Check for updates

Digital applications unlock remote sensing AI foundation models for scalable environmental monitoring

Benjamin Strong[1]*, Edward Boyda[1,2], Cameron Kruse[1], Tom Ingold[1] and Mikel Maron[1]

[1]The Earth Genome, Los Altos, CA, United States, [2]AI-Journalism Resource Center, Oslo Metropolitan University, Oslo, Norway

Remote sensing AI foundation models, which are large, pre-trained models adaptable to various tasks, dramatically reduce the resources required to perform environmental monitoring, a central task for developing ecosystem technologies. However, the unique challenges associated with remote sensing data necessitate the development of digital applications to effectively utilize these models. Here, we discuss early examples of user-centered digital applications that enhance the impact of remote sensing AI foundation models. By simplifying model training and inference, these applications open traditional machine learning tasks to a range of users, ultimately resulting in more locally-tuned, accurate, and practical data.

1 Introduction

AI foundation models have recently emerged as powerful tools for analyzing large environmental datasets, particularly in the field of remote sensing (Zhu et al., 2024; Zhang et al., 2024; Tao et al., 2023). These foundation models are pre-trained on large, frequently unlabeled datasets, enabling efficient and accurate fine-tuning on diverse downstream tasks (Bommasani et al., 2021; Han et al., 2022). For example, a remote sensing foundation model may be pre-trained through self-supervised learning on a large corpus of unlabeled satellite imagery, learning general representations that capture patterns across the pre-training data distribution (Klemmer et al., 2023; Wang D. et al., 2024). A user can then fine-tune this model for a specific downstream task, such as monitoring restoration of a specific habitat. Because of the efficient pre-training of the foundation model, a relatively small amount of task-specific training data is needed to reach a high level of performance (Allen et al., 2023). Another underappreciated advantage of foundation models is their potential to dramatically simplify the technical requirements and engineering work needed to build and deploy state of the art technologies (Bommasani et al., 2021).

Foundation models for remote sensing data show much promise within ecosystem technology, where ecosystem sensing is of critical importance. Ecosystem technology (ecotech) is an emerging field seeking to systematize and integrate natural-system inspired, technological interventions to promote climate resilience and biodiversity. All place-based ecotechnologies require up-to-date, contextualized monitoring data to inform (1) fundamental science for initial technology development (2) site selection and project planning and (3) ongoing monitoring, reporting, and verification. In particular, rewilding, landscape restoration, landscape engineering, afforestation, regenerative agriculture, and, along the coasts, cultivation of wetlands, dune ecosystems, mangroves and seagrasses may be effectively studied via remote

sensing. Given the diversity and place-specificity of anticipated interventions, new methods of analysis will be required, more precise and adaptable to local conditions than traditional, broad land-use—land-cover surveys. Remote sensing AI foundation models offer a promising pathway for developing these techniques.

Current research on foundation models has focused on the development of the models themselves (Cong et al., 2022; Jakubik et al., 2023; Tseng et al., 2023; Szwarcman et al., 2025), often in concert with substantial training datasets (Bastani et al., 2023; Wang et al., 2023), and on benchmarking these models on increasingly varied tasks (Lacoste et al., 2021; Lacoste et al., 2023; Marsocci et al., 2024). The introduction of climate-relevant benchmarks, such as forest and burn-scar datasets, and initiatives like the ESA-sponsored project Foundation Models for Climate and Society, with its focus on ice, drought, and flood-zone mapping, anticipate a shift toward more applications-oriented research. We note some experimental real-world use cases of foundation models below.

We argue that purpose-driven digital applications can unlock the potential of foundation models for environmental decision making. Building on sophisticated foundation model features that follow from extensive pre-training, a digital application can automate subsequent complex machine learning workflows and greatly simplify downstream model training and inference. Digital applications can thus allow environmental scientists, policymakers, and other stakeholders to more easily interact with these powerful models, and they can constitute critical components of future ecosystem technologies.

An analogous example is GPT, a foundation model for text (OpenAI, 2023). A state-of-the-art large language model, GPT only gained significant attention with the release of ChatGPT, a tuned chat-bot interface accessible through a web browser or smartphone app. The usability of ChatGPT drove record-setting adoption, making it the fastest digital application in history to reach 100 million users (Hu, 2023).

Importantly, digital applications for remote sensing AI foundation models provide value beyond simply visualizing model outputs. In addition, as we explain here, they can greatly simplify the fine-tuning process for foundation models. This capability opens the door for human-in-the-loop or active learning approaches for rapid fine tuning of models, greatly accelerating the process of developing accurate models and high quality datasets.

In this article, we share our perspective from having developed an early digital application for environmental monitoring that leverages remote sensing foundation models. This kind of application is not just a technical innovation but also a concrete example of ecotech—demonstrating how ecological monitoring and management are evolving through technology. We first provide general technical design considerations for developing these kinds of applications and then consider some existing prototypes. These examples point to compelling new possibilities for utilizing high-quality remote sensing datasets, and, we hope, can guide fruitful collaboration between application developers and research scientists.

2 Key considerations for developing user interfaces for remote sensing foundation models

In this section, we discuss several of the key considerations for building digital applications that leverage remote sensing AI

foundation models. Among these considerations, four have stood out in our work: (1) choice of task, (2) prompting strategy, (3) support for machine learning, and (4) integration with other tools, including other forms of AI.

2.1 What task is supported?

Many modern remote sensing foundation models are generally trained to be *multitask*, such that they support some combination of classification, segmentation, regression, object detection, and change detection (Lu et al., 2024). For example, the foundation model Skysense has been shown to perform well on 16 benchmark datasets across seven different kinds of tasks (Guo et al., 2024).

Development of a digital application for environmental monitoring requires understanding user needs and appropriately tying these needs back to these tasks. For instance, a user interested in tracking deforestation might be interested in creating an alert for new logging roads in a given area (change detection and classification). Alternatively, the user might be interested in quantifying the footprint of the deforested area (segmentation).

2.2 What is the prompting strategy?

Whereas traditional remote sensing machine learning models might require as input a curated dataset with defined labels and a single sensor image or pixel time series, a user can interact with a pre-trained foundation model in a more flexible, nuanced manner. Varied prompting strategies can often be directly incorporated into user interfaces built on these foundation models.

2.2.1 Natural language prompting

Given the popularity of ChatGPT and Large Language Models, a natural target for development is language prompting of remote sensing foundation models. Several "vision language models" (VLMs) have been developed specifically targeted for this mode of interaction with remote sensing data (Hu et al., 2023; Irvin et al., 2024; Kuckreja et al., 2024; Liu et al., 2024). These models allow a user to ask text-based questions such as, "Where has infrastructure development taken place in this landscape?" The models take this text query and a remote sensing image as input and provide a text string as output.

Natural language prompting simplifies interactions with remote sensing foundation models. However, in many cases natural language prompting can be limited. For example, training VLMs requires large geospatial datasets that associate imagery with textual labels (Irvin et al., 2024). While many significant advances have been made to create these training datasets, they often rely on geographically biased or otherwise incomplete sources like OpenStreetMap (Li et al., 2020).

2.2.2 Location-based prompting

An alternative to natural language prompting of foundation models is location-based prompting, where a user provides a point, bounding box, or other geographic reference as input to a model. Meta's Segment Anything Model (SAM) attracted significant attention when it was released in 2023, partly for the ease of use of its interactive point or bounding box prompting. This capability has been embraced

by the geospatial community, with projects like samgeo, a geospatial toolbox for SAM that includes a python notebook map-based user interface (Wu and Osco, 2023).

2.3 How are model training and inference incorporated?

To be useful for downstream applications, foundation models are often fine-tuned with application-specific training data. But this is not necessarily the case. For example, users of ChatGPT derive value from interactive chat without needing to fine tune the GPT model. In general there are four possibilities for how a technology might interact with a pre-trained foundation model for training and inference:

a) Pure inference: in this case, the pre-training task of the foundation model is directly useful for downstream applications. For example, a super resolution model can directly output super-resolved images.

b) Zero shot learning: LLMs have popularized the idea of zero shot learning, where a model is used to classify a target that was not present in its training set by providing a semantically-rich description of the target to the model at inference time. For example, the VLM TEOChat demonstrates zero-shot capability to perform question answering on novel remote sensing datasets that were not included in its training set (Irvin et al., 2024).

c) Few shot learning: given the extensive pre-training of foundation models, a handful of task-specific training examples can be used to fine-tune a foundation model or a projection head. Techniques used in few-shot learning include partial weight retraining (e.g., LoRA, Hu et al., 2021) or training projection heads that use the outputs of a foundation model encoder as inputs.

d) Full fine tuning: full fine tuning of the foundation model may achieve high levels of performance, but this usually requires more extensive labeled datasets and computation.

An example of an application incorporating full fine tuning is provided by IBM, NASA, and others with their Prithvi foundation model (Jakubik et al., 2023; Szwarcman et al., 2025). They have released a digital application to support fine-tuning workflows, with online dataset management, computational resources, and visualization tools (IBM Research, 2024). Fine-tuned models in public release include instances tuned for flood inundation mapping, burn scar mapping, and multi-temporal crop classification, and a recent study investigated fine-tuning for prediction of locust breeding grounds (IBM Research and NASA, 2025; Yusuf et al., 2024).

The zero shot and few shot learning cases are particularly well-suited for user interfaces, because these modes of interaction are enhanced by user feedback. Human-in-the-loop strategies may be employed to fine-tune models interactively, as we will discuss for two cases below.

2.4 How might these systems integrate future powerful AI tools?

Integration between remote sensing AI foundation models and other forms of AI technology offers substantial opportunities for enhancing environmental monitoring and ecosystem technologies. As discussed above, natural language prompting is one promising possibility. It points toward an array of interactions by which general-purpose AI models, including large language models like GPT, can complement and extend the specialized functions of remote sensing models.

In the future, general-purpose AI will streamline large-scale data analysis, interpreting outputs from remote sensing models into actionable insights (e.g., Planet Labs PBC, 2025). Calling on AI models' language capabilities, users without specialized technical backgrounds can easily query complex datasets, request detailed analyses, or explore scenarios through straightforward prompts. We can also look to generalized AI to optimize environmental monitoring workflows. These models may be used to dynamically adjust monitoring parameters, determine optimal sensor deployment, or prioritize analysis regions based on predictive modeling outcomes.

Importantly, integration with generative AI models introduces risks that go beyond the statistical errors expected of a discriminative classifier. Without careful engineering, a generative model like GPT may "hallucinate" nonfactual information or may behave in ways contrary to the goals of the digital application. These risks can be mitigated by managing the scope of integration, for instance, by constraining data retrieval to specific trusted sources and passing those trusted sources directly to the user. Active attention from the user and verification of generative outputs is warranted.

3 Case study: Earth Index

Our experimental user interface, called Earth Index, is a tool for few-shot classification of satellite imagery that functions via location-based prompting (Figure 1). In many respects it can be considered a search tool for the planet: A user indicates a few objects of interest on a web map, and in real time Earth Index surfaces other similar example objects from the area of interest (AOI).

Earth Index is a production web application that offers point and brush tools for labeling on the web map, options for searching and validating search results, a selection of satellite image basemaps, and logic for state and project management. The application refers user queries to a geospatially-enabled relational database, where pre-computed model embeddings are stored and indexed. An embedding is the vector output of the AI model on an input image patch. As such, it gives a mathematical representation of the objects and terrain on that part of the Earth's surface. For technical details, please refer to the Supplementary material. The basic workflow is intended to be almost self-explanatory for intended users, who include researchers with environmental domain expertise but no presumed exposure to machine learning.

Through close reading of proposed detections and adding positive and negative labels, the human user iteratively refines the search. The search is open-ended. Under the hood the tool constructs a representative embedding vector from the labeled vectors and runs an approximate-nearest-neighbor search for other embedding vectors in its vicinity, up to a user-defined radius. Mathematically, this is essentially a recommendation system. Its efficacy rests on the foundation model's power to usefully organize the embedding vector space.

As well as being capable discriminators of land cover, remote sensing foundation models offer good representations of *objects* on Earth's surface, which means they can be used to detect industrial infrastructure and signs of illicit resource extraction. To give one concrete example, the
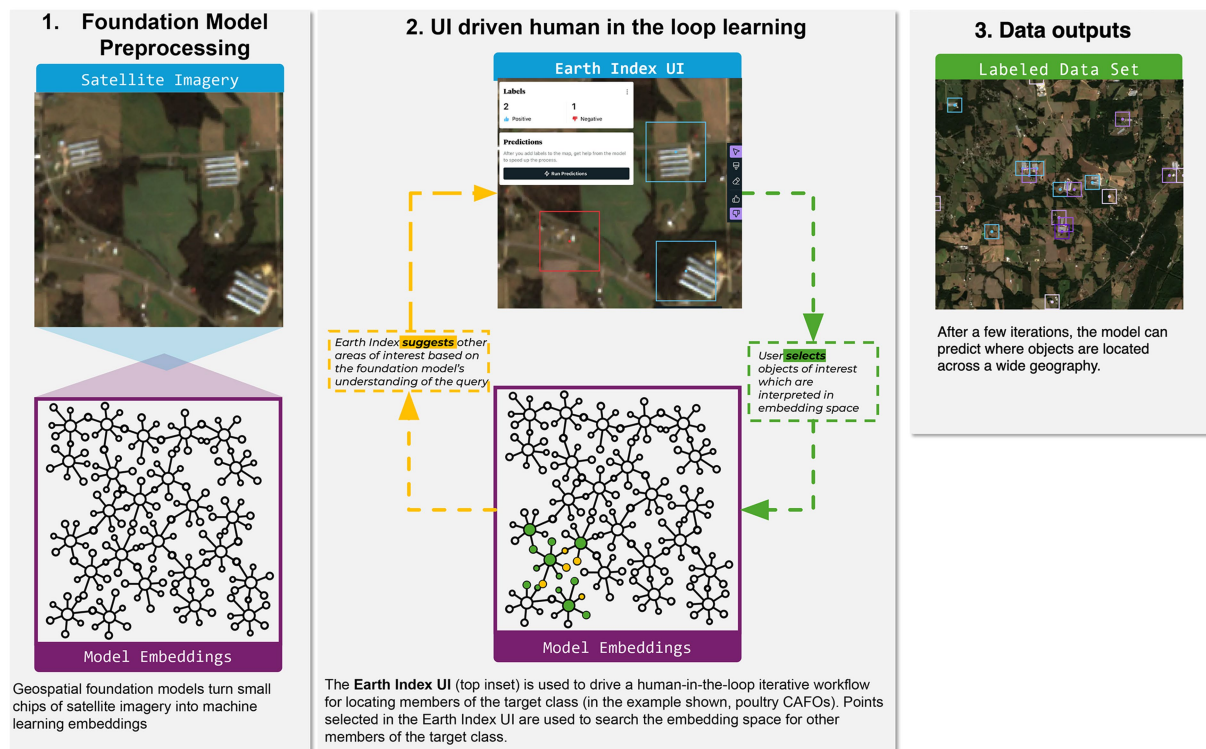
FIGURE 1
Earth Index is a user interface and associated set of services that is built on remote sensing foundation models. Using Earth Index, non-experts can search satellite imagery interactively, building on the knowledge of the pre-trained foundation model.

Pulitzer Center Rainforests Investigation Network is using Earth Index to investigate small-scale open-pit gold mining in the Mekong delta region. A data journalist at the Pulitzer Center and one of the present authors each began Earth Index searches with points identified from a mine field previously reported in the media. By iterative search and validation against external data sources, we learned the unique features of mine scars in this region. Notably, houses of miners and their families are often interspersed among the toxic mud-sludge pools left after the mining. We used this understanding to formulate successively more precise searches, finally producing detailed maps of five significant mining areas across a 200 kilometer-wide zone. These detections guided field reporting trips of a local reporter, who interviewed both miners and members of community groups who patrol the forest against illegal logging activities. The reporter conveyed amazement at having traveled hours on dirt road by motorbike, only to stop at a river crossing, walk up-river, and come upon the promised mines.

Challenges for sustained adoption relate to the difficulties of working with the satellite imagery itself: framing problems appropriate to the resolution of the imagery, interpreting findings, and persevering to disambiguate target objects from a complex background. We now include training materials for new users with worked examples of iterative search.

3.1 Streamlined machine learning

Geographic search seeks to surface sites which are similar to those queried, to be evaluated by a user according to greater or lesser relevance. To exhaustively catalog a phenomenon, one typically needs recourse to machine learning with labeled data. This workflow requires a higher level of data fluency from a user than open-ended Earth search.

We use a Jupyter notebook with a Leaflet map-based interface, a raw version of the Earth Index application, to train logistic regression or Multilayer Perceptron (MLP) models atop the embedding vectors output by the foundation model encoder. The Leaflet satellite web map supports labeling of sites on the fly, via clicks on the map. Code for model training, validation, ensembling, and inference follow in successive sections of the notebook. By building atop the foundation model, the complexity of subsequent modeling is cut by orders of magnitude versus end-to-end training. On a MacBook Pro, MLP training and inference on an AOI of 100,000 km^2 each take only a few minutes. Experiments with street-level photography indicate that the combination of a pre-trained foundation model with a projection head can meet or exceed performance of models trained end-to-end with labeled data (Chen et al., 2020).

In a pilot study, we applied this streamlined process to build a dataset of poultry animal feeding operations across the American southeast, including the largest poultry producing states. Concentrated Animal Feeding Operations (CAFOs) are industrial animal agriculture facilities where large numbers of animals are raised in confined conditions. They pose significant threats to the environment and to the health and livability of nearby communities (Crippa et al., 2021; Nicole, 2013; Wing et al., 2000).

Even though the facilities share notable visual similarities across states, naive attempts to extrapolate a machine learning model from one state to another can be frustrated by differences in background geography. This can be approached as a challenge in local fine-tuning, where the integrated labeling and training interface speeds model
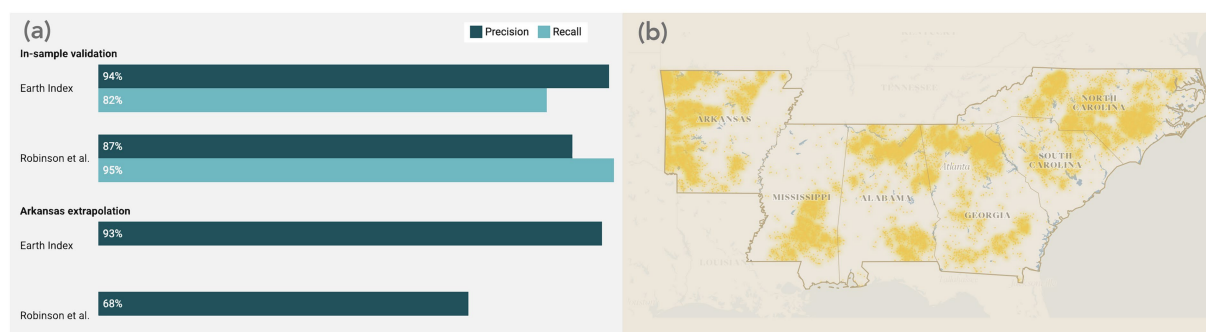
FIGURE 2
Local fine-tuning with the Earth Index machine learning workflow leads to precise detections of poultry CAFOs in the American southeast. **(a)** Earth Index and a prior machine learning inventory (Robinson et al., 2022) perform comparably in validation in the areas where training data were sampled (North Carolina and the Delmarva Peninsula, respectively); with fine-tuning, Earth Index maintains precision in extrapolation to Arkansas, while the naively extrapolated models falls from 87% to an estimated 68% precision. Algorithmic differences make this an imperfect comparison, but it nonetheless highlights the value of fine-tuning. For more details, see the supplementary materials. **(b)** A collection of state-tuned models from Earth Index yields a comprehensive, high-fidelity map of industrial poultry facilities, with 16,372 detections across the six states of interest.

iteration. Beginning with a model based on labeled data from North Carolina (Handan-Nader and Ho, 2019), and proceeding through negative sampling and iterative training, we derived a collection of state-based models that yields, to our knowledge, the most comprehensive available map of poultry facilities for the region (Figure 2). No federal agency maintains a comprehensive registry of CAFO locations, and patchwork regulations leave significant gaps in existing state-level data (Miller and Muren, 2019).

More details on the modeling process can be found in the Supplementary materials.

4 Conclusion

The development of novel digital applications for remote sensing foundation models signals a transformative shift in environmental monitoring. These applications simplify and accelerate labeling, model fine tuning, and intuition building, enabling human-in-the-loop workflows that generate accurate, locally fine-tuned data. By bridging the gap between data and domain experts, these technologies allow downstream users like policymakers, journalists, and environmental advocates to build impactful datasets with minimal technical expertise.

Digital applications built on AI foundation models for remote sensing data are powerful enablers of ecotechnologies. They can process massive amounts of environmental data, offer real-time monitoring, and generate predictive analytics that enhance our ability to manage and restore ecosystems. Paired with experts through iterative analyses, these tools can serve as a bridge from raw data to actionable insights that drive more integrated, ecologically informed interventions.

As foundation models and digital applications continue to mature, we envision a future where the technology stack for environmental monitoring is increasingly modular and simplified. Downstream users who want to make use of remote sensing data will no longer be required to learn specific technical skills for machine learning model development. Instead, technologies driven by simple interfaces will enable a collaborative future where domain experts directly participate in the model development process. Digital applications will be key to realizing the full potential of foundation models, ultimately enabling

critically needed environmental monitoring and stewardship at a global scale.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://github.com/earthrise-media/cafo-explorer/tree/main/data.

Author contributions

BS: Conceptualization, Methodology, Supervision, Writing – original draft. EB: Conceptualization, Investigation, Methodology, Writing – original draft. CK: Conceptualization, Visualization, Writing – review & editing. TI: Conceptualization, Software, Writing – review & editing. MM: Conceptualization, Supervision, Writing – review & editing.

Funding

Acknowledgments

including Sentinel-2 satellite imagery, the Zhu lab's foundation models, and the poultry CAFO dataset of Handan-Nader and Ho.

Conflict of interest

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.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fclim.2025.1520242/full#supplementary-material

References

Allen, M., Dorr, F., Gallego-Mejia, J. A., Martínez-Ferrer, L., Jungbluth, A., Kalaitzis, F., et al. (2023). Fewshot learning on global multimodal embeddings for earth observation tasks. Available at: https://arxiv.org/abs/2310.00119 (Accessed May 7, 2025).

Bastani, F., Wolters, P., Gupta, R., Ferdinando, J., and Kembhavi, A. (2023). "SatlasPretrain: a large-scale dataset for remote sensing image understanding," In *2023 IEEE/CVF international conference on computer vision (ICCV)*, 16726–16736.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). On the opportunities and risks of foundation models. Available at: https://arxiv.org/abs/2108.07258 (Accessed May 7, 2025).

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. (2020). "A simple framework for contrasting learning of visual representations." In *ICML '20: Proceedings of the 37th international conference on machine learning* 149, 1597–1607.

Cong, Y., Khanna, S., Meng, C., Liu, P., Rozi, E., He, Y., et al. (2022). SatMAE: pre-training transformers for temporal and multi-spectral satellite imagery. In NeurIPS 2022. *Adv. Neural Inf. Proces. Syst.* 35, 197–211.

Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F. N., and Leip, A. (2021). Food systems are responsible for a third of global anthropogenic GHG emissions. *Nat. Food.* 2, 198–209. doi: 10.1038/s43016-021-00225-9

Guo, X., Lao, J., Dang, B., Zhang, Y., Yu, L., Ru, L., et al. (2024). Skysense: a multi-modal remote sensing foundation model towards universal interpretation for earth observation imagery. *Proceed. IEEE/CVF Confer. Computer Vision Pattern Recog.*, 27672–27683. doi: 10.1109/CVPR52733.2024.02613

Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., et al. (2022). A survey on vision transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* 45, 87–110. doi: 10.1109/TPAMI.2022.3152247

Handan-Nader, C., and Ho, D. E. (2019). Deep learning to map concentrated animal feeding operations. *Nat. Sustain.* 2, 298–306. doi: 10.1038/s41893-019-0246-x

Hu, K. (2023). ChatGPT sets record for fastest-growing user base - analyst note. Available online at: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ (Accessed Oct. 29, 2024)

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., et al. (2021). Lora: Low-rank adaptation of large language models. Available at: https://arxiv.org/abs/2106.09685 (Accessed May 7, 2025).

Hu, Y., Yuan, J., Wen, C., Lu, X., and Li, X. (2023). RSGPT: a remote sensing vision language model and benchmark. Available at: https://arxiv.org/abs/2307.15266 (Accessed May 7, 2025).

IBM Research. (2024). Introducing IBM geospatial studio: A new AI-powered tool for geospatial analysis. IBM research blog. Available online at: https://research.ibm.com/blog/img-geospatial-studio-think (Accessed Mar. 6, 2025)

IBM Research and NASA (2025). IBM-NASA Prithvi Models Family. Available online at: https://huggingface.co/ibm-nasa-geospatial (Accessed Mar. 6, 2025)

Irvin, J. A., Liu, E. R., Chen, J. C., Dormoy, I., Kim, J., Khanna, S., et al. (2024). TEOChat: A Large Vision-Language Assistant for Temporal Earth Observation Data. Available at: https://arxiv.org/abs/2410.06234 (Accessed May 7, 2025).

Jakubik, J., Roy, S., Phillips, C. E., Fraccaro, P., Godwin, D., Zadrozny, B., et al. (2023). Foundation models for generalist geospatial artificial intelligence. Available at: https://arxiv.org/abs/2310.18660 (Accessed May 7, 2025).

Klemmer, K., Rolf, E., Robinson, C., Mackey, L., and Rußwurm, M. (2023). Satclip: Global, general-purpose location embeddings with satellite imagery. Available at: https://arxiv.org/abs/2311.17179 (Accessed May 7, 2025).

Kuckreja, K., Danish, M. S., Naseer, M., Das, A., Khan, S., and Khan, F. S. (2024). Geochat: grounded large vision-language model for remote sensing. *Proceed. IEEE/CVF Confer. Computer Vision Pattern Recog.*, 27831–27840. Available at: https://openaccess.thecvf.com/content/CVPR2024/html/Kuckreja_GeoChat_Grounded_Large_Vision-Language_Model_for_Remote_Sensing_CVPR_2024_paper.html

Lacoste, A., Lehmann, N., Rodriguez, P., and Sherwin, E. (2023). GEO-bench: toward foundation models for earth monitoring. In NeurIPS 2023. *Adv. Neural Inf. Proces. Syst.* 36, 51080–51093.

Lacoste, A., Sherwin, E., Kerner, H., Alemohammad, H., and Lütjens, B., (2021). "Toward foundation models for earth monitoring: proposal for a climate change benchmark." In *NeurIPS 2021: Workshop on tackling climate change with machine learning*, 73.

Li, H., Dou, X., Tao, C., Wu, Z., Chen, J., Peng, J., et al. (2020). RSI-CB: a large-scale remote sensing image classification benchmark using crowdsourced data. *Sensors* 20:1594. doi: 10.3390/s20061594

Liu, F., Chen, D., Guan, Z., Zhou, X., Zhu, J., Ye, Q., et al. (2024). Remoteclip: a vision language foundation model for remote sensing. *IEEE Trans. Geosci. Remote Sens.* 62, 1–16. doi: 10.1109/TGRS.2024.3390838

Lu, S., Guo, J., Zimmer-Dauphinee, J. R., Nieusma, J. M., Wang, X., Van Valkenburgh, P., et al. (2024). AI foundation models in remote sensing: A survey. Available at: https://arxiv.org/abs/2408.03464 (Accessed May 7, 2025).

Marsocci, V., Jia, Y., le Bellier, G., Kerekes, D., Zeng, L., Hafner, S., et al. (2024). PANGAEA: A global and inclusive benchmark for geospatial foundation models. Available at: https://arxiv.org/abs/2412.04204 (Accessed May 7, 2025).

Miller, D. L., and Muren, G. (2019). CAFOs: what we Don't know is hurting us. New York: Natural Resources Defense Council. Available online at: https://www.nrdc.org/sites/default/files/cafos-dont-know-hurting-us-report.pdf (Accessed Oct. 29, 2024)

Nicole, W. (2013). CAFOs and environmental justice: the case of North Carolina. *Environ. Health Perspect.* 121, a182–a189. doi: 10.1289/ehp.121-a182

OpenAI (2023). Gpt-4 technical report. Available at: https://arxiv.org/abs/2303.08774 (Accessed May 7, 2025).

Planet Labs PBC. (2025) Planet and anthropic partner to use Claude's advanced AI capabilities to turn geospatial satellite imagery into actionable insights. Business Wire. Available online at: https://www.businesswire.com/news/home/20250306606139/en/ (Accessed Mar. 6, 2025).

Robinson, C., Chugg, B., Anderson, B., Ferres, J. M. L., and Ho, D. E. (2022). Mapping industrial poultry operations at scale with deep learning and aerial imagery. *IEEE J Sel. Top. Appl. Earth Obs. Remote Sens.* 15, 7458–7471. doi: 10.1109/JSTARS.2022.3191544

Szwarcman, D., Roy, S., Fraccaro, P., Gíslason, Þ. E., Blumenstiel, B., and Ghosal Moreno, J. B. (2025). Prithvi-EO-2.0: A versatile multi-temporal foundation model for Earth observation applications. Available at: https://arxiv.org/abs/2412.02732 (Accessed May 7, 2025).

Tao, C., Qi, J., Guo, M., Zhu, Q., and Li, H. (2023). Self-supervised remote sensing feature learning: learning paradigms, challenges, and future works. *IEEE Trans. Geosci. Remote Sens.* 61, 1–26. doi: 10.1109/TGRS.2023.3276853

Tseng, G., Cartuyvels, R., Zvonkov, I., Purohit, M., Rolnick, D., and Kerner, H. (2023). "Lightweight, pre-trained transformers for remote sensing Timeseries." In *NeurIPS 2023: Workshop on tackling climate change with machine learning*, 58.

Wang, Y., Braham, N. A. A., Xiong, Z., Liu, C., Albrecht, C., and Zhu, X. X. (2023). SSL4EO-S12: a large-scale multimodal, multitemporal dataset for self-supervised

learning in earth observation [software and data sets]. *IEEE Geosci. Remote Sensing Magazine.* 11, 98–106. doi: 10.1109/MGRS.2023.3281651

Wang, D., Zhang, J., Xu, M., and Liu, L. (2024). MTP: advancing remote sensing foundation model via multi-task pretraining. *IEEE J Sel. Top. Appl. Earth Observ. Remote Sens.* 17, 11632–11654. doi: 10.1109/JSTARS.2024.3408154

Wing, S., Cole, D., and Grant, G. (2000). Environmental injustice in North Carolina's hog industry. *Environ. Health Perspect.* 108, 225–231. doi: 10.1289/ehp.00108225

Wu, Q., and Osco, L. P. (2023). Samgeo: a Python package for segmenting geospatial data with the segment anything model (SAM). *J. Open Source Softw.* 8:5663. doi: 10.21105/joss.05663

Yusuf, I. S., Yusuf, M. L., Panford-Quainoo, K., and Pretorius, A. (2024). A Geospatial Approach to Predicting Locust Breeding Grounds in Africa. Available at: https://arxiv.org/abs/2403.06860 (Accessed May 7, 2025).

Zhang, M., Yang, B., Hu, X., Gong, J., and Zhang, Z. (2024). Foundation model for generalist remote sensing intelligence: potentials and prospects. *Sci. Bull.* 69:3652. doi: 10.1016/j.scib.2024.09.017

Zhu, X. X., Xiong, Z., Wang, Y., Stewart, A. J., Heidler, K., Wang, Y., et al. (2024). On the Foundations of Earth and Climate Foundation Models. Available at: https://arxiv.org/abs/2405.04285 (Accessed May 7, 2025).