD/mt or $140.6 billion USD in 2017/18) is susceptible (World Bank Group, 2017; Food and Agriculture Organization of the United Nations (FAO), 2018). This fungus (a wind-dispersed, obligate biotroph that only infects and survives in a living host) has yellow urediniospores during its asexual infection cycle and is able to disperse over long distances across continents. It is also adapting and overcoming resistance genes via rapid stepwise evolution (Lei et al., 2017; Schwessinger, 2017). There are different types of resistance depending on host wheat plant growth stages and environment (laboratory or field), such that stripe rust resistances can be separated into all-stage (also called seedling) resistance (ASR) verses adult-plant resistance (APR); greenhouse resistance and field resistance; temperature insensitive resistance verses temperature sensitive resistance; and non-durable resistance vs. durable resistance (Wang and Chen, 2017). With the constant evolution of new rust strains, and their adaptation to higher temperatures, consistent and durable disease resistance is a key challenge. The dual/split application of fungicide, with half rates applied early and later, can reduce disease intensity (AUDPC metric) close to that of a single, full application, based on field trials of fungicide effectiveness (Braithwaite, 1998). Crop rotation likely does not prevent the spread of Pst, given its rate of spread is so fast across large areas (Xi et al., 2015). Nonetheless, delayed planting, reduced irrigation, avoidance of excessive nitrogen use, and elimination of volunteer and grass plants can reduce stripe rust severities—but these cultural practices are often not profitable, conflict with conservation farming, and/or reduce yield potential (Chen, 2007). A consideration of host resistance, pathogen survival and dynamics, alongside best management or cultural practices, enabling farmers to use fungicides more judicially, is a long-term, ultimate need and goal to minimize the risk of this disease (Xi et al., 2015).

Most areas of the United States are not suitable for Pst survival in both summer and winter, and only the Pacific Rim states (California, Oregon, and Washington) have favorable areas where the disease survives in summer and winter (i.e., oversummering and overwintering), based on summer/winter survival indices linked to climate conditions (Sharma-Poudyal, 2012; Sharma-Poudyal et al., 2014). The major source of stripe rust inoculum for Alberta (Western Canada) is considered to be from this Pacific Northwest region (hereafter PNW) (Xi et al., 2015). Pst occurrence is generally associated with higher elevations, northern latitudes or cooler years (Newberry et al., 2016). For regions north of latitude 40°N, it infects both winter wheat and spring wheat (winter wheat is generally more susceptible than spring wheat), surviving in *cool summers*, with

hot summers substantially decreasing the chance of its survival, and severe winters preventing its survival (Xi et al., 2015). Pst development becomes dormant for longer durations (slower cycle time) as night-time temperatures cool. In this way, indices based on average daily mean temperatures may overestimate infection risk, especially in early spring and late summer, when cooler nights are more frequent. Because different races can infect different wheat varieties/cultivars having different resistance to the pathogen, the reaction of different resistance genes on selected host varieties/cultivars (differentials) is used to determine the "race spectrum" for a given pathogen population (e.g., within a single field). As with most other wheat-growing regions, urediniopores (asexual cycle) are the only inoculum source for the initial and recurrent infection of wheat, with an infection cycle time that varies through the growing season. New and older rust isolates (i.e., pathogen isolated in field samples that are geographically or location specific) within Western Canada have similar urediniospore rates of germination, occurring between 2 and 20°C, and highest/optimal close to 5°C, with cooler temperatures favoring spore germination (Tran and Kutcher, 2015). While Pst outbreaks have been documented in Alberta since 1925, the variable virulence of Pst has enabled it to overcome the resistance of wheat cultivars, with increasing epidemics occurring in the 1990s, whereby the older population of races (i.e., pathotypes) in the United States have been replaced by a new population since 2000 with germination occurring at higher temperatures between (16–18°C) and being more tolerant of higher summer temperatures (Chen, 2010; Xi et al., 2015).

Warm chinook winds create milder winter weather conditions across southern Alberta and change snow cover. Snow cover is beneficial to urediniospore survival during the winter, whereby areas in the vicinity of Olds are more conducive for spore survival than areas near Lethbridge. This enables Pst to overwinter in this region, and cause outbreaks in early spring when the weather is cool and wet (Conner et al., 1988; Phillips and Newlands, 2011; Xi et al., 2015). Significant snow cover (>7.6 cm) has an insulative effect, enabling Pst to infect wheat within 4–6 h and survive at temperatures down to −10°C. With no snow cover and with temperatures less than −5°C, Pst goes dormant (Sharma-Poudyal et al., 2014). Newer isolates have thus adapted to warmer temperatures and have higher germination rates at higher temperatures than older isolates. The definition of latent period is the approximate time taken for an infection to result in new spores and is temperature-dependent; within an optimal temperature range between 12 and 20°C. Latent period is 10–14 days (Anonymous, 2018) and is cultivar-dependent; a susceptible cultivar AC Bellatrix of red winter wheat (first released by AAFC in Lethbridge in 1999) was shorter at higher temperature for new isolates, with a higher disease intensity over time for new isolates, compared to older ones (i.e., measured as the area under the disease progress curve or AUDPC) (Tran and Kutcher, 2015). This evidence supports the hypothesis that new stripe rust populations continue to adapt to warming temperatures, with increased aggressiveness and explains its expansion into Alberta, including other Canadian Prairie Provinces (i.e., Western Saskatchewan). For Eastern Saskatchewan and Manitoba the source of Pst is from the Mississippi Valley, recent field surveys conducted during July, August, and September 2016 on winter and spring wheat indicate that Pst is found with varying levels of infection depending on spatial location. Winter wheat lines, depending on location and cultivar, can have upwards of 70% infection (2016 Cereal Disease Situation Report, Western Committee on Plant Disease, WCPD. Unpubl.). A recent global analysis of Pst outbreaks involving 887 genetically diverse isolates across 35 countries (2009–2005) reveals that a few, highly divergent genetic races are driving its epidemics and that its populations are being largely shaped by invasion across geographical areas (Ali et al., 2017). With such high epidemic potential, there is greater urgent need for improved predictability of its emergence and dynamics.

Study Region (Southern Alberta, Canada)

The region of interest (hereafter, ROI) is southern Alberta (within Western Canada), a major area of agricultural production with a growing season of about 123 days (May-August). Wheat is the largest crop, followed by barley and canola. Crops are irrigated in this region due to reduced rainfall (semi-arid conditions: 300–450 mm/year). In 2015, adverse weather conditions (i.e., dry spring, low night-time temperatures and frost in fields, hot summer with limited moisture) led to poor growing conditions with yields being lower than long term averages. It first affects winter wheat fields before spreading into spring wheat, as winter wheat (e.g., AC Bellatrix and Radiant varieties) is direct-seeded in early September and harvested several weeks earlier than spring wheat the next year. Pst immigrates into this region from the south (i.e., PNW and areas in the vicinity of Portland, Oregon), as well as, from the north where it overwinters in central Alberta. Backward, diagnostic trajectories (5-day time frames) using analyzed wind fields, indicate that Pst within the NWR of the United States is the main Pst source region for southern Alberta, Canada (AAFC Cereal Rust/Wind Trajectory Event Update Report (Summer 2015) by Turkington et al. Unpubl.). Also, forward, prognostic trajectories using NOAA's HYSPLIT model (April-May, 1995) and forecast wind fields (discrete fields 700–850 hPa) having starting points within the PNW region (vicinity of Portland, Oregon USA) show an immigration zone for high potential for spore dispersion into southern Alberta (not included here for brevity, Newlands unpubl.). These simulations use archived (2-hourly) weather data from Regional Analysis and Forecast System (RAFS)'s Nested Grid Model (NGM) (US National Centers for Environmental Prediction, NCEP). The atmospheric deposition of spores onto the ground still needs to be accounted for, using models such as the one developed by Chamecki et al. (2012) to know more accurately where and when they fall.

Airborne inoculum sample data (weekly, June-October) was obtained for 2015 at Lethbridge (Fairfield site: N 49° 42.493/ W 112° 41.738) from the first year of sampling (**Figure 2**) with sticky microscope slides placed on a Burkard cyclone spore collector. This is considered a passive method of spore trapping/collection, instead of active sampling of the air. The slides were attached on the cap of the Burkard cyclone instrument, just below the collection orifice, so that the slide was always facing the prevailing wind (i.e., 180° incidence angle). Double-sided adhesive tape (#M

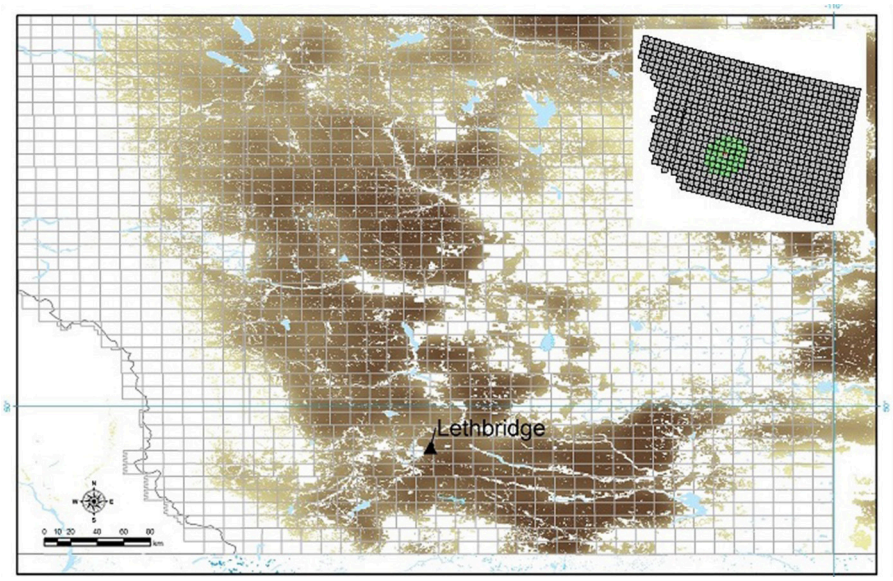


FIGURE 2 | The southern Alberta, Canada region of interest (ROI): The Lethbridge site (located within township TW008R21W4) is indicated; areas where wheat (winter, spring) are grown are indicated (brown) superimposed on a grid of Alberta townships (∼10 km). A sub-region of 40 municipalities was selected surrounding the Lethbridge site for model computations (upper inset).

Scotch® Removable Poster Tape 3/4′ (199 mm) wide, clear) was used. The adhesive tape covered almost the whole translucid surface of the slide, covering an area of 19 × 50 mm. The slides were kept inside a slide box at room temperature for a few days until they were analyzed under a light microscope (Laroche et al., 2018). While it could have been located anywhere in the study region for the entire growing season, it was fixed in its location in an agricultural field (Fairfield) located in Lethbridge, given the high Pst visual occurrence historically detected in fields near this site. Microscopy, Polymerase chain reaction (PCR), and multiplex qPCR molecular techniques were used to identify Pst urediniospores and quantify the concentration of spores collected by the cyclone collector instrument (Araujo et al., 2016) (PCR is a technique to make many copies of a specific DNA region *in vitro*).

Satellite measurements of major disease risk variables (i.e., canopy temperature and liquid water on the canopy surface) (Laroche et al., 2018), regional-scale climate reanalysis (Kobayashi et al., 2015), and quality-controlled weather station network data (hourly scale), spatially-interpolated to the regional municipality scale were provided by the Alberta Climate Information Service (ACIS) (1961-2016)[1]. The ACIS interpolation method linearly weights station-based air temperature estimates of up to 8 closest neighboring stations by inverse-distance, within a correlation radius of 60 km. Precipitation is inversely-weighted by distance (i.e., cube of the inverse distance within a correlation radius of 200 km), with the inverse distance monthly totals redistributed proportionally, relative to the nearest station with a complete monthly record. The JRA-Year Reanalysis (JRA-55) high-resolution, climate reanalysis dataset was used, being among the most sophisticated reanalyses currently available. It also includes pathology-relevant variables with a spatial resolution of 0.55° × 0.55° (∼55 km) and 3-hourly (and 6-hourly) temporal resolution for years 1990-2015 (Bebber et al., 2016). **Figure 3** compares the distribution of potential disease risk based simply on climate variables (i.e., thresholds in mean daily temperature and humidity range) illustrating how the importance of finer resolution in revealing spatial trends and correlation patterns, and the need to model such spatial impacts to predict disease risk rather than relying on climate threshold-based information alone. Simple thresholds are typically assumed in many current operational, weather-based disease risk forecasting systems, avoiding the use of models with assumptions of pathogen-host-environmental interaction that adds additional uncertainty to risk forecasts. The finer-scale pattern of potential Pst pathogen infection risk revealed across the townships (∼10 km) varies considerably from the host (i.e., wheat) distribution pattern, pointing to the need and importance of disease risk modeling to explain and better predict differences in host and pathogen distribution and variability in relation to environmental (e.g., climate) uncertainty (**Figure 2**). Also, information on fungicide efficacy in relation to spray timing and varietal response for stripe rust control is limited in central Alberta, contributing additional uncertainty (Xi et al., 2015).

Hourly leaf wetness duration (LWD) and canopy temperature (T_c) data was obtained from the JRA-55 reanalysis dataset. Canopy moisture (kg m^{-2}) and temperature (°C) was converted[2] from the JRA-55 variables named "moisture storage on canopy" (code 223) (m) and "canopy temperature" (code 114) (K) from

[1] Alberta Climate Information Service (ACIS): https://agriculture.alberta.ca/acis/.

[2] Water weight to water column: 1 kg m^{-2} = 1 mm = 0.001 m, T(°C) = T(K) − 273.15.

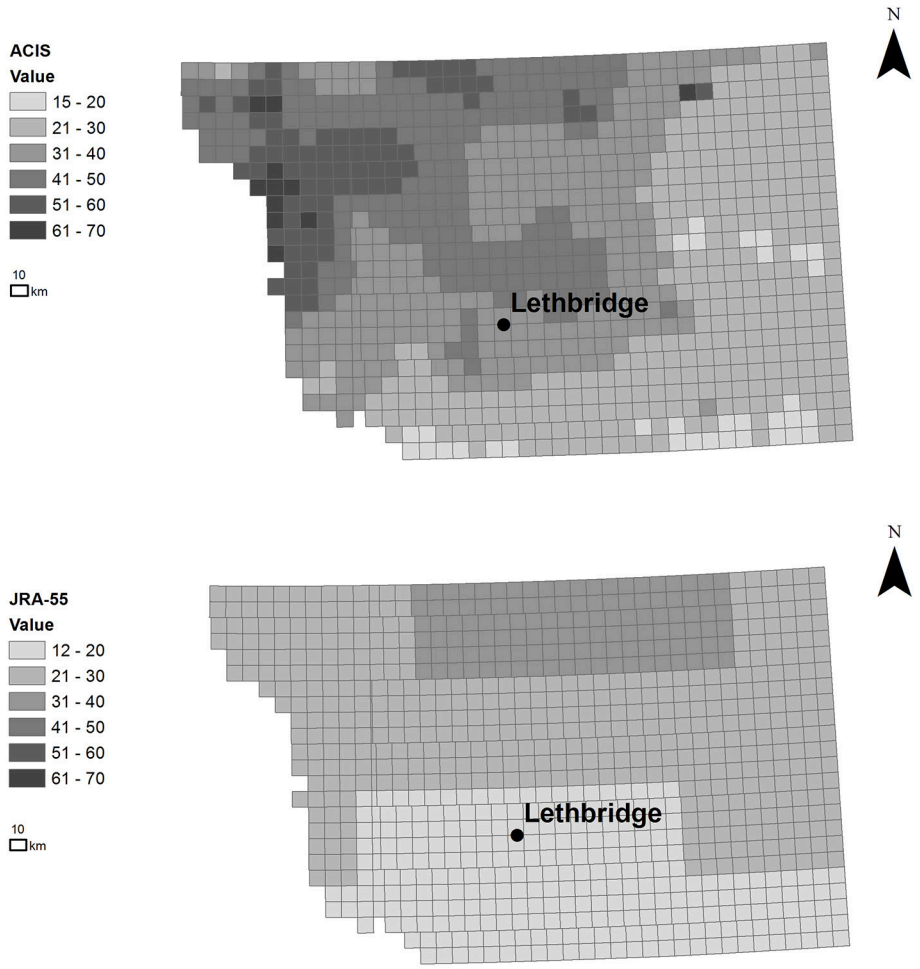


FIGURE 3 | Spatial distribution of potential Pst infection risk between June 17 and Sept 1 in 2015 (in units of total number of growing season days having conditions favorable to infection). The total number of growing season days was based on threshold range of mean daily temperature 3–25^{o}C, average humidity >57% (ACIS) and hourly canopy moisture >0 (wet days) (JRA-55), illustrating the effect of spatial resolution and importance of capturing spatial dependence (in finer-scale climate data) for improving model-based disease prediction. (ESRI ArcGIS 10.4, Canada Albers Equal Area Conic Projection).

the ground/land surface forecast fields (fcst_land125) produced every 3 h (at 00, 03, 06, 09, 12, 15, 18, and 21UTC) (**Figure 4**). Leaf wetness duration (LWD) is highly skewed, with many dry days (**Figure 4A**). Leaf wetness is the presence of free water on the surface of a crop canopy comprising canopy-intercepted rainfall/fog, irrigation, and dew (dewfall/dew-rise) that forms on leaves where water vapor condenses on a surface; it is triggered when the temperature of a canopy surface drops below the dew point temperature of the surrounding air (Rolandson et al., 2015). In June-Oct of 2015, coinciding with the airborne sampling measurement, the majority of wetness events were below 0.001 m (or 0.10 kg m^{-2}) early in the season (June), but increases through the season reaching 0.30–0.35 kg m^{-2} at end-of-season (August-Oct) (**Figure 4B**). There were 54 days with rainfall (of 77 total of airborne sampling in 2015). Canopy temperature showed considerable variability through the 2015 growing season (**Figure 4C**).

CLR Model (Site-Specific)

A site-specific model was first evaluated in predicting Pst disease. This model has previously been used to predict disease risk and a historical 2008-2011 outbreak of Coffee Leaf Rust (CLR, *Hemileia vastatrix*) in Colombia (Bebber et al., 2016) (hereafter, CLR model). The CLR model assumes infection by germinated fungal spores occurring on leaves that are wet for longer than a critical leaf wetness duration (W_{crit} or LWD_{crit}) and specifies a temperature response function [(Yan and Hunt, 1999) Type] of germination and infection based on pathogen-specific minimum, maximum and optimum temperature (θ_{min}, θ_{opt}, θ_{max}). Sensitivity analysis of a generic fungal model for Pst using these variables lends further support for their use in predicting Pst disease dynamics (Bregaglio et al., 2012). The model assumes: temperature is assumed constant over each hourly interval, an equal size of new spore cohorts that being germinating at the start of each wet hour, and no

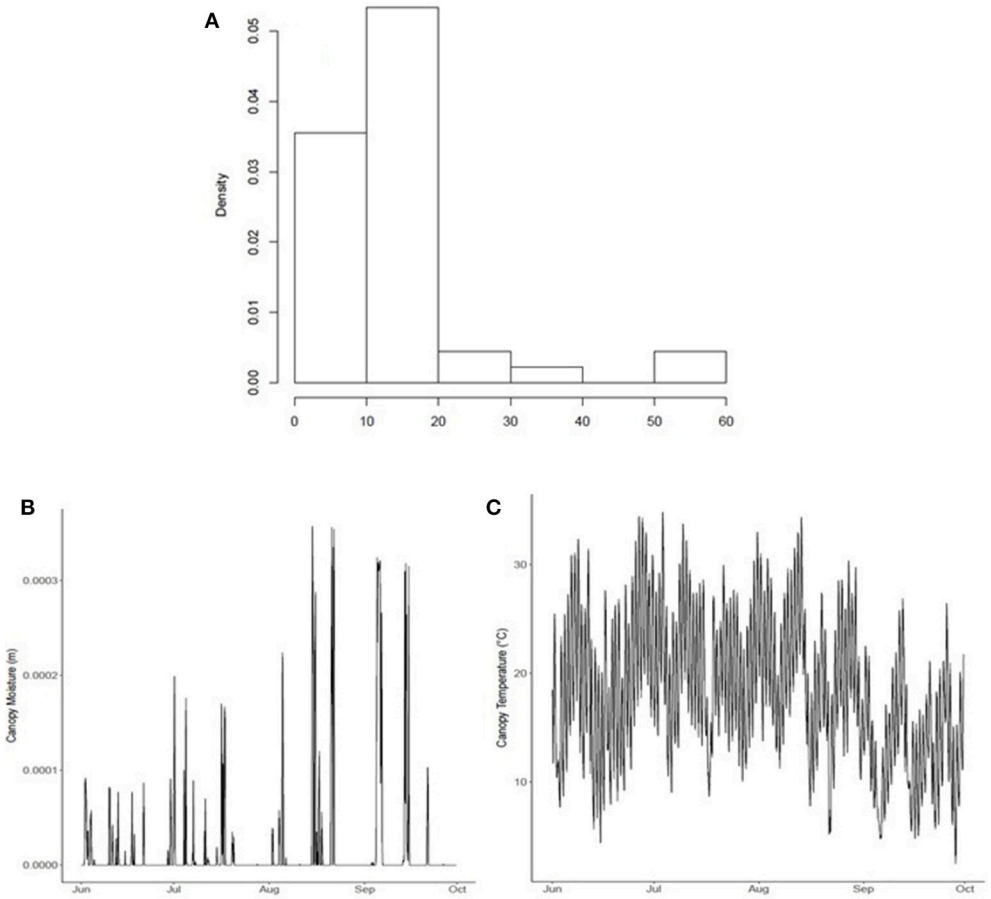


FIGURE 4 | (A) Growing season wetness duration (h) distribution, **(B)** canopy wetness (m), and **(C)** temperature (^{O}C) within the modeled ROI comprising 40 townships within the 2015 growing season, surrounding the Lethbridge site.

germination during dry periods, with no neighbor infections. The process of spore germination and appressoria formation (i.e., infection) over time is modeled as a Weibull-distributed survival process. This generates a cumulative hazard function having germination and infection processes that are time-varying and random, with rates that are greatest at θ_{opt} and decline to zero outside of the temperature range (θ_{min}, θ_{max}). Disease development is thus assumed to be dormant until the temperature moves back within the required range. The CLR model was implemented using validated R code [(Bebber et al., 2016; R Core Team, 2017) (includes supplement and example R code)] with the critical wetness, temperature-response, and germination/infection process (Weibull-distribution) parameters estimated for Pst, using the JRA-55 reanalysis 3-hourly climate data.

hhh4 Model (Spatial)

The hhh4 spatio-temporal endemic-epidemic model having spatial dependence assumptions was selected to compare with the CLR site-specific model predictions. This model has been implemented in the *surveillance* R package (Held and Paul, 2012; Paul and Meyer, 2016; Meyer S. et al., 2017). Separate runs of the model using JRA-55 reanalysis (∼55 km), ACIS station-based township (∼10 km) climate data as input, and both datasets combined, were performed to benchmark the effect of the spatial resolution of input climate data on disease model accuracy. The hhh4 model is a multivariate time-series model for disease incidence, Y_{it} involving multiple, geographical sub-regions (e.g., units of townships or fields), $i = (1,\ldots,I)$, across multiple time periods $t=(1,\ldots,T)$. It assumes a negative-binomial (i.e., clustered) distribution of spore counts, with an additive mean (Meyer S. et al., 2017),

$$\mu_{it} = e_{it}\,\nu_{it} + \lambda_{it} Y_{i,t-1} + \phi_{it} \sum_{j \neq i} w_{ji} Y_{j,t-1} \qquad (1)$$

and the over-dispersion parameter, $\psi_{i.}$ The additive mean μ_{it} consists of an endemic ($e_{it}\ \nu_{it}$) and epidemic ($\lambda_{it} Y_{i,t-1} + \phi_{it} \sum_{j \neq i} w_{ji} Y_{j,t-1}$) component, where e_{it} is the expected counts (i.e., a multiplicative offset to the endemic mean ν_{it}). The *epidemic* component consists of two autoregressive spatial effects, namely: disease reproduction within region *i*, and a neighborhood or spatial-temporal interaction effect involving spore transmission to other region *j*. The disease endemic and epidemic contributions can be assumed identical across regions,

vary within each region, random, or correlated between regions, by representing them as log-linear functions having intercept, α_i and associated predictor variables, whereby,

$$\log(\nu_{it}) = \alpha_i^{(\nu)} + \boldsymbol{\beta}^{(\nu)T}\mathbf{z}_{it}^{(\nu)}, \quad (2)$$

$$\log(\lambda_{it}) = \alpha_i^{(\lambda)} + \boldsymbol{\beta}^{(\nu)T}\mathbf{z}_{it}^{(\lambda)}, \quad (3)$$

$$\log(\phi_{it}) = \alpha_i^{(\phi)} + \boldsymbol{\beta}^{(\nu)T}\mathbf{z}_{it}^{(\phi)}, \quad (4)$$

where T denotes the transpose of a weight vector $\boldsymbol{\beta}^{(\nu)}$ for covariate vector $\mathbf{z}_{it}^{(\nu)}$.If the epidemic parameters, $\lambda = \exp(\alpha^{(\lambda)})$ and $\phi = \exp(\alpha^{(\phi)})$ are assumed homogeneous across all sub-regions, and constant over time (i.e., $\lambda_{it} = \lambda, = \phi, \forall_{i,t}$), an underlying, seasonal temporal trend effecting all regions equally at time t with annual frequency, $\omega = 2\pi/52$. Infection transmission is assumed to occur only between directly adjacent townships ($w_{ji} = I$), where I is the identity matrix, then Equation (1) and its component log-linear predictor (Equation 2), becomes (Meyer S. et al., 2017),

$$\mu_{it} = \mathrm{e}_{it}\,\nu_{it} + \lambda \mathrm{Y}_{i,t-1} + \phi \sum\nolimits_{j \neq i} w_{ji} Y_{j,t-1} \quad (5)$$

$$\log(\nu_t) = \alpha^{(\nu)} + \beta_t t + \gamma \sin(\omega t) + \delta \cos(\omega t) \quad (6)$$

For this homogeneous version of the hhh4 model, weather/climate variables are used to determine the initial infection profile in each township (Equation 5) and then scaling adjustments to these township profiles are made as the model simulation proceeds, whereas the heterogeneous model inputs the weather/climate variables in the covariate vector of the log-linear equations (Equations 2–4). A susceptibility correction to this homogeneous version of the hhh4 model was considered that assumes a township-specific proportion $(1 - \nu_i)$ as a proxy for the susceptible population. Model simulations (i.e., independent runs) were performed with and without a susceptibility correction (i.e., a specific percentage or population susceptible to disease) to gauge how susceptibility assumptions affect model accuracy. This susceptible proportion can be accounted for either as an offset to the endemic population i.e., $(1 - \nu_i)$ (i.e., resulting in a form of the model having a multiplicative offset and log-linear covariates), or as an offset to the autoregressive component of the model (i.e., resulting in a model form that has endemic and/or autoregressive effects). Susceptibility modifies the endemic effect through the substitution of this component with this offset (refer to Equation 1), whereby,

$$\mathrm{e}_{it}\,\nu_{it} \leftarrow (1 - \nu_i)^{\beta_s}\mathrm{e}_i \quad (7)$$

Alternatively, it can be considered as an offset to the epidemic component (i.e., an autoregressive/covariate effect),

$$\begin{aligned}\lambda_{it}\mathrm{Y}_{i,t-1} \leftarrow\ & \exp(\alpha^{(\lambda)} + \beta_s \log(1 - \nu_i))\mathrm{Y}_{i,t-1} \\ &= \exp(\alpha^{(\lambda)})\,(1 - \nu_i)^{\beta_s}\mathrm{Y}_{i,t-1}\end{aligned} \quad (8)$$

where, $\beta_s \geq 1$ is a power effect of high proportion of susceptible populations in sub-region *i* which boost new infections (Meyer S. et al., 2017). The arrow in Equations (7, 8) indicates replacement of the left side terms by the right side terms when susceptibility is considered as a model parameter. Both of these candidate models were evaluated in accounting for wheat cultivar/host susceptibility for Pst in the ROI, with the best-performing model (same dataset) selected by maximizing the likelihood/minimizing the Akaike Information Criterion (AIC) that corrects for variance due to the total estimated number of model parameters. The hhh4 model (homogeneous) has the seven parameters: $(\lambda,\ \phi,\ \nu,\ \exp(\beta_t),\ A, \varphi,\ \psi_i)$, where the two sinusoidal terms of the seasonality-adjustment in Equation (6) are combined into a sinusoidal wave of amplitude A and phase shift φ. Considering susceptibility adds one more model parameter, whether included as an offset ($\beta_s = 1$), or as a covariate (β_s estimated).

The Akaike Information Criterion (AIC) was used in measuring model accuracy and performance (e.g., best and worst cases) of the hhh4 model (R stats library). Infection within each township was modified both by changes in the hhh4 model parameters and subregion climate variability. Five competing cases were evaluated:

Case 1 (JRA hourly reanalysis climate)
Case 2 (ACIS daily station-based climate input)
Case 3 (combined JRA and ACIS)
Case 4 (Case 3 with susceptibility offset correction)
Case 5 (Case 3 with susceptibility covariate correction)

The climate variables selected to drive the model differed depending in each of the cases above, depending on which dataset was used (JRA coarse-scale and/or ACIS fine-scale) and whether the climate variables was hourly or daily. For case 1, JRA hourly mean, minimum and maximum canopy temperature, and leaf wetness duration (LWD) were used to drive the model, For case 2, ACIS daily mean, minimum and maximum temperature and daily relative humidity were used. For case 3, JRA hourly mean, minimum and maximum canopy temperature and leaf wetness duration were used to determine the infection profile for a given township using the temperature response function and assumptions of the CLR site-specific model. This was in addition to ACIS daily mean temperature and humidity for determining initial infection profiles in each township before the model was simulated and dynamical scaling adjustments made. As the JRA data was hourly and ACIS data was daily, this model case after initialization was then simulated at a common (i.e., weekly) aggregation scale, with the two datasets combined using simple, non-weighted averaging to avoid introducing any aggregation bias. Cases 4 and 5 used the same climate variables as in case 3, but with additional susceptibility corrections [i.e., case 4 with susceptibility correction as an offset (Equation 7) and case 5, with it introduced as an additional covariate (Equation 8)]. For the models and their various cases, **Table 1** provides a summary of input datasets, variable inputs, fixed parameters, data, and model spatial and temporal resolutions.

For the best-case, One-Step-Ahead forecasting was performed to approximate forecast error covariance. This was performed assuming a lead-time of a partial (1–4 weeks) and full season

TABLE 1 | A summary of input datasets, climate variable inputs, fixed parameters, data and model spatial, and temporal resolutions.

Model	Cases	Resolutions	Dataset	Variables	Parameters
CLR (non-spatial)	–	Hourly, 55 km	Jra	Canopy temperature (T_c) leaf wetness duration (LWD)	Temperature (θ_{min}, θ_{opt}, θ_{max}) $W_{crit,}$ scale (α) and shape (γ)
hhh4 (spatial) (homogeneous)	1	hourly, 55 km	JRA	T_c, LWD	λ, ϕ, ν, $\exp(\beta_t)$, A, φ, ψ_i, no susceptibility offset/covariate
	2	daily, township	ACIS	air temperature, relative humidity	(as above)
	3	hourly, 55 km	JRA + ACIS	T_c, LWD, air temperature, relative humidity	(as above)
	4	daily, township	(as above)	(as above)	λ, ϕ, ν, $\exp(\beta_t)$, A, φ, ψ_i, β_S offset
	5	(as above)	(as above)	(as above)	λ, ϕ, ν, $\exp(\beta_t)$, A, φ, ψ_i, β_S covariate

(1–11 weeks) time-window. The partial time-window could apply when forecasting disease within a growing season, while the longer time-window could apply when using data from a previous season in forecasting a future season. Four different statistical metrics or "scores" were evaluated in measuring the discrepancy between a model's predictive (i.e., "future" prediction is also termed a "forecast") distribution, μ_P, and future observed value, y, namely: squared-error score (ses), logarithmic score (logs), Dawid-Sebastian (dss), and ranked-probability score (rps), given by,

$$\begin{aligned}
&\text{ses}: (y-\mu_P)^2 \\
&\text{logs}: \frac{(y-\mu_P)^2}{\sigma_p^2} + 2\log\sigma_p \\
&\text{dss}: \frac{(y-\mu_P)^2}{2\sigma_p^2} + 2\log\sigma_p + \frac{1}{2}\log 2\pi \\
&\text{rps}: E_{cdf,P}\left|Y-y\right| - \frac{1}{2}E_{cdf,P}\left|Y-Y'\right| \qquad (9)
\end{aligned}$$

Y and Y' are independent random variables associated with the distribution function p, y is a future "observed" or measured value, and μ_P and σ_p^2 are the mean and variance of the predictive (i.e., forecast) distribution, p. Both the location and spread of the forecast distribution are taken into account by the logs, dss and rps scores in judging how close the distribution is to the observed value. The rps score uses the predictive cumulative density function (cdf) and reduces to absolute error if p is a point-forecast rather than a distribution-forecast. It measures how well probability distribution-based forecasts match observed outcomes. These scores are summary measures of the predictive performance that allow for the joint assessment of calibration and sharpness are reviewed by Gneiting and Katzfuss (2014) and were computed using the surveillance R library package (Meyer S. et al., 2017). Lower scores indicate a model that has better predictive power, with mean scores used to identify a model with the best (i.e., minimal) forecast accuracy.

RESULTS

Temperature response function parameters (θ_{min}, θ_{opt}, θ_{max}) for germination and appressorium formation (i.e., infection) were estimated using data, as (5.91, 15.41, 33.94) for germination, and (5.57, 15.57, 32.11) for infection (de Vallavieille-Pope et al., 1995). This optimal germination temperature range lies within the reported range (i.e., 16–18°C) of post-2000 Pst survival/occurrence at warmer temperatures (Chen, 2005; Xi et al., 2015) and previous reported estimates of (2.6,8.5,18) (Bregaglio et al., 2011). Critical wetness duration (W_{crit}) was set at 4 h, as a lower bound to the Pst reported range of 5–8 h (Bregaglio et al., 2011; Rolandson et al., 2015). The 4 h wetness duration estimate coincides with the timing of a rapid increase in infection for Pst, confirmed by experimental data under controlled conditions (de Vallavieille-Pope et al., 1995). This estimate also is close to observed mean wetness duration distribution peak for the ROI (**Figure 4C**). Risk distribution parameters (i.e., Weibull-distribution) of scale (α) and shape (γ) were estimated (α, γ), for germination as (13.36, 1.29) and for infection as (19.1, 2.14). The CLR (site-specific) model predictions (**Figure 5**) (scaled) are compared to observed Pst spore profile, collected at the Lethbridge site (Fairfield) during the 2015 wheat growing season [AIC = −749.47, root-mean-squared-error (RMSE) = 312.92].

Best-fit estimates (and associated Standard Error) of the spatial model parameters are summarized for the 5 cases considered (**Table 2**). AIC values are provided under each case number in brackets. The predicted Pst spore population (for township region that contains the Lethbridge sampling site) through June-October for the 2015 growing season is shown against the observed Pst/airborne inoculum profile collected at the Lethbridge site (**Figure 6A**). Variability in these

model predictions (case 3) for neighboring regions surrounding the Lethbridge township is shown for this particular season (2015) and subregion, to be strongly endemic, with low autoregressive and transmission contributions from neighboring regions (**Figure 6B**). Estimates forecast scores obtained from one-step-ahead model forecasting are summarized in b. The two different lead-times assume: (i) only Pst population information for the first few weeks (i.e., weeks 1–4 occurring before the main infection peak) is available, and (ii) full data for the entire season is available.

DISCUSSION

The site-specific model (CLR) with the coarser JRA reanalysis input was able to reproduce the general shape of the observed Pst detection (2015 at Lethbridge) based on the assumptions of critical thresholds of wetness and temperature-response, and independent, Weibull-distributed germination/infection processes (**Figure 5**). The model predicts the rate of infection (slope of mid-season peak) well, but predicts a narrower peak width (timing of main infection peak rise and fall) and predicts peak infection a week earlier than it occurred. Predictions of the model are more variable during the early- and late-season, which is attributed to higher variability in the weather conditions (i.e., fluctuations in canopy wetness and temperature shown in **Figures 4B,C**). Using the same coarse JRA climate input data, the spatial hhh4 model with spatial transmission assumptions (case 1), while not the best fit obtained to the observed infection curve, like the CLR site-specific model, does also predict a narrow peak. However, the hhh4 model predictions are less variable in early/late season and it predicts peak infection time correctly (**Figure 6**). This indicates that assumptions on how Pst immigrates, overwinters and moves between subregions, are important for accurately determining both infection variability in early/late season as well as the timing of peak infection. Changing the spatial resolution of climate input data (i.e., from JRA regional-scale in case 1 to finer ACIS township-scale in case 2), improved the prediction of width of the infection peak, which is especially crucial, as it is at this time that disease dynamics can switch from an endemic, to an epidemic disease occurrence pattern.

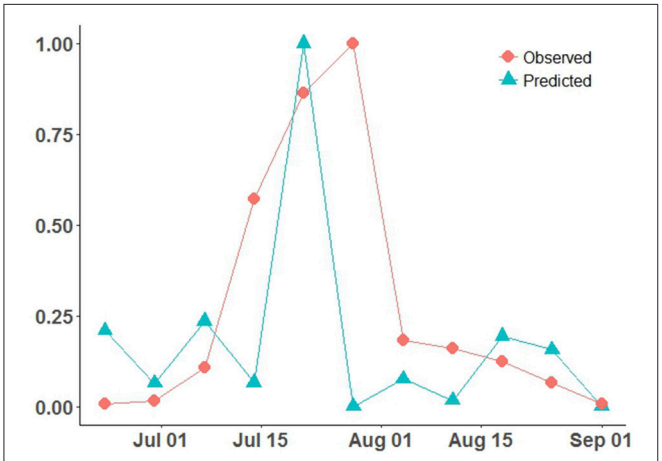


FIGURE 5 | CLR model site-specific predicted Pst disease incidence (scaled 0–1), compared to the observed Pst airborne inoculum counts/profile collected at Lethbridge during the 2015 wheat growing season.

Finer scale climate input improved the ability of the spatial hhh4 model to predict Pst disease dynamics (case 2 vs. case 1) (**Figure 6**). The best-performing model (hhh4 case 3 with lowest AIC) (**Table 2**) had the strongest endemic contribution (initial deposition followed by weak transmission of Pst through the growing season, measured at the Lethbridge site). Comparing AIC values for the various cases of the hhh4 model (see **Table 2**) in terms of relative gain in accuracy [i.e., ($|AIC_{old}-AIC_{new}|/|AIC_{old}|$) x100%, where AIC_{old} denotes the model with the higher AIC, and AIC_{new} the improved model with the lower AIC] provides quantification of the various improvements (spatial and temporal resolution, and inclusion of a susceptibility correction as an offset or covariate). Changing spatial resolution (55 km to township/10 km) and temporal resolution (hourly to daily) led to a relative improvement in accuracy (relative reduction in AIC) of 64% (case 2 vs. case 1). A further relative accuracy gain of 89% was achieved by combining information at different spatial and temporal scale (i.e., hourly and daily, 55 km and township/10 km) (multi-scale case 3 vs. case 2). Correcting for susceptibility as an offset led to a relative accuracy gain of <1% , and no change as a covariate, but these quantified changes are more unreliable due to associated increases in the

TABLE 2 | Best-fit estimates of the hhh4 homogeneous model parameters (Equations 5, 6) and associated Standard Errors (SE) for endemic, ν, and epidemic [i.e., autoregressive, $\lambda = \exp(\alpha^{(\lambda)})$, and spatio-temporal, $\phi = \exp(\alpha^{(\phi)})$] contributions are shown for each of the cases considered: case 1—JRA reanalysis only (~55 km), case 2—ACIS climate only (~10 km), and case 3—multiscale with both JRA and ACIS, and two susceptibility corrections of the best-fitting homogeneous model i.e., case 3 (indicated by *) (case 4—offset type, case 5—covariate type).

Model (AIC)	λ	ϕ	ν	$\exp(\beta_t)$	A	φ	ψ_i	β_s
Case 1 (10106)	0.291 ± 0.051	0.010 ± 0.009	$1.06 \times 10^6 \pm 2.41 \times 10^5$	1.08 ± 0.047	2.98 ± 0.022	−0.839 ± 0.033	0.263 ± 0.017	
Case 2 (6160)	<0.001 ± 0.000	1.02 ± 0.471	74.8 (error > 100%)	9.49 ± 4.18	10.3 ± 1.46	−4.26 ± 0.068	32.4 ± 2.64	
Case 3* (667.22)	<0.001 ± 0.000	<0.001 ± 0.000	0.003 ± 0.093	3.82 ± 1.84	7.20 ± 1.20	−0.786 ± 0.146	20.2 ± 3.84	
Case 4 (667.94)	<0.001 ± 0.000	<0.001 ± 0.000	0.003 ± 20.9	3.79 ± 1.62	5.03 ± 1.19	−5.12 ± 1.08	20.4 ± 3.88	1.000 ± 0.000
Case 5 (669.22)	<0.001 ± 0.000	<0.001 ± 0.000	0.003 ± 20.6	3.82 ± 1.62	5.09 ± 1.18	−5.10 ± 1.08	20.2 ± 3.84	0.039 (error > 100%)

Estimated AIC associated with each model simulation run is provided in brackets.

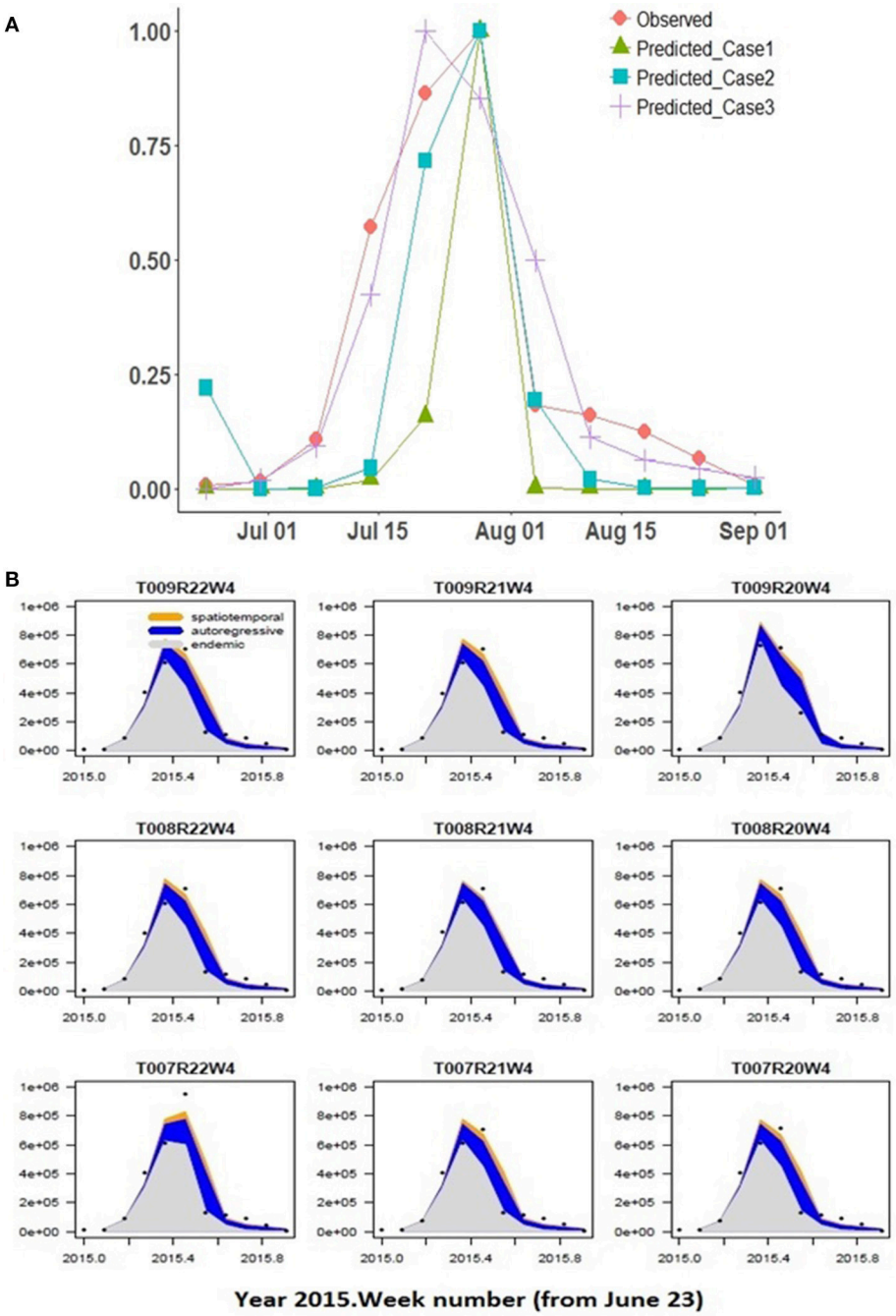


FIGURE 6 | (A) Spatial model predictions of the hhh4 model for cases 1–3 with different climate input data are shown compared to observed at the Lethbridge collection site, **(B)** Predicted Pst disease incidence (scaled 0–1) within the Lethbridge township (TW008R21W4) and its neighboring townships for the best-fitting model (case 3) vs. number of weeks. Variability in the endemic, autoregressive, and spatio-temporal contributions are shown driven by climate variability across the regions.

standard error of model parameters. Overall all cases, the best case (Case 3) was identified or defined as the model case that had the lowest AIC value or highest forecast accuracy) used multi-scale JRA and ACIS data and assumed no correction for susceptibility. The worst-case (Case 1) was identified as the model case that had the highest AIC value or lowest forecast accuracy using only JRA data and assuming no correction for sustainability. Case 3 not only obtained the lowest AIC, but also produced more accurate estimates of the endemic parameter (smaller standard error). This case explains the data the best

(with the lowest AIC) as having weak epidemic components, but strong endemic component, even though the standard error in some model parameters relative to their estimated value increased. Case 3, with the lowest AIC value (best case model), when compared to the highest AIC value (Case 1 worst case model) has a relative overall gain of 93.4% in model accuracy.

The small accuracy improvement offered by a susceptibility correction could be, in part, due to the homogeneous modeling assumptions. Accounting for susceptibility in the best-fit model (case 3), reduced model accuracy only slightly. More observed airborne inoculum sampling data would considerably reduce the high standard error (SE) of the intercept and susceptibility coefficient parameters ($\alpha^{(\nu)}, \beta_s$). Considering susceptibility as an endemic offset, rather than an epidemic autoregressive/covariate effect, produced a model with slightly higher accuracy. This was largely determined, however, by weather variability. Accounting for susceptibility increased the uncertainty in the endemic parameter considerably, pointing to the need for heterogeneous assumptions, alongside a larger data from Pst airborne surveillance in the region to reduce it. This also indicates that the susceptibility parameter is a useful metric for gauging disease risk and determining where to optimally monitor for Pst across the entire ROI. Additional susceptibility effects could be considered if heterogeneous assumptions were considered in the hhh4 model, whereby a multi-variate, "susceptibility function" could be defined (i.e., that further modifies Equations 7, 8). This function could integrate specific-cultivar attributes and/or fungicide control/spray data, and potential canopy changes due to disease detected from available vegetation index satellite data (Davidson, 2015; Devadas et al., 2015). More complex forms of the hhh4 spatial model with spatial-interaction assumptions, heterogeneity driven by long-range transmission, and higher-order neighbors/transmission across sub-regions (power law or second-order model) could be considered. Also, independent random effects uncertainty could be included that typically results from unobserved heterogeneity due to under-reporting of disease occurrence. The hhh3 case 3 model (relying on multi-resolution climate data and spatial dependence assumptions) had reasonable ranges of the predictive assessment scores (**Table 3**). Based on a test period of 4 weeks (i.e., in June-July in advance of peak infection) the best-fit model is able to forecast with 50% accuracy (using the full window of 11 weeks as a benchmark). This forecast accuracy was achieved using airborne inoculum data for a single site and season, alongside disease, environment and susceptibility assumptions. A larger dataset of Pst infection data would enable a more reliable determination of the forecast accuracy of this model and reliable attribution of forecast error to environmental (i.e., climate/weather) variability, and/or host-pathogen disease dynamics.

In summary, the integrated framework proposed offers a feasible way to combine diverse datasets and models with a wide range of assumptions to explain variability and uncertainty in observed disease incidence patterns and to forecast risk. Current findings show that the more complex, stochastic model (hhh4) using weather station network data with susceptibility correction provides sufficient accuracy and reliability. Machine-learning may further improve model-based disease risk forecasting under highly unpredictable weather or management regimes (Liao and Ji, 2009; Wen et al., 2017). Testing the feasibility of this framework and modeling approach relied on limited airborne inoculum data for Pst (i.e., available and high-quality controlled data from the recent 2015 season at Lethbridge) and homogeneous assumptions. Nonetheless, this is the first time such data has been collected for Pst in this region of interest. Current findings are supported by the integration of knowledge, parameter estimates, multiple evidence sources (i.e., published empirical data on Pst, included latest available, quality-controlled climate, satellite and airborne inoculum).

TABLE 3 | Performance of model-based forecasting (4 scores) for Pst disease (i.e., disease risk) within southern Alberta in 2015 for a partial and complete season lead-time, using the best training case (case 3) of the hhh4 model.

Test period	ses	logs	dss	rps
Partial (weeks 1–4)	2155	1.666	−0.414	16.87
Full (All weeks)	1097	0.8851	3.191	7.815

Estimates for squared-error (ses), logarithmic (logs), Dawid-Sebastian (dss), and ranked-probability (rps) are provided (rounded).

An expanded evaluation of model-based forecast accuracy will require a large, seasonal airborne surveillance program and heterogeneous assumptions. Heterogeneous assumptions would help to fine-scale incidence and risk variability based on measured changes in endemic-epidemic transition time, maximum infection potential, infection peak timing, and width between regions. The most accurate, efficient, and cost-effective airborne surveillance monitoring plan may be a multi-resolution sampling design that could sample disease across the full extent of a large ROI at a coarser resolution, while sampling denser, at a finer resolution within known disease hotspots. This is supported by the current findings; the spatial model (case 1) using coarser climate input predicted an underlying pattern of Pst occurrence that is under-dispersed ($\psi_i < 1$), while cases 2 and 3 indicated it is over-dispersed ($\psi_i > 1$). Given case 3 fit the observed data better, a more clumped, concentrated pattern is inferred for Pst based within the 2015 growing season. Hotspots could be identified as townships or subregions where a disease overwinters/oversummers across a sufficient number of sampling seasons, or higher risk clusters of townships within which Pst disease is first or early-detected. Susceptibility offsets and covariation provides an important spatial-based sustainability metric for gauging subregions where disease risk may be highest and where to more intensively sample (Kouadio and Newlands, 2015). Extensions of the spatial model could also include viability of inoculum (i.e., variability in the different isolates that are present into a "natural" inoculum). In laboratory studies, often a single isolate is used. Field studies can involve both a single or composite inoculum; when a composite, one can consider a number of different characterized isolates or assume that a field isolate that was present in a previous growing season has been amplified (multiplied) and is re-inoculating. Also, in further expanded modeling and with the availability of spore data for

2016-2018, the models could be run for the 2019 growing season to predict disease risk, using hourly/daily climate data across all past 4 years where spore data was being collected (i.e., 2015-2018). The 2016-18 time period could be then used to provide training/calibration data for the models, and uncertainty in the model fixed parameters and current season disease risk could be compared to an uncertainty range estimated using all past climate information. This would also provide a historical prediction range to compare against a current-season prediction range. A multi-scale airborne surveillance design that provides data to support operational model-based disease risk forecasting, may, in the future, enable more reliable, timely and cost-effective decisions in sustaining crop yields against multiple disease threats. Developing an integrated understanding of disease risks, impacts, consequences (whether anticipated or unanticipated), alongside decision trade-offs, could provide crucial, cornerstone insights to controlling crop disease, and increasing crop yields sustainably.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

NN was supported by Growing Forward II Federal Research Program (Agriculture and Agri-Food Canada, AAFC) funding (Project No. J-000179.001.02). My thanks to AAFC colleagues: Mr. Weixun Lu (AAFC-Summerland) for providing R language coding assistance, Mr. Kieran Forge for help in preparing disease incidence maps in ArcGIS (ESRI Inc., Ver. 10.4). I thank Dr. André Laroche (AAFC-Lethbridge) for providing the 2015 season wheat stripe rust inoculum data and comments on an earlier draft of this manuscript. I thank Dr. T. A. Porcelli for providing helpful feedback and editorial assistance. I thank two reviewers for feedback that helped to improve this manuscript.

REFERENCES

Ali, S., Rodriguez-Algaba, J., Thach, T., Sørensen, C. K., Hansen, J. G., Lassen, P., et al. (2017). Yellow rust epidemics worldwide were caused by pathogen races from divergent genetic lineages. *Front. Plant Sci.* 8:1057. doi: 10.3389/fpls.2017.01057

Anonymous (2018). *Government of Western Australia (Department of Primary Industries and Regional Development)*. Available online at: https://www.agric.wa.gov.au/grains-research-development/managing-stripe-rust-and-leaf-rust-wheat-western-australia (Accessed May 5, 2018).

Araujo, G. T., Amundsen, E., Gaudet, D., Selinger, B. L., and Laroche, A. (2016). "Incidence of important fungal wheat pathogens in western Canada," in *50th Annual Prairie University Biology Symposium* (Lethbridge, AB: University of Lethbridge).

Audsley, E., Milne, A., and Paveley, N. (2005). A foliar disease model for use in wheat disease management decision support systems. *Ann. Appl. Biol.* 147, 161–172. doi: 10.1111/j.1744-7348.2005.00023.x

Bebber, D. P., Castillo, A. D., and Gurr, S. J. (2016). Modelling coffee leaf rust in Columbia with climate reanalysis data. *Philos. Trans. R. Soc. B* 371:20150458. doi: 10.1098/rstb.2015.0458

Braithwaite, M., Cromey, M. G., Saville, D. J., and Cookson, T. (1998) "Effects of fungicide rates and timing on control of stripe rust in wheat Arable Crops," in *Proc. 51st. N.Z. Plant Protection Conf,* 66–70. Available online at: https://www.researchgate.net/publication/266356024

Bregaglio, S., Cappelli, G., and Donatelli, M. (2012). Evaluating the suitability of a generic fungal infection model for pest risk assessment studies. *Ecol. Modell.* 247, 58–63. doi: 10.1016/j.ecolmodel.2012.08.004

Bregaglio, S., Donatelli, M., Confalonieri, R., Actuis, M., and Orlandini, S. (2011). Multi-metric evaluation of leaf wetness models for large-area application of plant disease models. *Agric. For. Meteorol.* 151, 1163–1172. doi: 10.1016/j.agrformet.2011.04.003

Bryant, R. R., McGrann, G. R. D., Mitchell, A. R., Schoonbeek, H.-J., Boyd, L. A., Uauy, C., et al. (2014). A change in temperature modulates defence to yellow (stripe) rust in wheat line UC1041 independently of resistance gene Yr36. *BMC Plant Biol.* 14:10. doi: 10.1186/1471-2229-14-10

Carisse, O., Bacon, R., and Lefebvre, A. (2009). Grape powdery mildew (*Erysiphe necator*) risk assessment based on airborne conidium concentration. *Crop Prot.* 28, 1035–1044. doi: 10.1016/j.cropro.2009.06.002

Chakraborty, S., and Newton, A. C. (2011). Climate change, plant diseases and food security: an overview. *Plant Pathol.* 60, 2–14. doi: 10.1111/j.1365-3059.2010.02411.x

Chamecki, M., Dufault, D. S., and Isard, S. A. (2012). Atmospheric dispersion of wheat rust spores: A new theoretical framework to interpret field data and estimate downwind dispersion. *J. Appl. Meteor. Climatol.* 51, 672–685. doi: 10.1175/JAMC-D-11-0172.1

Chen, W., Wellings, C., Chen, X., Kang, Z., and Liu, T. (2014). Pathogen profile: wheat stripe (yellow) rust caused by *Puccinia striiformis f.* sp. tritici. *Mol. Plant Pathol.* 15, 433–446. doi: 10.1111/mpp.12116

Chen, X. M. (2005). Review: epidemiology and control of stripe rust (*Puccinia striiformis* f. sp. *tritici)* on wheat. *Can. J. Plant Pathol.* 27, 314–337. doi: 10.1080/07060660509507230

Chen, X. M. (2007). Challenges and solutions for stripe rust control in the United States. *Austr. J. Agric. Res.* 58, 648–655. doi: 10.1071/AR07045

Chen, X. M. (2010). *Race Summary of Puccinia striiformis f.* sp. *tritici (Wheat Stripe Rust) and P. striiformis f.* sp. *hordei (Barley Stripe Rust) in the United States in 2010*. Available online at: http://striperust.wsu.edu/races/data/ (Accessed June 15, 2018).

Conner, R. L., Thomas, J. B., and Kuzyk, A. D. (1988). Overwintering of stripe rust in southern Alberta. *Can. Plant Dis. Surv.* 68, 153–155.

Contreras-Medina, L. M., Torres-Pacheco, I., Guevara-González, R. G., Romero-Troncoso, R. J., Terol-Villalobos, I. R., and Osornio-Rios, R. A. (2009). Mathematical modeling tendencies in plant pathology. *Afr. J. Biotechnol.* 8, 7399–7408. Available online at: https://www.ajol.info/index.php/ajb/article/view/77754/68176

Davidson, A. (2015). *An Operational Canadian Ag-Land Monitoring System (CALMS): Near-Real-Time Agricultural Assessment From Space*. Agriculture and Agri-Food Canada, 60. Available online at: http://www.agr.gc.ca/atlas/supportdocument_documentdesupport/aafcModisNdvi/en/Background_to_the_MODIS_NDVI_Composites.pdf (Accessed Nov 15, 2017).

de Vallavieille-Pope, C., Huber, L., Leconte, M., and Goyeau, H. (1995). Comparative effects of temperature and interrupted wet periods on germination, penetration, and infection of *Puccinia recondita* f. sp. tritici and *P. striiformis* on wheat seedlings. *Ecol. Epidemiol.* 85, 409–415.

Dennis, J. I. (1987). Temperature and wet-period conditions for infection by *Puccinia striiformis* f.sp *tritici* race 104E137A+. *Trans. Br. Mycol. Soc.* 88, 119–121.

Devadas, R., Lamb, D. W., Blackhouse, D., and Simpfendorfer, S. (2015). Sequential application of hyperspectral indices for delineation of stripe rust infection and nitrogen deficiency in wheat. *Precision Agric.* 16, 477–491. doi: 10.1007/s11119-015-9390-0

El Jarroudi, M., Kouadio, A., Bock, C. H., El Jarroudi, M., Junk, J., Pasquali, M., et al. (2017). A threshold-based weather model for predicting stripe rust infection in winter wheat. *Plant Dis.* 101, 693–703. doi: 10.1094/PDIS-12-16-1766-RE

Food and Agriculture Organization of the United Nations (FAO) (2018). *World Food Situation: FAO Cereal Supply and Demand Brief*. Available

online at: http://www.fao.org/worldfoodsituation/csdb/en/ (Accessed, May 11, 2018).

Gneiting, T., and Katzfuss, M. (2014). Probabilistic forecasting. *Ann. Rev. Stat. Appl.* 1, 125–151. doi: 10.1146/annurev-statistics-062713-085831

Haran, M., Bhat, K. S., Molineros, J., and De Wolf, E. (2010). Estimating the risk of a crop epidemic from coincident spatio-temporal processes. *J. Agric. Biol. Environ. Stat.* 15, 158–175. doi: 10.1007/s13253-009-0015-9

Held, L., and Paul, M. (2012). Modeling Seasonality in space-time infectious disease surveillance data. *Biometr. J.* 54, 824–843. doi: 10.1002/bimj.201200037

Held, L., Höhle, M., and Hofmann, M. (2004). "A statistical framework for the analysis of multivariate infections disease surveillance data," *Paper 402 Ludwig-Maximilians-Universität Munchen, Insitute Fur Statistik Sonderforschungsbereich* 386. Available online at: http://epub.ub.uni-muenchen.de/ (Accessed Nov 15, 2017).

Höhle, M., Meyer, S., Paul, M., Held, L., Burkom, H., Correa, T., et al. (2017). *Temporal and Spatio-Temporal Modeling and Monitoring of Epidemic Phenomena*. Surveillance R package version 1.15.0. Available online at: http://surveillance.r-forge.r-project.org/ (Accessed Nov. 15, 2017).

Hovmøller, M. S. (2001). Disease severity and pathotype dynamics of *Puccinia striiformis* f.sp. tritici in Denmark. *Plant Pathol.* 50, 181–189. doi: 10.1046/j.1365-3059.2001.00525.x

Isard, S. A., Barnes, C. W., Hambleton, S., Ariatti, A.,Russo, J. M., Tenuta, A., et al. (2011). Predicting soybean rust incursions into the North American continental interior using crop monitoring, spore trapping, and aerobiological modeling. *Plant Dis.* 95, 1346–1357. doi: 10.1094/PDIS-01-11-0034

Juroszek, P., and von Tiedemann, A. (2013). Climate change and potential future risks through wheat diseases: a review. *Eur. J. Plant Pathol.* 136, 21–33. doi: 10.1007/s10658-012-0144-9

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteorol. Soc. Jpn. Ser. II.* 93, 5–48. doi: 10.2151/jmsj.2015-001

Kouadio, L., and Newlands, N. K. (2015). Building capacity for spatial-based sustainability metrics in agriculture. *Decis. Anal.* 2:2. doi: 10.1186/s40165-015-0011-9

Kuang, W., Liu, W., Ma, Z., Wang, H. (2013). "Development of a web-based prediction system for wheat stripe rust," in *6th Computer and Computing Technologies in Agriculture (CCTA), Oct 2012, Zhangjiajie, China. Springer, IFIP Advances in Information and Communication Technology, AICT-392 (Part I), Computer and Computing Technologies in Agriculture VI,* 324–335. Available online at: https://hal.inria.fr/hal-01348115 (Accessed Nov 15, 2017).

Laroche, A., Araujo, G. T., Lu, W., Amundsen, E., Newlands, N. K. Aboukhaddour, R., et al. (2018). "Aerobiological surveillance of wheat pathogens," in *Conference Proceedings of Agronomy Update 2018* (Red Deer, AB), 52.

Lei, Y., Wang, M., Wan, A., Xia, C., See, D. R., and Zhang, M., Chen, X. (2017). Virulence and molecular characterization of experimental isolates of the stripe rust pathogen (*Puccinia striiformis*) indicate stomatic recombination. *Phytopathology* 107, 329–344. doi: 10.1094/PHYTO-07-16-0261-R

Liao, W., and Ji, Q. (2009). Learning Bayesian network parameters under incomplete data with domain knowledge. *Pattern Recognit.* 42, 3036–3056. doi: 10.1016/j.patcog.2009.04.006

Mahlein, A.-K. (2016). Plant disease detection by imaging sensors – parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE

Meyer, M., Cox, J. A., Hitchings, M. D. T., Burgin, L., Hort, M. C., Hodson, D. P., et al. (2017). Quantifying airborne dispersal routes of pathogens over continents to safeguard global wheat supply. *Nat. Plants Lett.* 3, 780–796. doi: 10.1038/s41477-017-0017-5

Meyer, S., Held, L., and Höhle, M. (2017). Spatio-temporal analysis of epidemic phenomena using the R Package surveillance. *J. Stat. Soft.* 77, 1–55. doi: 10.18637/jss.v077.i11

Minchinton, E. J., Auer, D. P. F., Thomson, F. M., Trapnell, L. N., Galea, V., Faggian, R., et al. (2013). Evaluation of the efficacy and economics of irrigation management, plant resistance and Brassica$_{spot}^{(TM)}$ models for management of white blister on *Brassica* crops. *Austral. Plant Pathol.* 42, 169–178. doi: 10.1007/s13313-012-0181-z

Newberry, F., Qi, A., and Fitt, B. D. L. (2016). Modelling impacts of climate change on arable crop diseases: progress, challenges and applications. *Curr. Opin. Plant Biol.* 32, 101–109. doi: 10.1016/j.pbi.2016.07.002

Newlands, N. K. (2016). *Future Sustainable Ecosystems: Complexity, Risk, Uncertainty*. Boca Raton, FL: Taylor & Francis (Chapman & Hall/CRC Applied Environmental Statistics Series), 418.

Nicolopoulou-Stamati, P., Maipas, S., Kotampasi, C., Stamatis, P., and Hens, L. (2016). Chemical pesticides and human health: the urgent need for a new concept in agriculture. *Front. Public Health* 4:148. doi: 10.3389/fpubh.2016.00148

Nopsa, J. F. H., and Pfender, W. F. (2014). A latent period duration model for wheat stem rust. *Plant Dis.* 98, 1358–1363. doi: 10.1094/PDIS-11-13-1128-RE

Ojiambo, P. S., Yuen, J., van den Bosch, F., and Madden, L. V. (2017). Epidemiology: past, present and future impacts on understanding disease dynamics and improving plant disease management – A summary of focus issue articles. *Phytopathology* 107, 1092–1094. doi: 10.1094/PHYTO-07-17-0248-FI

Parsons, D. J., and Te Beest, D. (2004). Optimising fungicide applications on winter wheat using genetic algorithms. *Biosyst. Eng.* 88, 401–410. doi: 10.1016/j.biosystemseng.2004.04.012

Paul, M., and Meyer, S. (2016). *hhh4: An Endemic-Epidemic Modelling Framework for Infectious Disease Counts*. Available online at: https://cran.r-project.org/web/packages/surveillance/vignettes/hhh4.pdf (Accessed Nov. 15, 2017).

Peshin, R., Bandral, R. S., Zhang, W., Wilson, L., and Dhawan, A. K. (2009). "Integrated pest management: a global overview of history, programs and adoption," in *Integrated Pest Management: Innovation-Development Process*, eds R. Peshin and A. K. Dhawan (Dordrecht: Springer), 1–46.

Phillips, A. J., and Newlands, N. K. (2011). Spatial and temporal variability of soil freeze-thaw cycling across southern Alberta Canada. *Agric. Sci.* 2, 392–405. doi: 10.4236/as.2011.23051

Pretty, J., and Bharucha, Z. P. (2015). Integrated pest management for sustainable intensification pf agriculture in Asia and Africa. *Insects* 6, 152–182. doi: 10.3390/insects6010152

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org/ (Accessed Nov 15, 2017).

Riley, S., Earnes, K., Isham, V., Mollison, D., and Trapman, P. (2015). Five challenges for spatial epidemic models. *Epidemics* 10, 68–71. doi: 10.1016/j.epidem.2014.07.001

Rolandson, T., Gleason, M., Sentalhas, P., Gellespie, T., Thomas, C., and Hornbuckle, B. (2015). Reconsidering leaf wetness duration determination for plant disease management. *Plant Dis.* 99, 310–319. doi: 10.1094/PDIS-05-14-0529-FE

Sadyś, M., Skjøth, C. A., and Kennedy, R. (2016). Forecasting methodologies for *Ganoderma* spore concentration using combined statistical approaches and model evaluations. *Int. J. Biometeor.* 60, 489–498. doi: 10.1007/s00484-015-1045-3

Savage, D., and Renton, M. (2014). Requirements, design and implementation of a general model of biological invasion. *Ecol. Modell.* 272, 394–409. doi: 1016/j.ecolmodel.2013.10.001

Schwessinger, B. (2017). Fundamental wheat stripe rust research in the 21st Century. *New Pathol.* 213, 1625–1631. doi: 10.1111/nph.14159

Sharma-Poudyal, D. (2012). *Prediction of Disease Damage, Determination of Pathogen Survival Regions, and Characterization of International Collections of Wheat Stripe Rust*. Ph.D. Thesis, Washington State University. Available online at: https://research.wsulibs.wsu.edu/xmlui/bitstream/handle/2376/4061/SharmaPoudyal_wsu_0251E_10339.pdf?sequence=1 (Accessed Nov. 15, 2017).

Sharma-Poudyal, D., Chen, X., and Rupp, R. A. (2014). Potential oversummering and overwintering regions for the wheat stripe rust pathogen in the contiguous United States. *Int. J. Biometeorol.* 58, 987–997. doi: 10.1007/s00484-013-0683-6

Tran, V. A., and Kutcher, H. R. (2015). "Temperature effects on the aggressiveness of *Puccinia striiformis* f. sp. tritici, stripe rust of wheat," *Soil and Crops Workshop*, University of Saskatchewan, 6. Available online at: http://www.usask.ca/soilsncrops/conference-proceedings/index.php (Accessed Nov. 15, 2017).

Wang, M., and Chen, X. (2017). "Stripe Rust Resistance," (Chapter 5) in *Stripe Rust*, eds X. M.Chen and J. Kang (Dordrecht: Science+Business Media B.V.), 353–558.

Wen, L., Bowen, C. R., and Hartman, G. L. (2017). Prediction of short-distance aerial movement of *Phakopsora pachyrhizi* urediniospores using machine

learning. *Phytopathology* 107, 1187–1198. doi: 10.1094/PHYTO-04-17-0138-FI

West, J., and Kimber, R. (2015). Innovations in air sampling to detect plant pathogens. *Ann. Appl. Biol.* 166, 4–17. doi: 10.1111/aab.12191

World Bank Group (2017). *Commodity Market Outlook*. Washington, DC: World Bank.

Xi, K., Kumar, K., Holtz, M. D., and Turkington, T. K., Chapman, B. (2015). Understanding the development and management of stripe rust in central Alberta. *Can. J.Plant Pathol.* 37, 21–39. doi: 10.1080/07060661.2014.981215

Yan, W., and Hunt, L. A. (1999). An equation for modelling the temperature response of plants using only the cardinal temperatures. *Ann. Bot*. 84, 607–614.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership