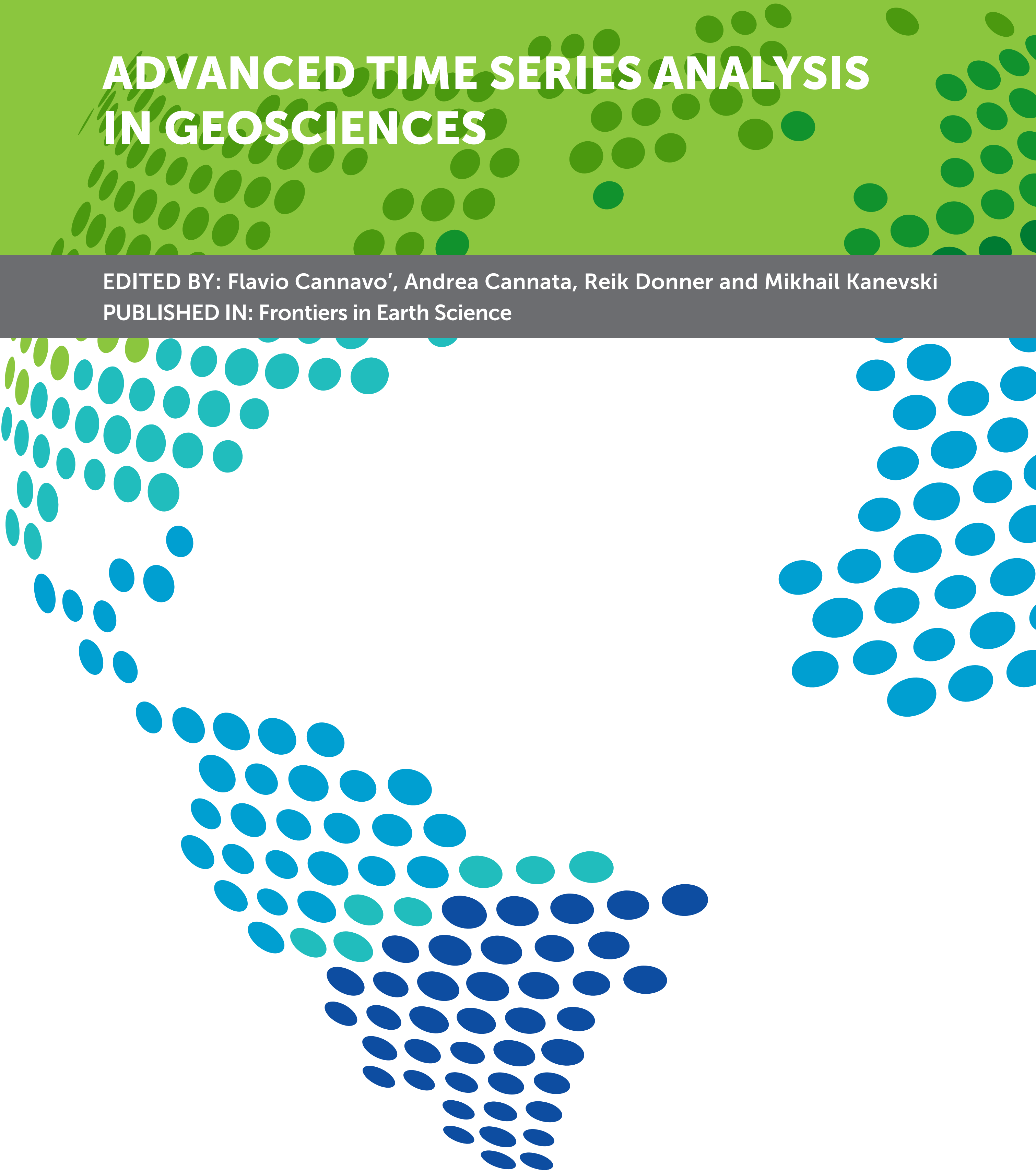
ADVANCED TIME SERIES ANALYSIS IN GEOSCIENCES

EDITED BY: Flavio Cannavo', Andrea Cannata, Reik Donner and Mikhail Kanevski
PUBLISHED IN: Frontiers in Earth Science



frontiers Research Topics

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: frontiersin.org/about/contact

ISSN 1664-8714
ISBN 978-2-88966-765-9
DOI 10.3389/978-2-88966-765-9

ADVANCED TIME SERIES ANALYSIS IN GEOSCIENCES

Topic Editors:
Flavio Cannavo', Section of Catania, Italy
Andrea Cannata, University of Catania, Italy
Reik Donner, Hochschule Magdeburg-Stendal, Germany
Mikhail Kanevski, University of Lausanne, Switzerland

Citation: Cannavo', F., Cannata, A., Donner, R., Kanevski, M., eds. (2021). Advanced Time Series Analysis in Geosciences. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-765-9

Table of Contents

frontiers in Earth Science

EDITORIAL
published: 22 March 2021
doi: 10.3389/feart.2021.666148

Editorial: Advanced Time Series Analysis in Geosciences

Flavio Cannavò[1]*, *Andrea Cannata*[1,2], *Reik V. Donner*[3] *and Mikhail Kanevski*[4]

[1]*Istituto Nazionale Di Geofisica e Vulcanologia, Osservatorio Etneo, Catania, Italy,* [2]*Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università Degli Studi Di Catania, Catania, Italy,* [3]*Department of Water, Environment, Construction & Safety, Magdeburg-Stendal University of Applied Sciences, Magdeburg, Germany,* [4]*University of Lausanne, Lausanne, Switzerland*

Keywords: time series, machine learning, data analytics, methods development, data mining

Editorial on the Research Topic

Advanced Time Series Analysis in Geosciences

A *time series* is an ordered sequence of data indexed by time. In other words, it is a sequence of discrete-time data, usually obtained at equally spaced points in time. Time series analysis is the attempt of extracting meaningful characteristics and statistical information from data organized in chronological order.

Nowadays, there are numerous types of data analysis approaches available for time series which are suitable for different purposes: diagnosing past behavior, prediction and forecasting, curve fitting, interpolation and extrapolation, classification and clustering, segmentation and decomposition, frequency characterization, etc.

The theoretical advances in time series analysis started early at the beginning of the last century with new developments in the field of stochastic processes. The first actual application of autoregressive models to time series can be identified in the work of Yule (1927) and Walker (1931). But it is since the pioneering book "Time Series Analysis" by Box and Jenkins in the 1970s (Box and Jenkins, 1970), that many lines of study in time series analysis have been developed.

Today we are witnessing a rapid increase in quantity, quality and importance of time series data in Earth Sciences. Across its vast number of subdisciplines, the massive production of data, e.g., through the growth of continuous monitoring networks and the availability of abundant remote sensing data, is making increasingly important the use of analysis tools capable of synthesizing information contained in large time series.

To deal with the increasing amount of available data in an automated way, the first emerging approaches of machine learning in time series analysis date back to the early 1980s (Nielsen, 2019). At present, although classical methods are still dominant, machine learning is rapidly emerging as a valid alternative approach to time series analysis, finding effectiveness especially in multivariate time series.

It is clear to everyone that, as continuous monitoring and data gathering become even more common in geosciences, the need for powerful time series analysis techniques, either classical/statistical or machine learning based techniques, will further increase. The impact of this need is proven by an exponential growth shown by the occurrence of the keyword "time series analysis" in papers published from 1985 to 2020, as indicated by both WoS (https://webofknowledge.com/) and Scopus (https://www.scopus.com/) databases (**Figure 1**).

In this context, this Research Topic collects some illustrative examples of state of the art research from across the world to delineate the dramatic and diverse nature of time series analysis in geosciences. With the analysis of data over time providing the basis of many modern scientific disciplines, this research covers a variety of applications in the field of geosciences.

Bueno et al. (https://doi.org/10.3389/feart.2019.00335/full) develop an innovative algorithm to detect volcanic explosions in infrasound time series. The algorithm makes extensive use of classical

Edited and reviewed by:
Peng Liu,
Institute of Remote Sensing and Digital Earth (CAS), China

****Correspondence:***
Flavio Cannavò
flavio.cannavo@ingv.it

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Earth Science

Received: *09 February 2021*
Accepted: *18 February 2021*
Published: *22 March 2021*

Citation:
Cannavò F, Cannata A, Donner RV and Kanevski M (2021) Editorial: Advanced Time Series Analysis in Geosciences. Front. Earth Sci. 9:666148. doi: 10.3389/feart.2021.666148

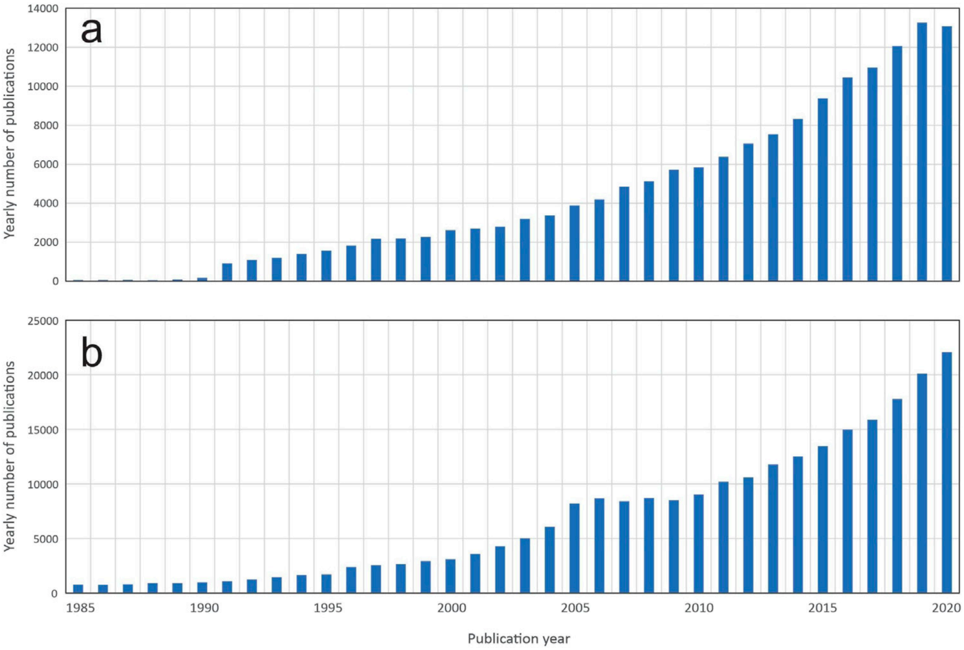


FIGURE 1 | Yearly number of publications with the keyword "time series analysis" per year during 1985–2020 from WoS **(A)** and Scopus **(B)** database.

signal processing techniques combined together to characterize continuous volcano-acoustic records and detect explosive events.

With a similar aim, and for a closely related type of observational signals, Sciotto and Montalto (https://doi.org/10.3389/feart.2020.579923/abstract) propose to use a subspace-based trigger algorithm to automatically detect also weak events and outperform the classical STA/LTA (short-time average/long-time average) detection algorithm, commonly used in seismic time series.

Staying on the same kind of acoustic infrasound data in volcanoes, De Angelis et al. (https://doi.org/10.3389/feart.2020.00169/full) propose a way to process the signals recorded through arrays of microphones to detecting and tracking acoustic sources from multiple craters and active vents associated with different types of volcanic activities. The proposed algorithm is based on least-squares beamforming, and allows to evaluate the direction and speed of propagation of acoustic waves between source and array, and includes a quantitative assessment of the uncertainty on array measurements based on error propagation theory.

In active volcanic areas, seismo-acoustic events represent a great source of information on volcano status. Giudicepietro et al. (https://doi.org/10.3389/feart.2020.581742/full) introduce a modern unsupervised neural network approach for clustering experimental seismo-acoustic events. The technique belongs to the growing field of machine learning applications, and the authors show its appropriateness for clustering natural events such as the seismo-acoustic transients accompanying Strombolian explosions. They also state that their parameterization strategy may be suitable to extract significant features of the seismo-acoustic signals linked to the physical conditions of the volcanic system.

Still focusing on a volcano-related topic, another interesting paper has been published by Rodríguez-Molina et al. (https://doi.org/10.3389/feart.2020.577588/full), where heterogeneous volcano deformation time series are analyzed to reconstruct the evolution of volcanic source volumes and understand the time scale of inter-eruptive processes. The adopted techniques draw inspiration from the classical regularized linear regression and Bayesian inversion.

In the context of oil extraction, Mancinelli (https://doi.org/10.3389/feart.2020.00285/full) show how a 3D forward density model can be used to compute the differential gravity signature through time, e.g., before and after a production period, thus helping to locate the causative source due to fluid production.

Moschella et al. (https://doi.org/10.3389/feart.2020.00114/full) analyze and put in a relation time series of microseism recorded by seismic coastal stations with time series of significant sea wave height measured offshore. They adopt some state-of-the-art machine learning algorithms to create a regression model able to link the two different kinds of signals.

As for the climate change topic, Diodato et al. (https://doi.org/10.3389/feart.2020.561148/full) use statistical techniques to build a model able to reconstruct the world longest time series of annual number of days of snow on the ground (1,681–2018).

Within the same field, Li et al. (https://doi.org/10.3389/feart.2020.00059/full) analyze the time series of net primary production of China, and two indices linked with droughts.

They employ statistical approaches to analyze the data both in spatial and time domains.

Guignard et al. (https://doi.org/10.3389/feart.2020.00255/full) present a more theoretical paper with broad possible applications. They propose the Fisher-Shannon information method as an effective data exploration tool able to give diverse insights into complex non-stationary time series. They also provide a *Python* toolbox for the Fisher-Shannon method.

Smaï and Wawrzyniak (https://doi.org/10.3389/feart.2020.00296/full) produced an open source *Python* library, called Razorback, implemented to handle, manipulate, and combine time series of synchronous data produced in magnetotelluric studies.

Finally, Loose et al. (https://doi.org/10.3389/feart.2020.537028/full) propose two machine learning-based methods (a generalized additive model and a long short-term memory neural network model) for bias correction of *in situ* sensor data. In particular, such methods are applied to data from submersible and shipboard mass spectrometers.

Of the great variety of techniques and applications in time series analysis, in this collection we have seen how the human ability to analyze data is still irreplaceable, despite the glimpse of the great potential in artificial intelligence approaches for scientific time series studies.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank all authors, reviewers, and editors that have contributed to this Research Topic.

REFERENCES

Box, G. E. P., and Jenkins, G. M. (1970). *Time series analysis: forecasting and control*. Hoboken New Jersey: Holden Day.

Nielsen, A. (2019). *Practical time series analysis: prediction with statistics and machine learning*. Sebastopol California: O'Reilly Media, Inc.

Walker, G. T. (1931). On periodicity in series of related terms. *Proc. R. Soc. Lond. Ser. A Containing Pap. Math. phys. Character.* 131 (818), 518–532. doi:10.1098/rspa.1931.0069

Yule, U. G. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Trans. R. Soc. Lond. Ser. A.* 226, 267–298. doi:10.1098/rsta.1927.0007

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.

frontiers in Earth Science

TECHNOLOGY AND CODE
published: 13 December 2019
doi: 10.3389/feart.2019.00335

VINEDA—Volcanic INfrasound Explosions Detector Algorithm

Angel Bueno[1]*, Alejandro Diaz-Moreno[2], Isaac Álvarez[1], Angel De la Torre[1], Oliver D. Lamb[3], Luciano Zuccarello[1,4,5] and Silvio De Angelis[2]

[1] Department of Signal Theory, Telematics and Communications, University of Granada, Granada, Spain, [2] Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom, [3] Department of Geological Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, [4] Department of Theoretical Physics and Cosmos, University of Granada, Granada, Spain, [5] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Pisa, Italy

Infrasound is an increasingly popular tool for volcano monitoring, providing insights of the unrest by detecting and characterizing acoustic waves produced by volcanic processes, such as explosions, degassing, rockfalls, and lahars. Efficient event detection from large infrasound databases gathered in volcanic settings relies on the availability of robust and automated workflows. While numerous triggering algorithms for event detection have been proposed in the past, they mostly focus on applications to seismological data. Analyses of acoustic infrasound for signal detection is often performed manually or by application of the traditional short-term average/long-term average (STA/LTA) algorithms, which have shown limitations when applied in volcanic environments, or more generally to signals with poor signal-to-noise ratios. Here, we present a new algorithm specifically designed for automated detection of volcanic explosions from acoustic infrasound data streams. The algorithm is based on the characterization of the shape of the explosion signals, their duration, and frequency content. The algorithm combines noise reduction techniques with automatic feature extraction in order to allow confident detection of signals affected by non-stationary noise. We have benchmarked the performances of the new detector by comparison with both the STA/LTA algorithm and human analysts, with encouraging results. In this manuscript, we present our algorithm and make its software implementation available to other potential users. This algorithm has potential to either be implemented in near real-time monitoring workflows or to catalog pre-existing databases.

Keywords: volcanic infrasound explosions, automatic detection, signal processing, characteristic function, sub-band processing

Edited by:
Andrea Cannata,
University of Catania, Italy

Reviewed by:
John J. Lyons,
U.S. Geological Survey, Alaska, United States
Stephen Arrowsmith,
Southern Methodist University, United States

***Correspondence:**
Angel Bueno
angelbueno@ugr.es

Specialty section:
This article was submitted to Volcanology, a section of the journal Frontiers in Earth Science

Received: *18 September 2019*
Accepted: *02 December 2019*
Published: *13 December 2019*

Citation:
Bueno A, Diaz-Moreno A, Álvarez I, De la Torre A, Lamb OD, Zuccarello L and De Angelis S (2019) VINEDA—Volcanic INfrasound Explosions Detector Algorithm. Front. Earth Sci. 7:335. doi: 10.3389/feart.2019.00335

1. INTRODUCTION

Seismic and acoustic signals are key in monitoring and characterizing volcanic unrest. Recent technological advances in sensor development, data transmission, and archival protocols have made the collection of large amounts of geophysical data commonplace at active volcanoes and other monitoring environments. The sheer amount of data recorded makes their manual analysis a challenging, frequently unfeasible, task. The implementation of automated tools to address this challenge is, thus, vital for effective monitoring operations.

Automatic event detection and classification work-flows applied to seismic data include an initial segmentation stage, commonly via the application of short-term average/long-term average (STA/LTA) algorithms in order to parse the continuous seismograms into individual earthquake waveforms with varied characteristics and sources (Allen., 1982). In more advanced processing work-flows, this is followed by automatic classification of the signals by different methods, including Neural Networks (Scarpetta et al., 2005), pattern recognition (e.g., Curilem et al., 2014), Hidden Markov Models (Ibáñez et al., 2009), Support Vector Machines (Giacco et al., 2009), or statistical properties (Bueno et al., 2019). A wealth of new algorithms are constantly published in the literature in order to improve the efficiency of automatic detection and classification procedures for different types of signals, including those associated with tectonic earthquakes (Di Stefano et al., 2006; Álvarez et al., 2013; Bhatti et al., 2016), low-frequency volcano-seismic events (Frank and Shapiro, 2014), avalanches (Marchetti et al., 2015), and debris flows (Schimmel and Hübl, 2016). Collectively, these algorithms represent an important toolbox for the creation of high-quality research databases.

Volcano infrasound is becoming increasingly popular as a monitoring tool (Johnson and Ripepe, 2011; Fee et al., 2013); among other applications, acoustic data are frequently used for detection and characterization of explosive volcanic activity (e.g., Garcés et al., 1999; Johnson et al., 2004; Vergniolle and Ripepe, 2008; Caplan-Auerbach et al., 2010; Fee and Matoza, 2013; Lamb et al., 2015; De Angelis et al., 2019). Volcanic explosions are commonly recorded by infrasound microphones in the 0.01–20 Hz frequency band as signals characterized by impulsive onsets followed by codas with variable duration, from few seconds to several minutes. Due to the increasing amount of acoustic infrasound data routinely collected on active volcanoes, the development of tools for automated signal detection is crucial for efficient monitoring. Here, we introduce an adaptative infrasound detector based on time- and frequency-domain characterization of volcanic explosion signals. We take advantage of advanced signal processing techniques, in combination with *a-priori* knowledge of recorded explosions, to implement a robust infrasound detector based on adaptive multi-band processing. The datasets selected for this study are obtained from two previous monitoring campaigns at Santiaguito (Lamb et al., 2019) and Mount Etna (Diaz-Moreno et al., 2019) volcanoes. On Mount Etna volcano, Diaz-Moreno et al. (2019) studied an inversion modeling workflow to derive infrasound acoustic sources and estimate rate and volume of erupted materials. On Santiaguito volcano, Lamb et al. (2019) provides a seismic and infrasound analysis of three eruptive phases in a multi-parametric monitoring framework, associating infrasound and seismic signals. Hence, the sheer volume of recorded explosions at both volcanoes makes an ideal ground for testing the VINEDA algorithm, as the geological properties are already known.

2. DETECTION OF INFRASOUND EXPLOSIONS

2.1. Algorithm Description

Detection of volcanic explosions in infrasound data streams is an important and challenging task. The algorithm detailed here, the **V**olcanic **IN**frasound **E**xplosions **D**etector **A**lgorithm (*VINEDA*), parses raw acoustic data streams, $x(n)$ into a normalized characteristic function (CF) within a given frequency range $[f_{low}, f_{high}]$; the frequency band is chosen to include the majority of energy transported by explosion infrasound. Explosion temporal boundaries are defined by the abrupt onset arrival and return to background amplitude, with expected average durations given by $D = \{D_{min} \dots D_{max}\}$, with D_{min} and D_{max} the minimum and maximum expected duration of infrasound explosions, respectively. The number of duration bands used in the discriminant analysis is defined as N_{db}. Background amplitude is defined as characteristic low-amplitude stationary signals (i.e., wind, noise, or lack of infrasound activity). The data processing pipeline is illustrated in **Figure 1**, key parameters for the algorithm are given in Algorithm 1, and **Figure 2** shows the outputs of each step of the workflow using infrasound recorded of an explosion at Santiaguito volcano, Guatemala.

As the first step, an anti-aliasing Finite Impulse Response (FIR) low-pass filter is applied with a corner frequency f_{high}, corresponding to the highest frequency component of interest in the explosion signal investigated. This stage eliminates noise outside the frequency band of interest for explosions, eases computational load and reduces signal complexity, while preserving important signal onset information. The sampling frequency, fs, is usually much higher than f_{high} (typical values for fs and f_{high} are of the order of 100 Hz and below 5 Hz, respectively); this allows, in the second processing step, downsampling of the input signal, $x(n)$, by a factor of $R = fs/fs'$, being $fs' = 2 \cdot f_{high}$. Downsampling is performed by a poly-phase anti-aliasing filter to compensate for potential delays due to re-sampling computations.

An adaptive de-trending filter is then applied to remove noise that could affect the lower frequency range of infrasound explosions, such as from wind or long-period instrument drift. Detrending is implemented as a zero-phase high-pass FIR filter. The order of this filter is given by the maximum duration of the explosion (D_{max}), whereas the cutoff frequency is fixed at f_{low} Hz. This filter also helps to mitigate the influence of low frequency noise, such as microbaroms, which is often located at very low frequency bands (i.e. [0.1–0.5] Hz) (Landès et al., 2012) A median filter is further applied to remove spikes in the signal, such as those associated with instrument or data transmission noise. This filter reduces extreme values and smooths the input signal (**Figure 2B**).

Since explosions are characterized by sharp onsets, a multi-band filter stage has been specifically designed to provide no edge delay (Álvarez et al., 2013). The design of the filter bank (number of bands, central frequencies and bandwidths) depends on the frequency content of the explosions to be detected. **Figure 2E** depicts an instance of the impulse response of the FIR filters designed. The center frequencies are uniformly distributed within the frequency band of the explosion signal $[f_{low}, f_{high}]$, according to:

$$fc_k = f_{low} + (k-1) \cdot \frac{(f_{high} - f_{low})}{N_{fb}} + \frac{1}{2} \cdot \frac{(f_{high} - f_{low})}{N_{fb}} \text{ with } k = 1, \dots, N_{fb} \qquad (1)$$

where N_{fb} is the number of sub-bands of the filter bank.

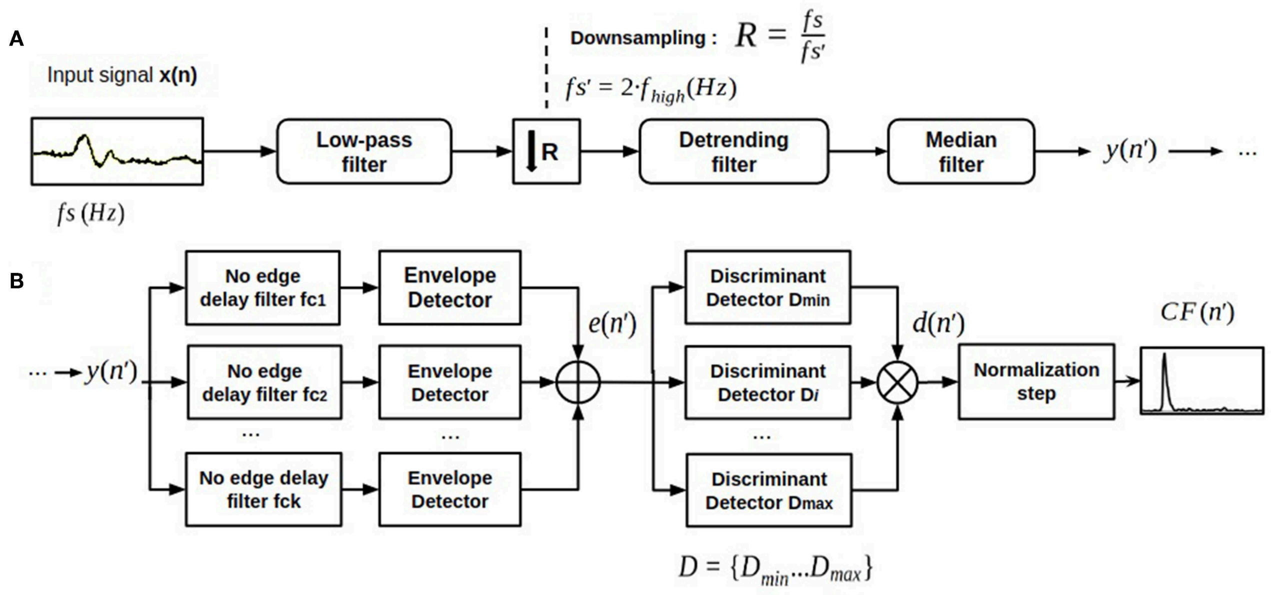


FIGURE 1 | VINEDA Algorithm block diagram. **(A)** The signal is decimated and filtered in order to suppress noise and avoid spike distortions. **(B)** Successive steps focus on the detection of infrasound explosions onset, using specialized designed filters and a discriminant detector based on a priori knowledge of the infrasonic explosions. The computed characteristic function *CF* contains the detected explosions.

Algorithm 1 VINEDA parameters definition.

Input:

x: Trace containing the infrasound waveform in data samples.

f_s: Sampling frequency of the infrasound waveform (Hz).

f_{low}: Lower frequency content of the explosions to be detected (Hz).

f_{high}: Upper frequency content of the explosions to be detected (Hz).

N_{fb}: Number of sub-bands used in the filter bank.

D_{min}: Minimum duration of the explosions to be detected (s).

D_{max}: Maximum duration of the explosions to be detected (s).

N_{db}: Number of duration bands used in the discriminant analysis.

β: Penalty factor for non-impulsive onsets.

Output:

CF: Characteristic function containing the detected explosions.

$f_{s'}$: Sampling frequency of the characteristic function (Hz).

Once the multi-bank filter is applied, signal envelopes are computed within each sub-band; the envelope, $e_k(n')$, for each sub-band is estimated as:

$$e_k(n') = \sqrt{s_k(n')^2 + \hat{s_k}(n')^2} \text{ with } k = 1, \ldots N_{fb} \tag{2}$$

where $\hat{s_k}(n')$ is the Hilbert Transform of $s_k(n')$ (Bracewell, 1999). All the sub-band envelopes are then added to obtain the signal $e(n')$. This global envelope allows characterizing how energy is delivered at the explosion onset (i.e., an abrupt rise followed by a slow decay). **Figure 2C** shows the detected envelope $e(n')$ of the explosion. The abrupt onset is detected, and the smooth decay is preserved.

The next step in the detection algorithm is the application of a discriminant filter to the signal $e(n')$ in order to compute a characteristic function, $CF(n')$, and detect explosion onsets (**Figure 2D**). The impulse response $h(n')$ of the proposed discriminant filter is shown in **Figure 2F**; this filter is designed to enhance signals with a sharp rise and gradual decay, such as infrasound explosions, with a duration in the order of D. Thus, the discriminant detector mitigates the effect of non-stationary noises while finding the best match of infrasonic signals (Álvarez et al., 2013). A penalty factor for non-impulsive onsets, β, is added to increase the robustness of the filter with respect to background noise. The value of β controls the impulsivity of the signal we are going to detect, and it should be selected on the basis of the expected onset. Larger β values are required to detect very impulsive onsets with respect to background noise.

In the final stage of the detection workflow, the characteristic function, $CF(n')$, is normalized following a non-linear companding method to emphasize onset arrivals without loosing amplitude information (Rabiner and Gold, 1975). The peak of $CF(n')$ corresponds to the onset of the detected explosion event (**Figure 2D**).

The sequence of filtering stages is summarized in **Figures 2A–D**. Note that the spike present in the raw signal $x(n)$ at ~18 s is suppressed during processing and the long-period trend is removed while the onset of the explosion signal is preserved. The $CF(n')$ for the original explosion waveform is shown in **Figure 2D**. The amplitude of the CF is proportional to the sharpness of the original explosion onset. Notice that the

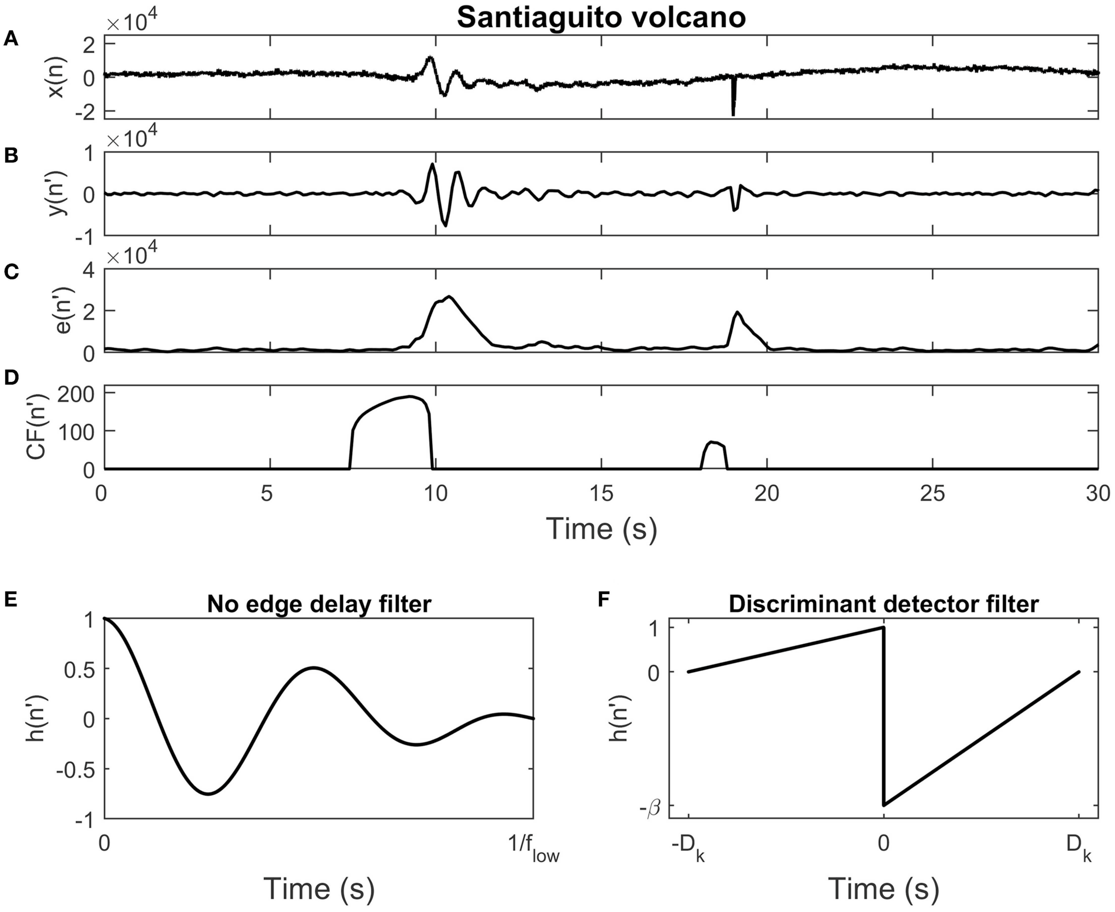


FIGURE 2 | Output example for all intermediate steps of the VINEDA workflow. **(A)** The original infrasound signal of an explosion at Santiaguito volcano, Guatemala, $x(n)$. **(B)** The signal $y(n')$ is the output after decimation, detrending and median filtering stages. **(C)** The signal $e(n')$ represents the output of the no edge delay filter bank. **(D)** A characteristic function $CF(n')$ is obtained from the signal, which can be used to perform the automatic detection. **(E)** Example of the impulse response of the bandpass filters. **(F)** Impulse response for the proposed discriminant detector (Álvarez et al., 2013).

discriminant detector has suppressed background noise, thus highlighting the explosion onset.

Finally, VINEDA is designed to be used with any single-station recordings. However, the flexibility of VINEDA allows its configuration in a parallel framework to work in a multi-station setting. Once the signals have been detected, derived CFs can be merged or interfaced with physical propagation models to locate the events of interest.

2.2. Multi-Station Application Example on Etna Volcano

VINEDA is a highly flexible algorithm that can be applied to network data to enhance detection of infrasound signals associated with volcanic explosions. Here, we demonstrate an application to network data recorded at Mt. Etna volcano by three infrasound sensors, during June 2017, with a frequency content between [1.0–3.0] Hz and maximum duration of $D_{max} = 5.0$ s (Diaz-Moreno et al., 2019). Each sensor is independently processed, as we aim to investigate how VINEDA detects infrasound explosions from stations installed at different locations from the volcanic vent.

Figure 3 shows the *CFs* along with signal envelopes ($e_1(n')$, $e_2(n')$, $e_3(n')$) for each of the input signals ($x_1(n)$, $x_2(n)$, $x_3(n)$). Note that, generally, the proposed pre-processing steps yield robust envelopes, as noise and long-period trends have been filtered out, and the no-edge delay filter along with the envelope detector characterize explosion onsets. However, when stations are closer to the vent, the recorded explosions are less attenuated with larger amplitudes and more obvious onsets ($ET01$, $ET04$, and $ET10$ located at 1.0, 1.3, and 6 km from the vent, respectively; Diaz-Moreno et al., 2019). While closer stations exhibit larger *CFs* as observed for $CF_1(n')$ and $CF_2(n')$, the detector is still able to observe a signal at the farthest station, $CF_3(n')$. The results presented in **Figure 3** highlight the capabilities of *VINEDA* to suppress much of the noise and still being able to detect attenuated explosions at greater distances. The robustness of the algorithm to perform multi-station detection permits a direct embedding with atmospheric propagation models to compute

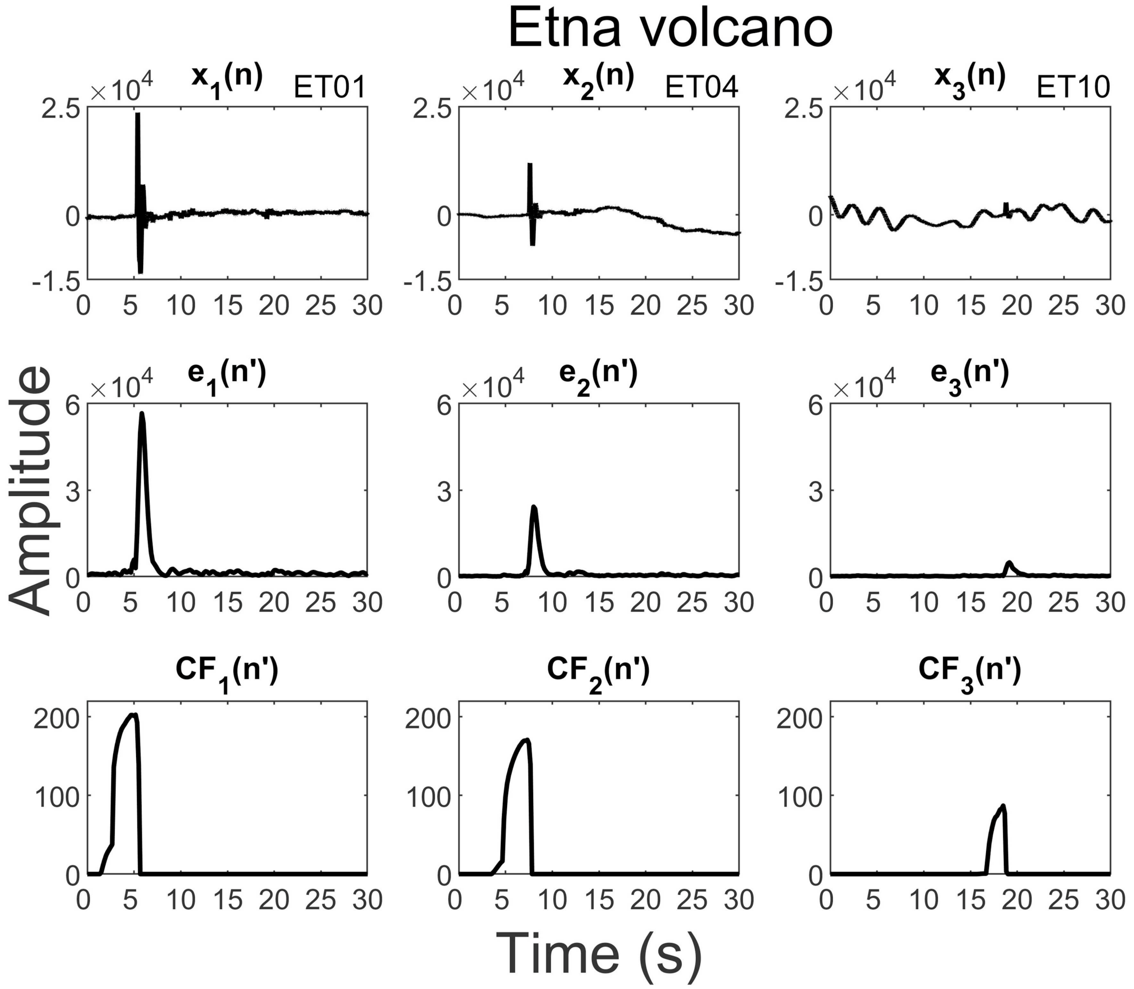


FIGURE 3 | Application example on a multi-station configuration at Etna volcano in three different stations (*ET*01, *ET*04, and *ET*10 located at 1.0, 1.3, and 6 km from the vent, respectively; Diaz-Moreno et al., 2019). Even if attenuated, the detector can recognize explosions at distant stations. When deployed on a single station, CF_3 could be considered as a false positive. However, when complemented with other *CFs* and propagation models, the number of false positives could be further reduced.

travel times based on VINEDA detections, thus reducing false positives and easing the location of infrasound sources.

3. DATA EXPERIMENTATION AND RESULTS

We applied the *VINEDA* algorithm to infrasound data from low-intensity Strombolian explosions recorded at Mt. Etna volcano during an experiment in the summer of 2017 (Diaz-Moreno et al., 2019). Additionally, the generalization capabilities of the algorithm across different volcanic settings are tested with infrasound data from Santiaguito (Lamb et al., 2019). A traditional STA/LTA algorithm is compared with *VINEDA* in order to assess its detection performance and robustness. In addition, 2 days of continuous data from both volcanoes were also manually evaluated by a group of 6 experts. For every waveform of a detected event, each individual expert checks the presence/absence of an event and assign a quality factor *Q*. Faced with the question *"Is this an explosion?"*, experts chose one of three quality assessments: "strongly agree" (Q = 1), "agree" (Q = 2), and "undecided" (Q = 3). **Figure 4** depicts three infrasound signals with different quality factors, as reviewed by experts. Notice that signal to noise ratio (SNR) along is the highest for Q1, where the explosion is clearly visible. For lower quality values, the segmented explosion is almost masked with background fluctuations. All our experiments were performed on a 64 bit computer with an i7-8700k CPU (3.70 GHz) processor, 16 GB RAM, and Ubuntu 16.04. On this machine, VINEDA takes and average of 0.72 s (over 10 runs) to process 1 day of infrasound data of both volcanoes.

The Receiver Operating Characteristic (ROC) curve is an excellent graphical technique to visualize the trade-off between the sensitivity and the specificity of any detection system, for a particular decision threshold (Fawcett, 2006). The *sensitivity* is a metric to evaluate the goodness of a model to detect true events and according to the common definition from Signal Detection Theory, the sensitivity can be calculated as:

$$sensitivity = \frac{TP}{FN + TP}$$

where infrasound explosions are successfully detected (making a true positive, TP), mistakenly ignored (making a false negative, FN), and vice versa: other events or noise can be mistakenly detected as explosions (making a false positive, FP).

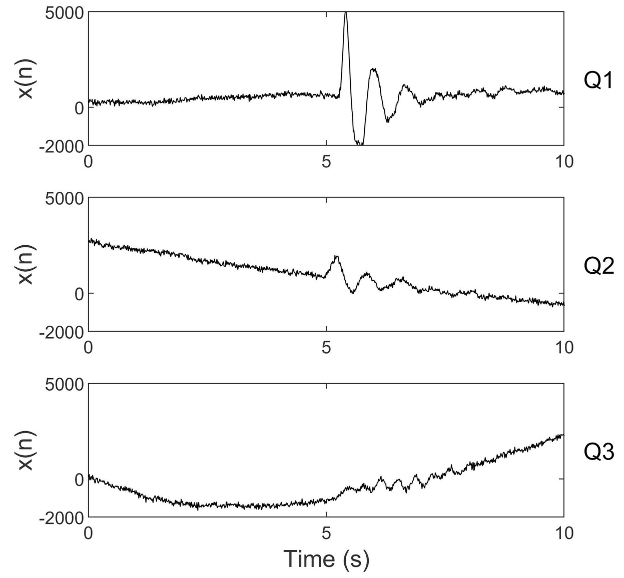


FIGURE 4 | Infrasound signals from Etna 2017 experiment, as presented to the experts during the annotation process. The assigned quality factors ($Q1, Q2, Q3$) is also shown. Notice that the highest quality of Q1 implies the best signal to noise ratio, with the explosion clearly distinguishable. For Q2, the event is detected by VINEDA, despite lower amplitude values. For Q3, atmospheric disturbances affect the infrasound signal, and even if the explosion is detected, a multi-station analysis would help to decrease false positives.

The *specificity* measures the proportion of actual negatives that are correctly identified as such. Since *VINEDA* is a detector (it detects infrasound explosions in an input signal in which an event could be found at any time) but not a classifier, the specificity of the detector is expressed as the rate of false positives per hour (FP/h).

Figure 5 illustrates the ROC curves of detection for Etna (left) and Santiaguito (right) volcanoes performed by *VINEDA* algorithm (red) when compared to STA/LTA algorithm (blue), and two quality criteria $Q \leq 2$ and $Q \leq 3$ as assessed by human analysts. For Santiaguito volcano, explosions were characterized by a duration of 4 s, with a frequency content in the range of [1.0–3.0] Hz. Mount Etna explosions were characterized by an average duration of 2 s, and a frequency content in the range of [1.0–3.0] Hz. In both settings the parameters β and N_{fb} were fixed to 3. Note that, from Equation (1), a $N_{fb} = 3$ yields a set of central frequencies that covers the range of infrasonic explosions for both volcanoes. The definition of the frequency band $[f_{low}, f_{high}]$ is essential to guarantee correct segmentation boundaries. For Etna and Santiaguito volcano with given band of 1.0–3.0 Hz., the computed frequencies for the multi-band analysis are $fc = [1.33, 2.00, 2.67]$ (see Equation 1), encompassing the range of frequencies of interest for all explosive activity analyzed. Similarly, a value of $\beta = 3$ helps to select events with sharp onsets mitigating the influence of the background noise. Given the mathematical design of the discriminant filter, larger values of β fit sharp onsets whilst decreasing the computed *CF* function for smaller (or attenuated) explosions. The vector range of durations D is defined, for both volcanoes and a delta time-step of 1 s, between $D_{min} = 2.0$ and $D_{max} = 5.0$ s. In practice, the selected parameters suffice for the range of frequencies in infrasonic explosions for these volcanoes, but this definition is highly flexible and is left to the analyst on a per-case basis.

All streams were sampled at $fs = 100$ Hz. The detection thresholds (thr) varied between 0 and 240. STA/LTA is a simple and effective method to detect transient events that defines a characteristic function for detection as the ratio between the STA and the LTA (the absolute values of the seismogram averaged over short and long windows, respectively). In such way, the LTA tracks the background seismic noise, while the STA/LTA ratio increases when there is a sudden increase in signal amplitude

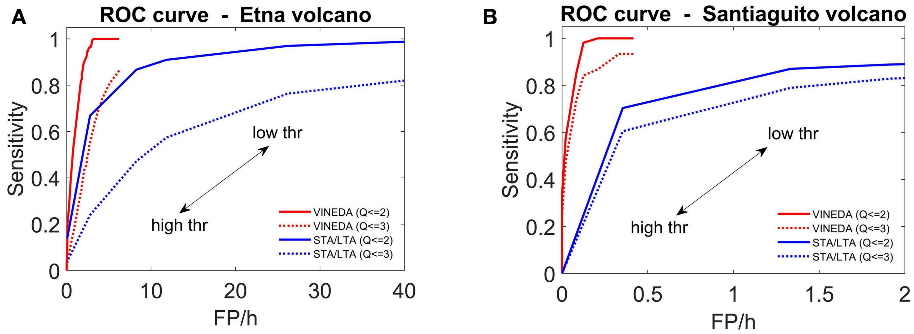


FIGURE 5 | ROC curve for ETNA **(A)** and Santiaguito **(B)** volcanoes comparing *VINEDA* and STA/LTA algorithms for different quality assessments of the experts (thr = thresholds).

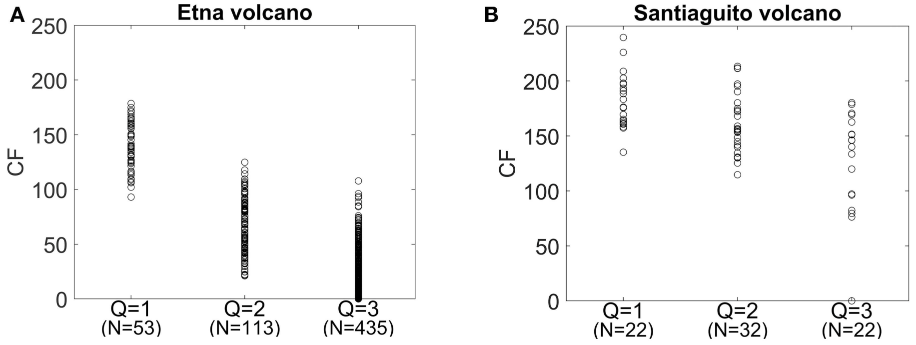


FIGURE 6 | Values of the characteristic function in the *VINEDA* algorithm for Etna **(A)** and Santiaguito **(B)** volcanoes, for the events manually assessed by the experts with the three different quality criterion considered: "strongly agree" (Q = 1), "agree" (Q = 2), and "undecided" (Q = 3). *N* indicates the total number of detected explosions by the algorithm with the assigned quality factor.

(Allen., 1982). Using the infrasound data gathered at Santiaguito and Mount Etna, STA/LTA was applied to the band of interest ([1.0–3.0] Hz for both volcanoes), with short and long windows of 1 and 10 s, respectively.

The area under the ROC curve could provide an idea of the benefits of any particular algorithm. For an algorithm with a good performance, the area under this curve would be maximum since the ideal detection threshold would be the one providing a result closest to Sensitivity = 1 and FP/h = 0. **Figure 5** shows that *VINEDA* detector outperforms STA/LTA for both volcanoes. The performance including events in which experts are "sure" ($Q \leq 2$) is better than that for "not sure" events ($Q \leq 3$) in all cases. The improvement in *sensitivity* and rate of *false positives per hour* for *VINEDA* compared to STA/LTA in all cases is associated to the specific processing of infrasound explosion signals that the algorithm *VINEDA* carries out. The maximum value of the *CF* function depends on the quality of the explosion signal. **Figure 6** shows the real value of the *CF* against the quality factor assigned by the experts, in combination with the total number of detected explosions (*N*). We observe that higher values on the *CF* function are obtained for events with better quality. These events would be detected using high values of the detection threshold, while keeping the number of detected explosions low. This is in line with the design of the algorithm, as restrictive detection thresholds will only retrieve very distinctive explosions with abrupt onsets, discarding the rest. By contrast, for less demanding thresholds, the number of detected explosions increases but the maximum value of the *CF* decreases. This behavior is expected, and should be taken into account to set the detection threshold depending on the particular needs or application.

The ever-increasing availability of infrasonic data requires the development of mathematical routines that can be used to detect events of geophysical interest. We have presented VINEDA; a generic and scalable multi-step algorithm designed to detect infrasound explosions. Our experimental evaluation with data from two volcanoes, Santiaguito and Etna, suggest that VINEDA improves performance over STA/LTA approaches. For both volcanoes, the refinement of the detector shown in the ROC curve, jointly with the value of the estimated *CF* function, confirms the capabilities of VINEDA to surpass STA/LTA. The discriminant detector helps to filter out non-stationary noises and acts as a penalty to temporally longer infrasound events, such as rockfalls, avalanches, or degassing. Further, the quality of detections is strengthened by VINEDA's capabilities to function in a multi-station setting. We suggest that the procedure described here can be used to annotate high-quality data from sequential infrasound streams for further post-processing, including the training of advanced machine learning models, picking algorithms or geo-statistical modeling.

4. CONCLUSIONS

Acoustic infrasound provides unique insights on the dynamics of erupting volcanoes. The detection and characterization of explosions from large streams of continuous, multi-channel, infrasound data is a challenging task. In this manuscript, we have introduced *VINEDA*, an infrasound detector which makes extensive use of signal processing techniques in order to characterize continuous volcano acoustic records and extract explosion signals. This algorithm stands as a middle point between the complete knowledge of the target signal beforehand, and incomplete knowledge of the explosions on the recorded infrasonic-stream as some prior knowledge of signal features suffice for VINEDA to detect target signals. *VINEDA* is suitable for deployment within volcano monitoring systems and offers a trade-off between quality and quantity of detections. We suggest that real-time implementations of algorithms like *VINEDA* are crucial to improve existing infrasound datasets and ultimately, increase our ability to monitor unrest at active volcanoes.

DATA AVAILABILITY STATEMENT

The Santiaguito and Etna infrasound datasets selected for this study are available via request to the authors. VINEDA code is open-source and can be downloaded at https://github.com/srsudo/vineda.

AUTHOR CONTRIBUTIONS

AB, AD-M, and IÁ developed and implemented the VINEDA algorithm, wrote the manuscript, and produced the figures. AD revised the signal processing part. OL, LZ, and SD contributed to data-gathering, data quality checking, software testing, and infrasound theoretical discussions. All authors contributed to the manuscript reviews and distribution of ideas.

FUNDING

This research was partially funded by KNOWAVES TEC2015-68752 (MINECO/FEDER), by NERC Grant NE/P00105X/1, by Spanish research grant MECD Jose Castillejo CAS17/00154 and by VOLCANOWAVES European Union's Horizon 2020 Research and Innovation Programme Under the Marie Sklodowska-Curie Grant Agreement no 798480.

ACKNOWLEDGMENTS

We would like to thank all the members at the Seismic Laboratory at University of Liverpool for continuous support and advice when developing this algorithm. We thank John Lyons and the other reviewer for their suggestions and comments which have improved VINEDA.

REFERENCES

Allen., R. (1982). Automatic phase pickers: their present use and future prospects. *Bull. Seismol. Soc. Am.* 72, S225–S242.

Álvarez, I., García, L., Mota, S., Cortés, G., Benítez, C., and Torre, Á. D. (2013). An automatic P-phase picking algorithm based on adaptive multiband processing. *IEEE Geosci. Rem. Sens. Lett.* 10, 1488–1492. doi: 10.1109/LGRS.2013.2260720

Bhatti, S. M., Khan, M. S., Wuth, J., Huenupan, F., Curilem, M., Franco, L., et al. (2016). Automatic detection of volcano-seismic events by modeling state and event duration in Hidden Markov Models. *J. Volcanol. Geotherm. Res.* 324, 134–143. doi: 10.1016/j.jvolgeores.2016.05.015

Bracewell, R. N. (1999). *The Fourier Transform and Its Applications. 3rd Edn.* New York, NY.

Bueno, A., Díaz-Moreno, A., De-Angelis, S., Benítez, C., and Ibáñez, J. M. (2019). Recursive entropy method of segmentation. *Seismol. Res. Lett.* 90, 1670–1677. doi: 10.1785/0220180317

Caplan-Auerbach, J., Bellesiles, A., and Fernandes, J. K. (2010). Estimates of eruption velocity and plume height from infrasonic recordings of the 2006 eruption of augustine volcano, alaska. *J. Volcanol. Geotherm. Res.* 189, 12–18. doi: 10.1016/j.jvolgeores.2009.10.002

Curilem, M., Vergara, J., San Martin, C., Fuentealba, G., Cardona, C., Huenupan, F., et al. (2014). Pattern recognition applied to seismic signals of Llaima volcano (Chile): an analysis of the events' features. *J. Volcanol. Geotherm. Res.* 282, 134–147. doi: 10.1016/j.jvolgeores.2014.06.004

De Angelis, S., Diaz-Moreno, A., and Zuccarello, L. (2019). Recent developments and applications of acoustic infrasound to monitor volcanic emissions. *Rem. Sens.* 11:1302. doi: 10.3390/rs11111302

Di Stefano, R., Aldersons, F., Kissling, E., Baccheschi, P., Chiarabba, C., and Giardini, D. (2006). Automatic seismic phase picking and consistent observation error assessment: application to the Italian seismicity. *Geophys. J. Int.* 165, 121–134. doi: 10.1111/j.1365-246X.2005.02799.x

Diaz-Moreno, A., Iezzi, A. M., Lamb, O., Fee, D., Kim, K., Zuccarello, L., et al. (2019). Volume flow rate estimation for small explosions at Mt. Etna, Italy, from acoustic waveform inversion. *Geophys. Res. Lett.* 46, 11071–11079. doi: 10.1029/2019GL084598

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recogn. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010

Fee, D., and Matoza, R. (2013). An overview of volcano infrasound: from Hawaiian to Plinian, local to global. *J. Volcanol. Geotherm. Res.* 249, 123–139. doi: 10.1016/j.jvolgeores.2012.09.002

Fee, D., McNutt, S. R., Lopez, T. M., Arnoult, K. M., Szuberla, C. A., and Olson, J. V. (2013). Combining local and remote infrasound recordings from the 2009 Redoubt Volcano eruption. *J. Volcanol. Geotherm. Res.* 259, 100–114. doi: 10.1016/j.jvolgeores.2011.09.012

Frank, W. B., and Shapiro, N. M. (2014). Automatic detection of low-frequency earthquakes (LFEs) based on a beamformed network response. *Geophys. J. Int.* 124, 8611–8625. doi: 10.1093/gji/ggu058

Garcés, M., Iguchi, M., Ishihara, K., Morrissey, M., Sudo, Y., and Tsutsui, T. (1999). Infrasonic precursors to a vulcanian eruption at sakurajima volcano, japan. *Geophys. Res. Lett.* 26, 2537–2540. doi: 10.1029/1998GL005327

Giacco, F., Esposito, A., Scarpetta, S., Giudicepietro, F., and Marinaro, M. (2009). Support vector machines and mlp for automatic classification of seismic signals at stromboli volcano. *Front. Artif. Intell. Appl.* 204:116. doi: 10.3233/978-1-60750-072-8-116

Ibáñez, J., Benítez, C., Gutiérrez, L., Cortés, G., García-Yeguas, A., and Alguacil, G. (2009). The classification of seismo-volcanic signals using hidden markov models as applied to the Stromboli and Etna volcanoes. *J. Volcanol. Geotherm. Res.* 187, 218–226. doi: 10.1016/j.jvolgeores.2009.09.002

Johnson, J. B., Aster, R. C., and Kyle, P. R. (2004). Volcanic eruptions observed with infrasound. *Geophys. Res. Lett.* 31:L14604. doi: 10.1029/2004GL020020

Johnson, J. B., and Ripepe, M. (2011). Volcano infrasound: a review. *J. Volcanol. Geotherm. Res.* 206, 61–69. doi: 10.1016/j.jvolgeores.2011.06.006

Lamb, O. D., De Angelis, S., and Lavallée, Y. (2015). Using infrasound to constrain ash plume rise. *J. Appl. Volcanol.* 4:20. doi: 10.1186/s13617-015-0038-6

Lamb, O. D., Lamur, A., Díaz-Moreno, A., De Angelis, S., Hornby, A. J., von Aulock, F. W., et al. (2019). Disruption of long-term effusive-explosive activity at Santiaguito, Guatemala. *Front. Earth Sci.* 6:253. doi: 10.3389/feart.2018.00253

Landès, M., Ceranna, L., Le Pichon, A., and Matoza, R. S. (2012). Localization of microbarom sources using the IMS infrasound network. *J. Geophys. Res. Atmos.* 117:D06102. doi: 10.1029/2011JD016684

Marchetti, E., Ripepe, M., Ulivieri, G., and Kogelnig, A. (2015). Infrasound array criteria for automatic detection and front velocity estimation of snow avalanches: towards a real-time early-warning system. *Nat. Hazards Earth Syst. Sci.* 15, 2545–2555. doi: 10.5194/nhess-15-2545-2015

Rabiner, L. R., and Gold, B. (1975). *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, Inc. (1975), 777pp.

Scarpetta, S., Giudicepietro, F., Ezin, E., Petrosino, S., Del Pezzo, E., Martini, M., et al. (2005). Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. *Bull. Seismol. Soc. Am.* 95, 185–196. doi: 10.1785/0120030075

Schimmel, A., and Hübl, J. (2016). Automatic detection of debris flows and debris floods based on a combination of infrasound and seismic signals. *Landslides* 13, 1181–1196. doi: 10.1007/s10346-015-0640-z

Vergniolle, S., and Ripepe, M. (2008). From strombolian explosions to fire fountains at Etna volcano (Italy): What do we learn from acoustic measurements? *Geol. Soc. Lond. Spec. Publ.* 307, 103–124. doi: 10.1144/SP307.7

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.

ORIGINAL RESEARCH
published: 17 March 2020
doi: 10.3389/feart.2020.00059

Responses of the Terrestrial Ecosystem Productivity to Droughts in China

Jianguo Li[1,2]*†, Yi Wang[2,3]*† and Lili Liu[1]

[1] School of Geography, Geomatics, and Planning, Jiangsu Normal University, Xuzhou, China, [2] Department of Geography, School of Global Studies, University of Sussex, Brighton, United Kingdom, [3] Department of Earth System Science, Institute for Global Change Studies, Tsinghua University, Beijing, China

Edited by:
Flavio Cannavò,
National Institute of Geophysics and Volcanology, Italy

Reviewed by:
Chengguang Lai,
South China University of Technology, China
Mei Huang,
Institute of Geographic Sciences and Natural Resources Research (CAS), China

***Correspondence:**
Jianguo Li
lijianguo@jsnu.edu.cn;
lijianguo531@126.com
Yi Wang
Yi.Wang@sussex.ac.uk

†These authors have contributed equally to this work

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

Received: *29 November 2019*
Accepted: *19 February 2020*
Published: *17 March 2020*

Citation:
Li J, Wang Y and Liu L (2020) Responses of the Terrestrial Ecosystem Productivity to Droughts in China. Front. Earth Sci. 8:59. doi: 10.3389/feart.2020.00059

The terrestrial ecosystem productivity (hereafter, TEP) is a key index of global carbon cycles and a fundamental constraint of carbon sequestration capacity, and also an important measure of ecosystem services and food security. However, the TEP has been significantly affected by the long-lasting droughts. Identifying the spatial relationship between droughts and the TEP is crucial for enhancing ecosystem services in China. Here the net primary production (hereafter, NPP) derived from the Carnegie-Ames-Stanford Approach model (CASA-NPP) and two drought indices, namely the Standard Precipitation Index (hereafter, SPI) and the Standard Precipitation Evaporation Index (hereafter, SPEI), are used to examine the spatial relationship between droughts and the NPP in China for the period of 1982–2012. Our main results have shown that: (1) China's annual NPP has increased slowly from 3.82 to 4.35 PgC per year (hereafter, PgC/yr), while droughts have become much severer from 1982 to 2012; (2) on the 3-month timescale, the NPP in arid and semi-arid ecosystems has decreased at a rate of 1.28 TgC per month with per "unit" decrease in the drought index (indicating drier conditions). (3) Overall, the NPP in China has increased 5.71 TgC per month with per "unit" increase in the drought index (indicating wetter conditions); the contribution of this NPP increase is mainly from forests and farmlands; (4) the SPEI is a relatively more effective and sensitive index in representing China's droughts. In southern China, the lagging period for the NPP response to droughts is about 3-month, while a 6-month lagging period is found in the arid and semi-arid ecosystems in northern China.

Keywords: terrestrial ecosystem productivity, the Chinese NPP trends, SPI/SPEI drought indices, the lagging responses of Chinese NPP to droughts, ecological effect

INTRODUCTION

The terrestrial ecosystem productivity (hereafter, TEP) is the fundamental indicator for ecosystem services, and an integrated component of global carbon cycles, biodiversity, and regional food security (e.g., Piao et al., 2005; Zhu and Pan, 2007; Luo et al., 2019) and is generally controlled by many interplaying factors (Li et al., 2018; Liu et al., 2019). Under the threats of global warming, both the broadness and devastation of droughts and floods will continue to intensify, which will greatly affect the TEP, especially under long-lasting droughts (e.g., Yu et al., 2007; Doughty et al., 2015; Lei et al., 2015; Huang et al., 2016; Su et al., 2018; Gherardi and Sala, 2019; Xu et al., 2019). Droughts are a comprehensive and frequently occurred natural disaster, involving both precipitation and temperature changes, and they also control the soil moisture and vapor pressure deficit that will greatly influence the plant growth (Eamus et al., 2013). Many studies have shown that long-lasting droughts can significantly constrain vegetation activity and reduce the net primary production (hereafter, NPP) (e.g., Zhao and Running, 2010; Mk et al., 2011; Pei et al., 2013; Hou et al., 2014; Lai et al., 2018; Anderegg et al., 2019; Li et al., 2019a). As one of the major agricultural countries in the world, China has suffered from the strong interannual variability of monsoonal climate changes (e.g., Liu and Wang, 2011), the higher frequency of floods and droughts, especially in northern China. According to the "China Flood and Drought Disaster Bulletin" (2016), on average, 2.17×10^5 km^2 farmlands were influenced by droughts each year from 1950 to 2007, resulting in a loss of nearly 15.8 billion kilograms of grain, accounting for 60% of the total loss caused by all-natural disasters (MWRPRC, 2016).

Droughts can significantly diminish plant growth by cutting back the availability of soil water and therefore reducing crop yields (e.g., Chen et al., 2013; Heyer et al., 2018). A previous study has revealed that the grain loss induced by droughts has contributed about 7% of the total reduction of the global grain yield (Lesk et al., 2016). During a drought, plants can survive by closing their stomata, stabilizing intracellular water potential, and reducing the rate of autotrophic respiration, which can remarkably decrease the gross primary production (e.g., Mk et al., 2011; Doughty et al., 2015; Su et al., 2018; Kannenberg et al., 2019). There are wide-range differences among the responses of the plant growth to droughts for different ecosystems and at different spatial–temporal scales.

As concluded from previous studies, the effect of droughts on NPP variation has generally been examined at annual timescale or specific events (e.g., Huang et al., 2016; Li et al., 2019a). For example, an earlier study has argued that the interannual NPP variation in semi-arid and arid ecosystems was driven mainly by droughts (Huang et al., 2016). Their study has shown that about 29% of the interannual variation of the global NPP is explained by droughts-dominated NPP in semi-arid ecosystems, and 33% of the interannual variation of the global NPP is contributed by droughts prone ecosystems in the Southern Hemisphere (Huang et al., 2016). Furthermore, Zhao and Running (2010) have reported that a total loss of 0.55 PgC/yr global NPP between 2000 and 2009 could be explained by droughts, occurred in the Southern Hemisphere. Compared with forest ecosystems, grass ecosystems are more sensitive to droughts (e.g., Lei et al., 2015; Fei et al., 2018). In addition, long-lasting and severe droughts can also significantly lower the regional NPP in semi-humid and humid areas (Vicente−Serrano et al., 2015). Many drought indices can be used to represent regional drought level and further to explore the impact of the droughts on ecosystem evolution, such as Palmer Drought Severity Index (hereafter, PDSI, reference-PDSI) and Standardized Precipitation–Evapotranspiration Index (hereafter, SPEI, reference-SPEI), Standardized Precipitation Index (hereafter, SPI, reference-SPI), and standardized evapotranspiration deficit index (hereafter, SEDI, reference-SEDI) (e.g., Li et al., 2016, 2019a; Peng et al., 2016). Vegetation growth is generally controlled by periodic/seasonal and interannual changes of major driving factors, including climatic factors and human activities. Therefore, it is crucial to investigate the NPP variation and meteorological droughts overall a wide range of timescales. Compared with the PDSI and SEDI, the multi-timescale indices of SPI and SPEI have a remarkable advantage in representing the impacts of droughts to consider their delayed impacts on ecosystems.

However, the relationships between the monthly variability of NPP and intensive droughts in China are still unclear, because the drought distribution and plant growth processes are controlled by inconclusive factors at different spatial–temporal scales. In particular, determining the interannual relationship between NPP and intensive droughts cannot reveal the actual influence of droughts on monthly ecosystem production, which may result in some misleading conclusions (e.g., Zhao and Running, 2010; Zarei and Eslamian, 2017; Piao et al., 2019). Most previous studies have paid more attention to responses of the forest NPP to annual droughts (Vicente−Serrano et al., 2015), and the grassland and farmland NPP to annual droughts (Ji and Peters, 2003). Other studies have mainly focused on the issues of different timescales and lagging effects of droughts (e.g., Vicenteserrano et al., 2013; Anderegg et al., 2015; Wu et al., 2015; Huang et al., 2016; Huang and Xia, 2019). In China, the terrestrial ecosystems are so diverse that the relationships between NPP and intensive droughts at multiple spatiotemporal scales need further detailed analyses. Moreover, how and to what extent droughts affect monthly NPP variation in China are still unclear (Pei et al., 2013). Precisely determining the effect of droughts on monthly NPP changes is beneficial to assess the impacts of the potential extreme climate on TEP in the near future.

Our main objectives here are: (1) to identify the hotspots where droughts have some significant impacts on the NPP, (2) to examine the response time or lagging effect of ecosystems to droughts, and (3) to determine a suitable drought index representing the relationship between the monthly variability of NPP and droughts well. The following paper is arranged as below. The section "Materials and Methods" describes briefly the datasets used and our methods. The section "Results and Analyses" presents our main results. The discussion is shown in the section "Discussion," followed by the section "conclusion."

MATERIALS AND METHODS

In this study, two drought indices, namely the SPI and the SPEI, are used to estimate the impact of droughts on NPP in China at monthly timescales (see **Supplementary Section S3** for more details). The NDVI images used to estimate the parameter, FPAR for CASA model (see **Supplementary Section S1** for more details), are from the Global Inventory Modeling and Mapping Studies-NDVI dataset (see GIMMS-NDVI[1]) (Ruimy et al., 1994). The GIMMS-NDVI dataset is constructed by Tucker et al. (2004), which includes the 31-year (372-month)

---

[1] https://ecocast.arc.nasa.gov/data/pub/gimms/3g.v1/

period from 1982 to 2012, and covers the whole China at a spatial resolution of 8 km × 8 km. The NPP in China during the period of 1982–2012 is estimated using the CASA model (Zhu et al., 2006). The meteorological data (including monthly precipitation, monthly mean temperature, and monthly total solar radiation, see **Supplementary Section S2** for more details) is downloaded from the China Meteorological Administration (CMA) Meteorological Data Center[2] (in total 823 meteorological stations, see **Figure 1** for locations of these meteorological stations). All meteorological data downloaded from the Data Center of CMA is imported into ArcGIS 10.2, and then spatially interpolated to a grid space of 8-km in order to obtain maps of monthly precipitation, monthly mean temperature, and monthly total solar radiation using the inverse distance weight method (Bartier and Keller, 1996). The detailed processes for all datasets and the CASA-NPP model can be found in **Supplementary Sections S1–S3**.

RESULTS AND ANALYSES

The Evaluation of Derived NPP

In this study, the CASA model shown in **Supplementary Section S1** is employed to derive the NPP in China from 1982 to 2012, which includes the annual NPP and monthly net photosynthesis. On average, our derived NPP has continuously increased from 3.7 PgC/yr in 1982 to 4.35 PgC/yr in 2012. Our result is similar to the work of Tao et al. (2003) and Yuan et al. (2014), but much larger than the results of Liu et al. (2013) and Piao et al. (2005). The different parameters in CASA model, namely the maximum efficiency of light energy utilization, and the more realistic classification of vegetation types, may explain the different results (Zhu and Pan, 2007). Our NPPs using the CASA model are generally higher than those derived from the MODIS products (MOD17 A3), which are produced by the Numerical Terradynamic Simulation Group (NTSG) of University of Montana (UMT) using MOD17 algorithm (Asrar et al., 1992; **Figure 2**). In fact, the MOD17 algorithm integrates maximum light energy utilization derived from Biome-BGC model, which does not consider many local factors (e.g., the water stress and vegetation types) that influence vegetation growth based on the high-precision vegetation classification. Therefore, the uncertainty of MOD17-derived NPP is relatively larger. To verify our estimation, mean NPPs in higher vegetated (e.g., humid zone) and lower vegetated (e.g., semi-arid and arid zones) regions have been extracted and compared with those derived from MOD17A3 dataset (**Figure 2**). Overall, our NPP in humid zones is larger than that derived from MOD17A3. In particular, from 2000 to 2006, our NPP is ∼400 gC/m^2/yr larger (**Figure 2A**). Using FPAR constrained from the satellite NDVI retrieval, CASA model calculates more realistic NPPs in densely vegetated regions (**Figure 2A**). On the other hand, in arid and semi-arid zones, our estimated NPP is on average ∼100 gC/m^2/yr smaller than that derived by MOD17A3 (**Figure 2B**). We speculate this may be caused by the different spatial resolutions between CASA and MOD17A3 calculations. At 1 km × 1 km resolution, the assignment of deserts in MOD17A3 is more realistic than that in CASA at 8 km × 8 km resolution. In China, deserts cover 1.33 million km^2, 13.6% of the total land area of China. Deserts have been assigned a zero NPP in MOD17A3 and CASA. However, at 8 km × 8 km resolution, the same assignment of the wide-spread deserts is much larger than that in the MOD17A3 model, which may directly result a potential reduction in the NPP calculation in CASA model for the arid and semi-arid regions.

Table 1 shows a comparison of NPP values in the year 2005 for different plant functional types within China. Most of our derived NPPs is similar to those observed, except for the farmland NPP. CASA model has underestimated the farmland NPP compared to observation (eddy covariance tower). This is because in the CASA model, factors involving agricultural productions (e.g., fertilization, agricultural management, and irrigation) have been excluded. In fact, these factors play an important role in promoting agricultural yield (Knapp and van der Heijden, 2018). Compared to other studies, our derived NPPs for evergreen broad-leaf forest and deciduous needle-leaf forest are very similar to observed NPP. However, our derived NPPs are significantly different from the result of Piao et al. (2005) and Shang et al. (2018). We speculate that this is mainly due to the difference in the maximum light use efficiency and other constraint factors, such as temperature and soil water content in CASA model.

The Characteristics of Annual Mean NPP in China From 1982 to 2012

Figure 3A shows the annual NPP in southern China is generally higher than in northern China. In addition, the annual NPP in western China is remarkably lower than in eastern China. The spatial distribution of Chinese NPP ranges from 0 to 2108 gC/m^2/yr with a mean value of 433.04 gC/m^2/yr. The lowest NPP values are found in the northern Tibet Plateau, southern Xinjiang, northern Gansu, and western Inner Mongolia. The highest NPP values are mainly located in the Qinling Mountains, southwestern Sichuan Mountains, Yarlun-Tzanpo River valley, the areas to south of the Yangtze River basin, and most of Hainan, Yunnan, and Taiwan. The NPP hotspots (locations with annual NPP values over 1600 gC/m^2/yr) are found in the Hainan province and southern Yunnan where the tropical rainforest is widely distributed. As shown in **Figure 3B**, the annual total NPP in China has increased substantially from 1982 (3.82 PgC/yr) to 2012 (4.35 PgC/yr), at an average rate of 16.97 TgC/yr.

The Characteristics of Droughts in China

The SPI and SPEI indices are characterized by multi-timescales, such as 3-, 6-, 9-, and 12-month (and longer) timescales. The drought status over a 3-month period (including current month) can be identified by using the SPI3 and SPEI3 indices. According to the SPEI ranking criteria (Yang et al., 2018), three classifications are defined: (1) SPEI < 0 (mild drought), (2) SPEI < -1 (moderate drought), and (3) SPEI < -1.5 (severe drought). In this study, our drought frequency has been defined as the ratio between the total month of drought's occurrences

---

[2] http://data.cma.cn/

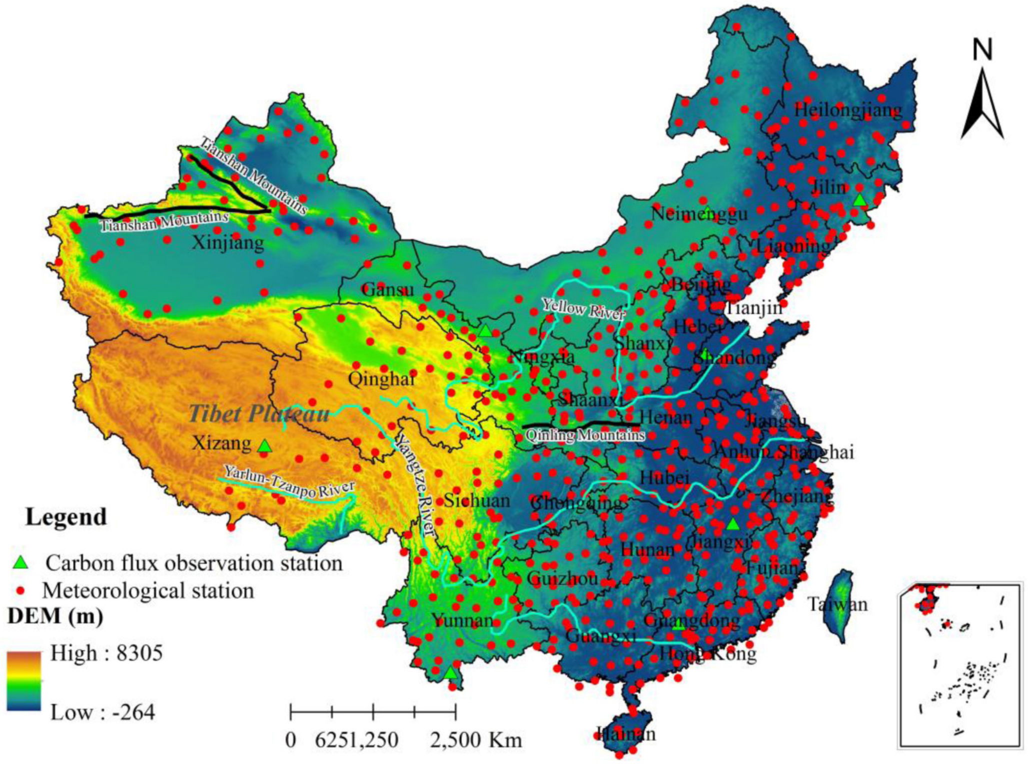


FIGURE 1 | The meteorological stations (red dots), carbon flux observation station (green triangles), and our study area including the DEM in color shades.

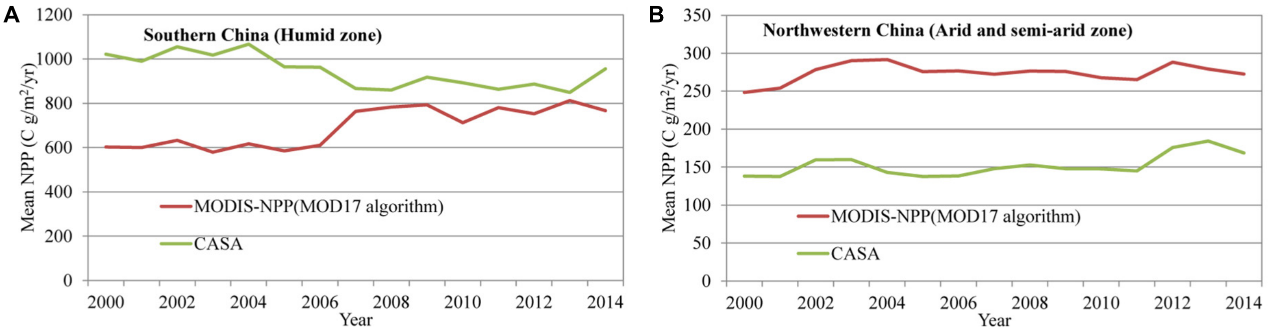


FIGURE 2 | A comparison of two NPP estimations across different vegetation zones. **(A)** Higher vegetation coverage zones. **(B)** Lower vegetation coverage zones.

TABLE 1 | A comparison of derived NPPs using CASA model of this study with previous studies and eddy covariance tower data for the year of 2005 (unit: gC/m^2/yr).

Plant functional types	Pixels	Our NPP	Observed NPP	Observations from station*	Piao et al. (2007)	Ni (2010)
Evergreen broad-leaf forest	6375	1046.32	1125.17	Qianyanzhou	525	945
Evergreen needle-leaf forest	14,579	1253.36	997.29	Changbai Mountain	354	439
Farmland	26,406	903.56	1403.16	Yucheng	216	N/A
Grassland	6596	458.13	462.17	Neimenggu	N/A	N/A

Data for observed NPP were obtained using eddy covariance tower in Chinese national observation network (Source: http://rs.cern.ac.cn/data/initDRsearch?classcode=SYC_A02), in which NPP is derived using this formula: NPP = NEE (Net Ecosystem Productivity) + R_{eco} (Ecosystem Respiration) − AR (Autotrophic Respiration) (Fei et al., 2018). The carbon flux stations are plotted in **Figure 1 as green triangles. The Chinese forestry data was listed in **Supplementary Section S6**, which was used to validate the precision of our CASA-NPP model.*

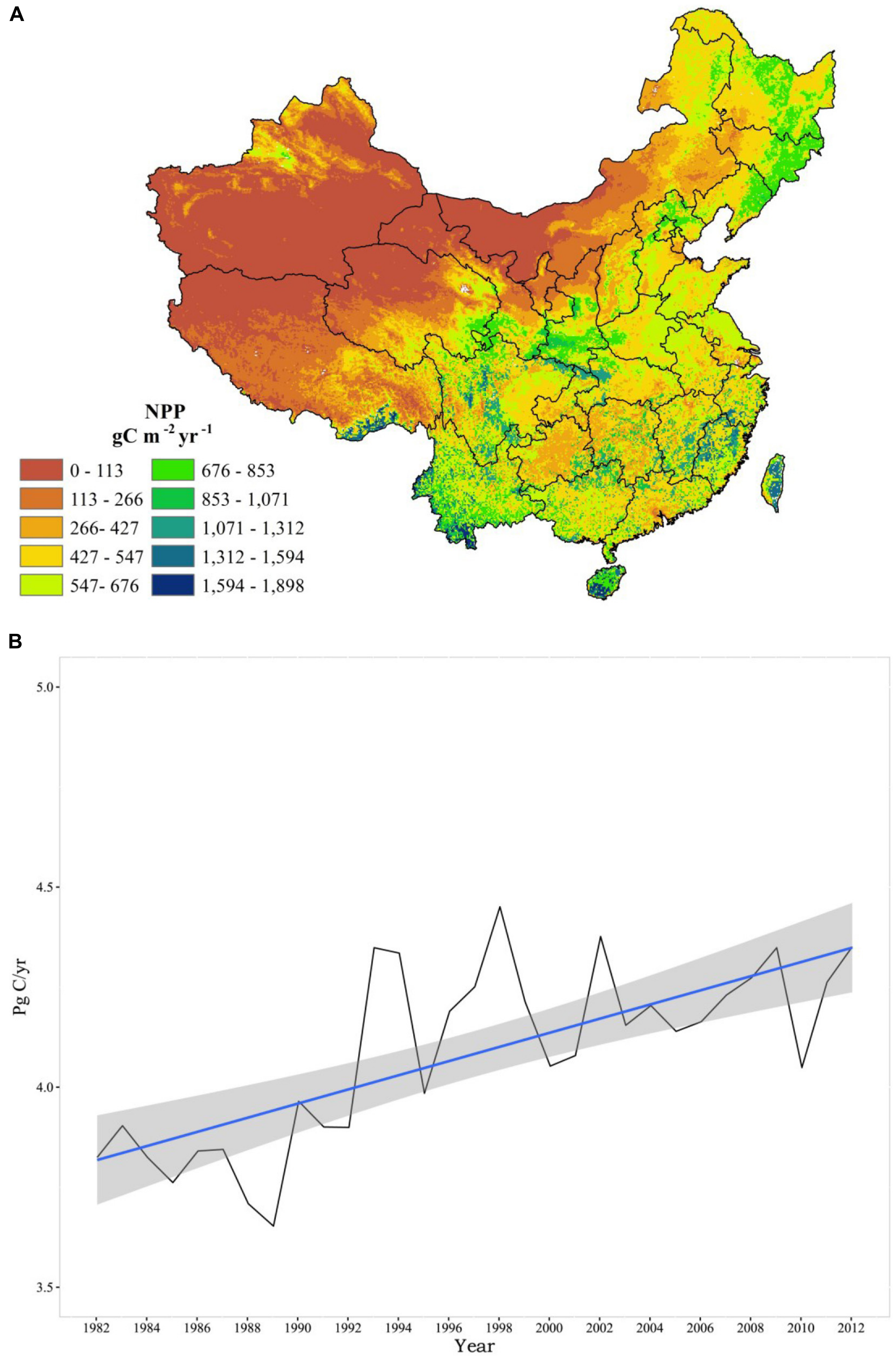


FIGURE 3 | (A) The spatial distribution of annual mean NPP in China from 1982 to 2012 and **(B)** the timeseries of total annual NPP in China.

and the whole study period (372 months from 1982 to 2012) (see **Supplementary Section S4.1** and Eq. 3). In this study, the datasets of SPI and SPEI were derived and validated based on the dataset at each meteorological station (see **Supplementary Sections S3**, **S5**). **Figure 4A** shows the mild drought is mainly located in the middle of the Tibet plateau. According to Eq. 3 in **Supplementary Section S4.1**, the frequency of mild drought is above 50% in most parts of China with a mean value of 56.25% (**Figure 4A**). The average frequency of moderate and severe droughts in China is 19.48 and 8.85%, respectively (**Figures 4B,C**). The areas to the north of Qinling Mountains-Huaihe River line generally have a higher severe drought

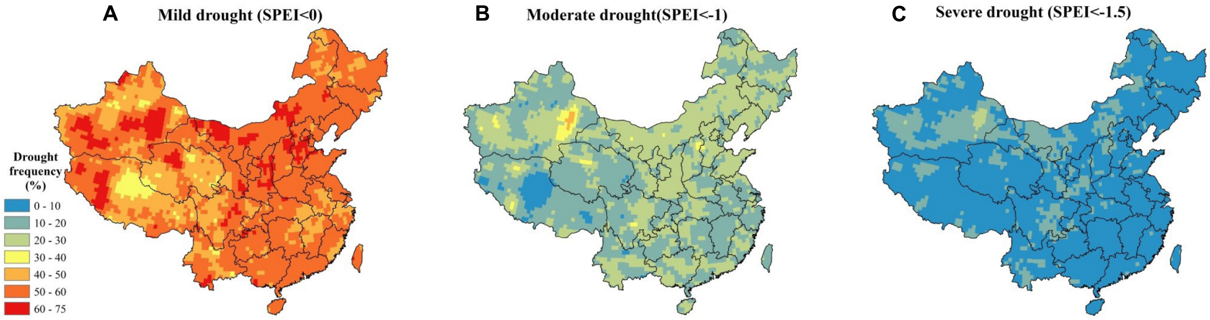


FIGURE 4 | The frequency of droughts in China from 1982 to 2012 (SPEI3). **(A)** Mild droughts (SPEI3 < 0), **(B)** moderate droughts (SPEI3 < –1), and **(C)** severe droughts (SPEI3 < –1.5).

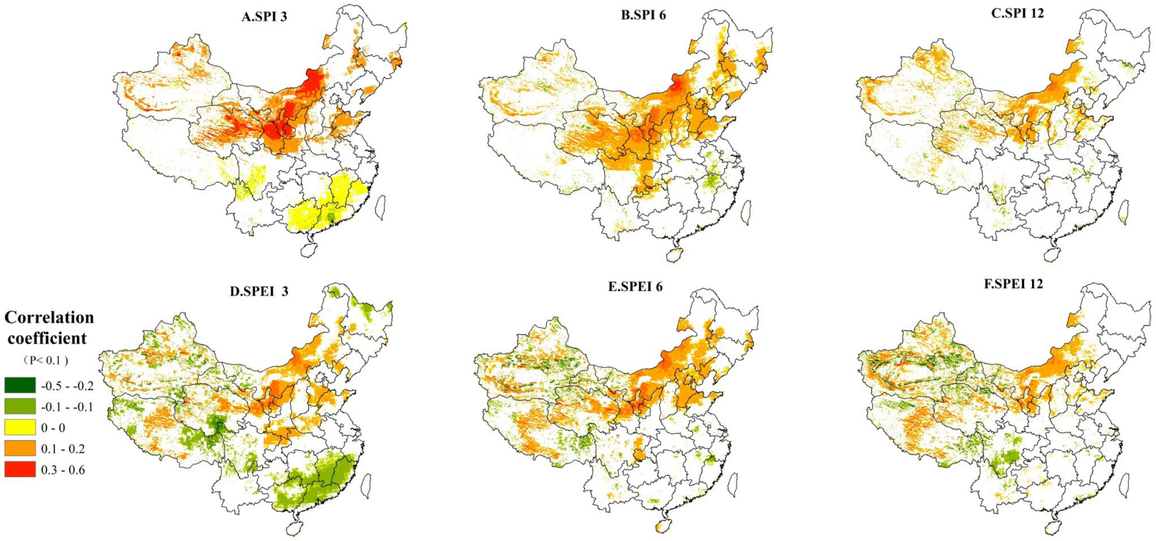


FIGURE 5 | The significant correlation coefficient (hereafter CC) between SPI/SPEI and monthly NPP (*P* < 0.1 or 90% significance level). **(A)** CC between SPI3 and monthly NPP, **(B)** CC between SPI6 and monthly NPP, **(C)** CC between SPI12 and monthly NPP, **(D)** CC between SPEI3 and monthly NPP, **(E)** CC between SPEI6 and monthly NPP, and **(F)** CC between SPEI12 and monthly NPP.

frequency (see **Figure 4B**, >20%), among which eastern Shanxi and western Hebei are the most pronounced areas. For moderate droughts, higher frequency (>40%) is mainly distributed in Turpan and Hami of Xinjiang province. Nevertheless, higher drought frequency (**Figure 4C**) appeared in farming-pastoral ecotones (around the Heihe-Tengchong line), where agriculture and animal husbandry are developed very rapidly, indicating a high risk to China's NPP.

The Correlation Between Droughts and Monthly NPP in China

At different timescales, the SPI and SPEI drought indices are used to examine the relationships between droughts and monthly NPP in China (i.e., net photosynthesis), respectively, and also to show the reliability of our correlation analyses using different drought indices (i.e., SPI and SPEI). Pixels with significant coefficient of correlation (*P* < 0.1 or 90% significance level) are plotted. **Figure 5** suggests that regardless which drought indices used, the results across different timescales (e.g., 3-, 6-, and 12-month) are quite similar. At the 3-month timescale, positive relationships between the monthly NPP and drought index are found in the Loess Plateau and the XilinGol League of Inner Mongolia (**Figures 5A,D**), in which corn planting and grassland are extensive. This indicates that the localized NPP (mainly from the agricultural production) would be significantly increased (decreased) under a wetter (drier) condition. The areas with negative relationships, mainly locate in Yangtze River Basin and its south, suggesting that excessive rainfall in these areas could cause a negative impact on the localized NPP at shorter timescales (i.e., 3 months). Our result is supported by a previous study about the heavy rainfall triggering saturated ground water and surface flooding, further decreases plant growth efficiency and results in hypoxia or death in some extreme cases (Tan et al., 2009). In southern China, very heavy rainstorms and typhoons often occur in the summer that will

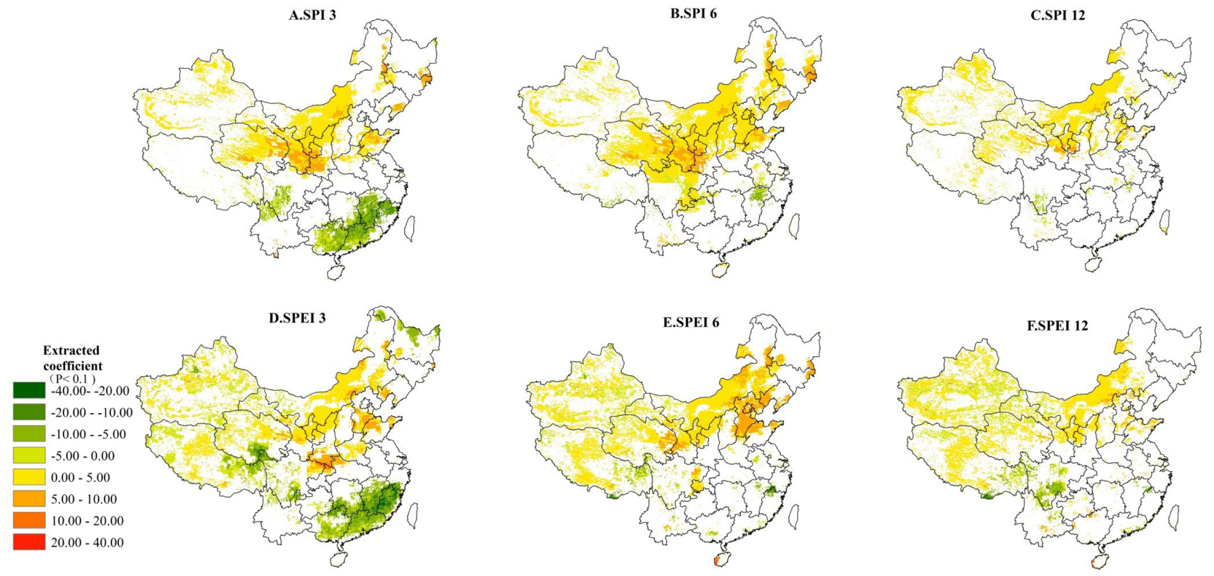


FIGURE 6 | The contributions (unit: gC/m^2 per unit of drought index) of SPI/SPEI to monthly NPP as defined by Eq. 6 in **Supplementary Section S4.4**. **(A)** SPI3 contributions, **(B)** SPI6 contributions, **(C)** SPI12 contributions, **(D)** SPEI3 contributions, **(E)** SPEI6 contributions, and **(F)** SPEI12 contributions. Only 90% statistically significant pixels are plotted here.

likely intensify this negative impact on the regional NPP. At the 6-month timescale (**Figures 5B,E**), the negative impact in southern China disappears. On the other hand, the positive correlation in northern China continues to increase, showing expanding influences of droughts at longer timescales (**Figure 5B**). The expanded areas now cover the whole Sanjiangyuan region, northern Sichuan, western Chongqing, and most parts of Shanxi and Gansu Provinces. Compared with the SPI, the more diverse relationships between the SPEI and monthly NPP have been identified, especially in the high mountains of western China. The positive relationship is mainly found in the northern and central China largely due to insufficient water supply in those arid and semi-arid regions. On the other hand, the negative relationship generally distributes in the shadow and nearby snow line (**Figure 5E**). At the 12-month timescale (**Figures 5C,F**), the statistically significant regions mainly concentrate in western China, especially in the Xinjiang and Sanjiangyuan regions. The main reason could be that everlasting droughts have significant impacts on arid or semi-arid ecosystems during the longer timescale (Chen et al., 2012). In summary, the areas with a significant relationship between drought indices and monthly NPPs at the 3-month timescale are much larger than those at 6- and 12-month timescales regardless which drought index is used.

The Contributions of the SPI/SPEI to Monthly NPPs Across Different Timescales

Contributions of the SPI/SPEI to the variability of NPP are determined using Eq. 5 in **Supplementary Section S4.5**, which represents the amount of NPP variation under the per-unit change of the drought index (unit: gC/m^2 per unit change of drought index). Our results (**Figures 6A,D**) show that significant contributions of droughts on monthly NPP have been identified in northern and southern China over shorter timescales (e.g., SPI3 and SPEI3). At longer timescales, positive contributions of droughts tend to become significant in the semi-arid and arid ecosystems in western and northern China. In areas with positive impacts between the drought index and monthly NPP, the mean values of contributions are 2.28, 3.55, and 1.75 (unit: gC/m^2 per unit change of drought index) at 3-, 6-, and 12-month timescales, respectively (**Figures 6A–C**). Our result shows that monthly NPP variation of arid and semi-arid ecosystems is driven mainly by the 6-month timescale droughts. On the 3-month timescale, droughts can only affect the monthly NPP in northern China, such as the North China Plain, the Chinese Loess Plateau, and the farming-pastoral ecotones. The strongest positive contribution is found in the catchment of River Wei and Guanzhong Plain in the southern boundary of the Chinese Loess Plateau (**Figures 6A,D**). Here, an increase of 20 gC/m^2 monthly NPP could occur under an increase of per unit of drought index (wetter conditions). Conversely, at 3-month timescale, the negative contributions are mainly found in southern China (**Figures 6A,D**), such as Jiangxi, Fujian, and Guangdong provinces, with a monthly NPP decrease of 2.54 gC/m^2 would result from per drought index unit increase (wetter conditions). The strongest negative contributions are found in the Sanjiangyuan Region and southern Sichuan Province. The probable reason for this phenomenon is that the ongoing increase in temperature (lowering the drought index) in these areas promotes the melting of snow and ice, causing a further increase of monthly NPP. On the 6-month timescale, the significant contributions between droughts and the NPP in southern China tend to disappear, while the same contributions are enhanced in semi-arid and arid regions in

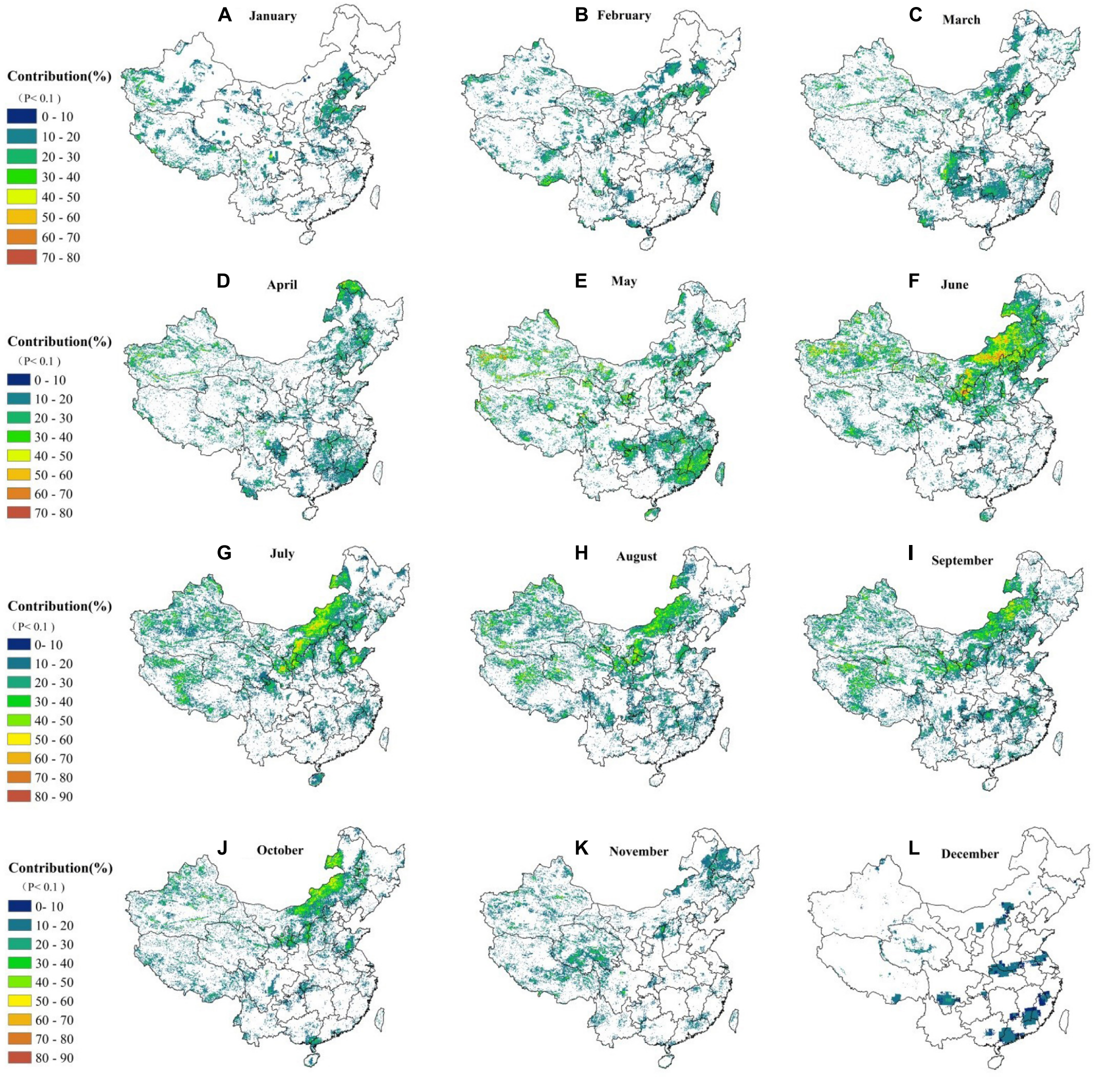


FIGURE 7 | The R^2 of monthly NPP and SPEI 3 in different months over a period of 1982–2012. From Images **(A)** to **(L)** stands from January to December.

northern China. At the 12-month timescale, the stronger positive contributions are found in semi-arid and arid regions in northern China (**Figures 6C,F**). In addition, the influences of droughts on monthly NPP variation in eastern China tend to become weaker at longer timescales, except for the North China Plain. Compared to the SPI, the SPEI index is more effective to contribute to monthly NPP variation in China due to a relatively higher coefficient and larger areas with statistically significant relationships (see **Figure 6**, $P < 0.1$ or 90% significance level).

Relative Contributions of Droughts (SPEI3) to Monthly NPP Changes in Individual Months

Based on the section "The Correlation Between Droughts and Monthly NPP in China," the SPEI3 is selected to examine the relative contribution (R^2, unitless, and in percentage) of droughts to monthly NPP in individual months (January–December) as explained in our Eq. 7 in **Supplementary Section S4.5**. **Figure 7** suggests that a unimodal tendency of R^2 (at first an increase followed by a decrease) is observed from January to December. Across different months, minimal impacts of droughts are found in December with a mean value of 12.25%. However, the greatest impacts are identified in June with a mean value of 22%. The greatest impact occurs between June and September, with a mean value of over 20%. In January–February, the strong relationships are mainly in the North China Plain. The average impacts are generally lower due to the non-growing season, with a mean value about 16%. Between March and May, the influence level of water tends to increase in southern China (to nearly 18%), especially in Sichuan, Hunan, Chongqing, Guizhou provinces, and a small portion of Jiangxi province. This is because the evergreen forests begin to grow in these areas and vegetables are sowing and germinating at that time. During this period, the northern winter wheat is in the tillering stage, and spring wheat is sown, which results in an increasing demand for water. Hence, the impacts continue to increase in northern China (however, the pixels with higher percentages are scatteredly distributed; see **Figure 7**). In June and July, the main grasses (Leymus chinensis, Agropyron grass, Stipa grandis, and Artemisia frigida) and corn in northern China begin to grow (Zhang and Yang, 2007); hence, the impact of droughts on monthly NPP increases to 50% in the Xilingol grassland and the Chinese Loess Plateau due to water deficits. In August and September, much more water is needed due to the growth of vegetables and upland crops in southern China as well as the grass and staple crops in northern regions. This causes a significant increase in the impact of drought, but the areas with a significant value of R^2 ($P < 0.1$ or 90% significance level) gradually decreases. Between October and December, the significant drought impact tends to decrease continuously. Interestingly, the period from May to September is the key season for vegetation growth in the Tibet Plateau and agricultural production of the isolated oasis farms in western China, which is significantly impacted by droughts.

DISCUSSION

Droughts can significantly limit plant growth and further decrease the NPP by restricting the availability of soil water (Xu et al., 2019). In this study, the strongest correlation between drought and monthly NPP is identified in the semi-arid ecosystems in the farming-pastoral ecotones in northern China (**Figure 5**). This information will help to promote the terrestrial ecosystem management in the area. At multi-timescales, the strategies for coping with droughts are different among individual biomes (Vicente−Serrano et al., 2015). The humid biomes are usually sensitive to droughts, and their NPP tends to sharply decrease under the water stress (Vicenteserrano et al., 2013). However, except for upland crops, the resistance of arid or semi-arid biomes to water deficits is much stronger when they change leaf and root morphology to reduce evaporation and retain water during long-term droughts. Therefore, the monthly NPP of semi-arid biomes tends to decrease gradually. The response time of biomes toward drought is called the "lagging time." Concerning interannual NPP variation, a significant lagging time usually exists in trees, and a period of 1–4 years is usual among the Pinaceae family (Anderegg et al., 2015). However, the lagging time in semi-arid ecosystems may be 16–19 months (Huang et al., 2016). In addition, the monthly NPP of grassland and farmland in semi-arid ecosystems is sharply reduced during droughts. Lei et al. (2015) reported that droughts could account for 51.75% of NPP loss in grasslands in Inner Mongolia of northern China. A total reduction of 154 gC/m^2/yr of the NPP can occur due to droughts in grasslands in China (Xiao et al., 2009). About 53% of the areas of China are occupied by arid or semi-arid ecosystems, of which 22% are semi-arid ecosystems where agriculture and animal husbandry are developed. However, severe droughts in these areas occur frequently (**Figure 4**), causing substantial fluctuations in the annual NPP. A previous study reported that the farmlands in China obviously have been moved to the north and across the Yellow River, indicating that there is no doubt that the future Chinese NPP will be threatened profoundly by droughts (Liu et al., 2009).

Over the past 20 years, the increasing annual NPP in the northern hemisphere, including China, has been verified by many studies (Liu, 2001; Lee and Veizer, 2003; Piao et al., 2007; Zhao and Running, 2010; Kim et al., 2012). In China, significant increasing trends in annual NPP (**Figure 8**) are mainly located in the southwestern China, including Sichuan, Guizhou, Yunnan, Guangxi, Guangdong, Hainan provinces, and the eastern Tibet Plateau, where an increase of 26.64 gC/m^2/yr of the annual NPP is identified during our study period (**Figure 8**). We also notice that the greening project led by the so-called "grain-for-green" policy in the Chinese Loess Plateau has resulted in a remarkable increase in vegetation activities. Our derived annual NPP has shown this significant increase, which has been confirmed by other studies (Peng et al., 2007; Persson et al., 2013; Jiang X. et al., 2019). However, in areas with strong human activities and/or an unreasonable combination of water and heat (e.g., urban zones and high mountains), the annual NPP tends to decrease in our study period. Our results have been supported by previous studies (Liu, 2007; Xu et al., 2017). The mean annual NPP decreasing rate of 42.82 gC/m^2/yr is found in those areas. On the 3-month timescale, increased SPEI3 index (indicating wetter conditions) can significantly enhance the monthly NPP in northern China, but significant heavy rain at shorter timescales (e.g., 3-month) in southern China can reduce the monthly NPP. Considering the extent of influence from droughts, the 3-month timescale is appropriate to evaluate the relationships between water shortage and monthly variability of NPP in China, which is consistent with the result of Yang et al. (2018). In addition, at the 6-month timescale, droughts can only be used to assess the relationships between water shortage and the monthly variability of NPP in northern and western China. Our results suggest that the 3- to 6-month drought indices (e.g., SPEI3 and SPEI6) are appropriate for evaluating the relationships between droughts and monthly variability of NPP in China.

According to **Figure 7**, the NPP in China is mostly sensitive to droughts in the March–October season, among which the period of May–July is the most important months. Here, our result agrees with Ji and Peters (2003). During May–July, the average contribution from droughts on monthly NPP is >22% (**Figure 7**). The greatest contributions of droughts to monthly NPP have been found in the Tibet Plateau and northwestern China, including the middle of the Tibet Plateau, the Yarlun-Tzanpo River Valley, and the Tianshan Mountains of western China, especially the variation of alpine meadow (**Figure 7**). More importantly, according to our results (**Figure 7**), the greatest impacting month of droughts on the NPP in southern China occurs in May, but in northern China, it occurs in June.

Previous studies have reported that longer and severe droughts, as represented by the PDSI and SEDI, can remarkably reduce the NPP in China over the past 30 years (e.g., Li et al., 2016, 2019a). Although soil moisture content and soil water loss have been considered in the PDSI, which is beneficial to reflecting

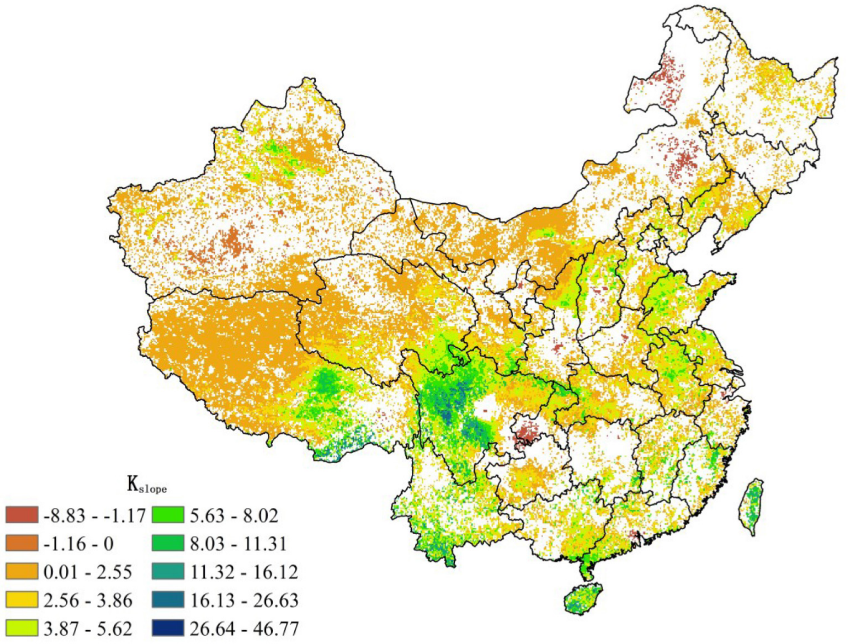


FIGURE 8 | The long-term trends of annual NPP in China from 1982 to 2012 ($P < 0.1$ or 90% significance level).

the actual status of droughts, lagged response from vegetation to soil water deficit can result in some unreliable results for quantitatively determining the relationship between PDSI and NPP. In this study, an outstanding 3–6-month lagging period of vegetation to droughts has been found; therefore, a multi-timescale drought index (e.g., 3–6-month drought index) can provide a reliable investigation about the lagged relationship between droughts and NPP changes. In previous studies, the effect of individual drought events on NPP has been focused broadly (e.g., Lai et al., 2018; Li et al., 2019a). However, compared with individual drought events, the evolution trend of droughts generally have greater impacts on vegetation activities at longer timescales (Piao et al., 2019). Thus, exploring the relationship between drought trends over longer timescales and their impacts on NPP changes has played a key role in revealing the mechanism of future TEP evolution within the context of extreme climate events.

More importantly, NPP variations are generally controlled by many inter-playing factors, including from both natural and anthropogenic sources. First, different tillage practices and grazing policies can all have profound impacts on farmland and grassland NPP changes, respectively. In general, ever-advancing tillage practices (fertilization, pesticides, plastic mulch, and irrigation) have significant and positive impacts on increasing NPP in farmlands. For instance, higher maize NPP (yield) can be achieved by improving tillage practices at a planting date ranging from late April to early May and an N-fertilizer input rate of 180–210 kgN ha^{-1} with two timing splits in northeastern China (Jiang R. et al., 2019). Furthermore, severe grazing can remarkably reduce the NPP in grass ecosystem and further result in desertification that has been verified in a case study of Xilingol Grassland, northern China (Chi et al., 2018). The study revealed that 94.6% of the area of Xilingol grassland showed a negative correlation between NPP residuals and grazing pressure (Chi et al., 2018). Second, at shorter timescales, human activities play some important roles in driving NPP variations. In the developing world, the rapid urbanization has an obviously negative impact on regional NPP changes (Li et al., 2019b). Under the processes of urbanization, large amounts of farmlands, wetlands, and forests are overall replaced with poorly permeable surfaces of cement and asphalt that result in the NPP reduction. Liu et al. (2019) have reported that the rapid urban expansion in the period of 2000–2010 has in turn reduced global terrestrial NPP, with a net loss of 22.4 Tg carbon per year (Tg C yr^{-1}), which can overall offset 30% of the climate-driven NPP increase (73.6 Tg C yr^{-1}) over the same period. Li et al. (2018) has shown that 63.02% of total NPP losses in China could be explained by the conversion from farmland to suburban land. Furthermore, different feedbacks and interactions between vegetation growth and human activities have been identified. For example, Piao's group has demonstrated that an increasing vegetation greenness (likely higher NPP) caused by human activities (land-use practice, CO_2 fertilization, and nitrogen deposition) has significantly changed regional water cycle and evapotranspiration, especially in dry regions. Data have revealed that the global leaf area index (LAI) has enhanced 8% between the early 1980s and 2010s, which will cause a global increase of 12.0 ± 2.4 mm yr^{-1} in evapotranspiration and 12.1 ± 2.7 mm yr^{-1} in precipitation representing about 55 ± 25 and 28 ± 6% of the observed increases in land evapotranspiration and precipitation, respectively (Zeng et al., 2018;

Piao et al., 2020). Chen et al. (2019) has demonstrated that human land use practice has contributed a larger proportion to Earth greening (over 33%). However, in our study, the anthropogenic factors mentioned above were not considered mainly due to factor of the large-scale nature of our study, and the lack of localized observations to match with policy-driven and LULC impacts on the NPP. Therefore, in the future, a process-based model (e.g., CLM4.5, ISAM, LPJGUESS, LPJ, and LPX) involving many interplaying factors can be used to accurately explore the relationships between the NPP and natural/anthropogenic factors (Ding et al., 2018).

CONCLUSION

The total annual NPP in China has increased from 1982 to 2012 with a mean rate of 433.04 gC/m^2/yr. In our study period, a total NPP increase of 16.97 TgC/yr is found mainly located in southern and southwestern China. Droughts in China continue to be severe, and the frequency of moderate drought is >20% in northern China. The frequency of severe drought is >10% across farming-pastoral ecotones. The strongest relationships between drought and monthly NPP in China are located in the farming-pastoral ecotones of semi-arid ecosystems. On the 3-month timescale, a negative contribution of droughts on the monthly NPP in southern China has been found, in which a mean NPP reduction of 2.54 gC/m^2 can be found with per-unit increases in the drought index (wetter conditions). However, the interannual NPP in the same areas has continued to increase due to the global warming, nitrogen deposition, and advances in agriculture that can compensate for the monthly NPP losses caused by flooding (higher SPEI index). On the 6-month timescale, the strongest correlations between drought and monthly NPP are located in farming-pastoral ecotones of arid and semi-arid ecosystems covering grassland and corn fields.

The SPEI is more effective in defining the relationships between droughts and monthly NPP in China. The response period (e.g., the lagging time) of monthly NPP to droughts in southern China is about 3 months. The lagging time in arid and semi-arid ecosystems of northern China is nearly 6 months. The drought accounts for >15% of monthly NPP variation in China each month. Over the whole China, the lowest impact season of droughts on monthly NPP has been found during the December–February period, but the greatest impact period has been determined in the months of May–July. Regionally, the greatest impact season for the middle of the Tibet Plateau, Yarlun-Tzanpo River valley, and the Tianshan Mountains of western China is the July–September period. While in northern China, the greatest impact month of droughts on monthly NPP is in June; this becomes May in southern China.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this article are not publicly available. Requests to access the datasets should be directed to lijianguo531@126.com.

AUTHOR CONTRIBUTIONS

JL worked on the implementation and modeling. YW and LL worked on the data processing and language polishing.

FUNDING

This work was supported by the National Science Foundation of China (No. 41701371), the Ministry of Education of Humanities and Social Science project (No. 17YJCZH085), the University Science Research Project of Jiangsu Province (No. 17KJB170006), and the Priority Academic Program Development of Jiangsu Higher Education Institutions to JL. This study was supported partially by the Sussex International Development Fund and School of Global Studies' Publication Grant awarded to YW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.00059/full#supplementary-material

REFERENCES

Anderegg, W. R., Anderegg, L. D., and Huang, C. Y. (2019). Testing early warning metrics for drought−induced tree physiological stress and mortality. *Glob. Chang. Biol.* 25, 2459–2469. doi: 10.1111/gcb.14655

Anderegg, W. R., Schwalm, C., Biondi, F., Camarero, J. J., Koch, G., Litvak, M., et al. (2015). Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science* 349, 528–532. doi: 10.1126/science.aab1833

Asrar, G., Myneni, R. B., and Choudhury, B. J. (1992). Spatial heterogeneity in vegetation canopies and remote sensing of absorbed photosynthetically active radiation: a modeling study. *Remote Sens. Environ.* 41, 85–103. doi: 10.1016/0034-4257(92)90070-Z

Bartier, P. M., and Keller, C. P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Comput. Geosci.* 22, 795–799. doi: 10.1016/0098-3004(96)00021-0

Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., et al. (2019). China and India lead in greening of the world through land-use management. *Nat. Sustain.* 2, 122–129. doi: 10.1038/s41893-019-0220-7

Chen, T., Werf, G. R. V. D., Jeu, R. A. M. D., and Wang, G. (2013). A global analysis of the impact of drought on net primary productivity. *Hydrol. Earth Syst. Sci.* 17, 3885–3894. doi: 10.5194/hess-17-3885-2013

Chen, Z. Q., Shao, Q. Q., Liu, J. Y., and Wang, J. B. (2012). Analysis of net primary productivity of terrestrial vegetation on the Qinghai-Tibet Plateau, based on MODIS remote sensing data. *Sci. China Earth Sci.* 55, 1306–1312.

Chi, D., Wang, H., Li, X., Liu, H., and Li, X. (2018). Assessing the effects of grazing on variations of vegetation NPP in the Xilingol Grassland, China, using a grazing pressure index. *Ecol. Indic.* 88, 372–383. doi: 10.1016/j.ecolind.2018.01.051

Ding, J., Yang, T., Zhao, Y., Liu, D., Wang, X., Yao, Y., et al. (2018). Increasingly important role of atmospheric aridity on Tibetan alpine grasslands. *Geophys. Res. Lett.* 45, 2852–2859. doi: 10.1002/2017GL076803

Doughty, C. E., Metcalfe, D. B., Girardin, C. A., Cabrera, D. G., Huasco, W. H., Silvaespejo, J. E., et al. (2015). Drought impact on forest carbon dynamics and fluxes in Amazonia. *Nature* 519, 78–82. doi: 10.1038/nature14213

Eamus, D., Boulain, N., Cleverly, J., and Breshears, D. D. (2013). Global change−type drought−induced tree mortality: vapor pressure deficit is more important than temperature per se in causing decline in tree health. *Ecol. Evol.* 3, 2711–2729. doi: 10.1002/ece3.664

Fei, X., Song, Q., Zhang, Y., Liu, Y., Sha, L., Yu, G., et al. (2018). Carbon exchanges and their responses to temperature and precipitation in forest ecosystems in Yunnan. *Sci. Total Environ.* 616–617, 824–840. doi: 10.1016/j.scitotenv.2017.10.239

Gherardi, L. A., and Sala, O. E. (2019). Effect of interannual precipitation variability on dryland productivity: a global synthesis. *Glob. Chang. Biol.* 25, 269–276. doi: 10.1111/gcb.14480

Heyer, J. P., Power, M. J., Field, R. D., and Van Marle, M. J. (2018). The impacts of recent drought on fire, forest loss, and regional smoke emissions in lowland Bolivia. *Biogeosciences* 15, 4317–4331. doi: 10.5194/bg-15-4317-2018

Hou, X., Ji, L., and Wang, Z. (2014). Responses of NPP and carbon accumulation in desert steppe and typical grassland to different rainfall and land use patterns. *Acta Ecol. Since* 34, 6256–6264.

Huang, K., and Xia, J. (2019). High ecosystem stability of evergreen broadleaf forests under severe droughts. *Glob. Chang. Biol.* 25, 3494–3503. doi: 10.1111/gcb.14748

Huang, L., He, B., Chen, A., Wang, H., Liu, J., Lû, A., et al. (2016). Drought dominates the interannual variability in global terrestrial net primary production by controlling semi-arid ecosystems. *Sci. Rep.* 6:24639. doi: 10.1038/srep24639

Ji, L., and Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sens. Environ.* 87, 85–98. doi: 10.1016/S0034-4257(03)00174-3

Jiang, R., He, W., Zhou, W., Hou, Y., Yang, J. Y., and He, P. (2019). Exploring management strategies to improve maize yield and nitrogen use efficiency in northeast China using the DNDC and DSSAT models. *Comput. Electron. Agric.* 166:104988. doi: 10.1016/j.compag.2019.104988

Jiang, X., Shen, W., and Bai, X. (2019). Response of net primary productivity to vegetation restoration in Chinese Loess Plateau during 1986-2015. *PLoS One* 14:e02192. doi: 10.1371/journal.pone.0219270

Kannenberg, S. A., Novick, K. A., Alexander, M. R., Maxwell, J. T., Moore, D. J., Phillips, R. P., et al. (2019). Linking drought legacy effects across scales: from leaves to tree rings to ecosystems. *Glob. Chang. Biol.* 25, 2978–2992. doi: 10.1111/gcb.14710

Kim, Y., Kimball, J. S., Zhang, K., and Mcdonald, K. C. (2012). Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: implications for regional vegetation growth. *Remote Sens. Environ.* 121, 472–487. doi: 10.1016/j.rse.2012.02.014

Knapp, S., and van der Heijden, M. G. (2018). A global meta-analysis of yield stability in organic and conservation agriculture. *Nat. Commun.*9:3632. doi: 10.1038/s41467-018-05956-1

Lai, C., Li, J., Wang, Z., Wu, X., Zeng, Z., Chen, X., et al. (2018). Drought-induced reduction in net primary productivity across mainland China from 1982 to 2015. *Remote Sens.* 10:1433. doi: 10.3390/rs10091433

Lee, D., and Veizer, J. (2003). Water and carbon cycles in the Mississippi River basin: potential implications for the Northern Hemisphere residual terrestrial sink. *Global Biogeochem. Cycles* 17, 109–124. doi: 10.1029/2002GB001984

Lei, T., Wu, J., Li, X., Geng, G., Shao, C., Zhou, H., et al. (2015). A new framework for evaluating the impacts of drought on net primary productivity of grassland. *Sci. Total Environ.* 536, 161–172. doi: 10.1016/j.scitotenv.2015.06.138

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature* 529, 84–87. doi: 10.1038/nature16467

Li, J., Wang, Z., and Lai, C. (2019a). Severe drought events inducing large decrease of net primary productivity in mainland China during 1982–2015. *Sci. Total Environ.* 703:135541. doi: 10.1016/j.scitotenv.2019.135541

Li, J., Wang, Z., Lai, C., Wu, X., Zeng, Z., Chen, X., et al. (2018). Response of net primary production to land use and land cover change in mainland China since the late 1980s. *Sci. Total Environ.* 639, 237–247. doi: 10.1016/j.scitotenv.2018.05.155

Li, J., Zou, C., Li, Q., Xu, X., Zhao, Y., Yang, W., et al. (2019b). Effects of urbanization on productivity of terrestrial ecological systems based on linear fitting: a case study in Jiangsu, eastern China. *Sci. Rep.* 9:17140. doi: 10.1038/s41598-019-53789-9

Li, Z., Chen, Y., Wang, Y., and Fang, G. (2016). Dynamic changes in terrestrial net primary production and their effects on evapotranspiration. *Hydrol. Earth Syst. Sci.* 20, 2169–2178. doi: 10.5194/hess-20-2169-2016

Liu, M. (2001). *Study on Carbon Storage and Vegetation Productivity in Chinese Terrestrial Ecosystem*. Ph.D. thesis, University of Chinese Academy of Sciences, Beijing.

Liu, W. (2007). *Study on Spatial and Temporal Variations of Net Primary Productivity and Carbon of Terrestrial Ecosystems in Xinjiang*. Ph.D. thesis, Xinjiang University, Ürümqi.

Liu, X., and Wang, Y. (2011). Contrasting impacts of spring thermal conditions over Tibetan Plateau on late-spring to early-summer precipitation in southeast China. *AtScL* 12, 309–315. doi: 10.1002/asl.343

Liu, X., Pei, F., Wen, Y., Li, X., Wang, S., Wu, C., et al. (2019). Global urban expansion offsets climate-driven increases in terrestrial net primary productivity. *Nat. Commun.* 10:5558. doi: 10.1038/s41467-019-13462-1

Liu, Y., Ju, W., He, H., Wang, S., Sun, R., and Zhang, Y. (2013). Changes of net primary productivity in China during recent 11 years detected using an ecological model driven by MODIS data. *Front. Earth Sci.* 7:112–127. doi: 10.1007/s11707-012-0348-5

Liu, Y., Wang, J., and Guo, L. (2009). Temporal and spatial dynamics of grain production and cultivated land change in China. *Sci. Agric. Sin.* 42, 4269–4274. doi: 10.3864/j.issn.0578-1752.2009.12.018

Luo, X., Croft, H., Chen, J. M., He, L., and Keenan, T. F. (2019). Improved estimates of global terrestrial photosynthesis using information on leaf chlorophyll content. *Glob. Chang. Biol.* 25, 2499–2514. doi: 10.1111/gcb.14624

Mk, V. D. M., Dolman, A. J., Ciais, P., Eglin, T., Gobron, N., Law, B. E., et al. (2011). Drought and ecosystem carbon cycling. *Agric. For. Meteorol.* 151, 765–773. doi: 10.1016/j.agrformet.2011.01.018

MWRPRC, (2016). *China Flood and Drought Disaster Bulletin*. Beijing: SinoMaps Press.

Pei, F., Li, X., Liu, X., and Lao, C. (2013). Assessing the impacts of droughts on net primary productivity in China. *J. Environ. Manage.* 114, 362–371. doi: 10.1016/j.jenvman.2012.10.031

Peng, D., Wu, C., Zhang, B., Huete, A., Zhang, X., Sun, R., et al. (2016). The influences of drought and land-cover conversion on inter-annual variation of NPP in the Three-North Shelterbelt Program zone of China based on MODIS data. *PLoS One* 11:e0158173. doi: 10.1371/journal.pone.0158173

Peng, H., Cheng, G., Xu, Z., Yin, Y., and Xu, W. (2007). Social, economic, and ecological impacts of the "Grain for Green" project in China: a preliminary case in Zhangye. *J. Environ. Manage.* 85, 774–784. doi: 10.1016/j.jenvman.2006.09.015

Persson, M., Moberg, J., Ostwald, M., and Xu, J. (2013). Chinese Grain for Green programme: assessing the carbon sequestered via land reform. *J. Environ. Manage.* 126, 142–146. doi: 10.1016/j.jenvman.2013.02.045

Piao, S., Fang, J., Zhou, L., Zhu, B., Tan, K., and Tao, S. (2005). Changes in vegetation net primary productivity from 1982 to 1999 in China. *Glob. Biogeochem. Cycles* 19, 1605–1622. doi: 10.1029/2004GB002274

Piao, S., Friedlingstein, P., Ciais, P., Viovy, N., and Demarty, J. (2007). Growing season extension and its impact on terrestrial carbon cycle in the Northern Hemisphere over the past 2 decades. *Glob. Biogeochem. Cycles* 21, 1148–1154. doi: 10.1029/2006GB002888

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., et al. (2019). Plant phenology and global climate change: current progresses and challenges. *Glob. Chang. Biol.* 25, 1922–1940. doi: 10.1111/gcb.14619

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., et al. (2020). Characteristics, drivers and feedbacks of global greening. *Nat. Rev. Earth Environ.* 1, 14–27. doi: 10.1038/s43017-019-0001-x

Ruimy, A., Saugier, B., and Dedieu, G. (1994). Methodology for the estimation of terrestrial net primary production from remotely sensed data. *J. Geophys. Res. Atmos.* 99, 5263–5283. doi: 10.1029/93JD03221

Shang, E., Xu, E., Zhang, H., and Liu, F. (2018). Analysis of spatiotemporal dynamics of the Chinese vegetation net primary productivity from the 1960s to the 2000s. *Remote Sens.* 10:860. doi: 10.3390/rs10060860

Su, B., Huang, J., Fischer, T., Wang, Y., Kundzewicz, Z. W., Zhai, J., et al., (2018). Drought losses in China might double between the 1.5 °C and 2.0 °C warming. *PNAS* 115, 10600–10605. doi: 10.1073/pnas.1802129115

Tan, S., Zhu, M., Zhang, K., Dang, H., and Zhang, Q. (2009). Response and adaptation of plants to submergence stress. *Chin. J. Ecol.* 28, 1871–1877.

Tao, B., Li, K., Shao, X., and Cao, M. (2003). The spatio-temporal characteristics of the net primary productivity based on the simulation in china. *Acta Geogr. Sin.* 58, 372–380.

Tucker, C., Pinzon, J., and Brown, M. (2004). *Global Inventory Modeling and Mapping Studies*. College Park, MD: Global Land Cover Facility.

Vicenteserrano, S. M., Gouveia, C., Camarero, J. J., Beguería, S., Trigo, R., Lópezmoreno, J. I., et al. (2013). Response of vegetation to drought time-scales across global land biomes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 52–57. doi: 10.1073/pnas.1207068110

Vicente−Serrano, S. M., Camarero, J. J., and Azorin−Molina, C. (2015). Diverse responses of forest growth to drought time−scales in the Northern Hemisphere. *Glob. Ecol. Biogeogr.* 23, 1019–1030. doi: 10.1111/geb.1218

Wu, D., Zhao, X., Liang, S., Zhou, T., Huang, K., Tang, B., et al. (2015). Time−lag effects of global vegetation responses to climate change. *Glob. Chang. Biol.* 21, 3520–3531. doi: 10.1111/gcb.12945

Xiao, J., Zhuang, Q., Liang, E., Mcguire, A. D., Moody, A., Kicklighter, D. W., et al. (2009). Twentieth-Century droughts and their impacts on terrestrial carbon cycling in China. *Earth Interact.* 13, 1–31. doi: 10.1175/2009EI275.1

Xu, B., Arain, M. A., Black, T. A., Law, B. E., Pastorello, G. Z., and Chu, H. (2019). Seasonal variability of forest sensitivity to heat and drought stresses: a synthesis based on carbon fluxes from North American forest ecosystems. *Glob. Chang. Biol.* 26, 901–918. doi: 10.1111/gcb.14843

Xu, X., Yang, G., Tan, Y., Tang, X., Jiang, H., Sun, X., et al. (2017). Impacts of land use changes on net ecosystem production in the Taihu Lake Basin of China from 1985 to 2010. *J. Geophys. Res. Biogeosci.* 122, 690–707. doi: 10.1002/2016JG003444

Yang, S., Meng, D., Li, X., and Wu, X. (2018). Multi-timescale responses of vegetation to SPEI in north China from 2001 to 2014. *Acta Ecol. Sin.* 38, 1028–1039.

Yu, D., Zhu, W., and Pan, Y. (2007). The role of atmospheric circulation system playing in coupling relationship between spring NPP and precipitation in East Asia area. *Environ. Monit. Assess.* 145, 135–143. doi: 10.1007/s10661-007-0023-6

Yuan, Q., Wu, S., Zhao, D., Dai, E., Li, C., and Lei, Z. (2014). Modeling net primary productivity of the terrestrial ecosystem in China from 1961 to 2005. *J. Geogr. Sci.* 24, 3–17. doi: 10.1007/s11442-014-1069-3

Zarei, A. R., and Eslamian, S. (2017). Trend assessment of precipitation and drought index (SPI) using parametric and non-parametric trend analysis methods (case study: Arid regions of southern Iran). *Int. J. Hydrol. Sci. Technol.* 7, 12–38. doi: 10.1504/IJHST.2017.080957

Zeng, Z., Piao, S., Li, L. Z., Wang, T., Ciais, P., Lian, X., et al. (2018). Impact of Earth greening on the terrestrial water cycle. *J. Clim.* 31, 2633–2650. doi: 10.1175/JCLI-D-17-0236.1

Zhang, C., and Yang, C. (2007). Comparison of growth dynamics of several different plants in typical steppe of Inner Mongolia of China. *Chin. J. Ecol.* 26, 1712–1718.

Zhao, M., and Running, S. W. (2010). Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* 329, 940–943. doi: 10.1126/science.1192666

Zhu, W., Pan, Y., He, H., Yu, D., and Hu, H. (2006). Simulation of maximum light use efficiency for some typical vegetation types in China. *Chin. Sci. Bull.* 51, 457–463. doi: 10.1007/s11434-006-0457-1

Zhu, W. Q., and Pan, Y. Z. (2007). Estimation of net primary productivity of Chinese terrestrial vegetation based on remote sensing. *Chin. J. Plant Ecol.* 31, 413–424.

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.

frontiers
in Earth Science

ORIGINAL RESEARCH
published: 05 May 2020
doi: 10.3389/feart.2020.00114

Insights Into Microseism Sources by Array and Machine Learning Techniques: Ionian and Tyrrhenian Sea Case of Study

Salvatore Moschella[1], Andrea Cannata[1,2]*, Flavio Cannavò[2], Giuseppe Di Grazia[2], Gabriele Nardone[3], Arianna Orasi[3], Marco Picone[3], Maurizio Ferla[3] and Stefano Gresta[1]

[1] Dipartimento di Scienze Biologiche, Geologiche e Ambientali – Sezione di Scienze della Terra, Università degli Studi di Catania, Catania, Italy, [2] Osservatorio Etneo – Sezione di Catania, Istituto Nazionale di Geofisica e Vulcanologia, Catania, Italy, [3] Centro Nazionale per la Caratterizzazione Ambientale e la Protezione Della Fascia Costiera, la Climatologia Marina e l'Oceanografia Operativa, Istituto Superiore per la Protezione e la Ricerca Ambientale, Rome, Italy

In this work, we investigated the microseism recorded by a network of broadband seismic stations along the coastline of Eastern Sicily. Microseism is the most continuous and ubiquitous seismic signal on Earth and is mostly generated by the ocean–solid earth interaction. On the basis of spectral content, it is possible to distinguish three types of microseism: primary, secondary, and short-period secondary microseism (SPSM). We showed how most of the microseism energy recorded in Eastern Sicily is contained in the secondary and SPSM bands. This energy exhibits strong seasonal patterns, with maxima during the winters. By applying array techniques, we observed how the SPSM sources are located in areas of extended shallow water depth: the Catania Gulf and a part of the Northern Sicily coastlines. Finally, by using the significant wave height data recorded by two buoys installed in the Ionian and Tyrrhenian Seas, we developed an innovative method, selected among up-to-date machine learning techniques (MLTs), able to reconstruct the time series of sea wave parameters from microseism recorded in the three microseism period bands by distinct seismic stations. In particular, the developed model, based on random forest regression, allowed estimating the significant wave height with a low average error (~0.14–0.18 m). The regression analysis suggests that the closer the seismic station to the sea, the more information concerning the sea state are contained in the recorded microseism. This is particularly important for the future development of an experimental monitoring system of the sea state conditions based on microseism recordings.

Keywords: microseism, machine learning, sea waves, array techniques, random forest

OPEN ACCESS

Edited by:
Susanne Buiter,
RWTH Aachen University, Germany

Reviewed by:
Luca De Siena,
Johannes Gutenberg University Mainz, Germany
Aurélien Mordret,
Université Grenoble Alpes, France

****Correspondence:***
Andrea Cannata
andrea.cannata@unict.it

Specialty section:
This article was submitted to Solid Earth Geophysics, a section of the journal Frontiers in Earth Science

Received: *09 January 2020*
Accepted: *26 March 2020*
Published: *05 May 2020*

Citation:
Moschella S, Cannata A, Cannavò F, Di Grazia G, Nardone G, Orasi A, Picone M, Ferla M and Gresta S (2020) Insights Into Microseism Sources by Array and Machine Learning Techniques: Ionian and Tyrrhenian Sea Case of Study. Front. Earth Sci. 8:114. doi: 10.3389/feart.2020.00114

INTRODUCTION

Microseism is the most continuous and ubiquitous seismic signal on Earth and is mostly generated by the ocean–solid earth interaction (Tanimoto et al., 2015). On the basis of its source mechanism and spectral content, it is classified as: primary microseism (hereafter referred to as PM), secondary microseism (SM), and short-period secondary microseism (SPSM) (Haubrich and McCamy, 1969). Concerning PM, it shares the same spectral content as the ocean waves (period band 13–20 s) and its source is associated with the energy transfer of ocean waves breaking/shoaling against the shoreline (Hasselmann, 1963; Ardhuin et al., 2015). As for SM, it is likely to be generated by

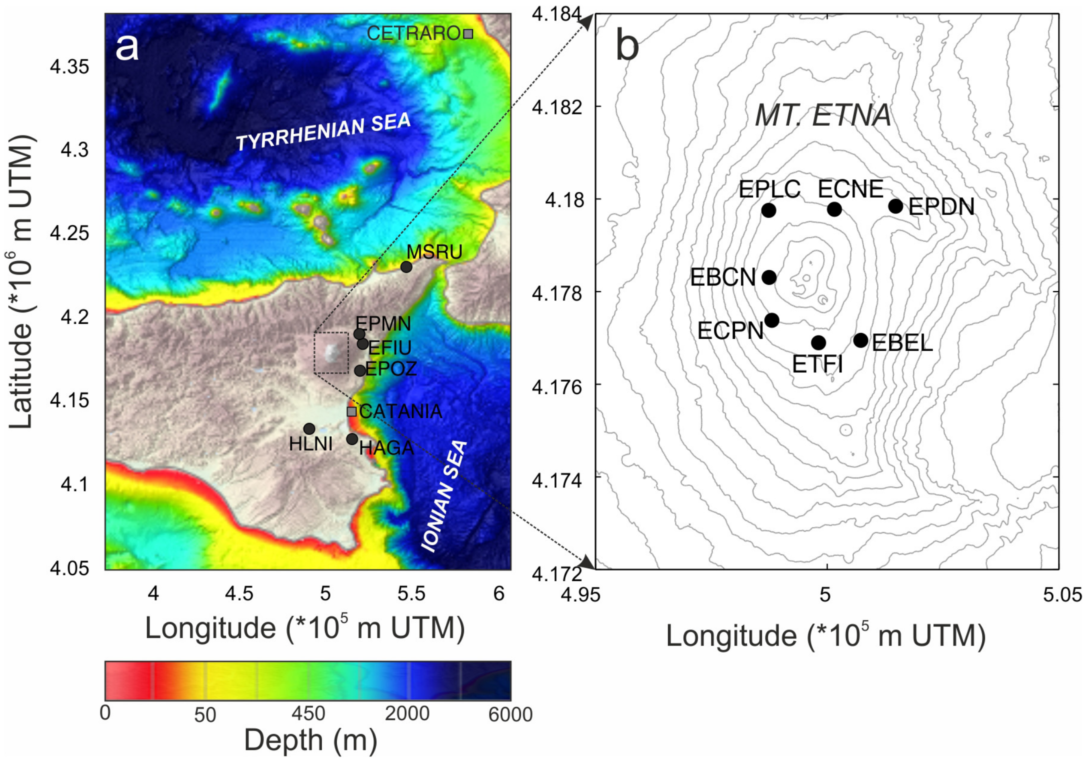


FIGURE 1 | (a) Bathymetric and topographic map (EMODnet Bathymetry Consortium, 2018), with the locations of the seismic stations (*black dots*), used to perform spectral and amplitude analysis of the microseism and to investigate its relationship with significant wave height, recorded by Catania and Cetraro buoy stations (*gray squares*). **(b)** Digital elevation model of Mt. Etna, with the locations of the seismic stations (*black dots*), used to perform array analysis.

interactions between waves of the same frequency traveling in opposite directions, has roughly twice the frequency of ocean waves (period band 5–10 s), and generally shows a higher amplitude than does PM (Longuet-Higgins, 1950; Oliver and Page, 1963; Ardhuin et al., 2012, 2015). Finally, SPSM is characterized by a period shorter than 5 s and is generated by local nearshore wave–wave interaction (Bromirski et al., 2005).

Because of its source mechanism, microseism has been used to make inferences on climate changes (e.g., Grevemeyer et al., 2000; Aster et al., 2008; Stutzmann et al., 2009). For instance, Grevemeyer et al. (2000) analyzed a 40-year-long record of wintertime microseism and observed an increase in the number of monthly days with strong microseism activity, hence inferring an increase over time in surface air temperatures and storminess of the northeast Atlantic Ocean.

Microseism amplitudes show strong seasonal modulation. Indeed, at temperate latitudes, microseism shows periodicity, with maxima during the winter seasons, when the oceans are stormier, and minima during the summers (Aster et al., 2008). This modulation is different along the coastlines of the Glacial Arctic Sea and the Southern Ocean where, during the winters, because of the sea ice, the oceanic waves cannot efficiently excite seismic energy (Aster et al., 2008; Stutzmann et al., 2009; Tsai and McNamara, 2011; Cannata et al., 2019).

Concerning the source location, microseism signals are non-impulsive, and the sources are generally diffuse and variable in time (e.g., Bromirski et al., 2013). Hence, the classical location algorithms, used in earthquake seismology and based on the picking of the different seismic phases, cannot be applied to locate microseism sources. Array processing techniques can overcome the above-mentioned difficulties and provide information on the microseism source areas that generally coincide with coastal regions and/or oceanic storm systems (e.g., Chevrot et al., 2007; Juretzek and Hadziioannou, 2017; Pratt et al., 2017; Lepore and Grad, 2018).

The link between microseism amplitudes and the ocean wave height has been empirically explored by several authors (e.g., Bromirski et al., 1999; Bromirski and Duennebier, 2002; Ardhuin et al., 2012; Ferretti et al., 2013, 2018). For instance, Bromirski et al. (1999) determined site-specific seismic-to-wave transfer functions in the San Francisco Bay area (California). Ferretti et al. (2013, 2018) found empirical relations to predict the significant wave height along the Ligurian coast (Italy). In addition, other authors have derived physics-based models of the generation of the different kinds of microseism from the sea state (e.g., Gualtieri et al., 2013; Ardhuin et al., 2015; Gualtieri et al., 2019).

Microseism investigations and, more generally, seismological studies are currently undergoing a rapid increase in dataset volumes (e.g., Kong et al., 2018; Jiao and Alavi, 2019). For this reason, nowadays, applications of machine learning techniques (hereafter referred to as MLTs) on seismological data are increasing in number day by day. Such techniques are used

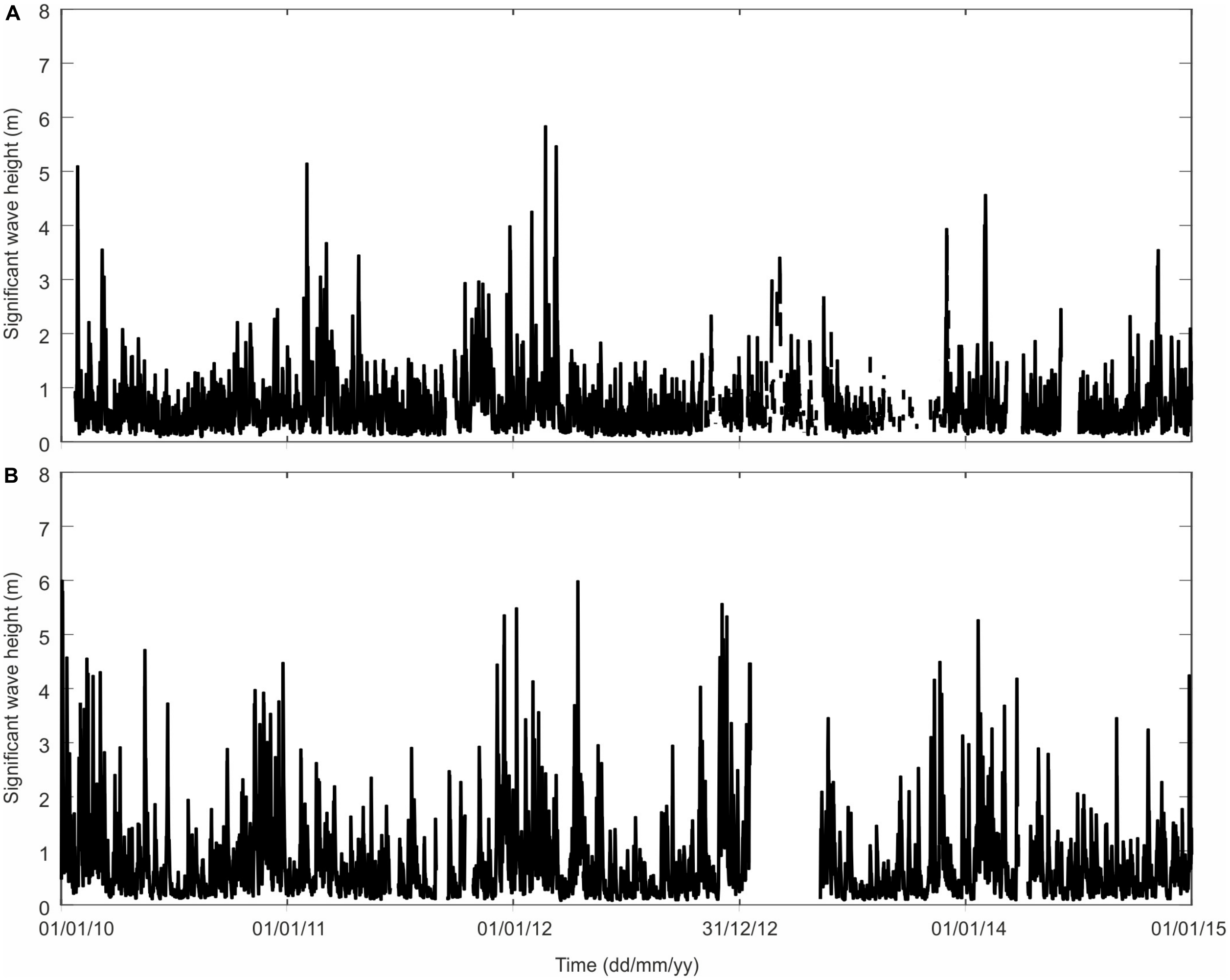


FIGURE 2 | Significant wave height time series recorded by the Catania **(A)** and Cetraro **(B)** buoys.

to extract information directly from data using well-defined optimization rules and help unravel hidden relationships between distinct parameters, as well as to build predictive models (e.g., Kuhn and Johnson, 2013; Kong et al., 2018). Examples of the applications of MLTs to seismology include earthquake detection and phase picking (e.g., Wiszniowski et al., 2014) and earthquake early warning (e.g., Kong et al., 2016).

In spite of the availability of seismic and buoy data in the Ionian and Tyrrhenian Seas and coastlines, the link between sea waves and microseism has never been explored in such areas. Furthermore, although the spectral features of the microseism recorded in this area have been studied (e.g., De Caro et al., 2014), the locations of its sources have never been constrained. In this work, we will study the microseism recorded along the coastline of Eastern Sicily in terms of spectral content, amplitude seasonal pattern, and source location. In addition, we will present a novel algorithm, based on up-to-date MLTs, able to reconstruct significant wave height time series in points located in both the Ionian and the Tyrrhenian Seas from the microseism recordings.

MATERIALS AND METHODS

Data

In order to investigate microseism, seismic signals recorded from 2010 to 2014 by the vertical component of six stations, belonging to the seismic permanent network run by Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo – Sezione di Catania (INGV-OE), were used (**Figure 1a**). These stations are equipped with broadband three-component Trillium 40-s seismometers (Nanometrics^{TM}) recording at a sampling rate of 100 Hz.

Moreover, to carry out array analysis, seismic signals recorded in January 2010–February 2012 by the vertical component of the seven stations (equipped with the same sensors as above), composing the summit ring of the Mt. Etna permanent seismic network, were used (**Figure 1b**). These stations were chosen because of: (i) the availability of continuously recorded data during the time interval 2010–beginning of 2012 (in February–March 2012, EBEL and ETFI stations were destroyed by lava

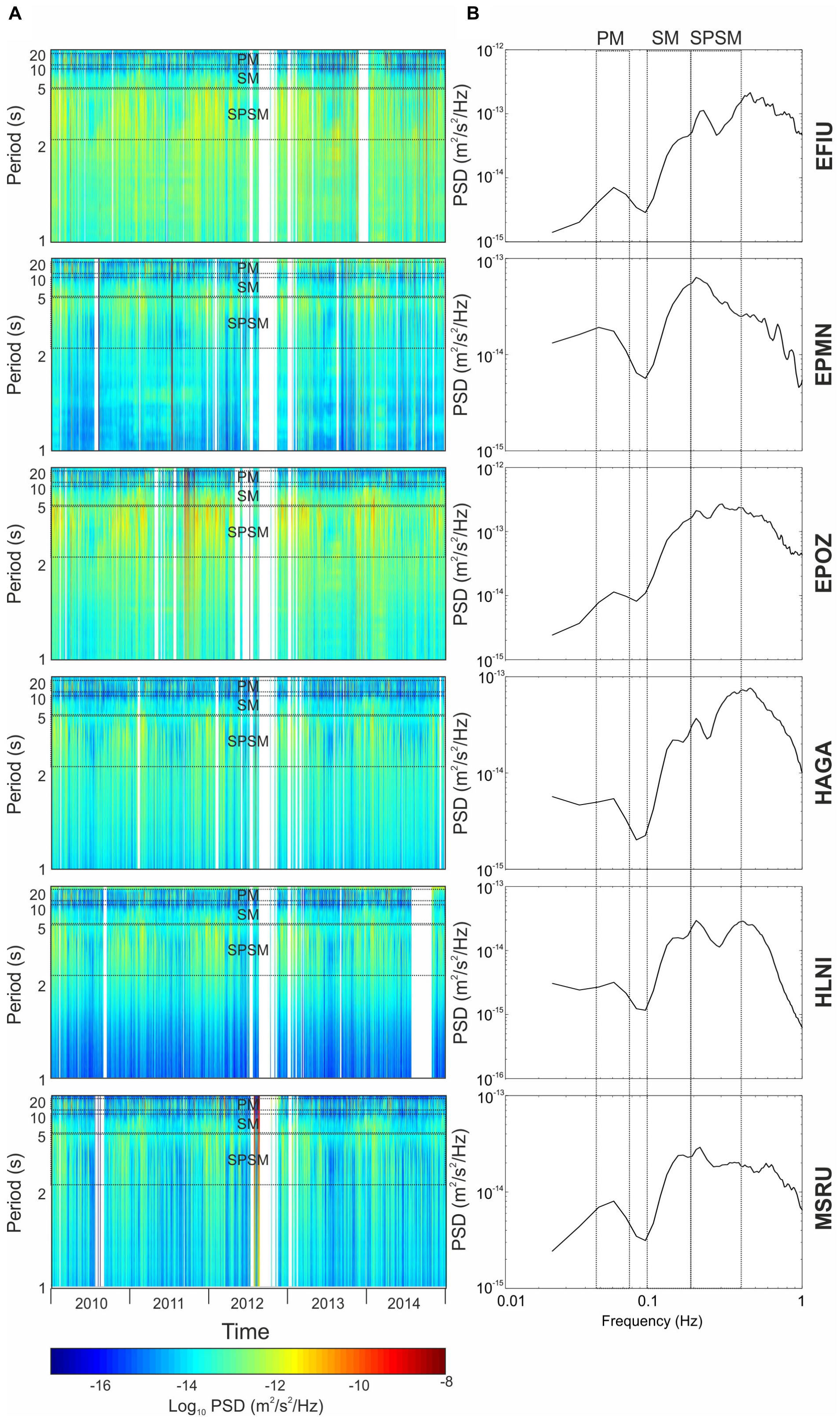


FIGURE 3 | (A) Spectrograms of the seismic signal recorded by the vertical component of the six considered stations. **(B)** Median spectra of the seismic signal recorded by the vertical component of the six considered stations. The acronyms *PM*, *SM*, and *SPSM* indicate primary microseism, secondary microseism, and short-period secondary microseism, respectively.

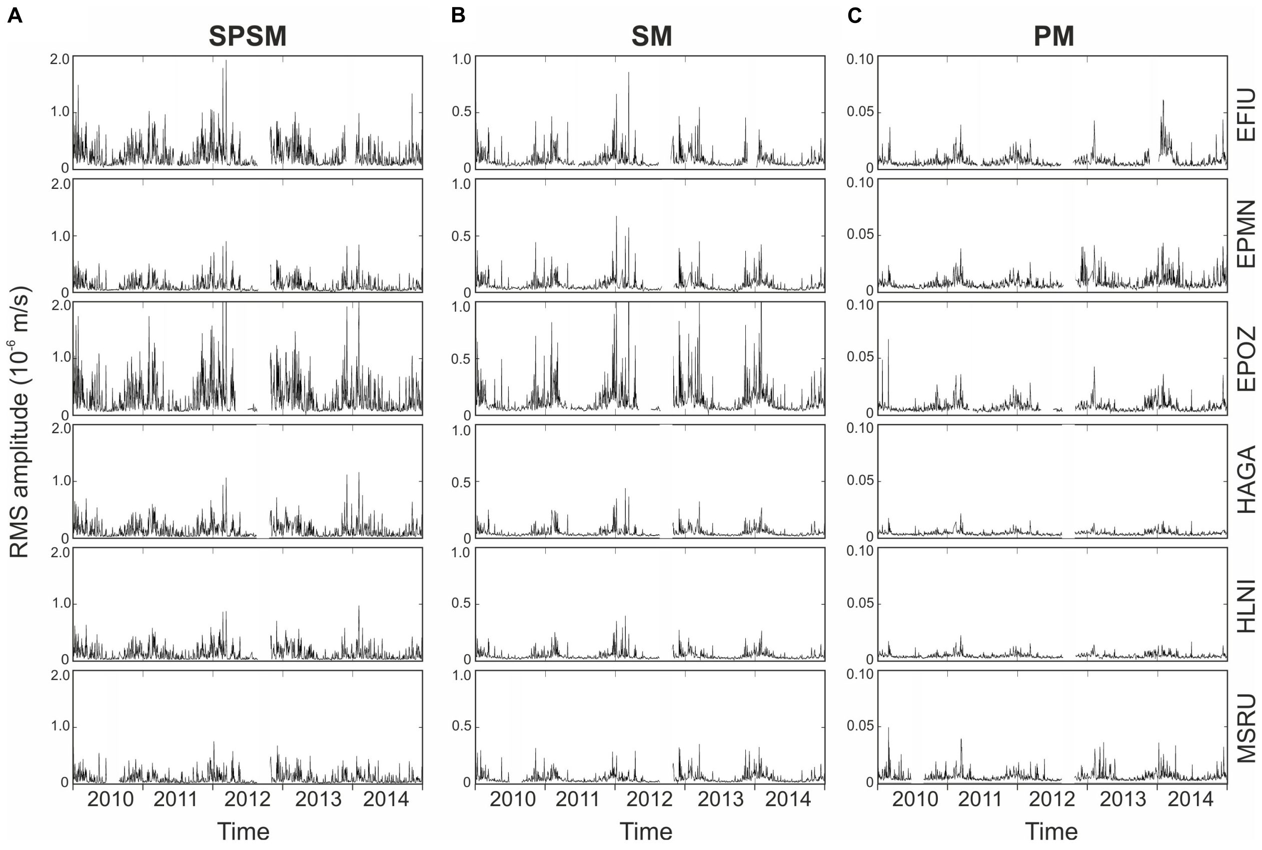


FIGURE 4 | RMS amplitude time series of the seismic signal recorded by the vertical component of the six considered stations and filtered in the bands **(A)** 2.5–5.0 s (*SPSM*, short-period secondary microseism); **(B)** 5–10 s (*SM*, secondary microseism); and **(C)** 13–20 s (*PM*, primary microseism).

flows); (ii) the ring-shaped geometry; and (iii) the distance from the coastline (and then from the prospective closest microseism sources associated with the nearshore wave–coast or wave–wave interaction).

Finally, to make quantitative comparisons between the microseism and wave height time series in the Ionian and the Tyrrhenian Seas, significant wave height data, recorded from 2010 to 2014 with a 30-min sampling step by two stations (Catania and Cetraro; see **Figure 1A**) belonging to the Italian Data Buoy Network, managed by Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), were used (Bencivenga et al., 2012; **Figure 2**). The significant wave height is defined as:

$$H_s = 4\sqrt{M_0} \tag{1}$$

where M_0 is the 0-moment of the auto-spectral correlation of the Fourier transformations of the buoy displacements in the frequency/time domain (Steele and Mettlach, 1993):

$$M_0 = \sum_{f_l}^{f_u} (S(f)d(f)) \tag{2}$$

where the sum of the spectral density $S(f)$ is over all frequency bands, from the lowest frequency f_l to the highest frequency f_u of the non-directional wave spectrum (calculated only for the elevation of the sea surface), and $d(f)$ is the bandwidth of each band.

Spectral and Amplitude Analysis

The spectral content of the seismic data, recorded by the vertical component of the six seismic stations shown in **Figure 1a**, was analyzed as follows: (i) spectra over non-overlapping 81.92-s-long sliding windows were computed; (ii) to obtain daily spectra (that is, spectra representing the frequency features of the signal acquired during a given day), all the spectra computed in (i) falling on the same day were averaged by Welch's segment averaging estimator (Welch, 1967); (iii) all the daily spectra were collected and visualized as spectrograms, which are 3D plots with time on the x-axis, frequency on the y-axis, and power spectral density (PSD) indicated by a color scale (**Figure 3A**).

Besides, to obtain information on the spectral features of the seismic signals recorded by the different stations during the whole investigated period, all the daily spectra composing the spectrograms were averaged. Hence, spectra showing the general spectral features of the 5-year-long seismic signals were shown (**Figure 3B**).

In addition, the time series of the root mean square (RMS) amplitude of the seismic signal, filtered in three period bands

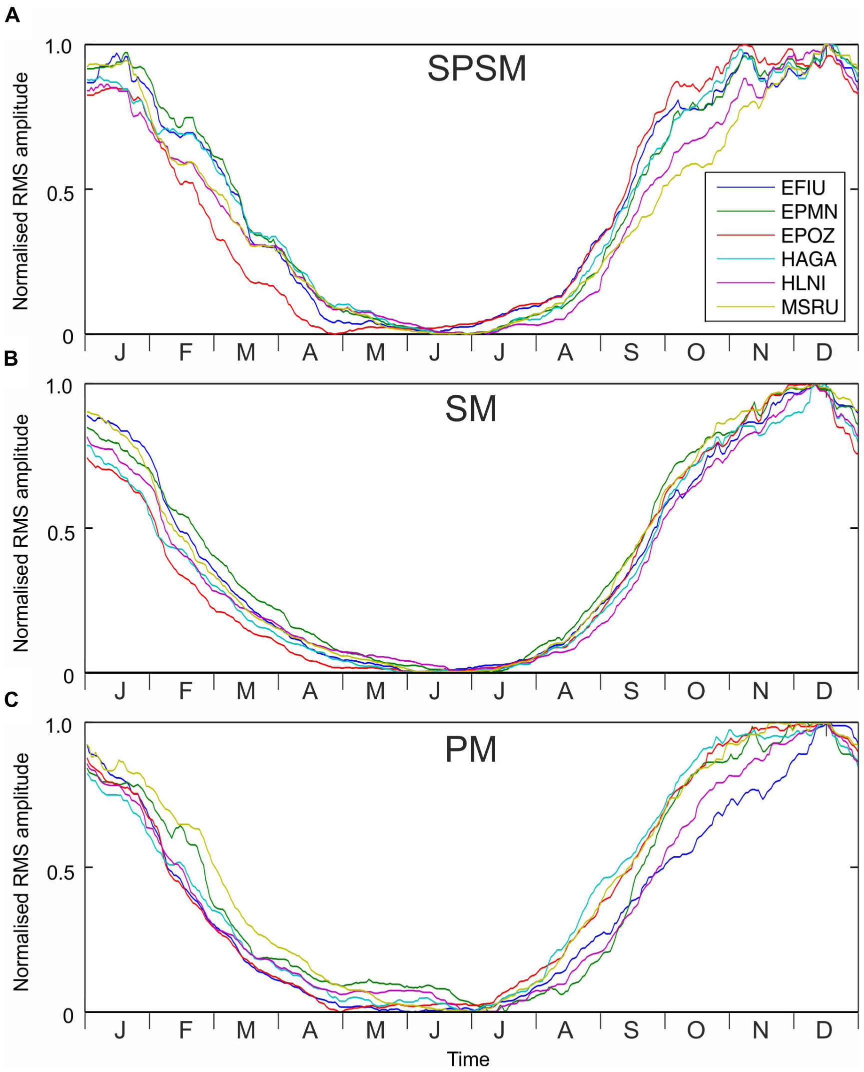


FIGURE 5 | RMS amplitude time series smoothed by a 90-day-long moving median, split into 1-year-long windows, stacked, and normalized for all the considered seismic stations (see the legends on the bottom right corner of **(A)**. In particular, regarding the period bands **(A)** 2.5–5 s (SPSM), **(B)** 5–10 s (SM), and **(C)** 13–20 s (PM). The time on the *x*-axis of **(A–C)** indicates the window onset of the 90-day-long moving median.

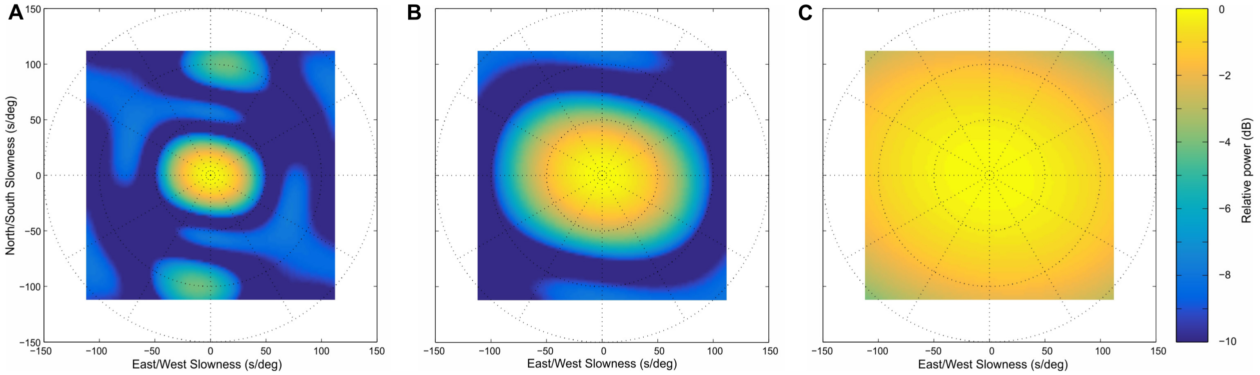


FIGURE 6 | Array response functions of the seven stations composing the summit ring of the Mt. Etna seismic permanent network (see **Figure 1b**) for a unit amplitude incident wave with slowness of 0 s deg^{-1} at periods of 2.5 s **(A)**, 5 s **(B)**, and 13 s **(C)**.

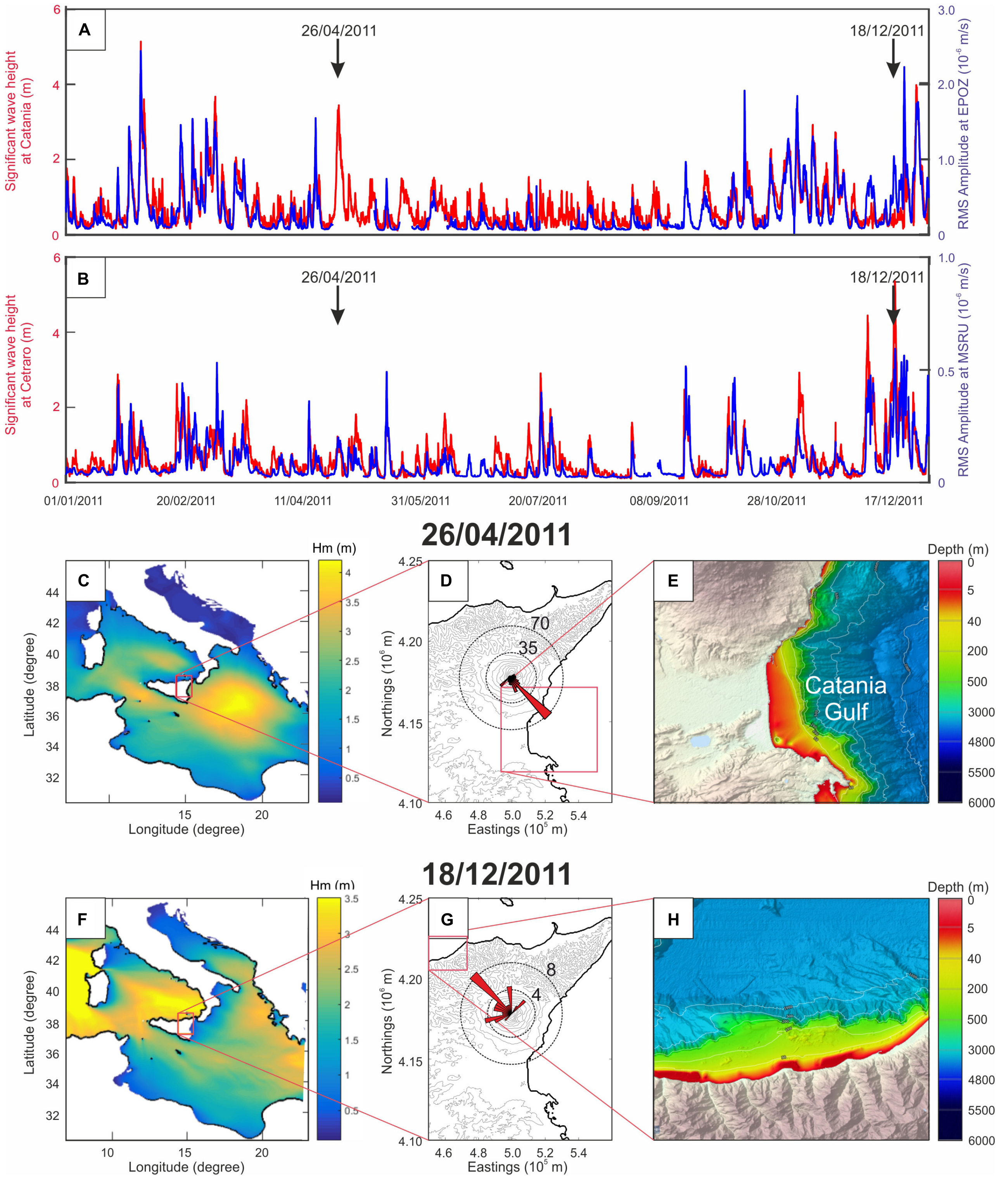


FIGURE 7 | (A,B) Time series of significant sea wave height recorded by the Catania and Cetraro buoys (*red lines*) and the RMS amplitude computed in the period band 2.5–5.0 s (SPSM) by EPOZ and MSRU stations (*blue lines*) in 2011. **(C,F)** Maps of a portion of the Mediterranean Sea showing the spatial distribution of the significant wave height on 26/04/2011 at 12:00 and on 18/12/2011 at 12:00, respectively (MEDSEA_HINDCAST_WAV_006_012 product from http://marine.copernicus.eu/services-portfolio/access-to-products/). **(D,G)** Digital elevation models of Eastern Sicily with rose diagrams, located at the center of the seismic summit ring of Mt. Etna (see **Figure 1b**), showing the distribution of the back azimuth values on 26/04/2011 and 18/12/2011, computed by *f–k* analysis. **(E,H)** Maps showing the bathymetry of portions of Sicily coastlines (EMODnet Bathymetry Consortium, 2018).

(PM, 13–20 s; SM, 5–10 s; and SPSM, 2.5–5.0 s), were computed with both daily and hourly rates. The daily RMS amplitude time series (**Figure 4**) were smoothed by a 90-day-long moving median, split in year-long windows, stacked, and rescaled between 0 and 1 (**Figure 5**).

Array Analysis

To get an idea on the locations of the main microseism sources surrounding the Eastern Sicilian coastlines, the seven stations composing the summit ring of the Mt. Etna seismic permanent network were used as a roughly circular array (**Figure 1b**). The array response functions (ARFs) were computed for the PM, SM, and SPSM for a plane wave arriving with a slowness of 0 s deg^{-1} (**Figure 6**). Such ARFs exhibit that only the SPSM case shows a fairly good resolution. This is due to the very long wavelength of PM and SM compared to the array aperture (~5 km). Indeed, taking into account a velocity of the S-waves (V_s) in the first kilometers of the crust equal to ~2 km/s (e.g., Hirn et al., 1991; Patanè et al., 1994), the wavelengths of PM, SM, and SPSM are ~26, 10, and 5 km, respectively. When the wavelength is much greater than the array aperture (as in the case of PM and SM), the array behaves like a single station (e.g., Schweitzer et al., 2012).

The portions of the Ionian and Tyrrhenian coastlines, where the microseism sources closest to the array could supposedly be located, are characterized by a minimum distance of ~20 and ~45 km, respectively, from the array center. Such distances are greater than two to three times the array aperture, and hence, on the basis of the synthetic tests performed by Almendros et al. (2002), the Etna summit ring array should be able to locate the microseism sources with a planar wavefront assumption.

Then, to apply array analysis, the following processing steps were carried out on the seismic signals: demeaning and detrending, correction for the instrument response, filtering within a 0.2- to 0.40-Hz band by a second-order Butterworth filter, and subdivision in 60-s-long windows, tapered with a Tukey window. The filter is also used to exclude volcanic tremor, whose energy at Mt. Etna is mainly radiated in the band 0.5–5.0 Hz (Cannata et al., 2010). Successively, the STA/LTA technique (acronym for short time average over long time average; e.g., Trnkoczy, 2012) was applied to detect prospective amplitude transients that could be related to volcano activity (i.e., long period events and very long period events). Windows containing amplitude transients were excluded from the array analysis. Finally, the frequency–wavenumber (*f–k*) analysis was carried out, allowing to calculate the power distributed among different slownesses and back azimuths (e.g., Capon, 1973; Rost and Thomas, 2002).

The array analysis was performed in January 2010–February 2012 on specific time intervals characterized by one of the following two conditions: (i) intense wave activity in the Ionian Sea, as shown by the Catania buoy data and/or by the high RMS amplitude values at EPOZ station; or (ii) intense wave activity in the Tyrrhenian Sea, as suggested by the Cetraro buoy data and/or by the high amplitude RMS values at MSRU station. Examples of the results for the days 26/04/2011 and 18/12/2011, exhibiting conditions (i) and (ii), respectively, are shown in **Figures 7**, **8**.

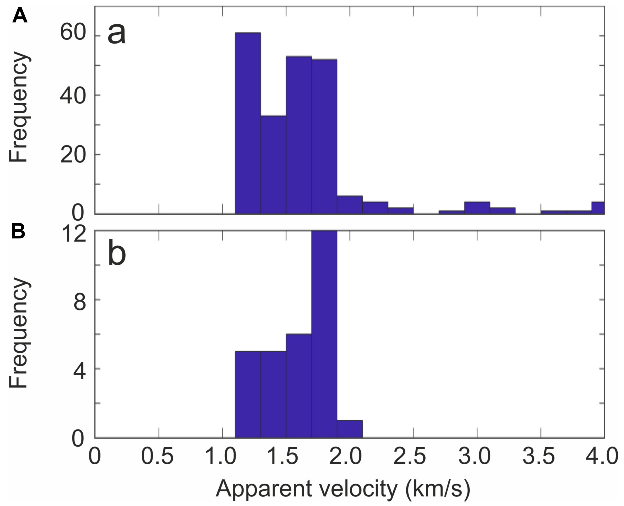


FIGURE 8 | Histograms showing the apparent velocity estimated by *f–k* analysis on 26/04/2011 **(a)** and on 18/12/2011 **(b)**.

To evaluate the error associated with the back azimuth estimation, the jackknife technique (Efron, 1982) was employed as follows. Firstly, the signal window was analyzed by the *f–k* technique by using all the seven stations composing the array. Successively, the analysis was repeated seven times, leaving one station out at a time, so providing further seven back azimuth values. An arithmetic mean of these estimates was assessed by the following equation:

$$\bar{P} = \frac{1}{n}\sum_{i=1}^{n} P_i \tag{3}$$

where P_i is the back azimuth value computed by omitting the i-th station and n is the number of stations composing the array. Then, it is possible to estimate the i-th so-called pseudovalue as:

$$J_i = n\hat{P} - (n-1)P_i \tag{4}$$

where $\hat{P}$ is the back azimuth value computed by considering all the seven array stations. The jackknife estimator of parameter P is given by:

$$J(\hat{P}) = \frac{1}{n}\sum_{i=1}^{n} J_i = n\hat{P} - (n-1)\bar{P} \tag{5}$$

The standard error of the jackknife estimates is given by:

$$\delta_{J(\hat{P})} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(J_i - J(\hat{P}))^2} \tag{6}$$

Finally, median error estimations were calculated separately for the back azimuths oriented toward the Ionian Sea and the Tyrrhenian Sea [the above-mentioned conditions (i) and (ii)].

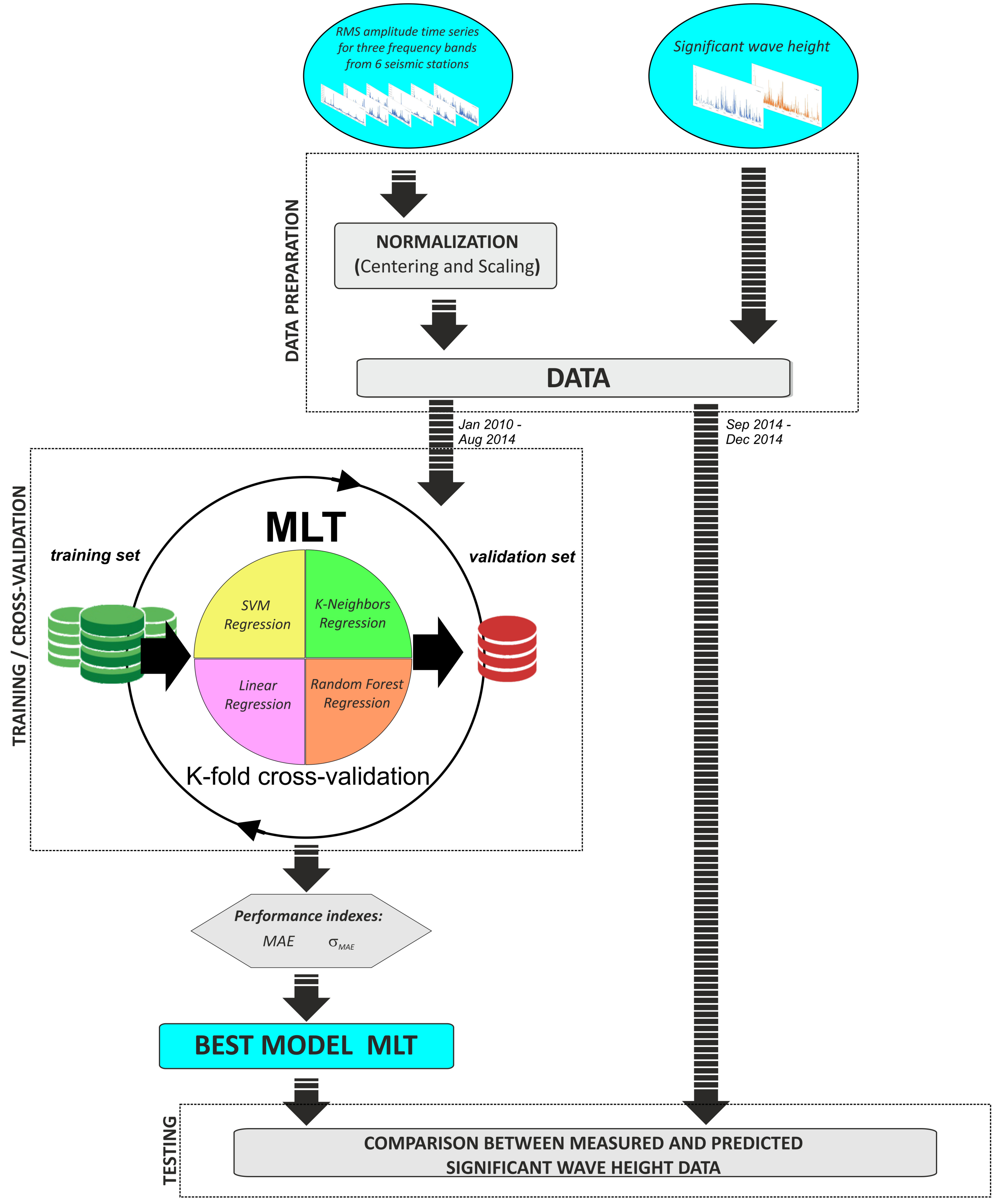


FIGURE 9 | Scheme of the modeling analysis to obtain the time series of significant wave height in the Catania and Cetraro buoy locations by using the microseism (see text for details). *MLT*, machine learning technique; *MAE*, mean absolute error; "σ_{MAE}", standard deviation computed on the mean absolute error; *SVM* support vector machine.

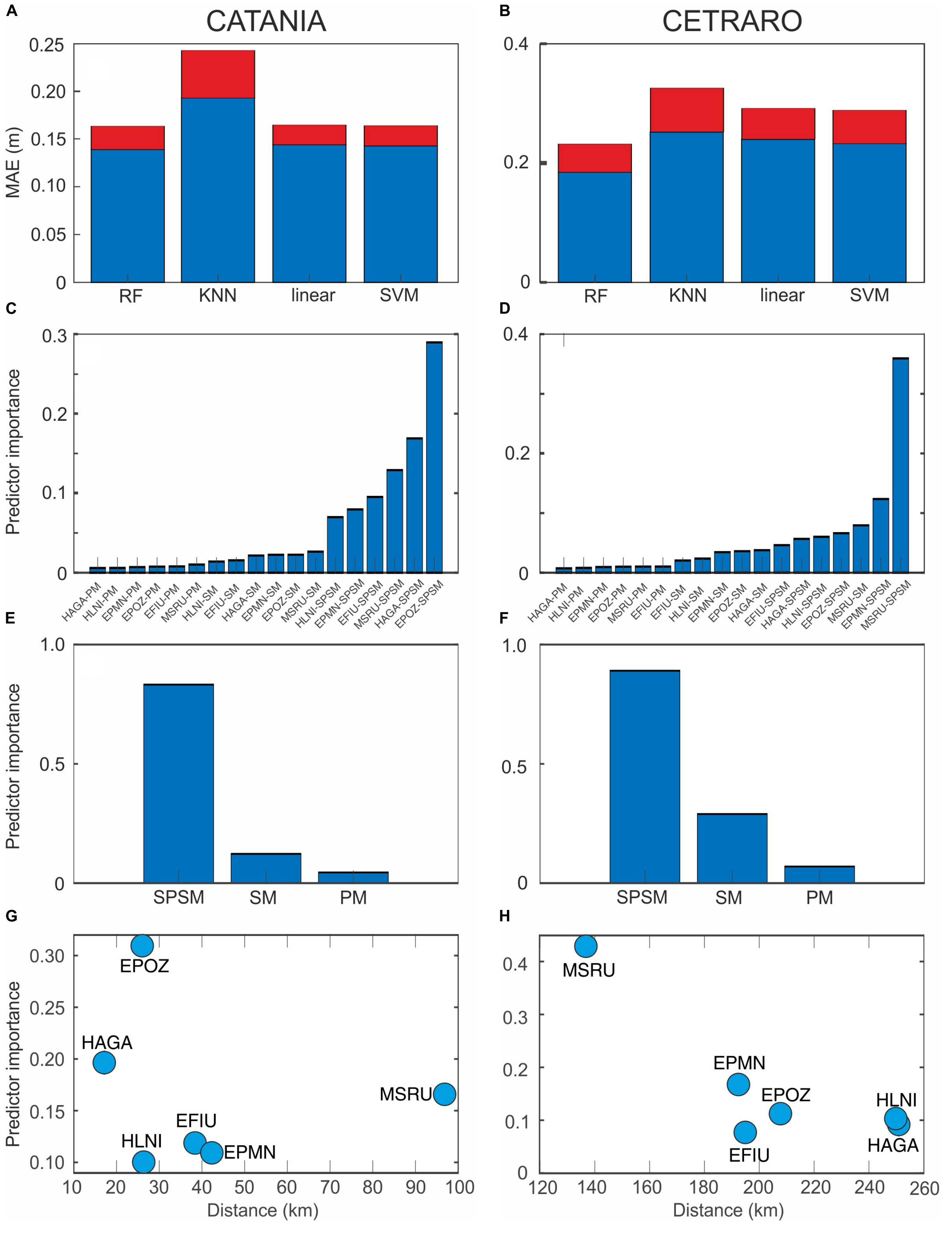


FIGURE 10 | Results of the machine learning analysis. **(A,B)** Average (*blue bars*) and standard deviation (*red bars*) of the mean absolute error (*MAE*), estimated by *k*-fold cross-validation, for the Catania and Cetraro buoy data, respectively. **(C,D)** Index of importance for all the input taken into account to model the Catania and Cetraro buoy data, respectively. **(E,F)** Aggregation through a summation of the input importance allowing to rank the microseism bands for the Catania and Cetraro buoy data prediction, respectively. **(G,H)** Aggregation through a summation of the station importance for the Catania and Cetraro buoy data prediction plotted *versus* the distance from the Catania and Cetraro buoys, respectively.

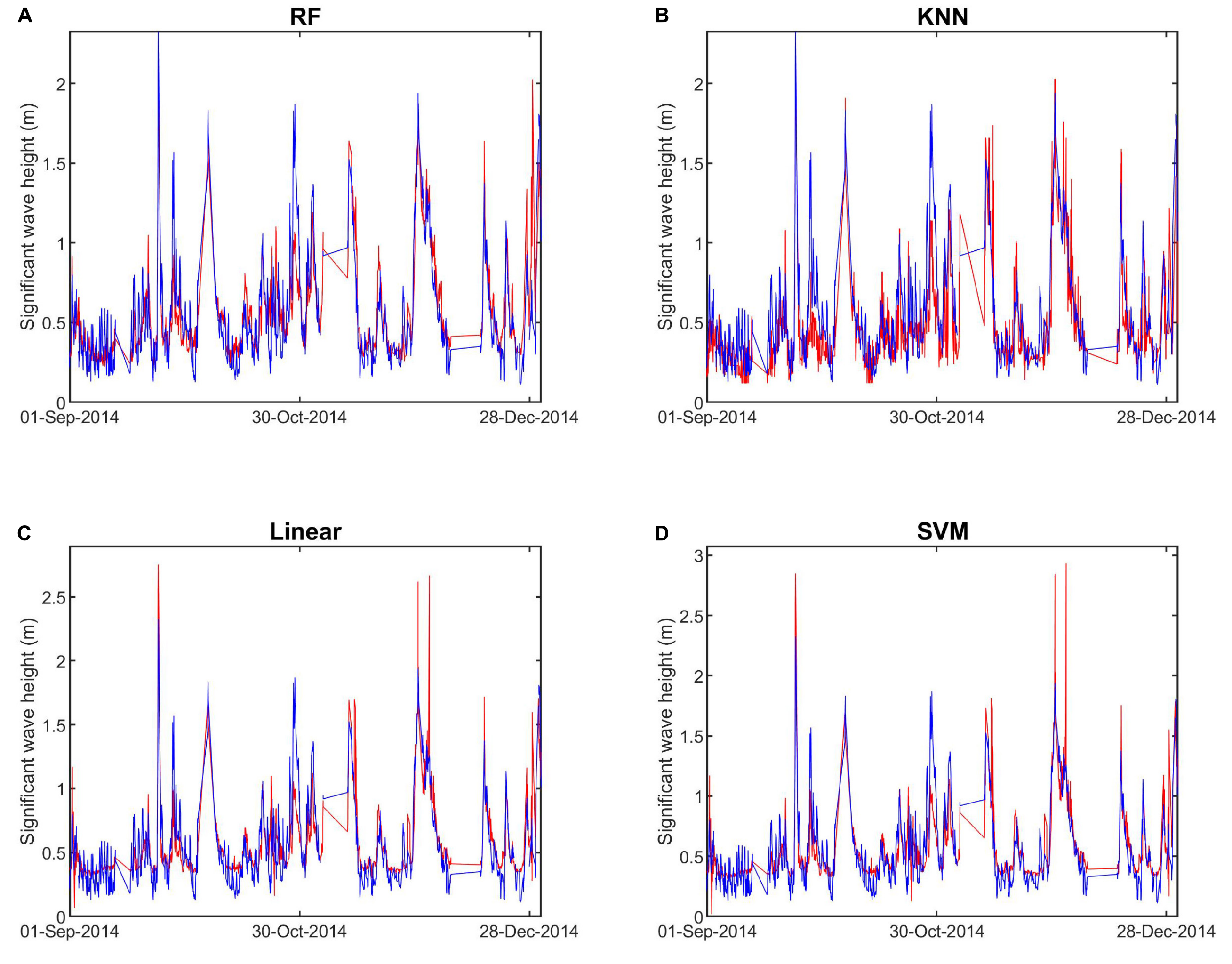


FIGURE 11 | Measured (*blue line*) and predicted (*red line*) significant wave height time series of the Catania buoy from 1 September to 31 December 2014. The prediction was carried out by RF regression **(A)**, KNN regression **(B)**, linear regression **(C)**, and SVM regression **(D)**.

Regression Analysis by Machine Learning

Modern MLTs have been tested to build reliable predictive models able to calculate the time series of significant wave height from microseism data. The method, similar to the one proposed by Cannata et al. (2019) to spatially and temporally reconstruct the sea ice distribution around Antarctica based on the microseism amplitudes, is composed of four main steps (summarized in **Figure 9**): (a) data preparation; (b) training; (c) cross-validation; and (d) testing.

Step (a) consisted of centering and scaling the predictor variables (Kuhn and Johnson, 2013), that is, the 18 time series of the microseism hourly RMS amplitudes from January 2010 to August 2014 (six stations by three frequency bands). The remaining data (September–December 2014) is used for testing step (d). To center the microseism predictor, the average is subtracted from all the values. Successively, to scale the data, each value of the microseism predictor is divided by its standard deviation. Hence, all the time series of the microseism RMS amplitudes share a common scale.

As for step (b), we made use of the following four MLTs to build predictive models: (i) random forest (RF) regression; (ii) K-nearest neighbors (KNN) regression; (iii) linear regression; and (iv) support vector machine (SVM) regression.

As for the RF technique, it is based on decision trees often used for classification and regression (Ho, 1995). One of the main problems with decision trees is the need to increase accuracy and avoid overfitting at the same time (Ho, 1998). RF overcomes such a limitation by generating many decision trees and aggregating their results (Liaw and Wiener, 2002). Recently, RF has had many applications in geosciences, such as geochemical mapping (Kirkwood et al., 2016) and the lithological classification of underexplored areas by geophysical and remote sensing data (Kuhn et al., 2018).

K-nearest neighbors is a non-parametric technique applied for both classification and regression tasks (Altman, 1992). KNN

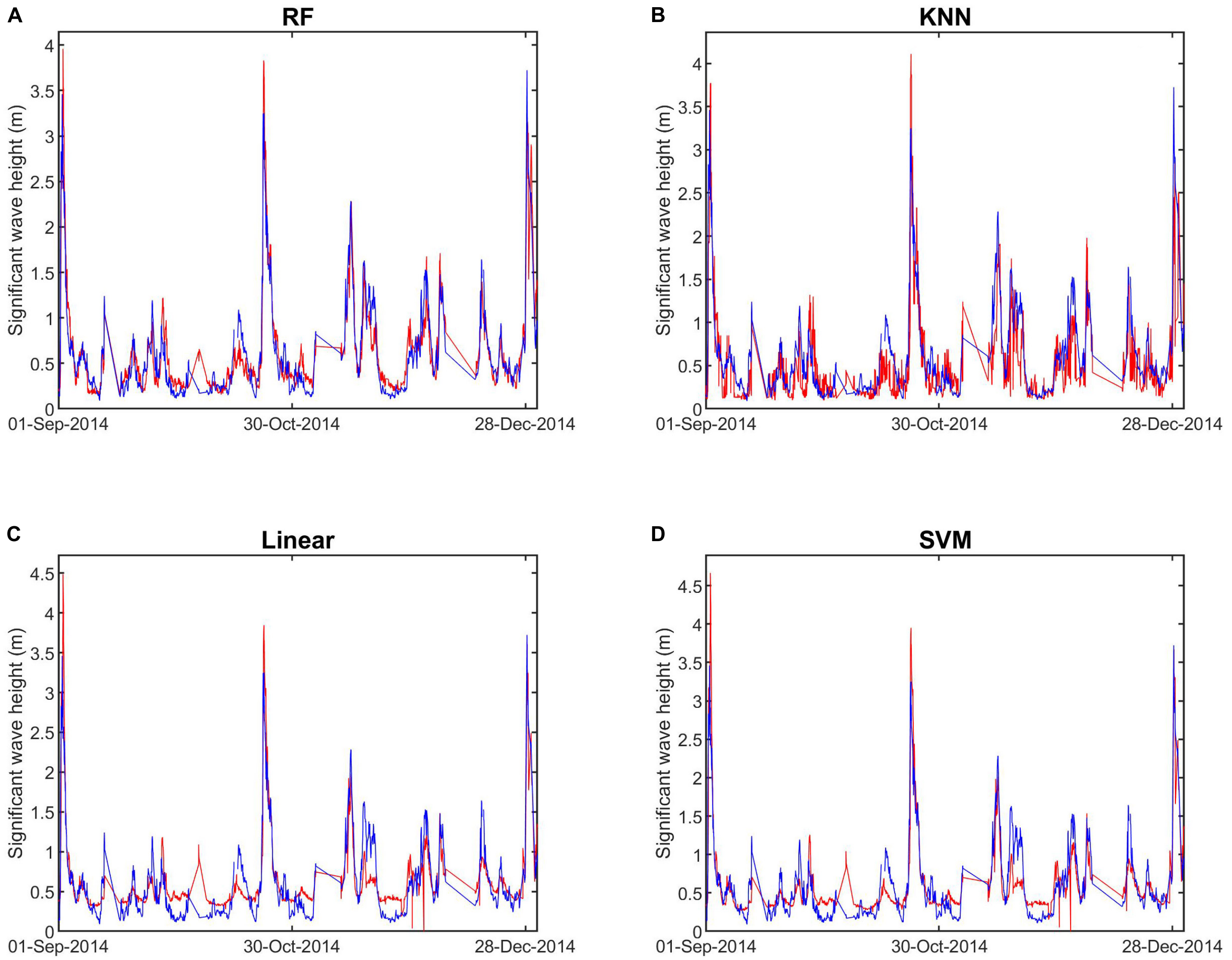


FIGURE 12 | Measured (*blue line*) and predicted (*red line*) significant wave height time series of the Cetraro buoy from 1 September to 31 December 2014. The prediction was carried out by RF regression **(A)**, KNN regression **(B)**, linear regression **(C)**, and SVM regression **(D)**.

regression simply predicts a new sample using the K-closest samples from the training set (Altman, 1992; Kuhn and Johnson, 2013). Hence, for a new input, the output is the average of the values of its K-nearest neighbors in the feature space of the training set. Such a method has been extensively used to classify remote sensing images (e.g., Li and Cheng, 2009; Noi and Kappas, 2018).

Concerning linear regressions, relationships are modeled using linear predictor functions; that is, the relationship between predictors and responses falls along a hyperplane (Kuhn and Johnson, 2013). Such linear relationships can be written as (Kuhn and Johnson, 2013):

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_n x_{in} + e_i \tag{7}$$

where y_i is the output for the i-th sample, b_0 is the estimated intercept, b_j represents the coefficient for the j-th predictor, x_{ij} represents the value of the j-th predictor for the i-th sample, and e_i represents random error for the i-th sample. Similar to the two previous machine learning methods, linear regressions have been used in many fields of Earth Sciences, such as iron mineral resource potential mapping (Mansouri et al., 2018) and catchment-level base cation weathering rates (Povak et al., 2014).

Finally, SVMs are supervised learning models for both classification and regression analysis (e.g., Drucker et al., 1997; Kuhn and Johnson, 2013). As for regression, the SVM's goal is to find a function that deviates from each training point by a value no greater than a chosen constant, and at the same time is as flat as possible (e.g., Vapnik, 2000; Kuhn and Johnson, 2013). Also, SVM has been applied in Earth Sciences, for instance to map landslide susceptibility (Reza Pourghasemi et al., 2013) and to classify remote sensing data (Jia et al., 2019).

Each of the four aforementioned techniques has its own advantages and disadvantages (e.g., Kuhn and Johnson, 2013; Yang et al., 2019). The main advantages of RF are its high accuracy and robustness to outliers and noise; also, RF parameter tuning does not have a drastic effect on performance. The disadvantages are the expensive training time and overfitting in the case of small

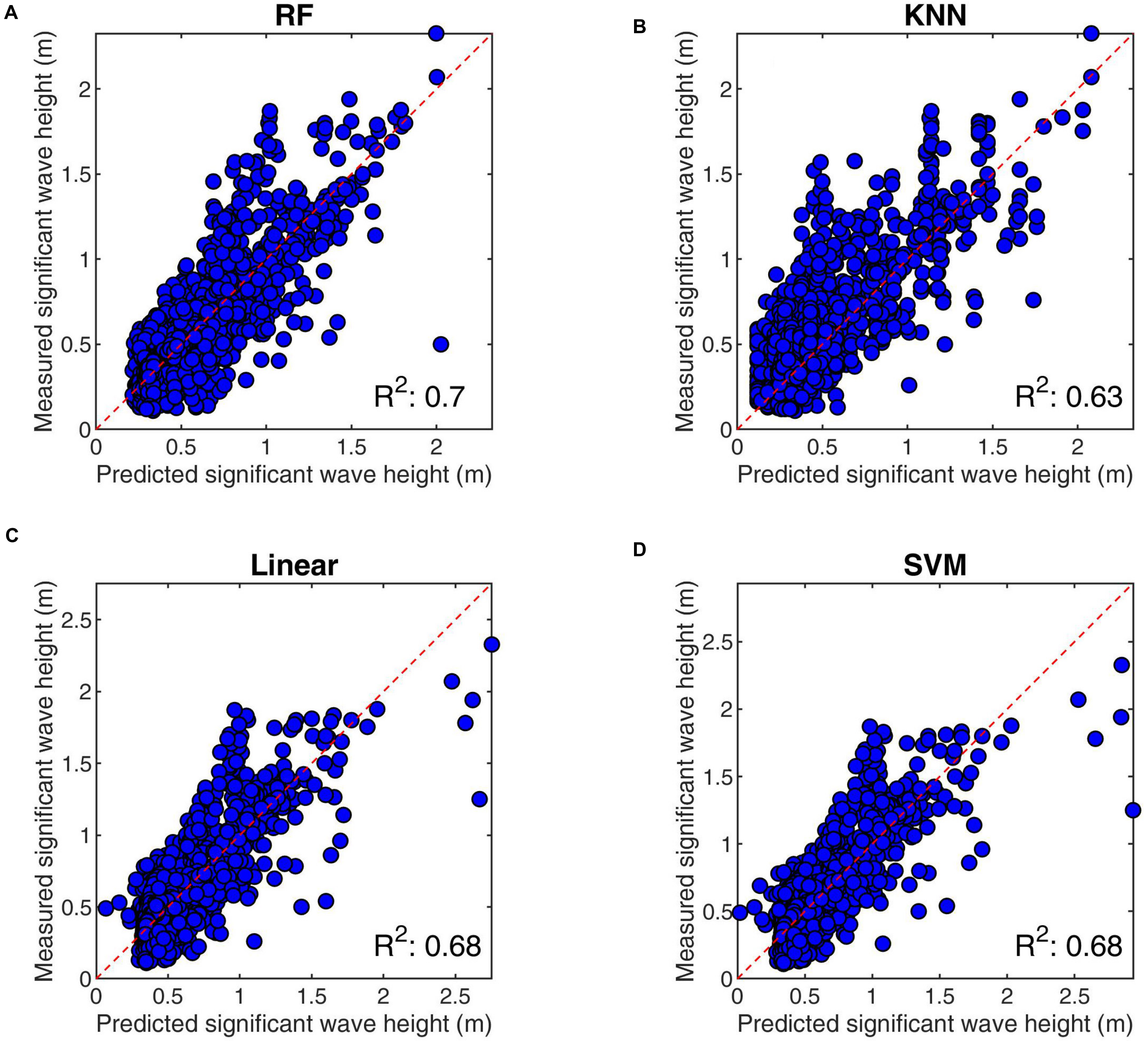


FIGURE 13 | Scatter plots showing the measured *versus* the predicted significant wave heights of the Catania buoy from 1 September to 31 December 2014. The prediction was carried out by RF regression **(A)**, KNN regression **(B)**, linear regression **(C)**, and SVM regression **(D)**. The *red dashed line* in **(A–D)** is the $y = x$ line. The value of the determination coefficient (R^2) is also reported in the *bottom right corner* of the plots.

datasets. KNN is effective and non-parametric, but it is not robust in the presence of noise and it is not easy to identify the best K value. As for linear regressions, they require short training times, and the results are easy to visualize and understand, but they are not suited to model non-linear relationships. Finally, SVMs are easy to implement and show good efficiency in training and generalization, but the tuning of parameters can be quite difficult.

For all the above-mentioned MLTs, the 18 time series of the centered and scaled seismic RMS amplitudes from January 2010 to August 2014 were used as the input, while the two time series of significant wave heights, recorded by the Catania and Cetraro buoys, were resampled by a sampling step of 1 h (the same rate as the seismic RMS amplitude time series) and considered as the output to build the regression models.

Step (c) consisted of evaluating the best MLT by carrying out the k-fold cross-validation (Kuhn and Johnson, 2013). The cross-validation implies partitioning the original input and output datasets into complementary subsets, constraining a model on one subset (called "training set"), and validating the model performance on the other subset (called "validation set"). In particular, in the performed k-fold cross-validation, the microseism amplitude and significant wave height samples are partitioned into 10 (k = 10) sets of consecutive samples. Ten models are trained by using all samples except one subset, which is used to validate the models. The parameters we used to estimate the model performance are: mean absolute error (MAE) between

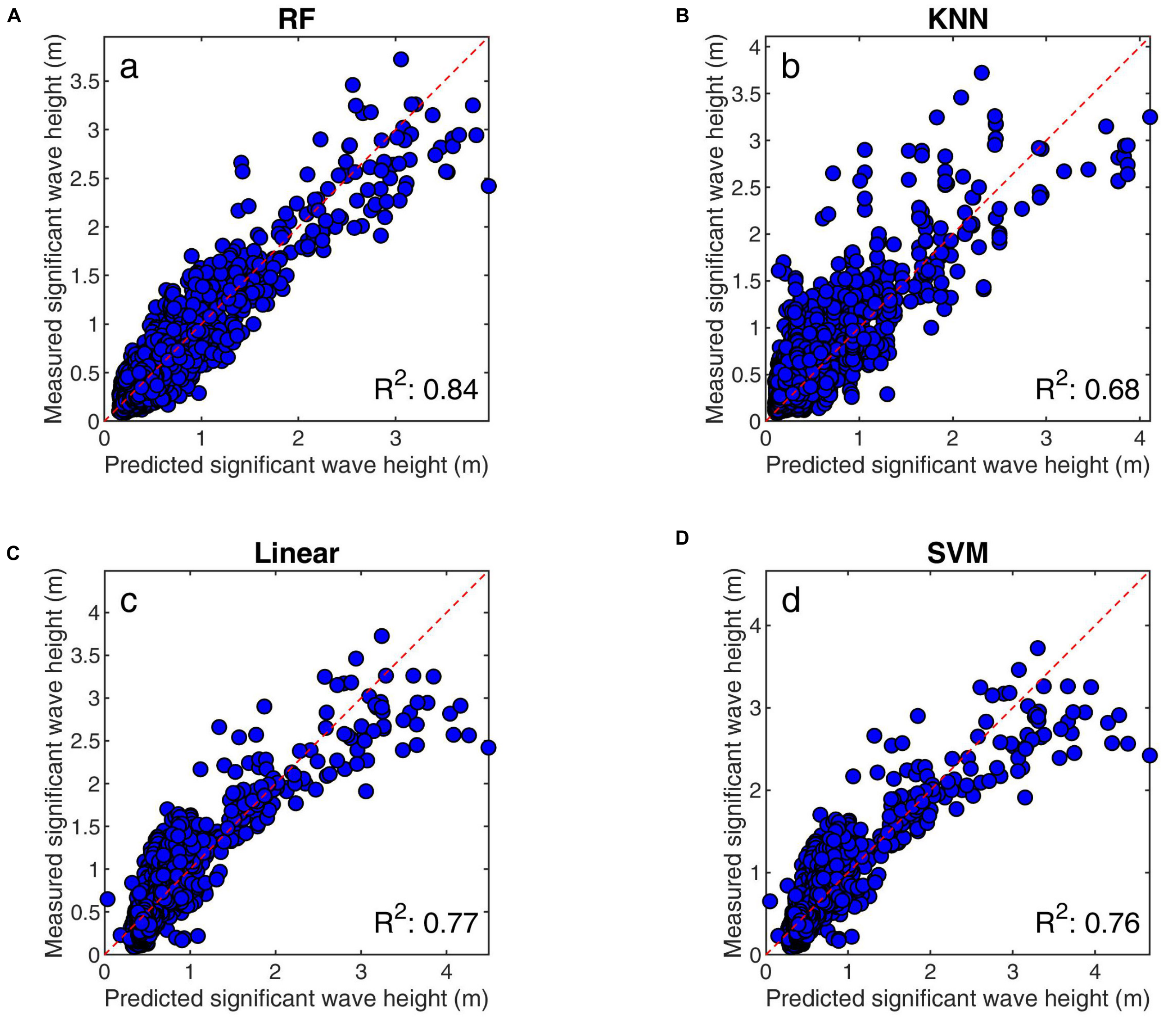


FIGURE 14 | Scatter plots showing the measured *versus* the predicted significant wave heights of the Cetraro buoy from 1 September to 31 December 2014. The prediction was carried out by RF regression **(A)**, KNN regression **(B)**, linear regression **(C)**, and SVM regression **(D)**. The *red dashed line* in **(A–D)** is the $y = x$ line. The value of the determination coefficient (R^2) is also reported in the *bottom right corner* of the plots.

the observed significant wave height and the predicted one and the corresponding standard deviation (σ_{MAE}). The former was estimated by the following equation:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} \left| y_i - x_i \right|}{n} \tag{8}$$

where x_i and y_i are the predicted and observed significant sea wave height values at the i-th time sample and n is the number of samples in x and y. The results are shown in **Figures 10A,B**.

The final model was trained with the whole dataset from January 2010–August 2014 and tested on the test set from September–December 2014 [testing step (d)]. The comparisons between the predicted and measured significant wave heights for the testing period are reported in **Figures 11–14**.

RESULTS

Microseism recorded in Eastern Sicily shows the highest amplitude in the bands 2.5–5.0 and 5–10 s (SPSM and SM, respectively) at all the considered stations (**Figures 3**, **4**). Moreover, evident amplitude seasonal modulation is shown in **Figures 4**, **5**, with maxima reached during the winter (December–February) and minima during the summer (June–August).

As for the array analysis, the summit ring of Mt. Etna seismic permanent network turned out to be effective in locating the microseism sources in the SPSM band (**Figure 6A**). During Ionian stormy days, the back azimuth values indicate the Catania Gulf, while during Tyrrhenian stormy days the back azimuth rotates, pointing north–westward. In both cases, the SPSM sources appear to be located in areas of extended shallow water depths (**Figure 7**). Concerning the median error in the back azimuth estimations obtained by the jackknife technique, it was equal to 21° and 12° for back azimuths oriented toward the Tyrrhenian and Ionian Seas, respectively. As for the apparent seismic velocity estimations, the histograms in **Figure 8** show values of ∼1.5–2.0 km/s.

Finally, MLTs have been able to reconstruct the time series of significant sea wave height on the basis of microseism data. The technique showing the best performance was RF regression (**Figures 10A,B**), allowing to get the minimum MAEs equal to 0.14 ± 0.02 m and 0.18 ± 0.05 m for the Catania (the Ionian Sea) and Cetraro (the Tyrrhenian Sea) data, respectively. It has to be underlined that the RF, linear, and SVM regressions show very similar MAE values, especially in the case of the Catania buoy. The RF approach has the advantage of easily supplying an index of predictor importance (**Figures 10C,D**), calculated by exploiting the random permutation of out-of-bag samples (Breiman, 2001). To get information on the importance of the different microseism bands in the prediction, aggregation through summation was performed (**Figures 10E,F**), showing how the SPSM band has the highest weight in reconstructing the significant wave height time series at the two buoys. In addition, aggregation through summation was performed also for the station importance and exhibited how the importance tends to decrease with increasing distance of station–buoy (**Figures 10G,H**).

Finally, the comparison between the measured and predicted significant wave height data during the testing period (September–December 2014; **Figures 11–14**) showed very similar patterns for the two time series, as also confirmed by the high values of determination coefficient equal to 0.7 and 0.84 for the Catania and Cetraro buoys, respectively, in the case of RF regressions.

DISCUSSION AND CONCLUSION

We investigated the microseism recorded close to the Eastern Sicily coasts and its relationship with the significant wave height recorded by two buoys installed in the Ionian and Tyrrhenian Seas. Concerning the microseism characterization, as measured in the seismic signals acquired worldwide (e.g., Aster et al., 2010), most of its energy is contained in the SPSM and SM bands (**Figures 3**, **4**). Also, the observed seasonal amplitude modulations (**Figures 4**, **5**) are a common feature of the microseism recorded at temperate latitudes, characterized by stormier seas during the winters (e.g., Aster et al., 2008; Stutzmann et al., 2009).

Taking into account the array analysis, performed by the seven seismic stations in **Figure 1b** by the f–k array technique in the SPSM band, we were able to obtain the slowness vector direction and, therefore, to get an idea on the locations of the microseism source in the SPSM band. It was observed that the SPSM sources appear to be located in areas of extended shallow water depths: the Catania Gulf and a part of the Northern Sicily coastlines (**Figure 7**).

The array analysis results are in agreement with Chen et al. (2011), who analyzed microseism data collected in Taiwan and showed how a stronger excitation in SPSM takes place in the narrow Taiwan Strait where the water depth is very shallow, while the excitations are relatively weak in the eastern offshore area, an open sea with water depth increasing rapidly off the coast. Although Juretzek and Hadziioannou (2017) focused on a different frequency band (PM), they also constrained the source locations of the microseism recorded in Europe in regions with extended shallow water areas, that is, Norwegian and Scottish coasts.

It has to be noted that the error associated with the microseism source locations is higher in the case of the Northern Sicily coastlines compared to the Catania Gulf. It derives from both the higher back azimuth error (21° for the Tyrrhenian Sea *versus* 12° for the Ionian Sea) as well as from the longer distance array–Northern Sicily coastlines (∼45 km) compared to the distance array–Catania Gulf (∼20 km).

The apparent seismic velocity estimations of 1.5–2.0 km/s in the SPSM band (**Figure 8**) are in agreement with the Rayleigh wave velocity calculated by using beamforming analysis, applied on the ambient seismic noise in New Zealand, by Brooks et al. (2009), as well as with the results obtained from investigating the seismic noise in the northeast of the Netherlands by Kimman et al. (2012). In addition, Rivet et al. (2015) also estimated comparable velocities (of 1.5 km/s at 1 Hz and 2.0 km/s at 0.5 Hz) by using a time–frequency analysis to measure the group velocity of Rayleigh wave on noise cross-correlation.

Finally, we propose an innovative method, based on up-to-date MLTs, able to reconstruct the time series of significant wave height by using microseism recorded in different period bands by distinct seismic stations. Such a method allows to reliably compute the significant wave height in two locations, coinciding with the two buoys in the Ionian and Tyrrhenian Seas, with fairly low error (MAE equal to ∼0.14 m for the Catania buoy and ∼0.18 m for the Cetraro buoy; **Figures 10A,B**). In particular, the MLT which showed the best performance was the RF regression. This can be related to several factors, such as: (i) the performance of the RF regression is not much affected by parameter selection (e.g., Li et al., 2011; Kuhn and Johnson, 2013); (ii) by making use of an ensemble of decision trees, RF regression does not overfit with respect to the source data (e.g., Li et al., 2011); and (iii) RF shows robustness to outliers and noise (Breiman, 2001). Finally, compared to linear regressions, RF regression is able to deal with non-linear relationships between the input and output. Indeed, according to Essen et al. (2003) and Craig et al. (2016), the relationship linking microseism amplitude and significant wave height is likely to be non-linear.

Focusing on the comparison between the highest measured and predicted (by RF regression) significant wave height data during the testing period, it is possible to note a slight underestimation and overestimation of the predicted values compared to the measured ones in the Catania and Cetraro cases, respectively (**Figures 13A**, **14A**). These different behaviors could be related to the different distances between the seismic stations and the buoys.

Although buoys are considered the most used and reliable instruments for *in situ* measurements of sea waves (Orasi et al., 2018), the high maintenance costs, together with the recurring damages and then lack of data, make the proposed microseism-based method a valid complementary tool for the monitoring of the sea state. Furthermore, once the regression model has been determined and if the seismic data are available, such a method could allow reconstructing the time series of sea wave height during periods prior to the buoy installation, with wide applications in many fields, first of all climate studies.

The RF regression also provides an index of importance of the distinct predictor variables, which are the seismic RMS amplitude time series. The aggregated importance of the different frequency bands exhibits how the SPSM band contains most of the information for the buoy data reconstruction (**Figures 10E,F**). According to the literature (e.g., Bromirski et al., 2005; Chen et al., 2011; Gualtieri et al., 2015), such a microseism band, characterized by high frequencies and then by quick attenuation with distance, is mostly generated by sources located in relatively shallow water close to the shelf break, close to the seismic stations. Such sources are likely related to local nearshore non-linear wave–wave interaction (e.g., Bromirski et al., 2005). This is in agreement with the location of the considered buoys, close to the coastlines, in shallow water conditions (90 and 100 m for Catania and Cetraro, respectively; Bencivenga et al., 2012). Both PM and SM turned out to have a much smaller importance for the buoy data reconstruction. Indeed, as for PM, its dominant source regions can be located thousands of kilometers away from the seismic stations (Gualtieri et al., 2019). Concerning SM, it has been shown how it can also have pelagic sources in deep ocean (e.g., Chevrot et al., 2007; Kedar et al., 2008).

In addition, the difference in the predictors with the maximum importance for the two buoys (EPOZ-SPSM for the Catania buoy and MSRU-SPSM for the Cetraro buoy; **Figures 10C,D**) reflects the different locations of the seismic stations. Indeed, EPOZ is very close to the coastline of the Ionian Sea, where the Catania buoy is installed, while MSRU is placed nearby the Tyrrhenian Sea, where the Cetraro buoy is located (**Figures 1**, **10G,H**). Hence, the closer the seismic station is to the sea, the more information concerning the sea state are contained in the recorded microseism. From a future perspective, this finding is important to build an experimental monitoring system of the sea conditions (mainly in terms of significant wave height) based on microseism recordings.

DATA AVAILABILITY STATEMENT

Seismic data are provided by the Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo-Sezione di Catania. Buoy data are provided by the Istituto Superiore per la Protezione e la Ricerca Ambientale (http://dati.isprambiente.it/dataset/ron-rete-ondametrica-nazionale/). Bathymetric data shown in **Figures 1**, **7** come from https://portal.emodnet-bathymetry.eu/. Hindcast maps of significant wave height shown in **Figure 7** are provided by http://marine.copernicus.eu/services-portfolio/access-to-products/.

AUTHOR CONTRIBUTIONS

SM and AC initiated the concepts. SM, AC, GD, and SG performed the seismic analyses. FC, SM, and AC performed the machine learning investigations. MF, GN, AO, and MP analyzed buoy data. All the authors wrote the manuscript and contributed to the interpretation of results.

FUNDING

This research was partially funded by Programma Nazionale di Ricerca in Antartide, grant no. PNRA14_00011, called ICE-VOLC project ("MultiparametrIC Experiment at antarctica VOLCanoes: data from volcano and cryosphere–ocean–atmosphere dynamics," www.icevolc-project.com).

ACKNOWLEDGMENTS

We thank the two reviewers and the Associate Editor for their critical reading of the manuscript and constructive comments, which helped us to improve the manuscript. We are indebted to the technicians of the Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo-Sezione di Catania for enabling the acquisition of seismic data. This study has been conducted using E.U. Copernicus Marine Service Information. The array *f–k* analysis was performed by "Seizmo – Passive Seismology Toolbox," version 0.6.16 Otgontenger 11-Jul-2014 (http://epsc.wustl.edu/~ggeuler/codes/m/seizmo/).

REFERENCES

Almendros, J., Ibáñez, J. M., Alguacil, G., and Del Pezzo, E. (2002). Array analysis using circular-wave-front geometry: an application to locate the nearby seismo-volcanic source. *Geophys. J. Int.* 136, 159–170. doi: 10.1046/j.1365-246x.1999.00699.x

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* 46, 175–185. doi: 10.1080/00031305.1992.10475879

Ardhuin, F., Balanche, A., Stutzmann, E., and Obrebski, M. (2012). From seismic noise to ocean wave parameters: general methods and validation. *J. Geophys. Res.* 117:C05002. doi: 10.1029/2011JC007449

Ardhuin, F., Gualtieri, L., and Stutzmann, E. (2015). How ocean waves rock the Earth: two mechanisms explain microseisms with periods 3 to 300 s. *Geophys. Res. Lett.* 42, 765–772. doi: 10.1002/2014GL062782

Aster, R. C., McNamara, D. E., and Bromirski, P. D. (2008). Multidecadal climate-induced variability in microseisms. *Seismol. Res. Lett.* 79, 194–202. doi: 10.1785/gssrl.79.2.194

Aster, R. C., McNamara, D. E., and Bromirski, P. D. (2010). Global trends in extremal microseism intensity. *Geophys. Res. Lett.* 37:L14303. doi: 10.1029/2010GL043472

Bencivenga, M., Nardone, G., Ruggiero, F., and Calore, D. (2012). The italian data buoy network (RON). *Adv. Fluid Mech.* 74, 321–332. doi: 10.2495/AFM120291

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Bromirski, P. D., and Duennebier, F. K. (2002). The near-coastal microseism spectrum: spatial and temporal wave climate relationships. *J. Geophys. Res.* 107:2166. doi: 10.1029/2001JB000265

Bromirski, P. D., Duennebier, F. K., and Stephen, R. A. (2005). Mid-ocean microseisms. *Geochem. Geophys. Geosyst.* 6:Q04009. doi: 10.1029/2004GC000768

Bromirski, P. D., Flick, R. E., and Graham, N. (1999). Ocean wave height determined from inland seismometer data: implications for investigating wave climate changes in the NE Pacific. *J. Geophys. Res.* 104, 20753–20766. doi: 10.1029/1999JC900156

Bromirski, P. D., Stephen, R. A., and Gerstoft, P. (2013). Are deep-ocean-generated surface-wave microseisms observed on land? *J. Geophys. Res. Solid Earth* 118, 3610–3629. doi: 10.1002/jgrb.50268

Brooks, L. A., Townend, J., Gerstoft, P., Bannister, S., and Carter, L. (2009). Fundamental and higher-mode Rayleigh wave characteristics of ambient seismic noise in New Zealand. *Geophys. Res. Lett.* 36:L23303. doi: 10.1029/2009GL040434

Cannata, A., Cannavò, F., Moschella, S., Gresta, S., and Spina, L. (2019). Exploring the link between microseism and sea ice in Antarctica by using machine learning. *Sci. Rep.* 9:13050. doi: 10.1038/s41598-019-49586-z

Cannata, A., Di Grazia, G., Montalto, P., Ferrari, F., Nunnari, G., Patanè, D., et al. (2010). New insights into banded tremor from the 2008–2009 Mount Etna eruption. *J. Geophys. Res.* 115:B12318. doi: 10.1029/2009JB007120

Capon, J. (1973). Signal processing and frequency-wavenumber spectrum analysis for a large aperture seismic array. *Methods Comput. Phys.* 13, 1–59. doi: 10.1016/b978-0-12-460813-9.50007-2

Chen, Y.-N., Gung, Y., You, S.-H., Hung, S.-H., Chiao, L.-Y., Huang, T.-Y., et al. (2011). Characteristics of short period secondary microseisms (SPSM) in Taiwan: the influence of shallow ocean strait on SPSM. *Geophys. Res. Lett.* 38:L04305. doi: 10.1029/2010GL046290

Chevrot, S., Sylvander, M., Benahmed, S., Ponsolles, C., Lefevre, J. M., and Paradis, D. (2007). Source locations of secondary microseisms in western Europe: evidence for both coastal and pelagic sources. *J. Geophys. Res.* 112:B11301. doi: 10.1029/2007JB005059

Craig, D., Bean, C., Lokmer, I., and Möllhoff, M. (2016). Correlation of wavefield-separated ocean-generated microseisms with North Atlantic Source regions. *Bull. Seism. Soc. Am.* 106, 1002–1010. doi: 10.1785/0120150181

De Caro, M., Monna, S., Frugoni, F., Beranzoli, L., and Favali, P. (2014). Seafloor seismic noise at Central Eastern Mediterranean sites. *Seismol. Res. Lett.* 85, 1019–1033. doi: 10.1785/0220130203

Drucker, H., Burges, C., Kaufman, L., Smola, A., and Vapnik, V. (1997). Support vector regression machines. *Adv. Neural Inform. Proc. Syst.* 28, 779–784.

Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia, PA: Soc. for Ind. and Appl. Math.

EMODnet Bathymetry Consortium (2018). “EMODnet Digital Bathymetry (DTM 2018)”, in *EMODnet Bathymetry Consortium*. Available at: https://doi.org/10.12770/18ff0d48-b203-4a65-94a9-5fd8b0ec35f6

Essen, H.-H., Krüger, F., Dahm, T., and Grevemeyer, I. (2003). On the generation of secondary microseisms observed in northern and central Europe. *J. Geophys. Res. Space Phys.* 2003, 1–15. doi: 10.1029/2002JB002338

Ferretti, G., Barani, S., Scafidi, D., Capello, M., Cutroneo, L., Vagge, G., et al. (2018). Near real-time monitoring of significant sea wave height through microseism recordings: an application in the Ligurian Sea (Italy). *Ocean Coast. Manag.* 165, 185–194. doi: 10.1016/j.ocecoaman.2018.08.023

Ferretti, G., Zunino, A., Scafidi, D., Barani, S., and Spallarossa, D. (2013). On microseisms recorded near the Ligurian coast (Italy) and their relationship with sea wave height. *Geophys. J. Int.* 194, 524–533. doi: 10.1093/gji/ggt114

Grevemeyer, I., Herber, R., and Essen, H. (2000). Microseismological evidence for a changing wave climate in the northeast Atlantic Ocean. *Nature* 408, 349–352. doi: 10.1038/35042558

Gualtieri, L., Stutzmann, E., Capdeville, Y., Ardhuin, F., Schimmel, M., Mangeney, A., et al. (2013). Modelling secondary microseismic noise by normal mode summation. *Geophys. J. Int.* 193, 1732–1745. doi: 10.1093/gji/ggt090

Gualtieri, L., Stutzmann, E., Capdeville, Y., Farra, V., Mangeney, A., and Morelli, A. (2015). On the shaping factors of the secondary microseismic wavefield. *J. Geophys. Res.* 120, 6241–6262. doi: 10.1002/2015jb012157

Gualtieri, L., Stutzmann, E., Juretzek, C., Hadziioannou, C., and Ardhuin, F. (2019). Global scale analysis and modelling of primary microseisms. *Geophys. J. Int.* 218, 560–572. doi: 10.1093/gji/ggz161

Hasselmann, K. A. (1963). Statistical analysis of the generation of microseisms. *Rev. Geophys. Space Phys.* 1, 177–210. doi: 10.1029/RG001i002p00177

Haubrich, R. A., and McCamy, K. (1969). Microseisms: coastal and pelagic sources. *Rev. Geophys. Space Phys.* 7, 539–571. doi: 10.2183/pjab.93.026

Hirn, A., Nercessian, A., Sapin, M., Ferrucci, F., and Wittlinger, G. (1991). Seismic heterogeneity of Mt. Etna: structure and activity. *Geophys. J. Int.* 105, 139–153. doi: 10.1111/j.1365-246x.1991.tb03450.x

Ho, T. (1995). “Random Decision Forests,” in *Proceedings of the Third Int’l Conf. Document Analysis and Recognition*, Montreal, 278–282.

Ho, T. (1998). The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.* 13, 340–354.

Jia, K., Liu, J., Tu, Y., Li, Q., Sun, Z., Wei, X., et al. (2019). Land use and land cover classification using Chinese GF-2 multispectral data in a region of the North China Plain. *Front. Earth Sci.* 13, 327–335. doi: 10.1007/s11707-018-0734-8

Jiao, P., and Alavi, A. H. (2019). Artificial intelligence in seismology: advent, performance and future trends. *Geosci. Front.* 19:30198. doi: 10.1016/j.gsf.2019.10.004

Juretzek, C., and Hadziioannou, C. (2017). Linking source region and ocean wave parameters with the observed primary microseismic noise. *Geophys. J. Int.* 211, 1640–1654. doi: 10.1093/gji/ggx388

Kedar, S., Longuet-Higgins, M. S., Graham, F. W. N., Clayton, R., and Jones, C. (2008). The origin of deep ocean microseisms in the North Atlantic Ocean. *Proc. R. Soc. A* 464, 777–793. doi: 10.1098/rspa.2007.0277

Kimman, W. P., Campman, X., and Trampert, J. (2012). Characteristics of seismic noise: fundamental and higher mode energy observed in the Northeast of the Netherlands. *Bull. Seismol. Soc. Am.* 102, 1388–1399. doi: 10.1785/0120110069

Kirkwood, C., Cave, M., Beamish, D., Grebby, S., and Ferreira, A. (2016). A machine learning approach to geochemical mapping. *J. Geochem. Explorat.* 167, 49–61. doi: 10.1016/j.gexplo.2016.05.003

Kong, Q., Allen, R. M., Schreier, L., and Kwon, Y.-W. (2016). MyShake: a smartphone seismic network for earthquake early warning and beyond. *Sci. Adv.* 2:e1501055. doi: 10.1126/sciadv.1501055

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., and Gerstoft, P. (2018). Machine learning in seismology: turning data into insights. *Seismol. Res. Lett.* 90, 3–14. doi: 10.1785/0220180259

Kuhn, M., and Johnson, K. (2013). *Applied Predictive Modeling*. New York: Springer.

Kuhn, S., Cracknell, M. J., and Reading, A. M. (2018). Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: a demonstration study from the Eastern Goldfields of Australia. *Geophysics* 82, 183–193. doi: 10.1190/geo2017-0590.1

Lepore, S., and Grad, M. (2018). Analysis of the primary and secondary microseisms in the wavefield of the ambient noise recorded in northern Poland. *Acta Geophys.* 66, 915–929. doi: 10.1007/s11600-018-0194-2

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. Softw.* 26, 1647–1659. doi: 10.1016/j.envsoft.2011.07.004

Li, Y., and Cheng, B. (2009). “An improved K-Nearest neighbor algorithm and its application to high resolution remote sensing image classification,” in *Proceedings of the IEEE Geoinformatics Conference*, Fairfax, VA, 1–4.

Liaw, A., and Wiener, M. (2002). Classification and regression by random Forest. *R News* 2, 18–22.

Longuet-Higgins, M. S. (1950). A theory of the origin of microseisms. *Philos. Trans. R. Soc. Lond. Ser. A* 243, 1–35. doi: 10.1098/rsta.1950.0012

Mansouri, E., Feizi, F., Rad, A. J., and Arian, M. (2018). Remote-sensing data processing with the multivariate regression analysis method for iron mineral resource potential mapping: a case study in the Sarvian area, central Iran. *Solid Earth* 9:373. doi: 10.5194/se-9-373-2018

Noi, P. T., and Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors* 18:18. doi: 10.3390/s18010018

Oliver, J., and Page, R. (1963). Concurrent storms of long and ultralong period microseisms. *Bull. Seismol. Soc. Am.* 53, 15–26.

Orasi, A., Picone, M., Drago, A., Capodici, F., Gauci, A., Nardone, G., et al. (2018). HF radar for wind waves measurements in the Malta-Sicily Channel. *Measurement* 128, 446–454. doi: 10.1016/j.measurement.2018.06.060

Patanè, D., Privitera, E., Ferrucci, F., and Gresta, S. (1994). Seismic activity leading to the 1991-1993 eruption of Mt. Etna and its tectonic implications. *Acta Vulcanol.* 4, 47–55.

Povak, N. A., Hessburg, P. F., McDonnell, T. C., Reynolds, K. M., Sullivan, T. J., Salter, R. B., et al. (2014). Machine learning and linear regression models to predict catchment-level base cation weathering rates across the southern Appalachian Mountain region, USA. *Water Resour. Res.* 50, 2798–2814. doi: 10.1002/2013wr014203

Pratt, M. J., Wiens, D. A., Winberry, J. P., Anandakrishnan, S., and Euler, G. G. (2017). Implications of Sea ice on Southern Ocean microseisms detected by a seismic array in West Antarctica. *Geophys. J. Int.* 209, 492–507. doi: 10.1093/gji/ggx007

Reza Pourghasemi, H., Jirandeh, A. G., Pradhan, B., Xu, C., and Gokceoglu, C. (2013). Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province. *Iran. J. Earth Syst. Sci.* 122, 349–369. doi: 10.1007/s12040-013-0282-2

Rivet, D., Campillo, M., Sanchez-Sesma, F., Shapiro, N. M., and Singh, S. K. (2015). Identification of surface wave higher modes using a methodology based on seismic noise and coda waves. *Geophys. J. Int.* 203, 856–868. doi: 10.1093/gji/ggv339

Rost, S., and Thomas, C. (2002). Array seismology: methods and applications. *Rev. Geophys.* 40:1008. doi: 10.1029/2000RG000100

Schweitzer, J., Fyen, J., Mykkeltveit, S., Gibbons, S. J., Pirli, M., Kühn, D., et al. (2012). "Seismic arrays," in *IASPEI New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, ed. P. Bormann (Potsdam: IASPEI). doi: 10.2312/GFZ.NMSOP-2_ch9

Steele, K. E., and Mettlach, T. (1993). "NDBC wave data – current and planned," in *Measurement and Analysis - Proceedings of the Second International Symposium* (Reston: ASCE), 198–207.

Stutzmann, E., Schimmel, M., Patau, G., and Maggi, A. (2009). Global climate imprint on seismic noise. *Geochem. Geophys. Geosyst.* 10:Q11004. doi: 10.1029/2009GC002619

Tanimoto, T., Heki, K., and Artru-Lambin, J. (2015). "Interaction of Solid Earth, Atmosphere, and Ionosphere," in *Treatise in Geophysics*, 2nd Edn, ed. G. Schubert (Amsterdam: Elsevier), 421–444. doi: 10.1016/b978-044452748-6.00075-4

Trnkoczy, A. (2012). "Understanding and parameter setting of STA/LTA trigger algorithm," in *IASPEI New Manual of Seismological Observatory Practice 2 (NMSOP-2)*, ed. P. Bormann (Potsdam: IASPEI), 1–20. doi: 10.2312/GFZ.NMSOP-2_IS_8.1

Tsai, V. C., and McNamara, D. E. (2011). Quantifying the influence of sea ice on ocean microseism using observations from the Bering Sea, Alaska. *Geophys. Res. Lett.* 38:L22502. doi: 10.1029/2011GL049791

Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory*. New York: Springer.

Welch, P. D. (1967). The use of Fast Fourier Transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 15, 70–73. doi: 10.1109/TAU.1967.1161901

Wiszniowski, J., Plesiewicz, B. M., and Trojanowski, J. (2014). Application of real time recurrent neural network for detection of small natural earthquakes in Poland. *Acta Geophys.* 62, 469–485. doi: 10.2478/s11600-013-0140-2

Yang, X. S., Sherratt, S., Dey, N., and Joshi, A. (2019). *Fourth International Congress on Information and Communication Technology: ICICT 2019*. London: Springer Nature.

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.

frontiers
in Earth Science

ORIGINAL RESEARCH
published: 28 May 2020
doi: 10.3389/feart.2020.00169

Uncertainty in Detection of Volcanic Activity Using Infrasound Arrays: Examples From Mt. Etna, Italy

Silvio De Angelis [1], Matthew M. Haney [2], John J. Lyons [2], Aaron Wech [2], David Fee [3], Alejandro Diaz-Moreno [4] and Luciano Zuccarello [5,6]*

[1] School of Environmental Sciences, University of Liverpool, Liverpool, United Kingdom, [2] Alaska Volcano Observatory, U.S. Geological Survey Volcano Science Center, Anchorage, AK, United States, [3] Alaska Volcano Observatory, Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK, United States, [4] School of Geography, Politics and Sociology, Newcastle University, Newcastle upon Tyne, United Kingdom, [5] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Pisa, Italy, [6] Departamento de Fisica Teorica y del Cosmos, University of Granada, Granada, Spain

The injection of gas and pyroclastic material from volcanic vents into the atmosphere is a prolific source of acoustic waves. Infrasound arrays offer efficient, cost-effective, and near real-time solutions to track the rate and intensity of surface activity at volcanoes. Here, we present a simple framework for the analysis of acoustic array data, based on least-squares beamforming, that allows to evaluate the direction and speed of propagation of acoustic waves between source and array. The algorithms include a new and computationally efficient approach for quantitative assessment of the uncertainty on array measurements based on error propagation theory. We apply the algorithms to new data collected by two 6-element infrasound arrays deployed at Mt. Etna during the period July–August 2019. Our results demonstrate that the use of two infrasound arrays allowed detecting and tracking acoustic sources from multiple craters and active vents associated with degassing and ash-rich explosions, vigorous and frequent Strombolian activity, opening of new eruptive fractures and emplacement of lava flows. Finally, we discuss the potential use of metrics based on infrasound array analyses to inform eruption monitoring operations and early warning at volcanoes characterized by episodic intensification of activity.

Keywords: volcano infrasound, infrasound arrays, volcanic degassing, ash explosions, lava flow, Mt. Etna

OPEN ACCESS

Edited by:
Andrea Cannata,
University of Catania, Italy

Reviewed by:
Stephen Arrowsmith,
Southern Methodist University,
United States
Stephen R. McNutt,
University of South Florida,
United States

****Correspondence:***
Silvio De Angelis
silvioda@liverpool.ac.uk

Specialty section:
This article was submitted to
Volcanology,
a section of the journal
Frontiers in Earth Science

Received: *13 February 2020*
Accepted: *01 May 2020*
Published: *28 May 2020*

Citation:
De Angelis S, Haney MM, Lyons JJ, Wech A, Fee D, Diaz-Moreno A and Zuccarello L (2020) Uncertainty in Detection of Volcanic Activity Using Infrasound Arrays: Examples From Mt. Etna, Italy. Front. Earth Sci. 8:169. doi: 10.3389/feart.2020.00169

1. INTRODUCTION

Over the past two decades infrasound has become an increasingly popular tool to monitor volcanoes (e.g., Fee and Matoza, 2013; McNutt et al., 2015). Because of its ability to detect and discriminate shallow and sub-aerial volcanic activity (Matoza et al., 2019) infrasound is a desirable complement to seismology for monitoring unrest, and to detect and track the evolution of eruptions in real- or near real-time (e.g., Ripepe et al., 2009, 2018; Cannata et al., 2013; Coombs et al., 2018) that is over time scales of the order of few seconds to few minutes. Acoustic waves are generated when the atmosphere is perturbed from equilibrium (e.g., Garcés et al., 2009). At volcanoes, small and large explosions, gas-and-ash jets and plumes, sector collapses, rockfalls, pyroclastic flows, and lahars, are likely to generate infrasound over a wide range of magnitudes, frequencies, and durations (Johnson and Ripepe, 2011; Fee and Matoza, 2013; Johnson and Palma, 2015; McNutt et al., 2015; Matoza and Fee, 2018). Infrasound is recorded by band-sensitive microphones at different scales

(Fee and Matoza, 2013) from local (<20 km) to regional (several 10s to few 100s of km), and even global (several 100s to 1,000s of km).

At the local scale, microphones are deployed either as individual sensors within distributed networks or as small-aperture arrays (i.e., tight clusters of three or more instruments at 50–150 m from one another) at distances of between few hundreds of meters to within 20 km of an active volcanic vent. Data from distributed networks have traditionally been used for absolute source location and event discrimination (e.g., Cannata et al., 2009), and to evaluate eruption source parameters (e.g., Caplan-Auerbach et al., 2010; Fee et al., 2017; De Angelis et al., 2019; Diaz-Moreno et al., 2019). On the other hand, the close spacing between sensors deployed as a small-aperture array allows detection of coherent infrasound even in low signal-to-noise ratio (SNR) conditions, characterization of the direction-of-arrival (DOA), and the apparent speed of propagation of acoustic waves. Notably, the use of infrasound arrays even at distances of a few tens of kilometers from active vents provides robust and efficient remote detection of eruptive activity and source discrimination (e.g., Fee et al., 2010), thus reducing risks for observatory personnel during field campaigns. Johnson (2004) and Ripepe et al. (2007) discussed applications of infrasound array methods to track degassing and eruptive activity from multiple active vents at Stromboli volcano (Italy). Ripepe et al. (2018) reported on over a decade of acoustic monitoring at Mt. Etna (Italy) with two small aperture infrasound arrays, demonstrating the ability to detect eruption onset in real-time, discriminate source position, and dispatch rapid notifications of the ongoing activity to local civil protection authorities. Matoza et al. (2007) analyzed array data during activity at Mt. St. Helens (USA) from October 2004 to March 2005; they discussed the key role of infrasound in separating surface from deeper processes and in identifying the timing and assessing the magnitude of eruptive events.

Previous studies focussing on (local scale) infrasound array applications to volcano monitoring do not provide quantitative estimates of the true uncertainty associated with their measurements. In this manuscript we tackle this issue presenting a simple framework for infrasound array processing based on least-squares beamforming (Olson and Szuberla, 2005; Haney et al., 2018) and introducing a new scheme for quantitative assessment of the uncertainty on estimates of DOA and apparent velocity. We test the method on data collected by two arrays deployed at Mt. Etna (Italy) during the summer of 2019 (**Figure 1**). We describe the array processing workflow, and show how infrasound data can be used to discriminate and track volcanic activity, from background degassing to individual ash-rich explosions, persistent Strombolian activity and lava effusion.

2. ACTIVITY AT MT. ETNA: JULY–AUGUST 2019

Mt. Etna, Italy, is one of the most active volcanoes in the world. Several eruptive episodes have taken place in the past two decades, at its summit, from the North East (NEC), Voragine (VOR), Bocca Nuova (BN), and New South East (NSEC) craters. Since 2011, more than 50 effusive events have taken place in the area, which is visited by thousands of tourists every year (e.g., De Beni et al., 2019; Sciotto et al., 2019). During the last decade lava fountaining has frequently been observed from the NSEC, as well as shorter episodes of Strombolian activity from BN, VOR, and the NEC (Sciotto et al., 2019). Eruption at Etna occasionally evolves into episodes of more intense activity referred to as paroxysms that is, a significant increment in the rate and intensity of explosions from one or multiple active vents accompanied by emplacement of lava flows and/or generation of significant ash plumes. Activity at Mt. Etna during July–August 2019 was marked by the occurrence of two paroxysms on 18 and 27–28 July, accompanied by the emplacement of lava flows from lateral vents in the NSEC area (**Figure 1**). The weeks preceding each paroxysm were characterized by both sustained and explosive degassing occasionally punctuated by ash-rich explosions (**Figures 2A,B**). The Istituto Nazionale di Geofisica e Vulcanologia (INGV) reported intense degassing and observations of four large explosions from the NEC on 2, 3, 8, and 13 July (**Figure 2B**). Degassing activity from VOR was low-to-moderate, whilst BN produced deep intra-crater gas explosions with minor amounts of ash (**Figure 2A**). Activity at the NSEC increased gradually from vigorous degassing to ash-rich Strombolian activity throughout the first week of July; degassing-only activity resumed on 7 July and escalated again into strong and nearly continuous Strombolian explosions on 18 July. Explosion rates varied from one every 1–2 min to several per minute, eventually leading to the opening of a fracture and subsequent emplacement of a lava flow on the NE flank of the NSEC (23:09 UTC on 18 July). During the night on 19 July, Strombolian activity shifted from NSEC to NEC, and gradually decreased until it completely halted—including lava flow effusion—in the morning of 20 July.

On 25 July, a new phase of eruption began at the NSEC with the onset of Strombolian activity, transitioning into nearly continuous lava fountaining early in the morning of 27 July. This activity was accompanied by sustained ash emissions (**Figures 2C,D**) just a few hours before the opening of two new vents on the southern flank of the NSEC from which two new lava flows developed (**Figure 1**). A strong explosion occurred on July 27 at 12:21 UTC, accompanied by an ash plume that reached nearly 4 km above the vent (**Figure 2C**). Strombolian activity ceased on 28 July at 03:40 UTC, while lava flows continued until late that day. The remainder of our deployment period was characterized by background levels of degassing from NSEC, VOR, and BN, while NEC continued producing episodic ash explosions during August.

INGV reported no major changes in deformation during the 2 months of our temporary deployment, while daily gas emissions increased notably during 15–29 July and then returned to background levels. Seismic tremor fluctuated with marked peaks on 6 July, coincident with the increase in Strombolian activity at NSEC. Seismic tremor remained stable, at high levels, during both paroxysms, and eventually returned to background during August. Finally, infrasound locations provided by INGV successfully identified the first paroxysm on July 18, but high

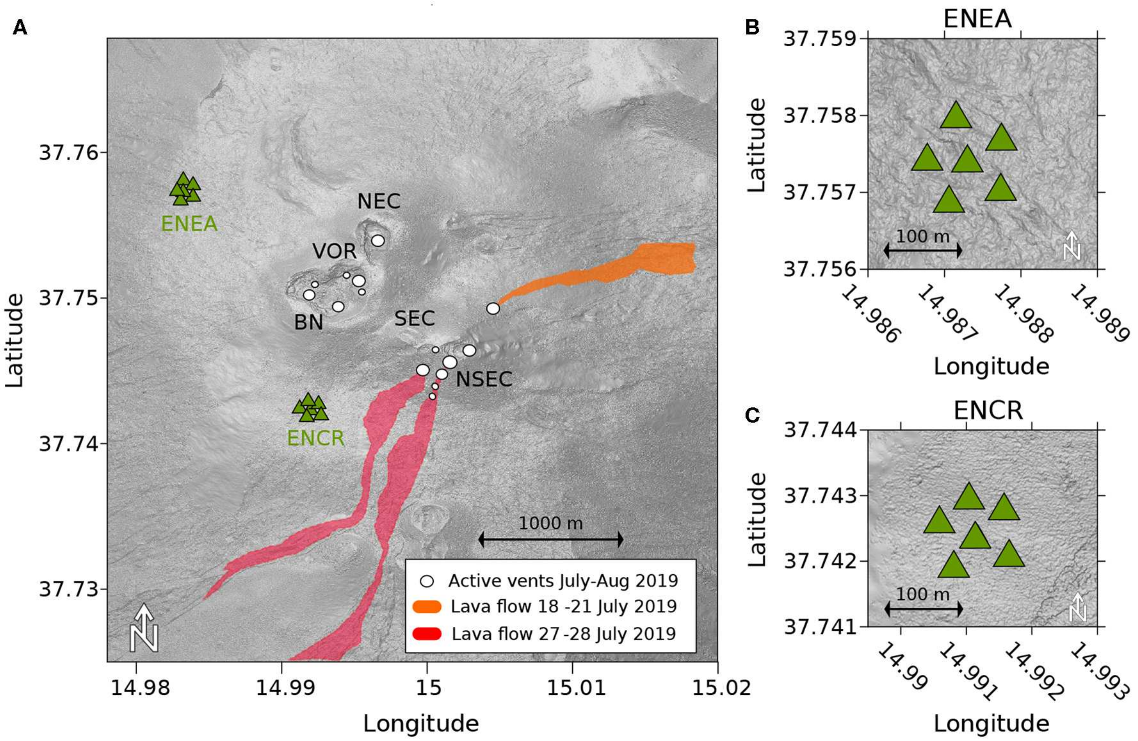


FIGURE 1 | (A) Map showing: (1) the locations of two infrasound arrays (green triangles), ENEA and ENCR, deployed at Mt. Etna, Italy, during July–August 2019; (2) approximate locations of vents active in the summit area (white circles of variable size according to vent dimensions as qualitatively inferred from visual observations) during the deployment period; (3) active lava flows during the deployment period. **(B,C)** Details of the geometry of the two infrasound arrays, ENEA and ENCR.

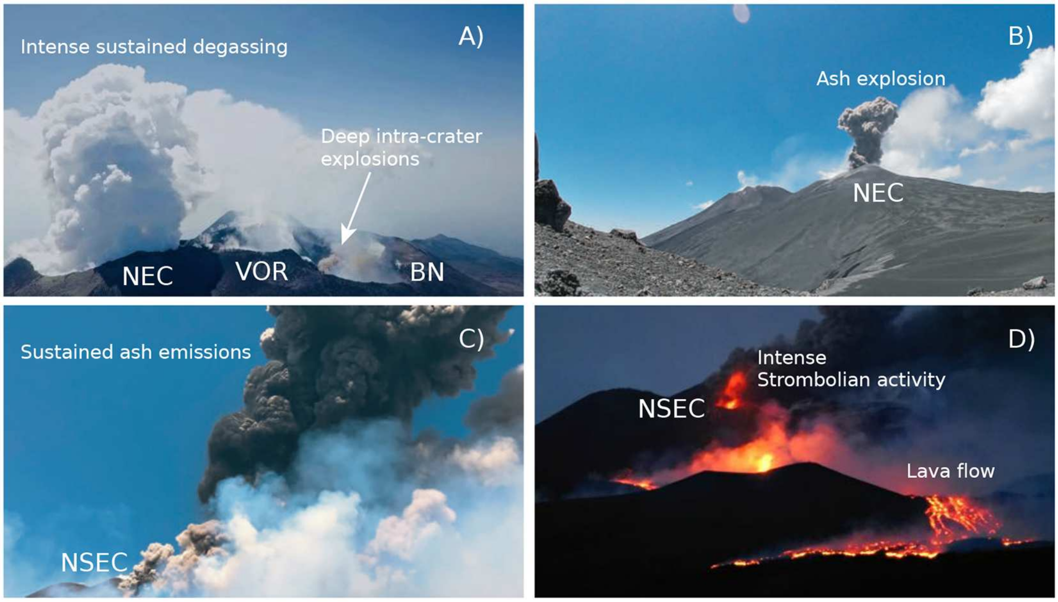


FIGURE 2 | Images of eruptive activity at Mt. Etna, Italy, during July–August 2019: **(A)** Intense sustained degassing activity from the North East Crater (NEC) and deep intra-crater gas-rich explosions from Bocca Nuova (BN) recorded during a UAV flight on 5 July 2019 at 08:00 UTC (Photo credit: A. Diaz-Moreno, University of Liverpool); **(B)** Ash-rich explosion from NEC on 3 July 2019 at 10:06 UTC (Photo credit: A. Diaz-Moreno, University of Liverpool); **(C)** Explosive activity accompanied by sustained ash emissions observed on 27 July 2019 at 12:21 UTC from the New South East crater (NSEC) area (Photo credit: Giuseppe Di Stefano/Marco Restivo, Etna Walk); **(D)** Intense Strombolian activity and lava flow originating from a vent in the NSEC area on 27 July 2019 (Photo credit: INGV-Osservatorio Etneo).

wind conditions on 27–28 July prevented the network from detecting the second paroxysm. The activity recorded during the summer of 2019 is typical at Mt. Etna; other eruptions with a similar fingerprint have been reported in the past (Corsaro et al., 2017; De Beni et al., 2019; Polacci et al., 2019).

3. DATA

In this study, we used data recorded by two 6-element infrasound arrays (ENEA and ENCR, **Figure 1**), deployed at Mt. Etna between 2 July and 25 August 2019. Changes in eruptive activity during this period, summarized in section 2 of this manuscript, are reported in the INGV activity bulletins and were confirmed by both Unmanned Aerial Vehicle (UAV) imagery (**Figure 2A**) and ground-based visual observations (**Figures 2B–D**) gathered during the deployment by the authors. The ENEA and ENCR arrays were installed at distances of between 1 and 1.8 km from the active vents (**Figure 1**). Each had similar configurations, deployed with an ∼100 m aperture (i.e., the largest distance between any two elements within an array) pentagon shape with a central element. Care was taken, considering the constraints imposed by topography, that the difference in elevation between any two microphones within each array was small; for both arrays, this difference did not exceed ∼30 m (see Methods section for additional details). The arrays were designed to have apertures large enough to discriminate acoustic arrivals and small enough to record coherent infrasound across all microphones (Ripepe et al., 2007). Data from both arrays were recorded on-site with a sampling frequency of 100 Hz at 24-bit resolution using DiGOS Data-cube digitizers (https://digos.eu/seismology-and-cubes/); ENEA was equipped with Chaparral Model 60 Ultra High Pressure microphones (full-scale range of 2000 Pa peak-to-peak, flat response between 0.03 and 245 Hz), and ENCR with IST2018 sensors (full-scale range of 480 Pa peak-to-peak, flat response between 0.1 and 40 Hz, as described in Grangeon and Lesage, 2019). Pressure amplitudes recorded by the two arrays ranged from few Pa for signals associated with both sustained and explosive degassing, to several tens of Pa during intense and persistent Strombolian activity. Examples of multi-channel data recorded by both ENEA and ENCR are shown in **Figures 3A–E**; these include an ash-rich explosion from NEC (**Figure 3C**) and repeated deep intra-crater explosions from BN (**Figure 3D**) on 2 July 2019, as well as Strombolian activity from NSEC on 18 July 2019 (**Figure 3E**). It is worth noting how infrasound waveforms in **Figures 3B,E** exhibit a marked asymmetry, although the investigation of this intriguing feature falls beyond the scopes of this manuscript.

4. METHODS

Here we discuss an array processing workflow to derive estimates of DOA and horizontal velocity from volcano infrasound array data. The method solves, via least-squares inversion, the problem of fitting a plane wave arrival traveling from an azimuth θ with horizontal velocity v to a vector of measured time delays between traces across the array (e.g., Claerbout, 1986; Olson and Szuberla, 2005; Haney et al., 2018). We also introduce a new method to quantify errors on the measurements of θ and v using standard propagation of error. For an array of n sensors, a data vector $\mathbf{d} = (\delta t_1, \delta t_2, \ldots, \delta t_N)$ of $N = n(n-1)/2$ delay times between the elements of the array can be estimated, for instance using waveform cross-correlation; a linear relationship exists between $\mathbf{d}$ and a model vector, $\mathbf{m} = (s_x, s_y)^T = (\sin\theta/v, \cos\theta/v)^T$ of slowness (defined as the inverse of velocity) in the East-West (s_x) and North-South (s_y) directions:

$$\mathbf{d} = \mathbf{Gm} \tag{1}$$

where $\mathbf{G}$ is a Nx2 matrix of horizontal distances between all pairs of array elements. When the measurements of delay times are affected by error, Equation (1) can be re-written to explicitly include it as:

$$\mathbf{d} = \mathbf{Gm} + \epsilon \tag{2}$$

where ϵ is the vector of errors in the estimates of time delays. These errors can be assumed to be normally distributed, with zero mean and variance σ_τ^2 (Olson and Szuberla, 2005). The solution to (2) for $\mathbf{m}$ is then found by minimizing the sum of squared errors:

$$E = \epsilon\epsilon^T = (\mathbf{d} - \mathbf{Gm})^T(\mathbf{d} - \mathbf{Gm}) \tag{3}$$

that is:

$$\mathbf{m} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{d} \tag{4}$$

Finally, DOA and apparent trace velocity across the array can be estimated from the solution vector $\mathbf{m}$ as:

$$\begin{cases} v = 1/\sqrt{s_x^2 + s_y^2} \\ \theta = \tan^{-1}(s_x/s_y) \end{cases} \tag{5}$$

It is important to note that the beamforming analysis in (1–5) is carried out in the horizontal plane (Edwards and Green, 2012), that is under the assumption that the contribution of differences in sensor elevations to the time delay measurements can be neglected; for this reason it is crucial that such differences are small.

For the most part, studies that apply array processing to volcano infrasound data either provide qualitative statements on the uncertainty associated with estimates of DOA and trace velocity or discuss the theoretical azimuthal resolution of the array in terms of its aperture in relation to the wavelength of the signals analyzed (e.g., Ripepe et al., 2018). The statistical confidence in the estimates of θ and v are rarely discussed (e.g., Szuberla and Olson, 2004); precise estimates of DOA and trace velocity are made difficult by the ubiquitous presence of noise in the data and the complex propagation of the acoustic wavefield, which affect measurements of time delays across arrays. Under the assumption that errors on the measurements of time delays across an array are normally distributed, with zero mean and variance σ_τ^2, the deviation of $\mathbf{m}$ from its expected value, due to

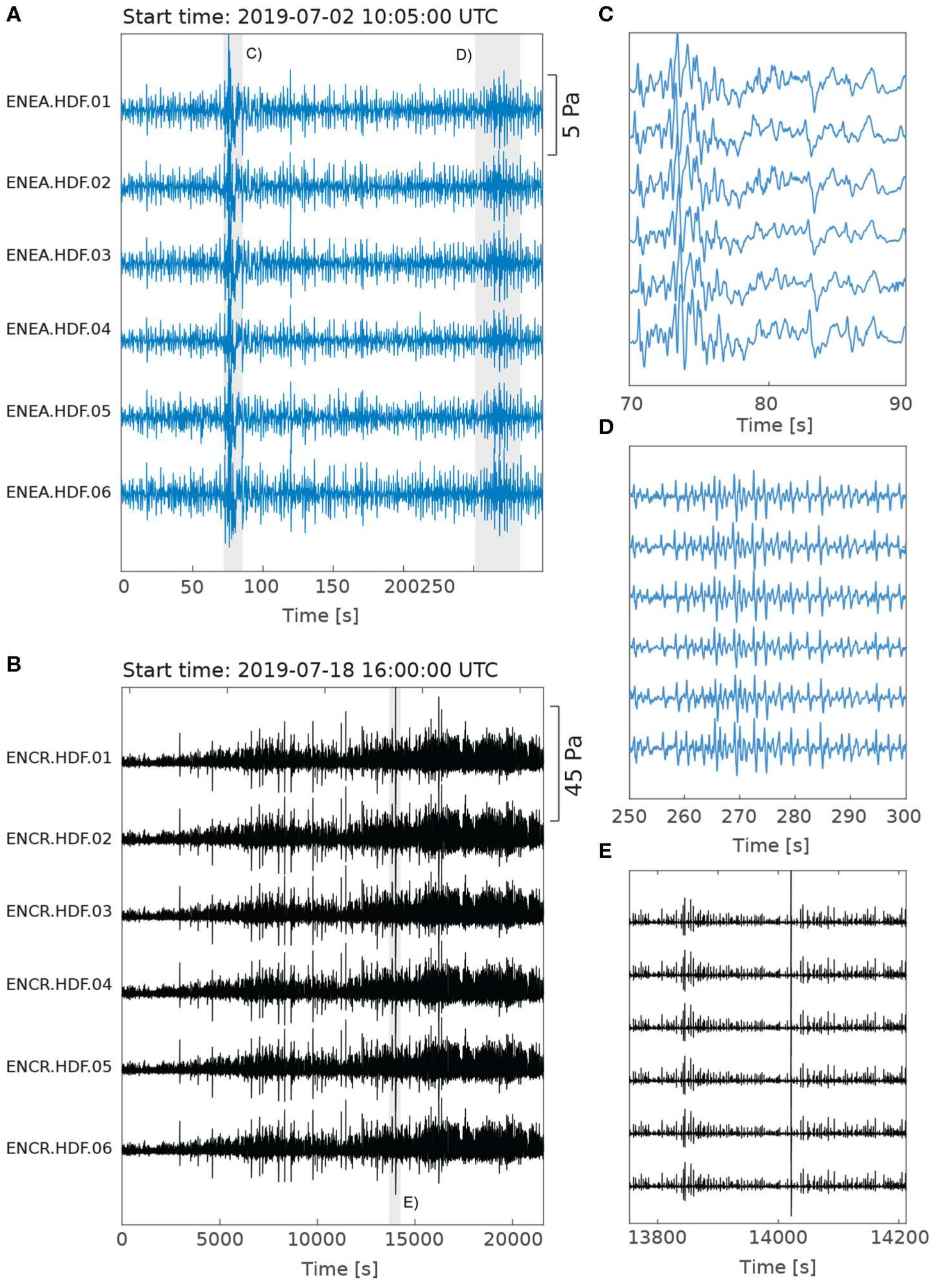


FIGURE 3 | Infrasound data recorded by the ENEA (blue) and ENCR (black) arrays during July 2019: **(A)** 5 min of continuous infrasound activity recorded by ENEA on 2 July 2019; **(B)** 6 h of intense Strombolian activity recorded by ENCR on 18 July 2019; **(C,D)** Detail (gray shaded areas in **A**) of ENEA infrasound associated with an ash-rich explosion from the NEC and deep intra-crater gas explosions from BN, respectively; **(E)** Detail (gray shaded area in **B**) of ENCR infrasound recorded during vigorous Strombolian activity on 18 July 2019. Note the difference in infrasound amplitudes between signals in **(A,B)**.

errors in the estimates of time delays, is described by the model covariance matrix, $\mathbb{C}(\mathbf{m})$, given by:

$$\mathbb{C}(\mathbf{m}) = \sigma_\tau^2(\mathbf{G}^T\mathbf{G})^{-1} = \sigma_\tau^2\mathbf{D} \tag{6}$$

The matrix $\mathbb{C}(\mathbf{m})$ is symmetrical and can be written in terms of the variances $(\sigma_{s_x}^2, \sigma_{s_y}^2)$ and covariance (σ_{s_x,s_y}^2) of s_x and s_y:

$$\mathbb{C}(\mathbf{m}) = \begin{bmatrix} \sigma_{s_x}^2 & \sigma_{s_x,s_y}^2 \\ \sigma_{s_x,s_y}^2 & \sigma_{s_y}^2 \end{bmatrix} \tag{7}$$

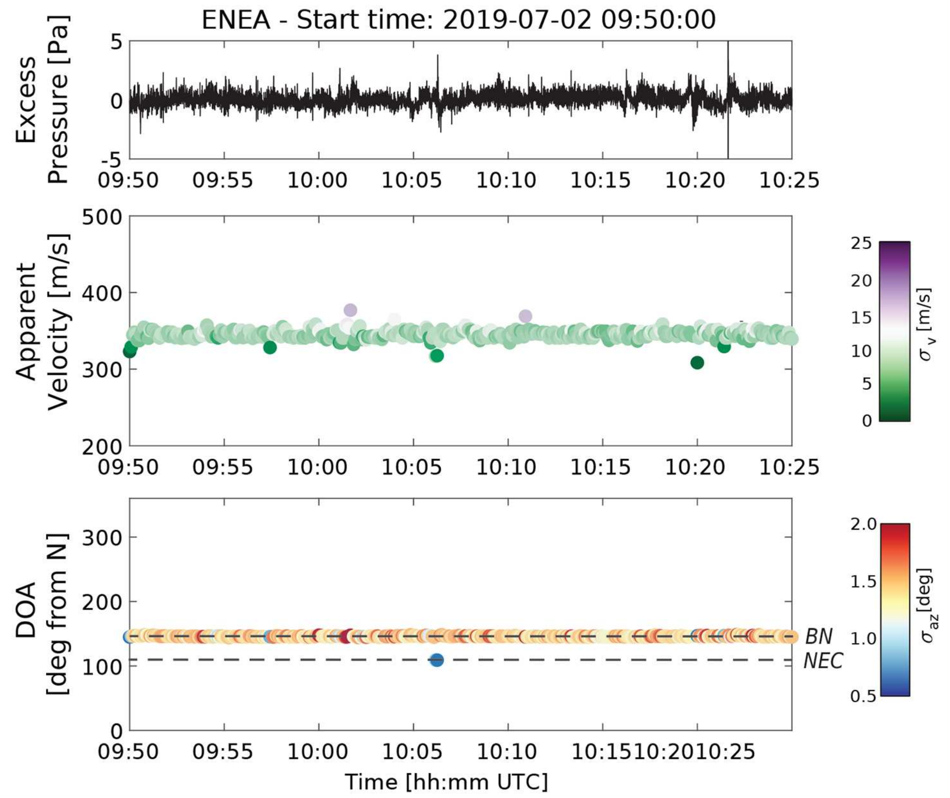


FIGURE 4 | Array processing of 35 min of data recorded at ENEA on 2 July 2019 between 9:50 and 10:25 UTC. Note (bottom panel) consistent detection of deep intra-crater activity from BN briefly interrupted by a larger ash-rich explosion from NEC (see section 2).

Equations (6) and (7) provide a link between the variances and covariance of s_x and s_y, and the variance of time delay measurements, that is:

$$\begin{cases} \sigma_{s_x}^2 = \sigma_\tau^2 D(1,1) \\ \sigma_{s_y}^2 = \sigma_\tau^2 D(2,2) \\ \sigma_{s_x,s_y}^2 = \sigma_\tau^2 D(1,2) \end{cases} \tag{8}$$

We are interested in error estimates on DOA and trace velocity, which are non-linear functions of s_x and s_y as shown in (5). In order to estimate the variances for θ and v (σ_θ^2 and σ_v^2, respectively) we apply standard propagation of errors theory (e.g., Vardeman and Jobe, 2001) to (5). Neglecting high-order terms we obtain:

$$\begin{cases} \sigma_v^2 \approx \sigma_{s_x}^2 \left(\frac{\partial v}{\partial s_x}\right)^2 + \sigma_{s_y}^2 \left(\frac{\partial v}{\partial s_y}\right)^2 + 2\sigma_{s_x,s_y}^2 \left(\frac{\partial v}{\partial s_x}\right)\left(\frac{\partial v}{\partial s_y}\right) \\ \sigma_\theta^2 \approx \sigma_{s_x}^2 \left(\frac{\partial \theta}{\partial s_x}\right)^2 + \sigma_{s_y}^2 \left(\frac{\partial \theta}{\partial s_y}\right)^2 + 2\sigma_{s_x,s_y}^2 \left(\frac{\partial \theta}{\partial s_x}\right)\left(\frac{\partial \theta}{\partial s_y}\right) \end{cases} \tag{9}$$

The terms in (9) are obtained by simply differentiating (5) and evaluating them at the values for s_x and s_y given by the solution of (4), that is:

$$\begin{cases} \sigma_v^2 \approx \sigma_{s_x}^2 s_x^2 v^6 + \sigma_{s_y}^2 s_y^2 v^6 + 2\sigma_{s_x,s_y}^2 s_x s_y v^6 \\ \sigma_\theta^2 \approx \sigma_{s_x}^2 s_y^2 v^4 + \sigma_{s_y}^2 s_x^2 v^4 + 2\sigma_{s_x,s_y}^2 s_x s_y v^4 \end{cases} \tag{10}$$

Finally, the standard deviations of trace velocity and DOA, σ_v and σ_θ, are obtained by taking the square root of (10).

5. RESULTS

We applied the array processing workflow described in section 4 of this manuscript to continuous infrasound data recorded at Mt. Etna between 2 July 2019 and 25 August 2019. Data collected by two infrasound arrays, ENEA and ENCR, were pre-processed by applying a band-pass (2-pole, zero-phase) filter within the frequency band of interest, which for the activity recorded during our experiment was between 0.7 and 15 Hz. The delay times between pairs of sensors across each array were determined using cross-correlation with sub-sample accuracy (e.g., Haney et al., 2018) within 10-s sliding windows overlapping by 5 s. DOA and trace velocity were calculated for all data windows by least-squares inversions according to (1–5), and their variances were calculated according to (7–9). The mean of the normalized cross-correlation maxima (MCCM) within each data window reflects the level of coherence between signals across the array; only values of DOA and trace velocity corresponding to MCCM > 0.5 were considered to ensure that inversion was performed for coherent signals across the arrays.

Examples of estimates of DOA and trace velocity, along with their uncertainties, are shown in **Figures 4**, **5**. In **Figure 4** we

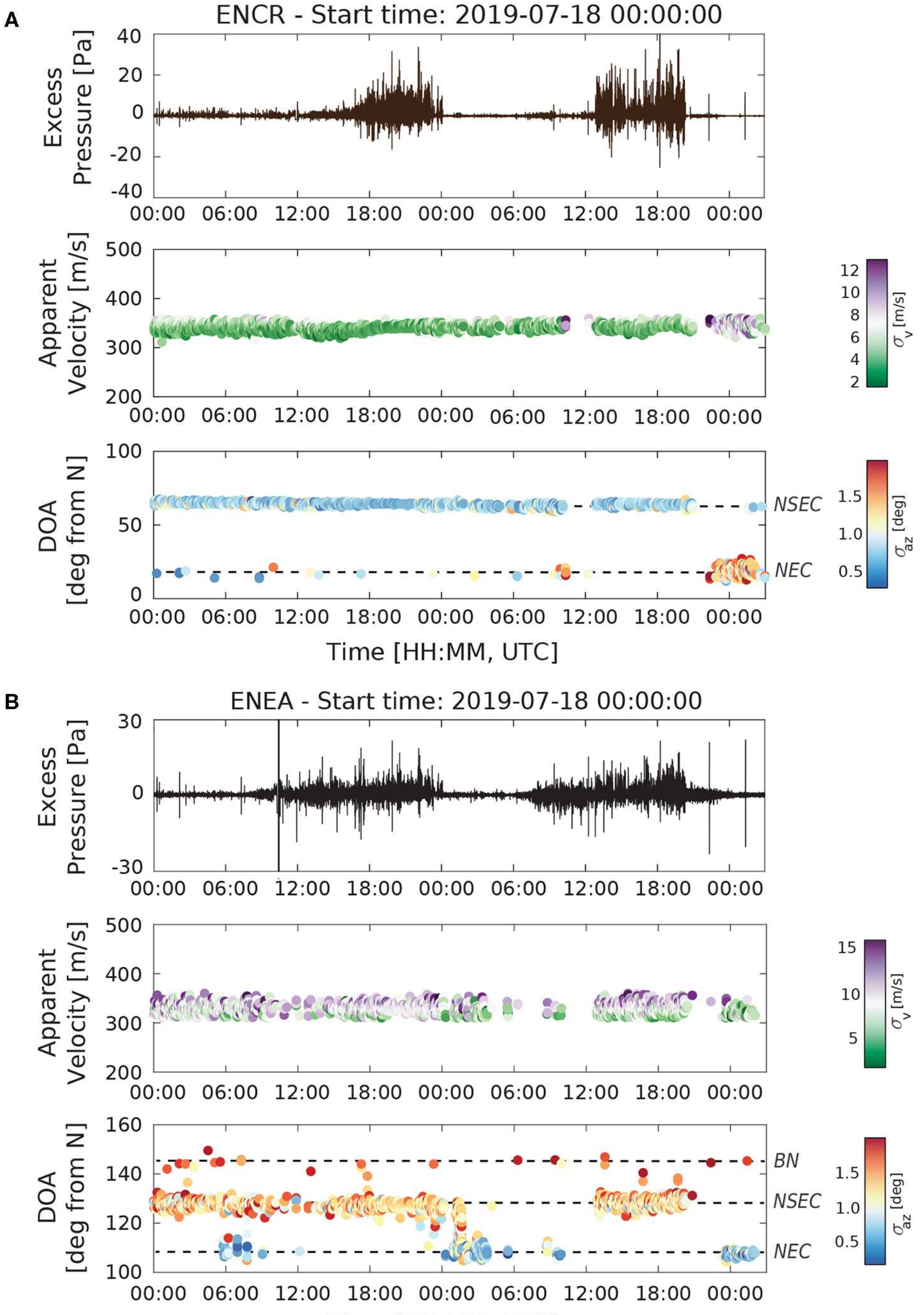


FIGURE 5 | Array processing of 51 h of data recorded at ENEA and ENCR, starting on 18 July 2019, 00:00 UTC: **(A)** ENCR shows stable detections during Strombolian activity (18:00–24:00 on 18 July and 12:00–22:00 on 19 July) as well as during a period of lava effusion from the NE flank of NSEC (23:09 on 18 July, see section 2 in this manuscript); **(B)** ENEA captures the most intense Strombolian activity from the NSEC but shows larger fluctuations in DOA (direction of arrival) than ENCR. This is due to less favorable signal-to-noise ratio for activity at the NSEC as well as its proximity to secondary sources within the BN crater. Both arrays detect a clear shift in activity from NSEC to NEC late during the night between 19 and 20 July. Note how uncertainties and scatter in locations are consistently smaller at the array closer to the active vent.

show data recorded at ENEA on 2 July 2019 (the only array installed at this time). **Figure 5** illustrates results from both ENEA and ENCR for just over 2 days of activity starting on 18 July 2019. These examples are representative of the variable activity, during July–August 2019, from multiple vents (**Figure 1**) and across all summit craters at Mt. Etna. **Figure 4** shows stable locations in the direction of BN associated with deep intra-crater explosions (**Figures 2A**, **3D**) and a sharp change in DOA at about 10:06 UTC, which corresponds to a larger ash-rich explosion from the NEC (**Figures 2B**, **3C**); trace velocities are of the order of 330–335 m/s with standard deviations of 5–15 m/s, and uncertainties on DOA estimates of 0.5–2° (≈10–40 m at the distance between ENEA and the vent). Locations in **Figure 5** are dominated by activity at NSEC, in particular between 18:00 and 24:00 on 18 July and during 12:00–19:00 on 19 July, when frequent and intense Strombolian activity was observed at this crater (**Figures 3B,E**). Values of trace velocities at ENEA during this period are of the order of 310–335 m/s, with uncertainties of 5–15 m/s and standard deviations on DOA between 1 and 2° (≈30–60 m); at ENCR, trace velocities are in 330–335 m/s range with uncertainties of 2–8 m/s, and uncertainties on DOA of 0.5–2° (≈10–40 m). Notably, both arrays detected a shift in activity in the direction of the NEC, also reported by INGV at about 23:00 on 19 July (section 2).

The two arrays were installed to provide the best coverage of all active craters at Mt. Etna (within limitations imposed by site access and safety). Owing to their positions relative to the active vents ENEA was ideally located to independently discriminate activity from NEC, whereas ENCR was positioned to optimize detection of activity from the NSEC area (**Figure 1**). In **Figure S1**, we have provided an additional example of array locations showing the stability and low uncertainties associated with detections at ENCR during a period of Strombolian activity, lava fountaining and sustained ash emissions at the NSEC on 27–28 July 2019 (see section 2 in this manuscript, **Figures 2C,D**).

6. DISCUSSION AND CONCLUSIONS

We have presented an algorithm for the inversion of infrasound array data that allows estimating DOA with apparent sound velocity, and their associated uncertainties. We have applied the workflow to data collected by two 6-element infrasound arrays deployed at Mt. Etna during the summer of 2019; continuous detections from the two arrays tracked the activity observed at the volcano, including shifts in degassing and eruptions across multiple summit craters and vents. The results presented in **Figures 4**, **5**, suggest that: (i) array location relative to a complex system of active vents is key to allow discrimination of variable activity across multiple craters, and (ii) the quality of detection and thus the final estimates of uncertainties are crucially influenced by SNR levels. **Figure 5**, for instance, shows higher quality (i.e., lower uncertainties) detections from ENEA for activity located at the NEC, while ENCR seems better suited to track eruptions in the NSEC area; ENEA is also able, albeit with comparatively large uncertainties, to track small intra-crater explosions from BN (**Figure 5B**) even during periods of elevated Strombolian activity at the NSEC. At the first order, this observation arguably reflects variable SNR at the two sites, resulting from the interplay between array proximity to the source (NEC and BN closer to ENEA; NSEC closer to ENCR) and wind noise levels (qualitatively observed to be consistently higher at ENEA). Wind strength and direction are factors that can potentially introduce a bias in the estimates of DOA and trace velocity from infrasound arrays (e.g., Schwaiger et al., 2020). At large source-receiver distances time reversal location of infrasound sources, propagating the acoustic wavefield in a windy atmosphere, has been used to assess the misfit between known source backazimuths and DOA estimates (e.g., Lonzaga, 2016). At the local scale, Johnson et al. (2012) demonstrated how temporal variations in acoustic parameters, such as infrasound travel times over short distances—<20 km—can be exploited to infer atmospheric conditions, including the strength of wind in a vertically stratified atmosphere. In theory, appropriate corrections for the effect of wind on infrasound measurements could be introduced in array inversion workflows; the main challenge for this lies in the fact that wind measurements or models with the required temporal and spatial resolution for the local scale are generally not available. In addition to wind noise, uncertainty on estimates of DOA and sound velocity are further linked to array configuration and its position relative to the sources; these factors control the degree to which the measured time delays across the array correspond to a physically realizable set of delays associated with the propagation of a plane wave across the array (Szuberla and Olson, 2004). Qualitatively, the plane wave approximation is considered valid at source-to-array distances much greater than the aperture of the array (e.g., Almendros et al., 1997), a condition met in our study for all source-array combinations. Finally, effects from topography, such as diffraction can affect acoustic propagation and introduce additional bias in array estimates of DOA and sound speed. At the local scale Fee et al. (2019) recently investigated the effects of topography on back-projection locations of volcano infrasound sources integrating improved travel time calculations from finite difference modeling of the acoustic wavefield. At Mt. Etna, specifically, Diaz-Moreno et al. (2019) performed 3D finite difference simulation of acoustic wavefield propagation; their results suggest that the effects of topography on acoustic propagation from sources in the area of the summit craters to the locations of the arrays used in this study are minor.

A quantitative estimate of how well the measured time delays correspond to a plane wave crossing the array is directly provided by the a priori variance of the time delay measurements, which is given by (Szuberla and Olson, 2004):

$$\sigma_\tau^2 = \frac{R_0^2}{N - r} \tag{11}$$

where (following the same notation of Equations 1–10) $R_0^2 = \mathbf{d}^{\mathbf{T}}(\mathbb{I} - \mathbf{R})\mathbf{d}$, with $\mathbb{I}$ being the identity matrix and $\mathbf{R} = \mathbf{G}[(\mathbf{G}^T\mathbf{G})^{-1} - 1]\mathbf{G}^T$; N is the number of station pairs and r is the rank of **R**. We calculated σ_τ^2 for periods with high MCMM > 0.5; values were of the order of 10^{-2} s, that is $\approx 1/f_s$ (with f_s being the data sampling frequency) as also reported in other studies (e.g., Szuberla and Olson, 2004). This gives confidence that the plane wave assumption does not introduce significant bias on the results of our array analysis. On the other hand for data with low MCMM, that is no coherent signal traveling across the array, the values of σ_τ^2 were very high, up to a few seconds.

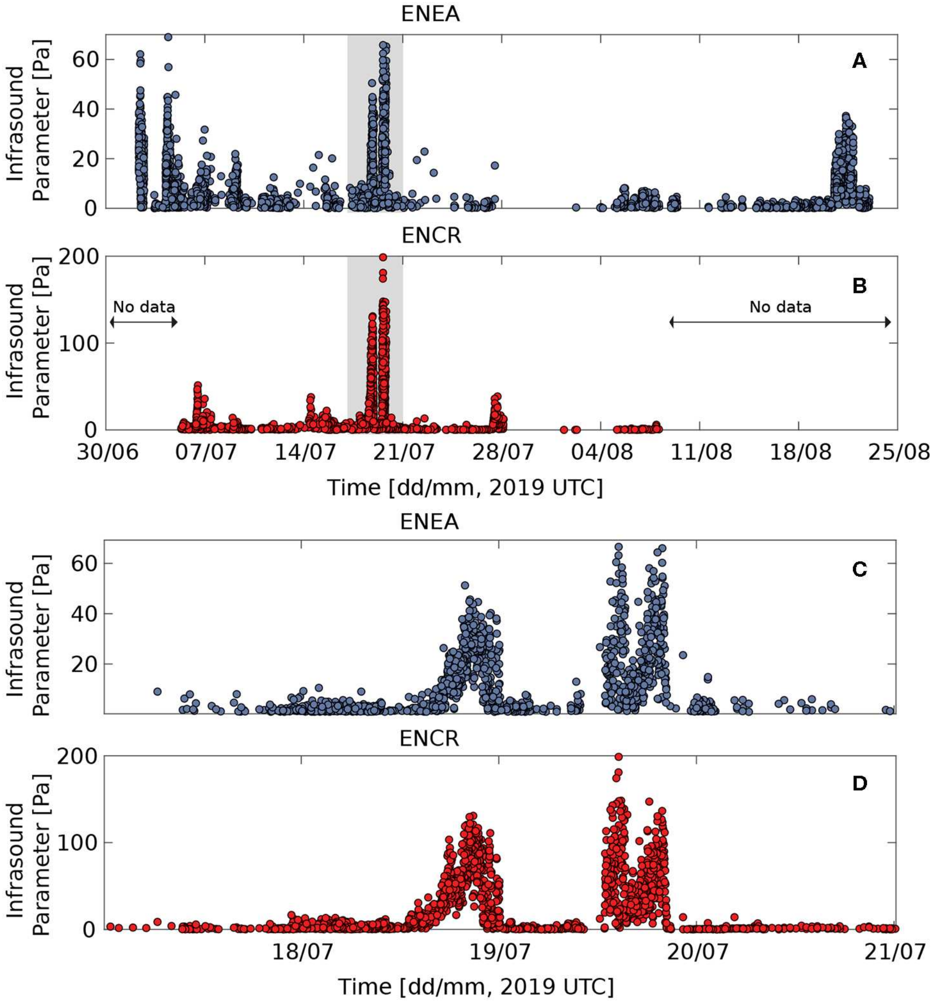


FIGURE 6 | One-minute Infrasound Parameter during July–August 2019 for arrays ENEA **(A)** and ENCR **(B)**; periods when no data were available are also indicated. Detail (zoom in of gray areas in **A,B**) of Infrasound Parameter values at ENEA **(C)** and ENCR **(D)** for the period 17–21 July 2019. The figure shows an obvious increase in both occurrence (i.e., increasing array detections per minute) and intensity (increasing signal amplitude) of Strombolian activity; the Infrasound Parameter peaks in coincidence with the opening of a new fracture and emplacement of a lava flow on the NE flank of the NSEC, late during the night on 18 July (see sections 2 and 6).

In our workflow, uncertainties on apparent velocity and DOA are evaluated using the theory of error propagation. While the variances estimated using this method are a first order approximation to the true uncertainty, this solution is easy to implement with very low computational overhead, an obvious advantage for deployment within real-time volcano monitoring systems. An exact solution for the statistical confidence in the estimates of DOA and signal velocity for planar arrays (i.e., without significant variations in elevation across the array) is discussed in Szuberla and Olson (2004). We benchmarked the results of our workflow against this solution and found good agreement between the two approaches (**Figure S2**).

Our study also reveals that combining estimates of DOA from multiple arrays improves the ability to discriminate multiple active vents. The ENEA and ENCR arrays were ideally positioned to independently discriminate sources within the NEC and NSEC, respectively (**Figure 1**). In the absence of complementary observations, however, resolving activity from variable vents within the other two summit craters (BN and VOR) frequently requires joint measurements from two arrays; combining the DOAs obtained from both arrays, a source location could be uniquely identified (**Figure S3**). Finally, we note that our workflow implicitly assumes a single arrival within each signal window analyzed. At volcanoes with multiple vents separated by a small distance different sources could be simultaneously active, and thus, produce multiple arrivals within a signal analysis window. This is discussed, for example, by Yamakawa et al. (2018) at Stromboli; the authors demonstrate how the MUSIC algorithm can help to resolve multiple active sources. We find that when active vent are separated by comparatively large distances and have different characteristic dimensions, such as Mt. Etna, appropriate selection of the signal analysis window and filters still allows effective source discrimination.

A key advantage of using infrasound arrays over individual microphones is that with a single data processing step multi-element arrays can provide information on source location, as well as the levels of volcanic activity in terms of both signal amplitude and rates of detection. This is an ideal scenario to inform early warning and assessment of alert levels during period of unrest accompanied by elevated surface activity at volcanoes. Ripepe et al. (2018) proposed an efficient way to combine the output of infrasound array processing into a single metric, the Infrasound Parameter, calculated as the product $IP = A_p \cdot N_d$ between the mean infrasound amplitude, A_p, and the number of array detections per minute, N_d. They calculated IP at Mt. Etna for a continuous period of about 8 years, and discussed the implementation of an alert color code used to dispatch early warnings of impending paroxysmal activity to the local civil protection authority. Here, we estimated a modified IP during July-August 2019 by taking the product between the mean of the maximum signal amplitude within each array detection window over a minute and the number of detections per minute. For this calculation we selected only high-quality detection windows, corresponding to MCMM > 0.5. The results for both arrays during the entire deployment period are shown in **Figures 6A,B**. **Figures 6C,D** show details of IP during 17–21 July 2019, when Strombolian activity escalated at the NSEC from one explosion every 1–2 min to several events within a minute, eventually leading to the opening of a fracture and emplacement of a lava flow on the NE flank of the NSEC late during the night on 18 July. The temporal evolution of the IP parameter represents changes in surface activity; Ripepe et al. (2018) reported that this behavior, typical of eruptions at Mt. Etna, reflects the transition from rapid Strombolian explosions, driven by the repeated ascent of gas slugs in the shallow conduit, to a churn flow regime when gas discharge increases and the eruption becomes sustained. The pattern observed in **Figures 6C,D** suggests that IP, when appropriately calibrated, may provide a valuable metric to monitor escalating surface activity, inform changes in alert color codes, and issue early warnings. While its use has been, thus far, limited to monitor activity at Mt. Etna, it may also be applicable at other basaltic volcanoes characterized by periodic occurrence of paroxysms.

In conclusion, we have presented a framework for the inversion of infrasound array data to provide rapid estimates of source location and apparent sound velocity during periods of elevated volcanic activity. Our algorithms include a new, and computationally efficient, procedure for quantitative assessment of the uncertainties on array measurements, which is particularly well-suited for real-time implementation. We applied the proposed workflow to data gathered at Mt. Etna during July–August 2019. Our results demonstrate that infrasound arrays allowed detection and tracking of variable activity from multiple active vents at Mt. Etna. Owing to the fact that infrasound propagates efficiently over large distances (e.g., Fee and Matoza, 2013), we suggest that this data analysis framework may also hold potential to monitor eruptive activity at the regional scale (i.e., source-array distances of up to several hundreds of km), in particular in remote areas where local monitoring of individual volcanoes is not viable (Fee et al., 2019). We have further discussed a simple metric derived from infrasound array analyses that may be suitable to inform monitoring operations and form the basis to issue early warnings of impending paroxysms at basaltic volcanoes. We surmise that infrasound offers a simple and effective tool to track the temporal evolution of volcanic activity and to assist with real-time volcano monitoring, as well as the potential to inform models of the processes that control degassing and eruption at volcanoes.

DATA AVAILABILITY STATEMENT

The data analyzed for this study are available via direct request to the authors or from the Natural Environment Research Council Geoscience Data Centre (https://www.bgs.ac.uk/services/ngdc/accessions/index.html?). A software implementation of the algorithms presented in this paper, along with example data, can be downloaded from: https://github.com/silvioda/Infrasound-Array-Processing-Matlab.

AUTHOR CONTRIBUTIONS

SD, AD-M, and LZ conducted the field experiment. SD, MH, AW, JL, and DF developed the workflow for inversion and uncertainty calculations. AD-M produced the map in **Figure 1** based on activity bulletins from INGV. SD produced all other figures with input from all co-authors and wrote the manuscript with input from all co-authors. All authors initiated the study and guided the investigation.

FUNDING

AD-M was supported by NERC grant NE/P00105X/1. LZ has also received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 798480. SD was partly supported by a Geoscientists without Borders grant from the Society of Exploration Geophysics.

ACKNOWLEDGMENTS

SD and LZ acknowledge the support from the EUROVOLC project under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 731070). The authors thank the staff of the Istituto Nazionale di Geofisica e Vulcanologia Sezione di Catania, in particular Salvo Rapisarda, Daniele Pellegrino, Mario Pulvirenti, and Danilo Contrafatto for their valuable support in the field.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.00169/full#supplementary-material

REFERENCES

Almendros, J., Ibanez, J. M., Alguacil, G., Del Pezzo, E., and Ortiz, R. (1997). Array tracking of the volcanic tremor source at deception island, antarctica. *Geophys. Res. Lett*. 24, 3069–3072. doi: 10.1029/97GL03096

Cannata, A., Di Grazia, G., Aliotta, M., Cassisi, C., Montalto, P., and Patané, D. (2013). Monitoring seismo-volcanic and infrasonic signals at volcanoes: Mt. Etna case study. *Pure Appl. Geophys*. 170, 1751–1771. doi: 10.1007/s00024-012-0634-x

Cannata, A., Montalto, P., Privitera, E., and Russo, G. (2009). Characterization and location of infrasonic sources in active volcanoes: Mount Etna, September–November 2007. *J. Geophys. Res*. 114:B08308. doi: 10.1029/2008JB006007

Caplan-Auerbach, J., Bellesiles, A., and Fernandes, J. K. (2010). Estimates of eruption velocity and plume height from infrasonic recordings of the 2006 eruption of Augustine Volcano, Alaska. *J. Volcanol. Geotherm. Res*. 189, 12–18. doi: 10.1016/j.jvolgeores.2009.10.002

Claerbout, J. F. (1986). Fundamentals of geophysical data processing, 2nd edn. *Geophys. J. R. Astron. Soc*. 86, 217–219. doi: 10.1111/j.1365-246X.1986.tb01085.x

Coombs, M. L., Wech, A. G., Haney, M. M., Lyons, J. J., Schneider, D. J., Schwaiger, H. F., et al. (2018). Short-term forecasting and detection of explosions during the 2016–2017 eruption of Bogoslof Volcano, Alaska. *Front. Earth Sci*. 6:122. doi: 10.3389/feart.2018.00122

Corsaro, R. A., Andronico, D., Behncke, B., Branca, S., Caltabiano, T., Ciancitto, F., et al. (2017). Monitoring the December 2015 summit eruptions of Mt. Etna (Italy): implications on eruptive dynamics. *J. Volcanol. Geotherm. Res*. 341, 53–69. doi: 10.1016/j.jvolgeores.2017.04.018

De Angelis, S., Diaz-Moreno, A., and Zuccarello, L. (2019). Recent developments and applications of acoustic infrasound to monitor volcanic emissions. *Remote Sens*. 11, 1–18. doi: 10.3390/rs11111302

De Beni, E., Cantarero, M., and Messina, A. (2019). UAVs for volcano monitoring: a new approach applied on an active lava flow on Mt. Etna (Italy), during the 27 February–02 March 2017 eruption. *J. Volcanol. Geotherm. Res*. 369, 250–262. doi: 10.1016/j.jvolgeores.2018.12.001

Diaz-Moreno, A., Iezzi, A., Lamb, O., Fee, D., Kim, K., Zuccarello, L., et al. (2019). Volume flow rate estimation for small explosions at Mt. Etna, Italy, from acoustic waveform inversion. *Geophys. Res. Lett*. 46, 11071–11079. doi: 10.1029/2019GL084598

Edwards, W. N., and Green, D. N. (2012). Effect of interarray elevation differences on infrasound beamforming. *Geophys. J. Int*. 190, 335–346. doi: 10.1111/j.1365-246X.2012.05465.x

Fee, D., Garces, M., and Steffke, A. (2010). Infrasound from tungurahua volcano 2006–2008: Strombolian to Plinian eruptive activity. *J. Volcanol. Geotherm. Res*. 193, 67–81. doi: 10.1016/j.jvolgeores.2010.03.006

Fee, D., Izbekov, P., Kim, K., Yokoo, A., Lopez, T., Prata, F., et al. (2017). Eruption mass estimation using infrasound waveform inversion and ash and gas measurements: evaluation at Sakurajima Volcano, Japan. *Earth Planet. Sci. Lett*. 480, 42–52. doi: 10.1016/j.epsl.2017.09.043

Fee, D., and Matoza, R. S. (2013). An overview of volcano infrasound: from Hawaiian to Plinian, local to global. *J. Volcanol. Geotherm. Res*. 249, 123–139. doi: 10.1016/j.jvolgeores.2012.09.002

Fee, D., Toney, L. D., Matoza, R. S., Sanderson, R. W., Haney, M. M., Lyons, J. J., et al. (2019). "A framework for improved infrasound back-projection on local and regional scales," in *AGU Fall Meeting Abstracts*, Vol. 2019, S41E-0577 (San Francisco).

Garcés, M. A., Fee, D., and Matoza, R. (2009). Volcano acoustics. *Model. Volcan. Process. Phys. Math. Volcan*. 9780521895439, 359–383. doi: 10.1017/CBO9781139021562.016

Grangeon, J., and Lesage, P. (2019). A robust, low-cost and well-calibrated infrasound sensor for volcano monitoring. *J. Volcanol. Geotherm. Res*. 387:106668. doi: 10.1016/j.jvolgeores.2019.106668

Haney, M. M., Van Eaton, A. R., Lyons, J. J., Kramer, R. L., Fee, D., and Iezzi, A. M. (2018). Volcanic thunder from explosive eruptions at Bogoslof Volcano, Alaska. *Geophys. Res. Lett*. 45, 3429–3435. doi: 10.1002/2017GL076911

Johnson, J. B. (2004). Source location variability and volcanic vent mapping with a small-aperture infrasound array at Stromboli volcano, Italy. *Bull. Volcanol*. 67, 1–14. doi: 10.1007/s00445-004-0356-8

Johnson, J. B., Anderson, J., Marcillo, O., and Arrowsmith, S. (2012). Probing local wind and temperature structure using infrasound from Volcan Villarrica (Chile). *J. Geophy. Res. Atmos*. 117:D17107. doi: 10.1029/2012JD017694

Johnson, J. B., and Palma, J. L. (2015). Lahar infrasound associated with Volcán Villarrica's 3 March 2015 eruption. *Geophys. Res. Lett*. 42, 6324–6331. doi: 10.1002/2015GL065024

Johnson, J. B., and Ripepe, M. (2011). Volcano infrasound: a review. *J. Volcanol. Geotherm. Res*. 206, 61–69. doi: 10.1016/j.jvolgeores.2011.06.006

Lonzaga, J. B. (2016). Time reversal for localization of sources of infrasound signals in a windy stratified atmosphere. *J. Acoust. Soc. Am*. 139, 3053–3062. doi: 10.1121/1.4953017

Matoza, R., and Fee, D. (2018). The inaudible rumble of volcanic eruptions. *Acoust. Today* 14, 17–25. Available online at: https://acousticstoday.org/inaudible-rumble-volcanic-eruptions/

Matoza, R. S., Arciniega-Ceballos, A., Sanderson, R. W., Mendo-Pérez, G., Rosado-Fuentes, A., and Chouet, B. A. (2019). High-broadband seismoacoustic signature of vulcanian explosions at popocatépetl volcano, mexico. *Geophys. Res. Lett*. 46, 148–157. doi: 10.1029/2018GL080802

Matoza, R. S., Hedlin, M. A., and Garcés, M. A. (2007). An infrasound array study of Mount St. Helens. *J. Volcanol. Geotherm. Res*. 160, 249–262. doi: 10.1016/j.jvolgeores.2006.10.006

McNutt, S. R., Thompson, G., Johnson, J., Angelis, S. D., and Fee, D. (2015). Seismic and infrasonic monitoring. *Encycl. Volcan*. 2015, 1071–1099. doi: 10.1016/B978-0-12-385938-9.00063-8

Olson, J. V., and Szuberla, C. A. L. (2005). Distribution of wave packet sizes in microbarom wave trains observed in Alaska. *J. Acoust. Soc. Am*. 117, 1032–1037. doi: 10.1121/1.1854651

Polacci, M., Andronico, D., de' Michieli Vitturi, M., Taddeucci, J., and Cristaldi, A. (2019). Mechanisms of ash generation at basaltic volcanoes: the case of Mount Etna, Italy. *Front. Earth Sci*. 7:193. doi: 10.3389/feart.2019.00193

Ripepe, M., De Angelis, S., Lacanna, G., Poggi, P., Williams, C., Marchetti, E., et al. (2009). Tracking pyroclastic flows at soufriére hills volcano. *EOS* 90, 229–230. doi: 10.1029/2009EO270001

Ripepe, M., Marchetti, E., Delle Donne, D., Genco, R., Innocenti, L., Lacanna, G., et al. (2018). Infrasonic early warning system for explosive eruptions. *J. Geophys. Res. Solid Earth* 123, 9570–9585. doi: 10.1029/2018JB015561

Ripepe, M., Marchetti, E., and Ulivieri, G. (2007). Infrasonic monitoring at stromboli volcano during the 2003 effusive eruption: insights on the explosive and degassing process of an open conduit system. *J. Geophys. Res. Solid Earth* 112:B09207. doi: 10.1029/2006JB004613

Schwaiger, H., Lyons, J., Iezzi, A., Fee, D., and Haney, M. (2020). Evolving infrasound detections from Bogoslof volcano, alaska: insights from atmospheric propagation modeling. *Bull. Volcanol*. 82:27. doi: 10.1007/s00445-020-1360-3

Sciotto, M., Cannata, A., Prestifilippo, M., Scollo, S., Fee, D., and Privitera, E. (2019). Unravelling the links between seismo-acoustic signals and eruptive parameters: Etna lava fountain case study. *Sci. Rep*. 9:16417. doi: 10.1038/s41598-019-52576-w

Szuberla, C. A. L., and Olson, J. V. (2004). Uncertainties associated with parameter estimation in atmospheric infrasound arrays. *J. Acoust. Soc. Am*. 115, 253–258. doi: 10.1121/1.1635407

Vardeman, S., and Jobe, J. (2001). *Basic Engineering Data Collection and Analysis. Statistics Series*. Cengage Learning.

Yamakawa, K., Ichihara, M., Ishii, K., Aoyama, H., Nishimura, T., and Ripepe, M. (2018). Azimuth estimations from a small aperture infrasonic array: test observations at Stromboli Volcano, Italy. *Geophys. Res. Lett*. 45, 8931–8938. doi: 10.1029/2018GL078851

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.

frontiers
in Earth Science

BRIEF RESEARCH REPORT
published: 09 July 2020
doi: 10.3389/feart.2020.00285

Four Dimensional Gravity Forward Model in a Deep Reservoir

*Paolo Mancinelli**

Dipartimento di Ingegneria e Geologia, Università G. D'Annunzio di Chieti-Pescara, Chieti, Italy

In this work, we calculate the gravity signature of small density changes in a real-case deep reservoir. Based on the 3D forward density modeling of constrained geometries and using parameters of the involved rocks and fluids, we compute the differential gravity signature before and after the production period of the Volve oil field in the North Sea. Causative sources of the retrieved residuals are spatially correlated with positions of the most productive wells, locating areas of maximum density change. Results show that the 4D gravity forward model is capable of resolving residual gravity signatures also for deep and small density changes. In particular, we locate $\sim$−13 μGal gravity minima over the 2750 m deep reservoir; this minima was caused by the −53 kg m^{-3} density change related to production and injection activity.

Keywords: 4D modeling, gravity forward modeling, production data constraints, deep source, small density contrast

Edited by:
Andrea Cannata,
University of Catania, Italy

Reviewed by:
Ingo Heyde,
Federal Institute for Geosciences and Natural Resources, Germany
Filippo Greco,
Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy

Correspondence:
Paolo Mancinelli
paolo.mancinelli@unich.it

Specialty section:
This article was submitted to Solid Earth Geophysics, a section of the journal Frontiers in Earth Science

Received: *04 May 2020*
Accepted: *19 June 2020*
Published: *09 July 2020*

Citation:
Mancinelli P (2020) Four Dimensional Gravity Forward Model in a Deep Reservoir. Front. Earth Sci. 8:285. doi: 10.3389/feart.2020.00285

INTRODUCTION

Using gravity data for sub-surface modeling of geological structures may represent a useful tool to address open questions about geological or geophysical processes at the crustal or local scale. Application of 2D or 3D forward or inverse techniques depends on real gravity data availability over the target area (e.g., Mancinelli et al., 2015; Dressel et al., 2018; Fedi et al., 2018; Mancinelli et al., 2019, 2020; Sobh et al., 2019) but often the smaller gravity signatures are concealed and hard to elaborate through filtering. In the last two decades, the increase in gravimeters' precision allowed the development of 4D gravity models monitoring fluid-related density changes at increasing resolution (e.g., Hare et al., 1999; Eiken et al., 2000, 2008; Ferguson et al., 2007, 2008; Vasilevskiy and Dashevsky, 2007; Alnes et al., 2008; Davis et al., 2008; Gasperikova and Hoversten, 2008; Stenvold et al., 2008; Jacob et al., 2010; Krahenbuhl et al., 2011; Krahenbuhl and Li, 2012; Wilson et al., 2012; Reitz et al., 2015; Elliott and Braun, 2016, 2017). It should be noted that in all these case histories, the top of the gravity source was always above 2500 m depth. The depth of the source is the most critical parameter affecting the resolvability of a gravity anomaly (e.g., Blakely, 1996). In the case of CO_2 storage, the depth of 2500 m is considered a threshold between shallower (thus resolvable) and deeper (very difficult to resolve) field applications (Cooper et al., 2009).

The evolution of gravity modeling techniques has closely followed the technological improvements that allowed significant increase in precision and accuracy of gravimeters over the last decades. A compelling revision is provided by Van Camp et al. (2017). The goal of lowering the threshold for stable measurements, together with the increased ability of identifying noise sources, has allowed to reach precisions below 1 nm s^{-2} – i.e., 0.1 μGal, also for transportable absolute gravimeters (Ménoret et al., 2018). Seafloor gravimeters can achieve measurements with time-lapse accuracy better than 5 μGal (Eiken et al., 2000; Sasagawa et al., 2003; Stenvold et al., 2008). However, the monitoring of fluid-induced density changes at production or storage sites through gravity 4D models has not evolved in a similar way. In fact, very few production and storage reservoirs have been investigated using 4D gravity methods. This is likely due to problems related to three main sources: (i) size of the production/injection process in terms of depth, thickness, volumes, and densities of the involved deposits and fluids; (ii) high-quality gravity data availability due to costs for gravimeters and data acquisition and knowledge about noise sources; and (iii) best familiarity with the 4D seismic method which is preferred to any other approach. If the first two are related to the non-uniqueness of the modeling solution, the third is related to the results offered by modern 3D and 4D seismic surveys, even if the method has also limitations in locating fluid content, porosity, and saturation changes (Devriese and Oldenburg, 2014).

In this work, we calculate the gravity response of small density changes in a small and deep reservoir with a size/depth ratio slightly lower than 1. This scenario represents the worst case for the application of gravity methods to locate density changes related to fluid production or injection. Through 3D forward modeling of the reservoir geometries as obtained from seismic and borehole data, we model production-related density changes to retrieve the gravity response before and after production. Modeling is constrained by volumes and densities of produced and injected fluids at reservoir conditions.

DATA AND METHODS

The Volve Field Database

The Volve field is located in the Central North Sea (**Figure 1A**), and it was discovered in 1993 by the well 15/9-19 SR. It is located ~5 km north of the Sleipner East field (e.g., Alnes et al., 2008, 2011) and production started in February 2008 (Production data report in the Volve field dataset). Oil and gas production ended in September 2016 after a total production of 1.5×10^9 Sm^3 including oil, gas, and formation water. Among the several wells drilled in the field, the majority of the production (~98%) was achieved by wells 15/9-11, 15/9-12, and 15/9-14 (**Figure 1**). The Volve reservoir is located in the Hugin Jurassic sandstones between 2750 and 3120 m total vertical depth (TVD) below sea level (Discovery report in the Volve field dataset). Hugin sands were deposited in a nearshore marine setting, show high total organic carbon (up to 80%), and are gas prone with the capability to generate also light liquid hydrocarbons (e.g., Isaksen et al., 2002). According to well data in the Volve field, the thickness of the Hugin formation can range between a few tens of meters up to ~150 m where all the 18 levels of the formation have been drilled. Production at Volve was sustained by formation water injection of a total volume of 3×10^7 Sm^3 from the wells 15/9-F-4 and 15/9-F-5 that were active between April 2008 and September 2016 (Production data report in the Volve field dataset).

After shut down and removal of the production equipment in summer 2017, the entire dataset regarding seismic volumes, well logs, and petrophysical and geological characterization of the field was disclosed by Statoil (now Equinor) in June 2018[1].

The Volve dataset encompasses over 5 Tbytes of data counting over 11000 files including both raw data and interpretations of the formation tops and faults. It includes also the 3D models used for production simulations as resulting from the seismic data interpretation and well logs. Locations of production and injection wells, daily and monthly reports on production, and injection activities are available as well. Finally, geochemical analyses of formation fluids at reservoir conditions of pressure and temperature are provided together with detailed characterization of the involved formations. The availability for the scientific community of such a detailed and complete database regarding a reservoir system and its production history represents a unique opportunity and allows the community for unprecedented attempts in testing modeling techniques.

Modeling Procedure

In order to provide an input geometry for the forward gravity model, we first created a 3D model of the reservoir using Hugin top and bottom surfaces as provided in the Volve dataset. The surfaces of the main faults in the area were depth-converted as they are provided as two-way travel-times (TWT) in the dataset only. This step was carried out using TWT-depth data provided in well logs and reports on well activities included in the dataset. Finally, the 3D model of the reservoir was obtained by intersection of the top and bottom of the Hugin formation with the bounding faults (**Figure 1B**). We chose to not include other structures in the model to keep the geometry as simple as possible; moreover, we assume that all bounding faults are sealing – i.e., all density changes are contained within the reservoir volume.

Once the geometry of the reservoir was defined, we discretized the $\sim 4.3 \times 10^9$ m^3 Hugin volume in 275642 cubic cells with size 25 m × 25 m × 25 m and included it in a larger 3D mesh with same cell size. In this geometric model, the only surface included together with the Hugin volume is the sea bottom representing the major density contrast between sea water and the underlying rock volume (**Figure 2**). Despite the seafloor effect is obviously removed by the 4D approach, it was included in the single 3D models to evaluate the gravity signature related to bathymetry. Seafloor is found between 85 and 100 m below sea level, and the bathymetry shows a main step (~10 m) in the eastern area of the modeled region and a gentle westward descending gradient across the entire area (**Figure 1B**).

Within the geometric model, we set density values according to the petrophysical and geochemical data recovered from the reports available in the Volve database. In particular, by averaging the density values provided for the units surrounding Hugin, we set a bulk reference density for the volume of 2670 kg $m^{?3}$. This value was used to calculate density contrasts of sea water volume (-1640 kg m^{-3}) and the Hugin volume (-70 kg m^{-3}) – i.e., we assume that seawater and Hugin have a density of 1030 and 2600 kg m^{-3}, respectively. The density contrasts were used as input for a pre-production 3D forward model.

---

[1] https://www.equinor.com/en/how-and-why/digitalisation-in-our-dna/volve-field-data-village-download.html

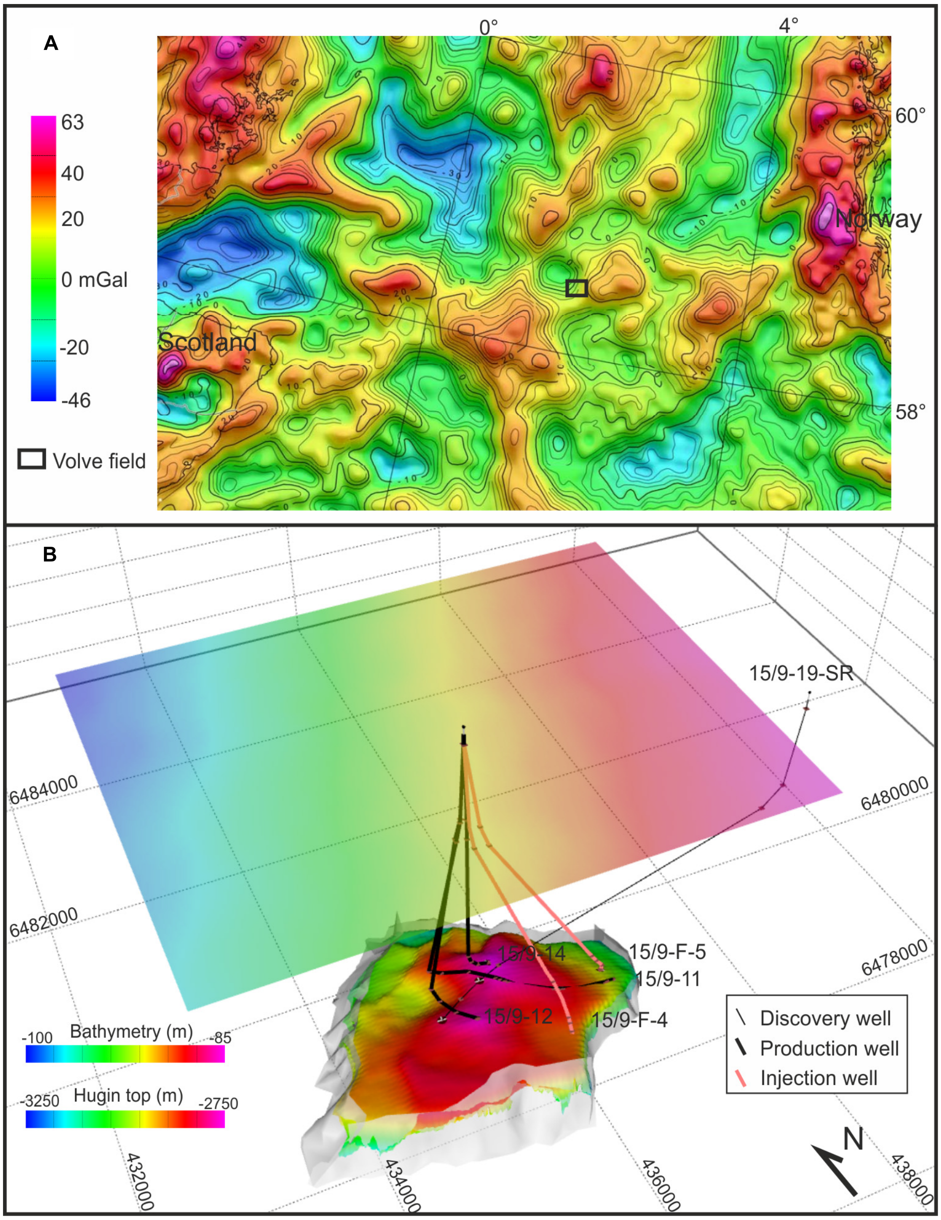


FIGURE 1 | (A) Free-air anomaly at sea and Bouguer anomaly on land over the study region (mod. from Olesen et al., 2010). **(B)** 3D model of the Volve field reservoir. Light gray surfaces represent bounding faults used to delimit Hugin top and bottom surfaces. Spacing between gray lines in the bounding box provides horizontal and vertical scales: spacing in the X–Y plane is 2 km while spacing in depth is 0.5 km. Coordinates are in ED50 UTM 31N.

The procedure to calculate the forward gravity models of the volumes is based on the algorithm proposed by Li and Oldenburg (1998).

The vertical component of the gravity field produced by the density $\rho(x,y,z)$ is given by:

$$g_z(r_0) = \gamma \int_V \rho(r) \frac{z - z_0}{|r - r_0|^3} \, dv \tag{1}$$

where V is the anomalous mass volume, $r_0 = (x_0, y_0, z_0)$ is a vector locating the observation point, $r = (x, y, z)$ locates the source, and γ is the gravitational constant.

Assuming a constant density contrast within each prismatic cell in a 3D orthogonal mesh, the gravity field at the ith observation location is given by:

$$g_z(r_{0i}) = \sum_{j=1}^{M} \rho_j \left\{ \gamma \int_{\Delta V_j} \frac{z - z_0}{|r - r_{0i}|^3} \, dv \right\} \tag{2}$$

where ρ_j and ΔV_j are the density and volume of the jth cell, respectively.

To evaluate density changes induced by fluid production and injection, we recovered data from the activity reports

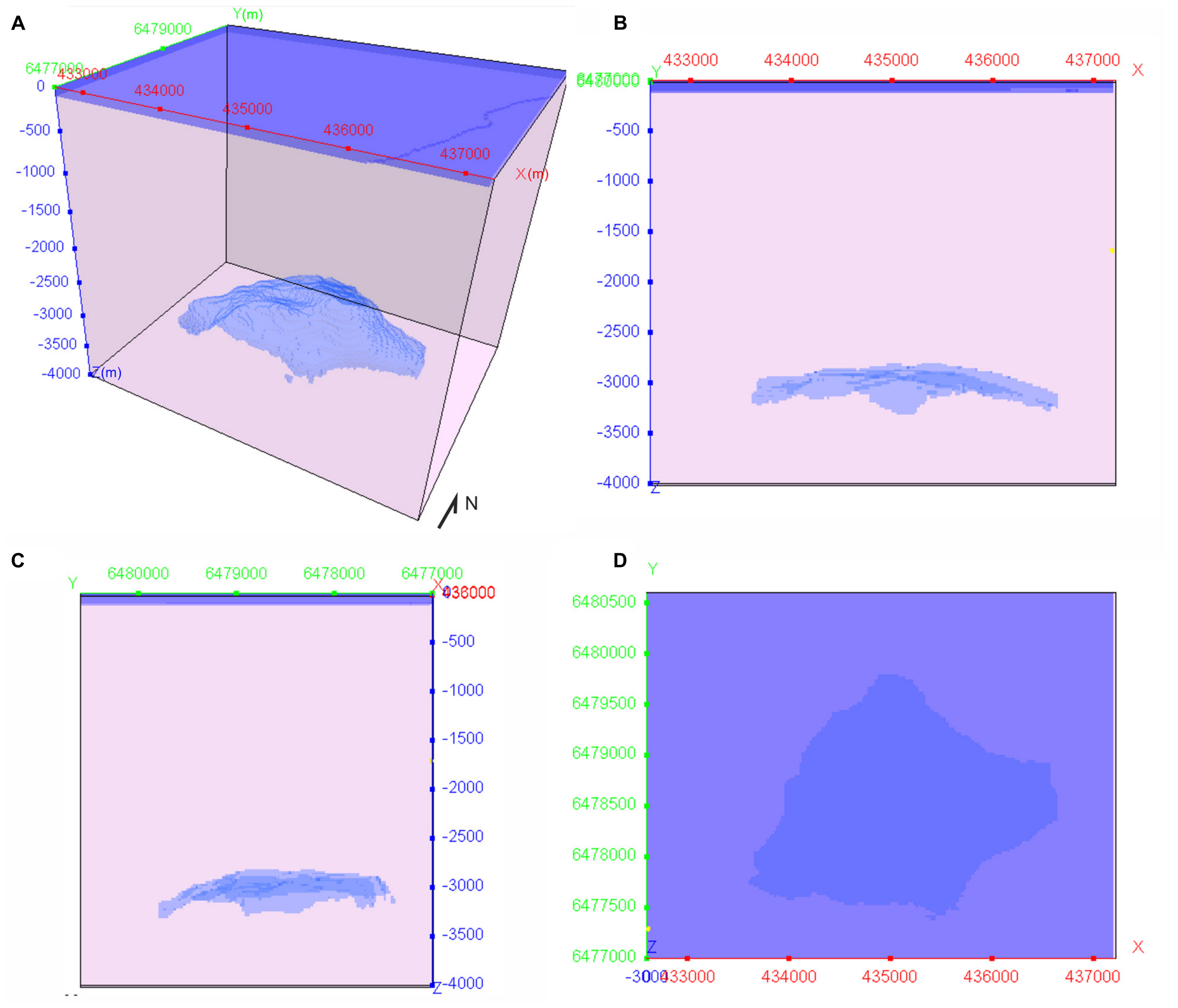


FIGURE 2 | Geometries of the Hugin reservoir and sea water (blue volumes) within the entire model. **(A)** Perspective, **(B)** northward, **(C)** eastward and **(D)** downward view of the mesh. All scales are in meters. This mesh was used for the pre- and post-production 3D forward models. Cell size is 25 m × 25 m × 25 m. *Y* axis locates geographic north. Coordinates are in the same system as in **Figure 1**.

provided in the Volve dataset. The total reported volume balance accounting for 1.5 × 10^9 Sm^3 of volumes produced and 30 × 10^6 Sm^3 of injected water is 1.47 × 10^9 Sm^3 (Sm^3 stands for standard cubic meters and is referred to 15°C and 1010 × 10^2 Pa) including oil, gas, and water production and injection. At reservoir conditions (106°C and 3.28 × 10^7 Pa), this volume becomes 3.21 × 10^8 m^3. At the same temperature and pressure conditions, the fluids in the reservoir have a density of 710 kg m^{-3} as reported in the geochemical analyses of samples from the discovery well 15/9-19-SR and included in the Volve dataset. Thus, the mass loss in the reservoir, accounting for production and injection activities and reservoir conditions, is of 2.28 × 10^{11} kg. Considering that injection of water was contemporaneous to production (wells 15/9-F-4 and 15/9-F-5 were active between April 2008 and September 2016), we assume zero compaction of the Hugin sandstones during this period.

If a homogeneous mass loss across the entire reservoir is assumed, after the production each cell in the discretized Hugin volume should weigh 8.27 × 10^5 kg less than before production – i.e., the total mass loss divided by the total number of cells. In other words, knowing the volume of each cell (1.56 × 10^4 m^3), the production and injection activities resulted in a reservoir formation density loss of ∼53 kg m^{-3}. We note here that this configuration of the post-production model provides no preferential "paths" for density variations within the reservoir because it is not including constraints regarding injection and production wells' locations. Finally, the achieved density loss value is used to correct the pre-production density of the Hugin volume and perform a post-production 3D gravity forward model of the total volume that is compared with the pre-production model to evaluate gravity changes related to the production activity.

RESULTS

Figure 3 shows the gravity signatures obtained after 3D forward calculation of the pre-production model (**Figure 3a**) and of the post-production model (**Figure 3b**) and the differential signature obtained by removing the forward-calculated gravity of the pre-production gravity from the post-production (**Figure 3c**). All these maps are produced with regular station spacing of 50 m located on sea surface, but similar results are obtained using a double station spacing of 100 m.

Despite the results of the pre- and post-production, forward models are extremely similar in the way that it would not be possible to qualitatively locate any local undulation; the difference between these two gravity data highlights a $\sim -13\ \mu\text{Gal}$ ($-130 \times 10^{-9}\ \text{m s}^{-2}$) minimum centered at 435256 E, 6478496 N (ED50, UTM 31N). Considering the modeled volume and the imposed density change, we interpret this minimum as representing the gravity effect of oil, gas, and water production and injection activities at the Volve field between February 2008 and September 2016.

The resulting residuals are in the same order of magnitude than those retrieved in the Sleipner field through gravity data acquisition (e.g., Alnes et al., 2008, 2011).

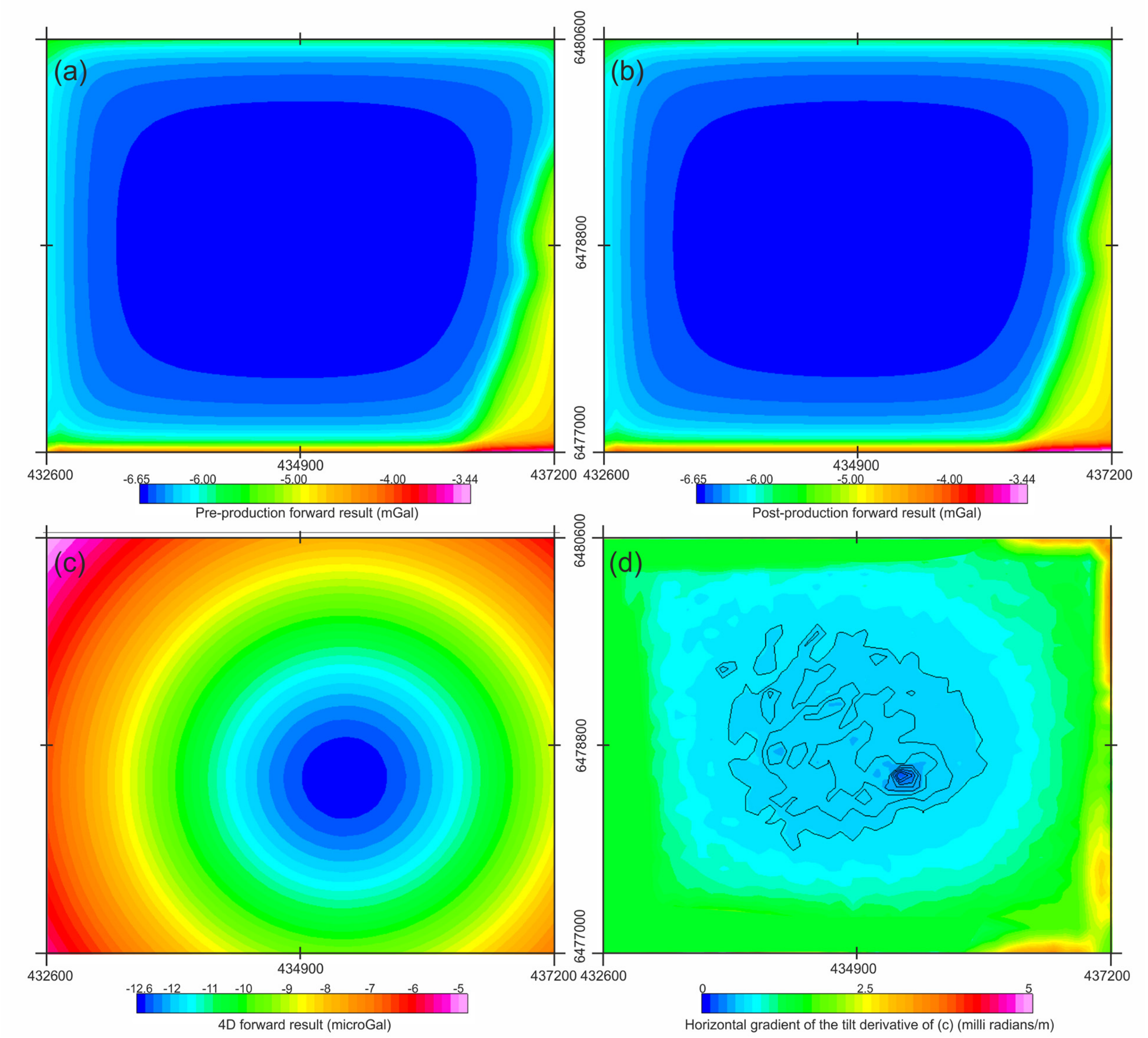


FIGURE 3 | Forward-calculated gravity signature of the pre-production **(a)** and post-production **(b)** models. **(c)** Differential gravity signature calculated by subtraction of pre-production **(a)** from post-production gravity **(b)**. Spacing between stations for forward calculations is 50 m. **(d)** Horizontal gradient of the tilt derivative (THDR) computed from **(c)**. Contours locate areas with value <1 milliradians m^{-1}.

In a final processing step, we calculate the tilt derivative (Miller and Singh, 1994a; Verduzco et al., 2004) of the anomaly in **Figure 3c**. The horizontal gradient of the tilt derivative (THDR) is computed in order to locate zero values of the THDR representing the boundary of the causative source (e.g., Miller and Singh, 1994b). Projecting the calculated THDR (**Figure 3d**) on the reservoir top (**Figure 4**), we note that the THDR values ≈0 are surrounded by the most producing wells – i.e., 15/9-11, 15/9-12, and 15/9-14 whose production accounts for more than 98% of the total volume (**Figure 4B**).

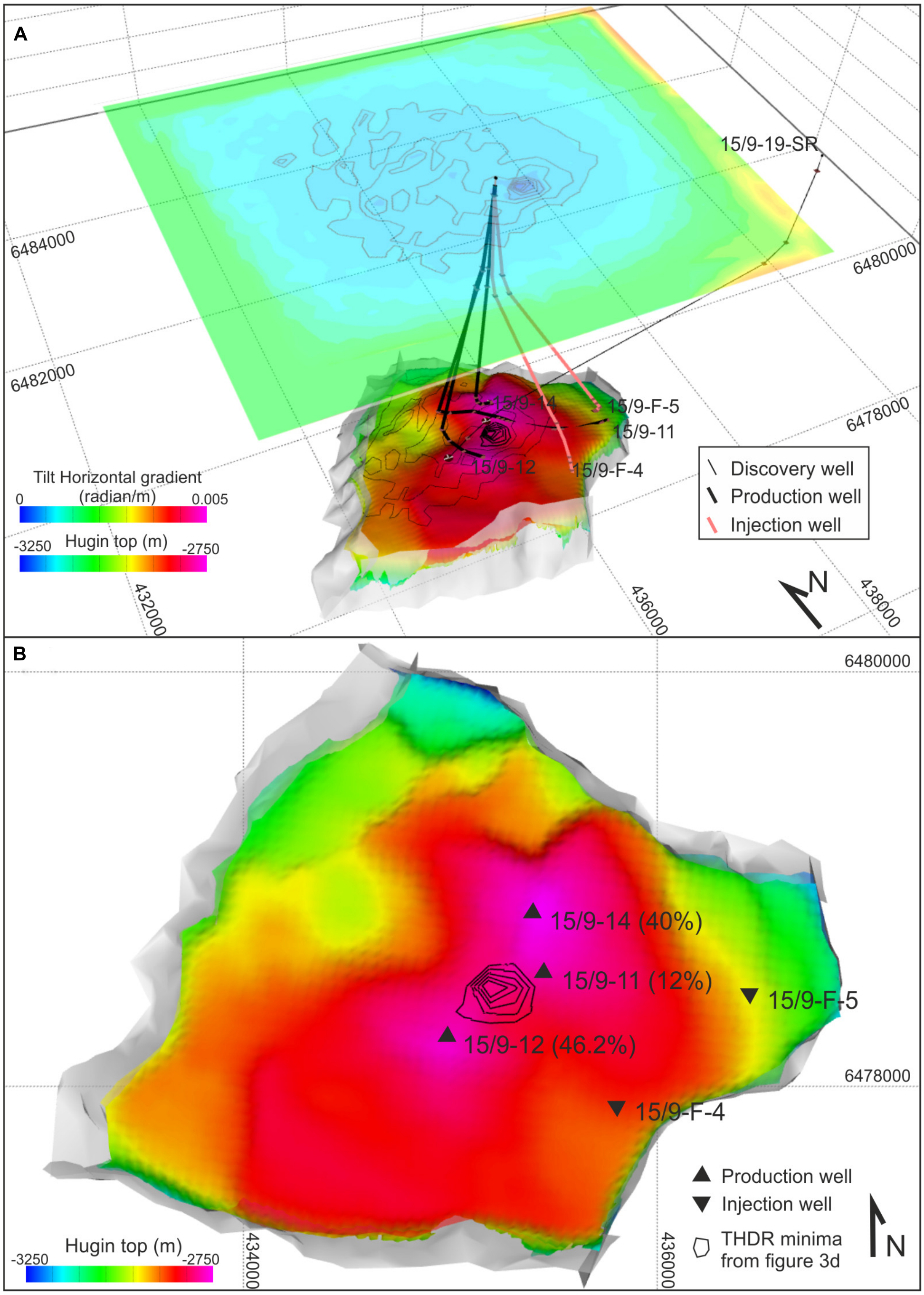


FIGURE 4 | (A) Three-dimensional view of the THDR over the Volve field. The contoured values (thin black lines) were projected to the top Hugin at depth for spatial comparison with reservoir geometry and wells locations. **(B)** Map view of the reservoir top (i.e., top of the Hugin formation) compared with THDR values <0.6 milliradians m^{-1} and locations of the contacts between well tracks and top Hugin (triangles). Percentage values within parentheses represent the production over total mass balance achieved by each well. Contour lines are spaced 0.1 milliradians m^{-1}. Horizontal and vertical scales are the same as in **Figure 1**.

This evidence supports the former interpretation concerning the gravity minima observed in **Figure 3c** being caused by reservoir production and injection activities.

DISCUSSION

The 4D gravity approach is not a new idea to monitor fluid-related density variations, but its application has always relied on high-precision gravity data acquisitions repeated in time and on source depth <2500 m (e.g., Ferguson et al., 2007; Vasilevskiy and Dashevsky, 2007; Alnes et al., 2008, 2011; Elliott and Braun, 2016). In this work, we applied a simple modeling procedure to evaluate the 4D gravity response computed from 3D forward models repeated at pre-production (February 2008) and post-production (September 2016) time steps. This procedure allowed to evaluate the gravity effects of pre-determined density changes in the >2750 m deep Volve reservoir without gravity data acquisition. The procedure was constrained by production data and geometry of the reservoir resulting from seismic and well data.

We assumed sealing conditions of the bounding faults and isotropy of the reservoir volume. Without detailed data about fault behavior and anisotropies within the reservoir levels, we made these assumptions to confine the mass variation within the modeled reservoir and to homogeneously distribute the mass change within the entire reservoir volume. However, the reference volume can be further parametrized in order to locate density changes according to available data or simulated scenarios. For example, to investigate eventual effects on the spatial distribution of density changes after introducing a leaking fault in the model, the reservoir volume should be extended beyond the investigated leaking fault and be parametrized with proper pre-production and post-production density contrasts. Furthermore, volume parametrization allows to model also eventual noise sources – e.g., topography/bathymetry, local density changes in surrounding volumes, as long as these are quantifiable by density or volume changes within the modeled period.

Based on the well-known algorithm proposed by Li and Oldenburg (1998), we modeled the gravity signature without filtering any regional or topographic components that are naturally removed by the differential approach. The only requirements are given by well-constrained geometries of the reservoir, availability of data about produced and injected volumes, and petrophysical properties of the fluids and reservoir formation. All these data are normally produced for oil and gas field characterization.

Interestingly, the procedure locates the causative source of the residual anomaly (**Figure 3c**), within the area of maximum production, where THDR values ≈ 0 are surrounded by the three most productive wells (**Figure 4**). In general, it is not surprising that a gravity low is centered over its causative source's top as we observed in the Volve field (**Figure 4**). However, what is surprising in this case is the observation *per se* of a minima being caused by ~-53 kg m^{-3} density change at reservoir conditions and whose source is below 2750 m depth. In fact, the Volve field represents a challenging study case for the gravity method because, with a size/depth ratio slightly below the critical value of 1, gravity signatures related to density changes within the modeled volume should not be detectable because of the rapid attenuation of the gravity signal with increasing distance between the observation point and the source – e.g., Stenvold et al. (2008) and references therein.

This outcome strengthens the conclusions obtained by Stenvold et al. (2008) by expanding the retrievable depth for fluid-related density changes below the threshold of 2500 m. Furthermore, the Volve case modeling demonstrates that even without a detailed parametrization of the reservoir units, the simple forward calculation based on reservoir geometry and produced and injected volumes may prove reliable. The applicability of this method depends on the size/depth ratio of the source and on the magnitude of the density change both in terms of the density contrast imposed in the single cell and of thickness of the modeled reservoir carrying a density contrast. Approximating the reservoir shape to a horizontal slice of a cylindrical source (Telford et al., 1990; Stenvold et al., 2008) with radius 1.5 km, the imposed density change of 53 kg m^{-3} at a depth of 2750 m should produce a 10 μGal gravity change with reservoir thickness of ∼40 m (eq. 1 in Stenvold et al., 2008). Following the same approximation, a reservoir thickness of 100 m with the same density change and the same radius should produce a gravity change of ∼10 μGal at a depth of ∼4000 m. Of course, these approximations are not valid when dealing with complex structures and morphologies such as those often found in real geological contexts like the Volve field.

Considering the position of the gravity minimum and that the post-production input model assumed a homogeneous density change within the entire reservoir volume without constraints regarding injection and production wells, this outcome strengthens the reliability of the retrieved gravity anomaly. Moreover, we expect this approach to retrieve higher amplitude differential gravity residuals when dealing with higher density contrasts than those modeled at the Volve field. In fact, a scenario where reservoir fluids (typically oil, water, and gas) are replaced by CO_2 sequestration or gas storage should produce higher density contrasts and in turn should allow to resolve density changes related to gas propagation better.

CONCLUSION

In this work, we successfully locate the causative source for -53 kg m^{-3} density changes due to fluid production from a reservoir by 3D forward calculation of gravity at pre- and post-production time steps. The Volve field was chosen as test site because it represents a challenging reservoir with a size/depth ratio slightly below 1, and in these conditions, gravity methods should not allow detection of fluid-induced small density changes. Moreover, the Volve field represents a unique opportunity because in this case, data that usually are kept confidential also after exploitation have been made available to the community allowing the parametrization of the models.

The density change used as input for the post-production model, once removed the pre-production gravity signature, resulted in differential gravity residuals with a -13 μGal minimum above the Volve field. Unsurprisingly, despite the input density variation was distributed within the entire reservoir volume, THDR minimum located on top of the reservoir, between the three most productive wells in the field, whose total production accounted for more than 98% of total mass balance.

This work demonstrates that the 4D gravity method, whether it is constrained to gravity data acquisitions or to forward modeling of known geometries and density changes, represents a reliable technique even in challenging cases. Interestingly, the production-related anomalies observed at 2750 m depth in the Volve field should be theoretically retrievable by seafloor gravimeters with 5 μGal accuracy. However, this goal is still difficult to achieve and likely requires significant efforts in terms of number of stations in order to cover small sources like the Volve field.

Finally, relying on the availability of data that are confidential in the case of oil fields or very expensive to produce for research purposes, this approach has intrinsic limitations in the possibility of being tested in more complex geological scenarios. However, the owners of such data should follow the example of Equinor and release data regarding fields that are no longer exploited; this will likely open new ways in understanding geological and geophysical processes.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: Volve data village page: https://www.equinor.com/en/how-and-why/digitalisation-in-our-dna/volve-field-data-village-download.html.

AUTHOR CONTRIBUTIONS

PM conceived the study, produced the models, and wrote the manuscript.

FUNDING

This work was supported by the Chieti-Pescara University funds to PM.

ACKNOWLEDGMENTS

The author thanks the Associate Editor and the two reviewers for their critical reading of the manuscript and constructive comments that improved the manuscript. The author warmly thanks Equinor for providing the Volve dataset and Maurizio Fedi for his constructive comments on a preliminary version of this work.

REFERENCES

Alnes, H., Eiken, O., Nooner, S., Sasagawa, G., Stenvold, T., and Zumberge, M. (2011). Results from sleipner gravity monitoring: updated density and temperature distribution of the CO2 plume. *Energy Procedia* 4, 5504–5511. doi: 10.1016/j.egypro.2011.02.536

Alnes, H., Eiken, O., and Stenvold, T. (2008). Monitoring gas production and CO2 injection at the Sleipner field using time-lapse gravimetry. *Geophysics* 73, WA155–WA161. doi: 10.1190/1.2991119

Blakely, R. J. (1996). *Potential Theory in Gravity and Magnetic Applications*. Cambridge: Cambridge University Press.

Cooper, C., Dodds, K., and McKnight, R. (2009). "Monitoring programs for CO2 storage," in *A Technical Basis for Cabon Dioxide Storage. CO2 Capture Project*, ed. C. Cooper (Amsterdam: Elesevier).

Davis, K., Li, Y., and Batzle, M. (2008). Time-lapse gravity monitoring: a systematic 4D approach with applications to aquifer storage and recovery. *Geophysics* 73, WA61–WA69.

Devriese, S. G., and Oldenburg, D. W. (2014). "Enhanced imaging of SAGD steam changers using broadband electromagnetic surveying," in *Proceedings fo the SEG Denver 2014 Annual Meeting* (Tulsa, OK: SEG) 765–769.

Dressel, I., Barckhausen, U., and Heyde, I. (2018). A 3D gravity and magnetic model for the Entenschnabel area (German North Sea). *Int. J. Earth Sci.* 107, 177–190. doi: 10.1007/s00531-017-1481-x

Eiken, O., Stenvold, T., Zumberge, M., Alnes, H., and Sasagawa, G. (2008). Gravimetric monitoring of gas production from the Troll feld. *Geophysics* 73, WA149–WA154.

Eiken, O., Zumberge, M. A., and Sasagawa, G. S. (2000). "Gravity monitoring of offshore gas reservoirs," in *Proceedings of the 70th Annual International Meeting, SEG, Expanded Abstracts* (Tulsa, OK: SEG) 431–434.

Elliott, E. J., and Braun, A. (2016). Gravity monitoring of 4D fluid migration in SAGD reservoirs – forward modeling. *CSEG Rec.* 41, 16–21.

Elliott, E. J., and Braun, A. (2017). On the resolvability of steam assisted gravity drainage reservoirs using time-lapse gravity gradiometry. *Pure Appl. Geophys.* 174, 4119–4136. doi: 10.1007/s00024-017-1636-1635

Fedi, M., Cella, F., D'Antonio, M., Florio, G., Paoletti, V., and Morra, V. (2018). Gravity modeling finds a large magma body in the deep crust below the gulf of Naples, Italy. *Sci. Rep.* 8:8229. doi: 10.1038/s41598-018-26346-z

Ferguson, J., Klopping, F., Chen, T., Seibert, J., Hare, J., and Brady, J. (2008). The 4D microgravity method for waterflood surveillance: Part III – 4D absolute microgravity surveys at Prudhoe Bay. Alaska. *Geophysics* 73, WA163–WA171.

Ferguson, J. F., Chen, T., Brady, J. L., Aiken, C. L. V., and Seibert, J. E. (2007). The 4D microgravity method for waterflood surveillance II — gravity measurements for the Prudhoe Bay reservoir, Alaska. *Geophysics* 72, I33–I43. doi: 10.1190/1.2435473

Gasperikova, E., and Hoversten, G. M. (2008). Gravity monitoring of CO2 movement during sequestration: model studies. *Geophysics* 73, WA105–WA112. doi: 10.1190/1.2985823

Hare, J. L., Ferguson, J. F., Aiken, C. L. V., and Brady, J. L. (1999). The 4-D microgravity method for waterflood surveillance: a model study for the Prudhoe Bay reservoir, Alaska. *Geophysics* 64, 78–87. doi: 10.1190/1.1444533

Isaksen, G. H., Patience, R., van Graas, G., and Jenssen, A. I. (2002). Hydrocarbon system analysis in a rift basin with mixed marine and nonmarine source rocks: the south Viking Graben. North Sea. *AAPG Bull.* 86, 557–591.

Jacob, T., Bayer, R., Chery, J., and Le Moigne, N. (2010). Time-lapse microgravity surveys reveal water storage heterogeneity of a karst aquifer. *J. Geophys. Res.* 115, 1–18. doi: 10.1029/2009JB006616

Krahenbuhl, R., and Li, Y. (2012). Time-lapse gravity: a numerical demonstration using robust inversion and joint interpretation of 4D surface and borehole data. *Geophysics* 77, G33–G43.

Krahenbuhl, R. A., Li, Y., and Davis, T. (2011). Understanding the applications and limitations of time-lapse gravity for reservoir monitoring. *Lead. Edge* 30, 1060–1068. doi: 10.1190/1.3640530

Li, Y., and Oldenburg, D. W. (1998). 3-D inversion of gravity data. *Geophysics* 63, 109–119.

Mancinelli, P., Pauselli, C., Fournier, D., Fedi, M., Minelli, G., and Barchi, M. R. (2020). Three dimensional gravity local inversion across the area struck by the 2016–2017 seismic events in central Italy. *J. Geophys. Res.* 125:e2019JB018853. doi: 10.1029/2019JB018853

Mancinelli, P., Pauselli, C., Minelli, G., and Federico, C. (2015). Magnetic and gravimetric modeling of the central Adriatic region. *J. Geodynamics* 89, 60–70. doi: 10.1016/j.jog.2015.06.008

Mancinelli, P., Porreca, M., Pauselli, C., Minelli, G., Barchi, M. R., and Speranza, F. (2019). Gravity and magnetic modeling of central Italy: insights into the depth extent of the seismogenic layer. *Geochem. Geophys. Geosyst.* 20, 2157–2172.doi: 10.1029/2018GC008002

Ménoret, V., Vermeulen, P., Le Moigne, N., Bonvalot, S., Bouyer, P., Landragin, A., et al. (2018). Gravity measurements below 10-9 g with a transportable absolute quantum gravimeter. *Sci. Rep.* 8:12300. doi: 10.1038/s41598-018-30608-30601

Miller, H. G., and Singh, V. (1994a). Semiquantitative techniques for the removal of directional trends from potential field data. *J. Appl. Geophys.* 32, 199–211. doi: 10.1016/0926-9851(94)90021-3

Miller, H. G., and Singh, V. (1994b). Potential field tilt-a new concept for location of potential field sources. *J. Appl. Geophys.* 32, 213–217. doi: 10.1016/0926-9851(94)90022-1

Olesen, O., Ebbing, J., Gellein, J., Kihle, O., Myklebust, R., Sand, M., et al. (2010). *Gravity Anomaly Map, Norway and Adjacent Areas. Scale 1:3 Million*. Trondheim: Geological survey of Norway.

Reitz, A., Krahenbuhl, R., and Li, Y. (2015). Feasibility of time-lapse gravity and gravity gradiometry monitoring for steam-assisted gravity drainage reservoirs. *Geophysics* 80, WA99–WA110. doi: 10.1190/GEO2014-0217.1

Sasagawa, G. S., Crawford, W., Eiken, O., Nooner, S., Stenvold, T., and Zumberge, M. A. (2003). A new sea-floor gravimeter. *Geophysics* 68, 544–553. doi: 10.1190/1.1567223

Sobh, M., Ebbing, J., Mansi, A. H., and Götze, H.-J. (2019). Inverse and 3D forward gravity modelling for the estimation of the crustal thickness of Egypt. *Tectonophysics* 752, 52–67. doi: 10.1016/j.tecto.2018.12.002

Stenvold, T., Eiken, O., and Landro, M. (2008). Gravimetric monitoring of gas-reservoir water influx – A combined flow- and gravity-modeling approach. *Geophysics* 73, WA123–WA131. doi: 10.1190/1.2991104

Telford, W. M., Geldart, L. P., and Sheriff, R. E. (1990). *Applied Geophysics*, 2nd Edn. Cambridge: Cambridge University Press.

Van Camp, M., de Viron, O., Watlet, A., Meurers, B., Francis, O., and Caudron, C. (2017). Geophysics from terrestrial time-variable gravity measurements. *Rev. Geophys.* 55, 938–992. doi: 10.1002/2017RG000566

Vasilevskiy, A., and Dashevsky, Y. (2007). "Feasibility study of 4D microgravity method to monitor subsurface water and gas movements," in *Proceedings of the 77th Annual International Meeting, SEG, Expanded Abstracts* (Tulsa, OK: SEG) 816–820.

Verduzco, B., Fairhead, J. D., Green, C. M., and MacKenzie, C. (2004). New insights into magnetic derivatives for structural mapping. *Lead. Edge* 2004, 116–119. doi: 10.1190/1.1651454

Wilson, C. R., Scanlon, B., Sharp, J., Longuevergne, L., and Vu, H. (2012). Field test of the superconducting gravimeter as a hydrologic sensor. *Ground Water* 50, 442–449. doi: 10.1111/j.1745-6584.2011.00864.x

Conflict of Interest: 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.

frontiers
in Earth Science

ORIGINAL RESEARCH
published: 14 July 2020
doi: 10.3389/feart.2020.00255

Advanced Analysis of Temporal Data Using Fisher-Shannon Information: Theoretical Development and Application in Geosciences

Fabian Guignard[1]*, **Mohamed Laib**[2], **Federico Amato**[1] **and Mikhail Kanevski**[1]

[1] Faculty of Geosciences and Environment, Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland, [2] Department of Information Technologies for Innovative Services, Luxembourg Institute of Science and Technology—LIST, Belvaux, Luxembourg

Complex non-linear time series are ubiquitous in geosciences. Quantifying complexity and non-stationarity of these data is a challenging task, and advanced complexity-based exploratory tool are required for understanding and visualizing such data. This paper discusses the Fisher-Shannon method, from which one can obtain a complexity measure and detect non-stationarity, as an efficient data exploration tool. The state-of-the-art studies related to the Fisher-Shannon measures are collected, and new analytical formulas for positive unimodal skewed distributions are proposed. Case studies on both synthetic and real data illustrate the usefulness of the Fisher-Shannon method, which can find application in different domains including time series discrimination and generation of times series features for clustering, modeling and forecasting. The paper is accompanied with Python and R libraries for the non-parametric estimation of the proposed measures.

Keywords: Fisher-Shannon complexity, Fisher-Shannon information plane, Shannon entropy power, Fisher information measure, statistical complexity, non-linear time series, dynamical behavior characterization, high frequency wind speed

Edited by:
Juergen Pilz,
Alpen-Adria-Universität Klagenfurt,
Austria

Reviewed by:
Hwa-Lung Yu,
National Taiwan University, Taiwan
Luis Gomez,
University of Las Palmas de Gran
Canaria, Spain

****Correspondence:***
Fabian Guignard
fabian.guignard@unil.ch

Specialty section:
This article was submitted to
Environmental Informatics and
Remote Sensing,
a section of the journal
Frontiers in Earth Science

Received: *04 December 2019*
Accepted: *09 June 2020*
Published: *14 July 2020*

Citation:
Guignard F, Laib M, Amato F and
Kanevski M (2020) Advanced Analysis
of Temporal Data Using
Fisher-Shannon Information:
Theoretical Development and
Application in Geosciences.
Front. Earth Sci. 8:255.
doi: 10.3389/feart.2020.00255

1. INTRODUCTION

The ubiquity and extensive growth of available temporal data requires the development of reliable techniques to extract knowledge from them and to understand multifaceted time-dependent phenomena. Over the last decades, an increasing attention was payed toward the use of Fisher-Shannon information as a measure to characterize the complexity and non-stationarity of non-linear time series. Originally proposed for statistical estimation purposes (Fisher, 1925), the Fisher information measure (FIM) has been extensively used in theoretical physics (Frieden, 1990). FIM and Shannon entropy power (SEP) (Shannon, 1948) are closely related, as shown by information theory (Dembo et al., 1991; Cover and Thomas, 2006). The Fisher-Shannon complexity (FSC)—the FIM and SEP product—was proposed as a possible definition of atom complexity (Angulo et al., 2008; Esquivel et al., 2010).

Following Frieden work, FIM has found applications in non-linear time-series analysis. Martin et al. (1999) analyzed complex non-stationary electroencephalographic signals and showed that FIM can have better discrimination performance than Shannon entropy. FIM was also used to detect behavior changes of dynamical systems (Martin et al., 2001). Vignat and Bercher (2003) showed that a joint analysis of both SEP and FIM can be required to perform effective discrimination of non-stationary signals.

The Fisher-Shannon method has been used to analyse complex dynamical processes in geophysics. Discrimination between the electric and magnetic components of magnotelluric signals is performed in Telesca et al. (2011). Tsunamigenic and non-tsunamigenic earthquakes were efficiently separated in the Fisher-Shannon information plane, using FSC (Telesca et al., 2013). Micro-tremors time series were identified depending on the soil characteristics of the measurement sites (Telesca et al., 2015b). Telesca et al. (2015a) proposed a classifier of (non-)tsunamigenic potential of earthquake build on several time series features, including FIM, SEP, FSC. Finally, FIM was also used dynamically with sliding window techniques in order to study precursory patterns in seismology (Telesca et al., 2009b) and volcanology (Telesca et al., 2010).

Many environmental processes have also been studied using the Fisher-Shannon method. Lovallo et al. (2013) and Pierini et al. (2011) studied climatic regimes identification in rainfall time series. Hydrological regimes discrimination have also been investigated (Pierini et al., 2015). Analyzing remotely sensed sea surface temperature, Pierini et al. (2016) have shown that the Fisher-Shannon method is able to clearly identify the Brazil-Malvinas Confluence Zone, which is known to be one of the most energetic area of oceans. Telesca and Lovallo (2011) analyzed more than 10 years of hourly wind speed data in the Fisher-Shannon information plane. The same authors studied yearly variation of the FIM, the SEP and the FSC on wind measurements (Telesca and Lovallo, 2013). Guignard et al. (2019b) have found correlations between daily variance of temperature and daily FSC of high-frequency wind speed records in urban area. Authors have also pointed out relationships between Fisher-Shannon analysis of wind speed daily means and topographical features—height and slope—in complex mountainous regions (Guignard et al., 2019a). Telesca et al. (2009a) discriminated some pollutants, including cadmium, iron, and lead, in the Fisher-Shannon plane. Similarly, Amato et al. (2020) have shown a relationship between the Fisher-Shannon analysis outputs of three air pollutants—Nitrogen dioxide, Ground level ozone and Particulate Matter—and measurement location in term of land use and of anthropogenic sources of pollutant emission.

The research involving Fisher-Shannon method is rather scattered and comes from various fields, e.g., information theory, physics, dynamical systems, machine learning, and statistics. Therefore, the present paper contributes to the methodological studies on Fisher-Shannon information measures along with some applications.

The main objectives of this research can be summarized as follows:

- discussing the state-of-the-art of Fisher-Shannon information measures and their applications,
- identifying FSC as a sensitivity measure of the SEP and as a scale-independent non-Gaussianity measure of data,
- presenting some new theoretical results on FIM and SEP,
- developing operational FIM and SEP tools for the nonlinear time-series analysis,
- demonstrating through two case studies, based on simulated (chaotic) and real data (high frequency wind speed measurements), the efficiency and usefulness of the proposed methods.

The remainder of the paper is organized as follows. Concepts of Fisher-Shannon analysis, including SEP, FIM, FSC, and information plane, are presented and reviewed in section 2. Section 3 provides analytical formula for such quantities in the particular cases of random variables following some positive skewed distributions, namely Gamma, Weibull, and log-normal ones. Then, a non-parametric kernel based estimation of the density function—for which Python and R packages are proposed—is presented in section 4. Experiments on simulated and real-world data are performed in section 5. Finally, section 6 concludes the paper.

2. FISHER-SHANNON ANALYSIS

2.1. Shannon Entropy Power and Fisher Information Measure

Let us consider a univariate continuous random variable X with its probability density function (PDF) $f(x)$, which is supposed to be sufficiently regular for the exposition of our purpose. Its *differential entropy* (Cover and Thomas, 2006) is defined as

$$H_X = \mathbb{E}\left[-\log f(X)\right] = -\int f(x)\log f(x)\,dx. \tag{1}$$

For example, if X is a centered Gaussian random variable of variance σ^2, a direct computation gives $H_X = \frac{1}{2}\log(2\pi e\sigma^2)$. However, it will be more convenient to work with the following quantity, called the *Shannon Entropy Power* (SEP) (Dembo et al., 1991),

$$N_X = \frac{1}{2\pi e}e^{2H_X}, \tag{2}$$

which is a strictly monotonically increasing transformation of H_X. The SEP is constructed such that in the Gaussian case we have $N_X = \sigma^2$. Very often, entropies H_X and N_X are interpreted as global measures of disorder / uncertainty / spread of $f(x)$. The higher the entropy, the higher the disorder.

The *Fisher Information Measure* (FIM) (Vignat and Bercher, 2003), also known as the *Fisher information of X with respect to a scalar translation parameter* (Dembo et al., 1991), is defined as

$$I_X = \mathbb{E}\left[\left(\frac{\partial}{\partial x}\log f(X)\right)^2\right] = \int \frac{\left[\frac{\partial}{\partial x}f(x)\right]^2}{f(x)}\,dx. \tag{3}$$

This quantity should not be confused with the Fisher information of a distribution parameter. In particular, the derivative of the log-density is relative to x and not to some parameter. However, the FIM is equivalent to the Fisher information of a location parameter of a parametric distribution (Cover and Thomas, 2006). Under mild regularity conditions, one has the following alternative formulation (Lehmann, 1999),

$$I_X = \mathbb{E}\left[-\frac{\partial^2}{\partial x^2}\log f(X)\right]. \tag{4}$$

The quantity I_X is sometimes interpreted as a measure of order / organization / narrowness of X. If X is Gaussian, $I_X = 1/\sigma^2$. It should be noted that H_X, N_X, and I_X only depend on the distribution $f(x)$.

2.2. Properties

The SEP and the FIM respect several properties. First, both quantities are positive. It is also easy to see the *scaling properties* of the SEP and the FIM (Rioul, 2011),

$$\begin{aligned} N_{aX} &= a^2 N_X, \\ I_{aX} &= a^{-2} I_X. \end{aligned} \tag{5}$$

for any real number $a \neq 0$, by change of variable. Notice also that the SEP and the FIM are invariant under additive deterministic constant, by the same argument. Harder to show are the *entropy power inequality* (Dembo et al., 1991) and its dual the *Fisher information inequality* (Zamir, 1998),

$$N_{X+Y} \geq N_X + N_Y, \tag{6}$$

$$I_{X+Y}^{-1} \geq I_X^{-1} + I_Y^{-1}, \tag{7}$$

for a random variable Y independent of X, with equality if X and Y are Gaussian.

Moreover, several relationships show that the FIM closely interact with the SEP and the differential entropy. Let Z be a random variable independent of X with finite variance σ_Z^2. The *de Bruijn's identity* (Cover and Thomas, 2006; Rioul, 2011) states that

$$\left. \frac{d}{dt} H_{X+\sqrt{t}Z} \right|_{t=0} = \frac{1}{2} \sigma_Z^2 I_X, \tag{8}$$

i.e., the variation of the differential entropy of a perturbed X is proportional to I_X. Therefore, a possible interpretation of the FIM is that it quantifies the sensitivity of H_X to a small independent additive perturbation Z. Using the entropy power inequality (6) and de Bruijn identity (8), one can show the *isoperimetric inequality for entropies*,

$$N_X I_X \geq 1, \tag{9}$$

with equality if and only if X is Gaussian. The proof and the nomenclature motivation of equation (9) can be found in Dembo et al. (1991), where a remarkable analogy is done with geometry. This shows that SEP and FIM are intimately interlinked.

2.3. Fisher-Shannon Complexity

The joint FIM/SEP analysis has been used as a statistical complexity measure, albeit there is no clear consensus about the definition of signal complexity (Esquivel et al., 2010). The *Fisher-Shannon Complexity* (FSC) is defined as $C_X = N_X I_X$ (Angulo et al., 2008). From the scaling properties (5), it is easy to show that the FSC is constant under scalar multiplication and addition. In particular, normalization or standardization of X has no effect on the FSC. Additionally, the isoperimetric inequality for entropies (9) states that $C_X \geq 1$, with equality if and only if X is Gaussian. An interpretation of this quantity is the following.

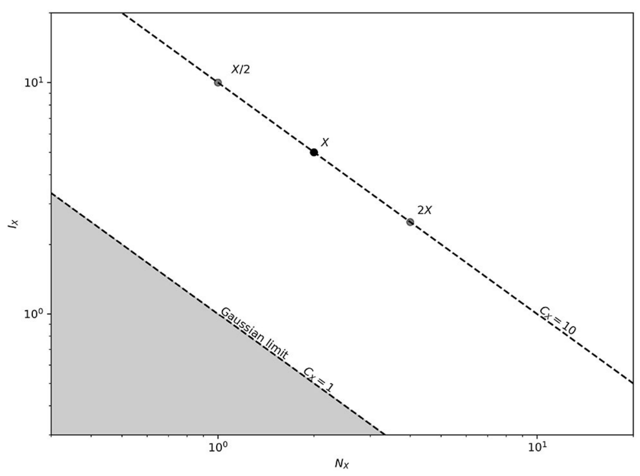


FIGURE 1 | The Fisher-Shannon information plane with a random variable X of FSC equal to 10. Scalar multiplication of X corresponds to a displacement along the iso-complex curve passing through X. The unreachable points are in gray. Note the logarithmic scale.

If Z is independent of X and has a finite variance σ_Z^2, one obtains the following relationship by using the de Bruijn identity (8),

$$\left. \frac{d}{dt} N_{X+\sqrt{t}Z} \right|_{t=0} = 2N_X \left. \frac{d}{dt} H_{X+\sqrt{t}Z} \right|_{t=0} = \sigma_Z^2 N_X I_X = \sigma_Z^2 C_X.$$

Hence, the FSC can be interpreted as a sensitivity measure of N_X to a small independent additive perturbation.

2.4. Fisher-Shannon Information Plane

The PDF of X can be analyzed displaying the SEP and FIM within the so-called Fisher-Shannon Information Plane (FSIP), see **Figure 1** (Vignat and Bercher, 2003). Although standard linear scale plot are very often used for the FSIP in the literature, a log-log plot is more adequate in practice. In the FSIP, the only reachable values are in the set $\mathcal{D} = \{(N_X, I_X) \in \mathbb{R}^2 | N_X > 0, I > 0 \text{ and } N_X I_X \geq 1\}$, due to (9). Vignat and Bercher (2003) showed that for any point $(N, I) \in \mathcal{D}$, it exists a random variable X (from an exponential power distribution) such that $N_X = N$ and $I_X = I$.

A curve in $\mathcal{D}$ is said to be *iso-complex* if the FSC along the curve is constant. As C_X is constant up to a multiplicative factor $a \neq 0$, and looking up at the scaling properties (5), one can move on any iso-complex curve by varying a. **Figure 1** shows the iso-complex curve of complexity $C_X = 10$ as an example. The boundary of $\mathcal{D}$ is the iso-complex curve with FSC equal to 1, and is reached if and only if X is Gaussian, as states by (9). On this boundary, the standard deviation σ (which plays the role of the scaling parameter in the Gaussian case) is equivalent to the multiplicative factor a. Hence, while a point in the FSIP is described by (N_X, I_X), one can also describe it by (a, C_X). In the light of this, one can also think of FSC as a scale-independent measure of non-Gaussianity of X.

3. ANALYTICAL SOLUTIONS FOR SOME DISTRIBUTIONS

In this section, we propose analytical formulas for the SEP, FIM and FSC for several parametric distributions. They can be used directly for parametric estimations. Vignat and Bercher (2003) obtained analogous results for the Student's t-distribution and the exponential power distribution (also known as generalized Gaussian distribution). The Gaussian case was already presented in section 2 as an example.

The differential entropy of the distributions proposed in this section have been computed by Lazo and Rathie (1978), from which the SEP is directly obtained. However, to our knowledge, the FIM-based calculations for Gamma, Weibull and log-normal distributions were never presented. Proofs can be found in the **Appendix**.

3.1. Gamma Distribution

The PDF of a Gamma random variable X is given by

$$f(x) = f(x;\theta,k) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k\Gamma(k)}, \qquad \text{for } x \geq 0,$$

and $f(x) = 0$, for $x < 0$, where Γ denotes the gamma function and $\theta, k > 0$ are, respectively, the scale and shape parameters.

Proposition 1. *The SEP of the Gamma distribution with scale $\theta > 0$ and shape $k > 0$ is*

$$N_X(\theta,k) = \frac{\theta^2\Gamma^2(k)}{2\pi e}e^{2[(1-k)\psi(k)+k]},$$

where ψ is the digamma function.

The FIM and the FSC of the Gamma distribution with scale $\theta > 0$ and shape $k > 2$ are, respectively,

$$I_X(\theta,k) = \frac{1}{(k-2)\theta^2},$$
$$C_X(k) = \frac{\Gamma^2(k)}{2\pi e(k-2)}e^{2[(1-k)\psi(k)+k]}.$$

3.2. Weibull Distribution

The PDF of a Weibull random variable is

$$f(x) = f(x;\mu,\lambda,k) = \frac{k}{\lambda}\left(\frac{x-\mu}{\lambda}\right)^{k-1}e^{-(\frac{x-\mu}{\lambda})^k}, \qquad \text{for } x \geq 0,$$

and $f(x) = 0$, for $x < 0$, where μ is the location parameter, $\lambda > 0$ is the scale parameter and $k > 0$ is the shape parameter.

Proposition 2. *The SEP of the Weibull distribution with location μ, scale $\lambda > 0$ and shape $k > 0$ is*

$$N_X(\lambda,k) = \frac{(1-\alpha)^2\lambda^2 e}{2\pi}e^{2\alpha\gamma},$$

where $\alpha = \frac{k-1}{k}$ and γ is the Euler-Mascheroni constant.

The FIM and the FSC of the Weibull distribution with location μ, scale $\lambda > 0$ and shape $k > 2$ are, respectively

$$I_X(\lambda,k), = \frac{\alpha^2}{(1-\alpha)^2\lambda^2}\Gamma(2\alpha-1),$$
$$C_X(k) = \frac{\alpha^2 e}{2\pi}\Gamma(2\alpha-1)e^{2\alpha\gamma}.$$

3.3. Log-Normal Distribution

The log-normal PDF with parameters μ and $\sigma > 0$ is

$$f(x) = f(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}}, \qquad \text{for } x > 0,$$

and $f(x) = 0$, for $x \leq 0$.

The notation of the parameters μ and σ are motivated by the fact that the logarithm of a log-normal random variable follows a normal distribution of mean μ and variance σ^2. However, μ and σ play, respectively, the role of the scale parameter and the shape parameter for the log-normal distribution.

Proposition 3. *The SEP, the FIM and the FSC of the log-normal distribution with μ and $\sigma > 0$ are given by*

$$N_X(\mu,\sigma) = \sigma^2 e^{2\mu},$$
$$I_X(\mu,\sigma) = \left(1 + \frac{1}{\sigma^2}\right)e^{2(\sigma^2-\mu)},$$
$$C_X(\sigma) = (1+\sigma^2)e^{2\sigma^2}.$$

4. DATA DRIVEN NON-PARAMETRIC ESTIMATION

Complex real-world data sets rarely follow parametric distributions. Providing enough data, it is also possible to carry out Fisher-Shannon analysis with a non-parametric estimation of density, which release parametric assumptions on the distribution (Telesca and Lovallo, 2017). In this paper, *integral estimates* of the SEP and the FIM are considered, which consist of substituting the kernel density estimators (KDE) of both $f(x)$ and its derivative in the integral forms of (1) and (3) (Bhattacharya, 1967; Dmitriev and Tarasenko, 1973; Prakasa Rao, 1983; Györfi and van der Meulen, 1987; Joe, 1989). Python and R implementations of this section content are proposed, see the software availability section at the end of this paper.

Following (Wand and Jones, 1994), let $X_1, \dots, X_n$ be a random sample of size n from a PDF $f(x)$. Consider also the *kernel* $K(u)$, a bounded PDF which is symmetric around zero, has a finite fourth moment and is differentiable. The KDE of $f(x)$ is

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right), \tag{10}$$

where $h > 0$ is the *bandwidth* parameter. In this paper, the Gaussian kernel defined by $K(u) = (2\pi)^{-1/2}\exp(-u^2/2)$ is used and the estimator (10) becomes

$$\hat{f}_h(x) = \frac{1}{\sqrt{2\pi}nh}\sum_{i=1}^{n}\exp\left\{-\frac{1}{2}\left(\frac{x-X_i}{h}\right)^2\right\}.$$

The integral estimate of (2) is

$$\hat{N}_X = \frac{1}{2\pi e} \exp\left\{-2\int \hat{f}_h(x)\log\hat{f}_h(x)\,dx\right\}.$$

Let us note f', the derivative of f with respect to x. Usually, f' is estimated by $\hat{f}_h'$. With the Gaussian kernel we obtain

$$\hat{f}_h'(x) = \frac{-1}{\sqrt{2\pi}\,nh^3}\sum_{i=1}^{n}(x-X_i)\exp\left\{-\frac{1}{2}\left(\frac{x-X_i}{h}\right)^2\right\}.$$

Then, the integral estimate of (3) is

$$\hat{I}_X = \int \frac{\left(\hat{f}_h'(x)\right)^2}{\hat{f}_h(x)}\,dx.$$

The FSC is estimated by multiplying $\hat{N}_X$ by $\hat{I}_X$.

Several techniques exist in order to automatize the bandwidth choice (Wand and Jones, 1994). In the following, the 2-stages direct plug-in method (Sheather and Jones, 1991) is used. This method estimates the optimal bandwidth regarding the asymptotic mean integrated squared error of $\hat{f}_h$. The interested reader can found further technical details in (Wand and Jones, 1994) and (Sheather and Jones, 1991).

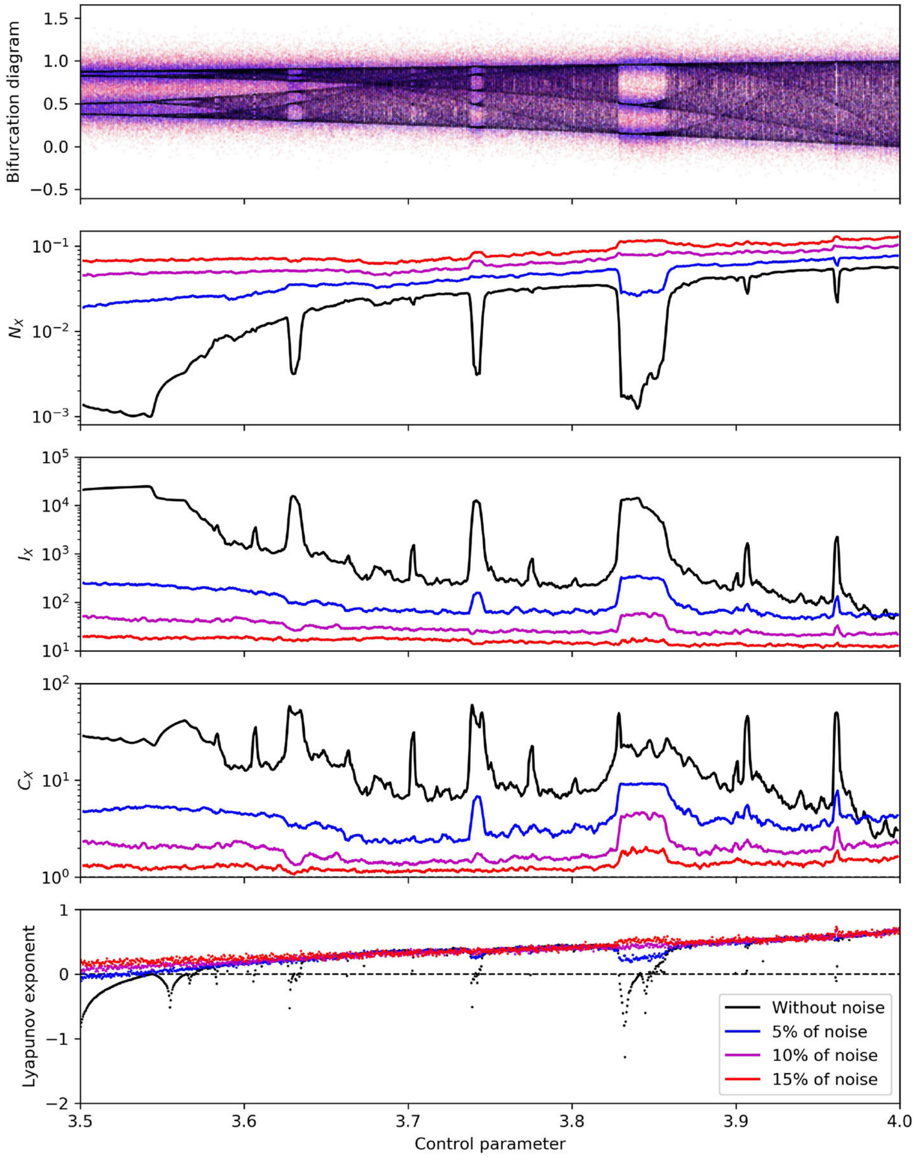


FIGURE 2 | Logistic map with different level of noise. From top to bottom : bifurcation diagram, SEP, FIM, FSC, and Lyapunov exponent sliding windows. Note the logarithmic scale on the y-axis for SEP, FIM, and FSC.

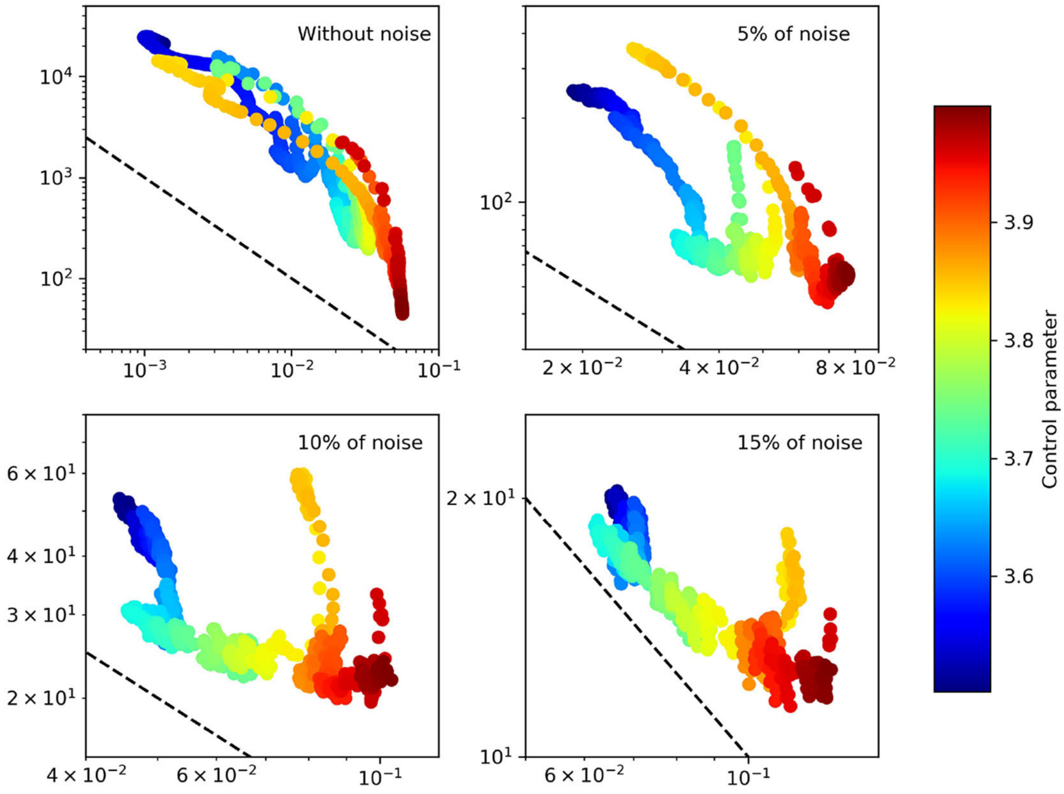


FIGURE 3 | Trajectory of the logistic map in the FSIP.

5. CASE STUDIES

In this section we explore two applications of SEP, FIM and FSC to time series. First, a synthetic experiment is used to show the usefulness of the method in detecting the dynamical behavior of chaotic systems. Then, an example of application of the proposed method to real complex environmental data is discussed.

5.1. Logistic Map

A synthetic experiment is designed to investigate how SEP, FIM, and FSC can be used to detect behavioral changes in non-linear dynamical systems. In the present research, the well-known logistic map is considered as a benchmark simulated case study, that illustrates and helps to understand important features of the considered measures.

Following the experiment proposed by Martin et al. (2001), the *logistic map* defined by

$$x_{n+1} = cx_n(1 - x_n), \qquad x_0 \in [0, 1], \quad c \in [0, 4],$$

where c is the control parameter, is analyzed using sliding window technique. Analysis within the sliding window pursues the goal of revealing dynamical evolution of properties of time series like Gaussianity and non-stationarity.

The sequence (x_n) is computed up to $n = 1,000$ for $c \in [3.5, 4]$. Centered Gaussian noise with different level of variance, 0.05, 0.10, 0.15, is added to x_n. The well-known bifurcation diagram of the logistic map is displayed in **Figure 2**. The SEP, FIM, and FSC are computed on data included in the overlapping windows of width $2.5 \cdot 10^{-3}$ along the control parameter, and the results are shown in the same figure. The Lyapunov exponent is also added for comparison reasons (Kantz and Schreiber, 2004). The results are also displayed in the FSIP, see **Figure 3**.

Analyzing the results obtained from the data without noise, it is easy to see how the SEP, FIM and FSC peak occurrences correspond to dynamic changes shown by the bifurcation diagram and the Lyapunov exponent. With the logarithmic scale on the y-axis, the behavior of the SEP is somewhat symmetric to the behavior of the FIM, i.e., the FIM seems to be inversely proportional to the SEP. However, this is not exactly the case, otherwise the FSC would be constant. In some sense, the perturbations in the FSC reflect the departure from the inverse proportionality between the SEP and the FIM. In the FSIP, perfect inverse proportionality corresponds to iso-complex curves. Indeed, the trajectory of the logistic map in the FSIP is stretched along iso-complex curves, see **Figure 3**.

Adding noise shows that most of the peaks become undetectable, see **Figure 2**. However, FSC seems to be the measure which suffers the least to noise in data. Note also, that FIM is less impacted than SEP. The noise effect is more interesting in the FSIP, see **Figure 3**. While the uncorrupted data is quite hard to interpret due to the superposition of the trajectory

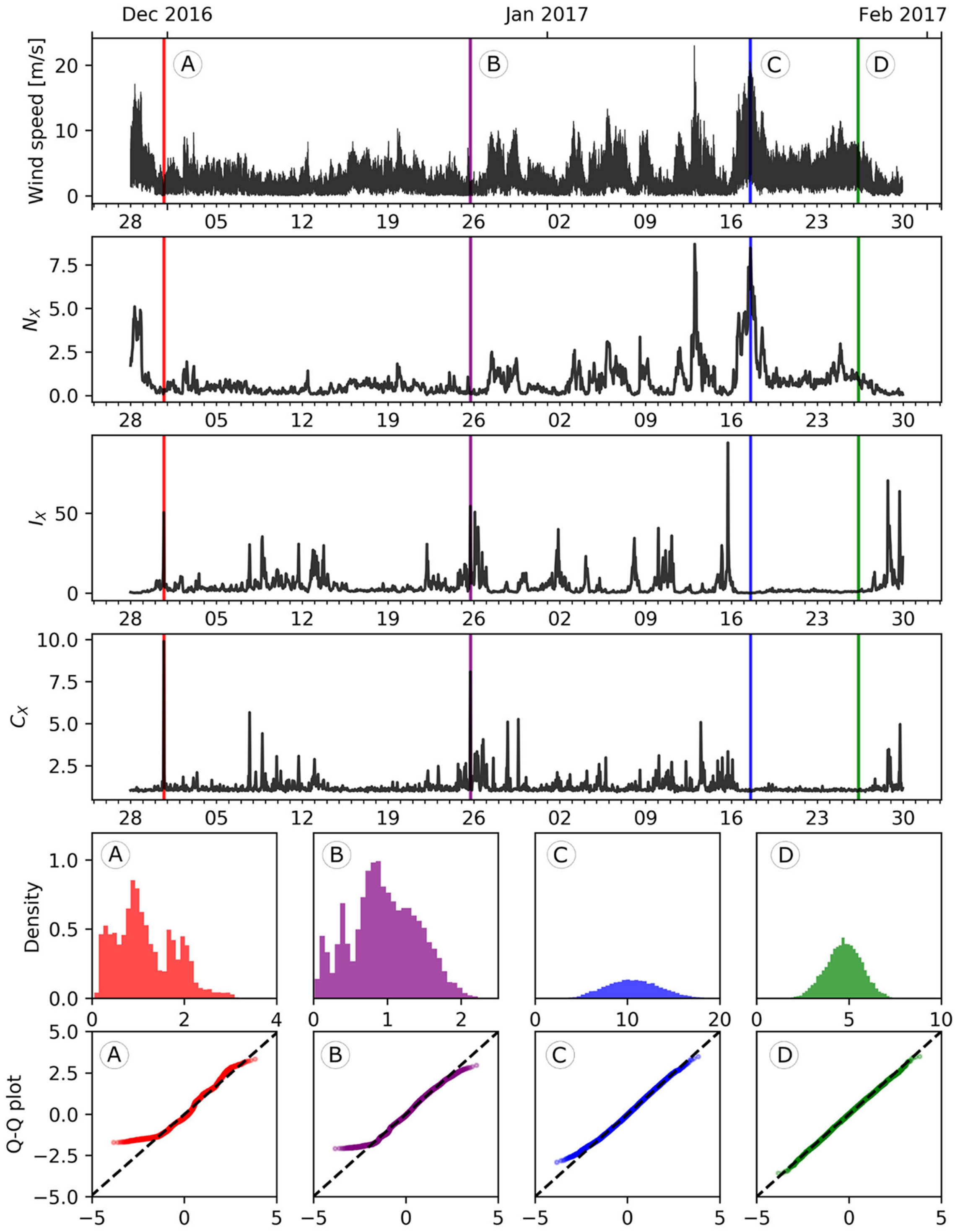


FIGURE 4 | High-frequency wind speed time series. From top to bottom : time series, SEP, FIM, and FSC moving windows, histograms and Q-Q plots of some time series subsets.

with itself, adding some noise seems to clarify complexity and trajectory behaviors in the FSIP. Noise stimulates the emergence of protuberances roughly corresponding to "islands of stability" of the (uncorrupted) bifurcation diagram, where Lyapunov exponent is negative. This emergence is due to the fact that FIM is less impacted than SEP, as it was seen above.

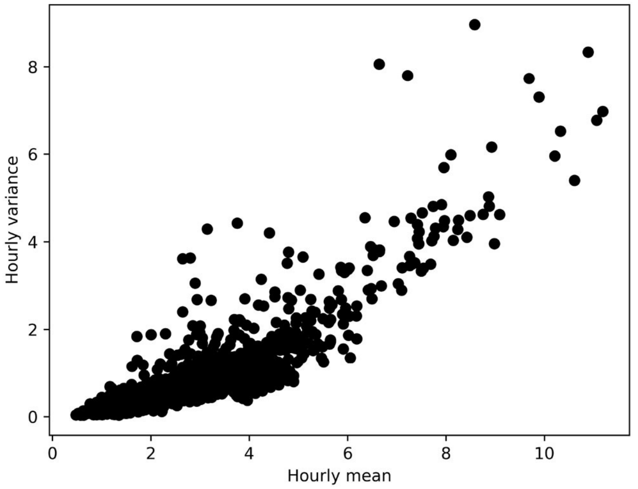


FIGURE 5 | Scatter plot of hourly variance against hourly mean of the high-frequency wind speed data.

5.2. Application to High Frequency Wind Data

The Fisher-Shannon information method can find a wide application in the geo-environmental domains. In the present section, we demonstrate how they can be applied to retrieve relevant knowledge from environmental time series. Specifically, high frequency wind speed data are analyzed. The time series consists of 1Hz frequency wind speed data, from 28 November 2016 to 29 January 2017 (**Figure 4**). The data (motus.epfl.ch) were measured at 25.5 m above the ground by a sensor which is placed on meteorological mast located on the campus of the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Notice that the mast is surrounded by a building layout of 10 m average height. More information on these measurements can be found in Mauree et al. (2017a,b).

The Fisher-Shannon quantities are computed with non-overlapping moving windows of 1 hour width along the time axis. Globally, all quantities vary with time, indicating non-stationarity, see **Figure 4**. The SEP seems to roughly replicate the behavior of the original time series. This is due to a proportional effect between the mean and the variance of the data, as shown in **Figure 5**. As for the logistic map case, the FIM is roughly inversely proportional to the SEP (not shown in logarithmic scale). The FSC is close to 1 during long period of time, e.g., between the 17th and the 27th January 2017. This should indicate a local behavior of wind speed close to a Gaussian one. During these periods, wind speed is not necessarily calm, e.g., the 17th January. Conversely, The FSC also exhibits some peaks where wind speed is rather low, which should indicate a more complex distribution of the data.

To verify this, a closer exploration of the data is required. To this aim, we considered four subsets of 3 hours length, denoted by A, B, C, and D and represented on **Figure 4** by color red, purple, blue, and green, respectively. Histograms and quantile-quantile (Q-Q) plots of these data subsets are also plotted with the corresponding colors. The subset D is chosen during the period of almost unitary FSC. The corresponding histogram and Q-Q plot confirms the very-close-to-Gaussian behavior of the data. The subset C is also chosen with a FSC close to 1, but centered on the maximum of SEP of the 17th January 2017 which corresponds also to a high wind speed activity. The histogram shows again a distribution close to a Gaussian one, but with a higher variance than C. This was an expected output, since for Gaussian distribution the SEP equals to the variance and C was chosen with a high SEP. The Q-Q plot shows little departure for the left tail, but the data are still relatively close to what was expected. The subset B is centered on a peak of FSC. The histogram shows a distribution which is very far from Gaussianity. It is clearly asymmetric and has at least two modes—maybe three. The Q-Q plot shows a strong departure from the Gaussian distribution, especially on the left tail. The subset A is centered on the highest FSC value. Its histogram shows three—maybe four—modes. The corresponding Q-Q plot shows how for this subset data are even farther from Gaussianity than for the previous subset.

These results show the high complexity of these data, whose behavior can rapidly change locally in time or even during calm weather. Further analysis on a larger set of these measurements using the FSC can be found in Guignard et al. (2019b), where authors analyzed wind speed and temperature data gathered by sensors similar to the one used to collect the data analyzed in this section, fixed along a mast located in a urban canyon. A FSC analysis suggested different wind dynamics induced by the building layout. The daily variation of temperature was also found to be an important predictor for high-frequency wind speed daily complexity. Moreover, FSC was used to show that wind speed and height are related by a non-linear relationship. More generally, this demonstrates the high versatility of analysis based on Fisher-Shannon information, which had yield numerous and various insight on these data.

6. CONCLUSIONS

The paper discusses the Fisher-Shannon information method as an effective data exploration tool able to give diverse insights into complex non-stationary time series. The Fisher-Shannon method was presented in a unified framework and new interpretations of FSC were pointed out. In particular, the FSC was identified as a sensitivity measure of the SEP and as a scale-independent non-Gaussianity measure, which both provide interpretation of this quantity as well. The detection of potential Gaussian behavior in the data was successfully showed on high-frequency wind speed data.

In the methodological part of the paper, FIM and FSC were computed in closed forms for several parametric distributions which are widely used in geo-environmental data analyses. Theoretical formulas for other random variables can be derived depending on the problem at hand. Furthermore, it was also shown—by injecting noise in the logistic map—how these information measures can be used to detect potential dynamic changes in a quite robust manner—especially with FSC. While SEP, FIM and FSC were presented as versatile information-based exploratory tools, they can also be used as time series discrimination or, more generally, to generate time series features for clustering, modeling, and forecasting.

The Fisher-Shannon method has been widely used in geosciences, as shown in the first part of this paper. However, according to our opinion, its full potential is still unexploited and underestimated. To simplify the access to this method for environmental data analysis and foster reproducibility, open source libraries written in R and Python for the computation of the three measures via a non-parametric kernel density estimation are provided.

From a theoretical point of view, future studies should involve generalization of the Fisher-Shannon method to the multivariate case. Several numerical investigations could be carried out for the KDE of the FIM. In particular, other estimates could be provided by re-substitution techniques as with entropy. Optimal bandwidth choice regarding to asymptotic mean squared error of FIM—or even FSC—could be derived. More practically, a challenging exploratory analysis of spatio-temporal data is planned.

SOFTWARE AVAILABILITY

A Python package is proposed on PyPI and GitHub (https://github.com/fishinfo/FiShPy), as well as an R package available on CRAN and GitHub (https://github.com/fishinfo/FiSh). They allow interested users to compute non-parametric estimation of the SEP, FIM and FSC.

DATA AVAILABILITY STATEMENT

The data presented in this article are not readily available as the authors are not their owners. Requests to access the data should be directed to the responsible parties of the MoTus experiment (motus.epfl.ch), Dasaraden Mauree (dasaraden.mauree@epfl.ch).

AUTHOR CONTRIBUTIONS

FG conceived the main conceptual ideas, conduct investigations, developed the theoretical formalism, performed the calculations, interpreted the computational results, and wrote the original draft. FG and ML developed Python and R packages. FG, FA, and MK wrote the final version of the paper. MK carried out the supervision, project administration, and funding acquisition. All authors discussed the results, provided critical feedback, commented, reviewed and edited the original manuscript, and gave final approval for publication.

FUNDING

FG and MK acknowledge the support of the National Research Programme 75 Big Data (PNR75) of the Swiss National Science Foundation (SNSF), project no. 167285.

ACKNOWLEDGMENTS

The authors are grateful to Dasaraden Mauree from the Ecole Polytechnique Fédérale de Lausanne, Switzerland, for providing the high-frequency wind speed data, and to Sylvain Robert for analytical computation checking. They also thank Luciano Telesca from Institute of Methodologies for Environmental Analysis of the Italian National Research Council for profitable discussions, and the reviewers for their constructive comments, which contributed to improve the paper.

REFERENCES

Amato, F., Laib, M., Guignard, F., and Kanevski, M. (2020). Analysis of air pollution time series using complexity-invariant distance and information measures. *Phys. A Stat. Mech. Appl.* 547:124391. doi: 10.1016/j.physa.2020.124391

Angulo, J., Antolín, J., and Sen, K. (2008). Fisher-Shannon plane and statistical complexity of atoms. *Phys. Lett. A* 372, 670–674. doi: 10.1016/j.physleta.2007.07.077

Bhattacharya, P. K. (1967). Estimation of a probability density function and its derivatives. *Sankhy = a* Indian J. Stat. Ser. A 29, 373–382.

Cover, T. M., and Thomas, J. A. (2006). *Elements of Information Theory*. New York, NY: Wiley-Interscience.

Dembo, A., Cover, T. M., and Thomas, J. A. (1991). Information theoretic inequalities. *IEEE Trans. Inform. Theory* 37, 1501–1518. doi: 10.1109/18.104312

Dmitriev, Y., and Tarasenko, F. (1973). On the estimation of functionals of the probability density and its derivatives. *Theory Probabil. Appl.* 18, 628–633. doi: 10.1137/1118083

Esquivel, R. O., Angulo, J. C., Antolín, J., Dehesa, J. S., Lopez-Rosa, S., and Flores-Gallegos, N. (2010). Analysis of complexity measures and information planes of selected molecules in position and momentum spaces. *Phys. Chem. Chem. Phys.* 12, 7108–7116. doi: 10.1039/b927055h

Fisher, R. A. (1925). Theory of statistical estimation. *Math. Proc. Cambridge Philos. Soc.* 22, 700–725. doi: 10.1017/S0305004100009580

Frieden, B. R. (1990). Fisher information, disorder, and the equilibrium distributions of physics. *Phys. Rev. A* 41, 4265–4276. doi: 10.1103/PhysRevA.41.4265

Guignard, F., Lovallo, M., Laib, M., Golay, J., Kanevski, M., Helbig, N., et al. (2019a). Investigating the time dynamics of wind speed in complex terrains by using the Fisher-Shannon method. *Phys. A Stat. Mech. Appl.* 523, 611–621. doi: 10.1016/j.physa.2019.02.048

Guignard, F., Mauree, D., Lovallo, M., Kanevski, M., and Telesca, L. (2019b). Fisher-Shannon complexity analysis of high-frequency urban wind speed time series. *Entropy* 21:47. doi: 10.3390/e21010047

Györfi, L., and van der Meulen, E. C. (1987). Density-free convergence properties of various estimators of entropy. *Comput. Stat. Data Anal.* 5, 425–436. doi: 10.1016/0167-9473(87)90065-X

Joe, H. (1989). Estimation of entropy and other functionals of a multivariate density. *Ann. Instit. Stat. Math.* 41, 683–697. doi: 10.1007/BF00057735

Kantz, H., and Schreiber, T. (2004). *Nonlinear Time Series Analysis, Vol. 7*. Cambridge, MA: Cambridge University Press. doi: 10.1017/CBO9780511755798

Lazo, A. V., and Rathie, P. (1978). On the entropy of continuous probability distributions (corresp.). *IEEE Trans. Inform. Theory* 24, 120–122. doi: 10.1109/TIT.1978.1055832

Lehmann, E. (1999). *Elements of Large-Sample Theory. Springer Texts in Statistics*. New York, NY: Springer-Verlag. doi: 10.1007/b98855

Lovallo, M., Shaban, A., Darwich, T., and Telesca, L. (2013). Investigating the time dynamics of monthly rainfall time series observed in northern Lebanon by means of the detrended fluctuation analysis and the Fisher-Shannon method. *Acta Geophys*. 61, 1538–1555. doi: 10.2478/s11600-012-0094-9

Martin, M., Pennini, F., and Plastino, A. (1999). Fisher's information and the analysis of complex signals. *Phys. Lett. A* 256, 173–180. doi: 10.1016/S0375-9601(99)00211-X

Martin, M., Perez, J., and Plastino, A. (2001). Fisher information and nonlinear dynamics. *Phys. A Stat. Mech. Appl*. 291, 523–532. doi: 10.1016/S0378-4371(00)00531-8

Mauree, D., Deschamps, L., Bequelin, P., Loesch, P., and Scartezzini, J. L. (2017a). "Measurement of the impact of buildings on meteorological variables," in *Building Simulation Application Proceedings* (Bolzano: BU Press).

Mauree, D., Lee, D. S. H., Naboni, E., Coccolo, S., and Scartezzini, J. L. (2017b). Localized meteorological variables influence at the early design stage. *Energy Proc*. 122, 325–330. doi: 10.1016/j.egypro.2017.07.331

Pierini, J. O., Lovallo, M., Gomez, E. A., and Telesca, L. (2016). Fisher-Shannon analysis of the time variability of remotely sensed sea surface temperature at the brazil-malvinas confluence. *Oceanologia* 58, 187–195. doi: 10.1016/j.oceano.2016.02.003

Pierini, J. O., Restrepo, J. C., Lovallo, M., and Telesca, L. (2015). Discriminating between different streamflow regimes by using the Fisher-Shannon method: an application to the colombia rivers. *Acta Geophys*. 63, 533–546. doi: 10.2478/s11600-014-0229-2

Pierini, J. O., Scian, B., Lovallo, M., and Telesca, L. (2011). Discriminating climatological regimes in rainfall time series by using the fisher-Shannon method. *Int. J. Phys. Sci*. 6, 7799–7804. doi: 10.5897/IJPS11.737

Prakasa Rao, B. (1983). Nonparametric Functional Estimation. Probability and Mathematical Statistics: A Series of Monographs and Textbooks. Academic Press.

Rioul, O. (2011). Information theoretic proofs of entropy power inequalities. *IEEE Trans. Inform. Theory* 57, 33–55. doi: 10.1109/TIT.2010.2090193

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J*. 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Sheather, S. J., and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Stat. Soc. Ser. B* 53, 683–690. doi: 10.1111/j.2517-6161.1991.tb01857.x

Telesca, L., Caggiano, R., Lapenna, V., Lovallo, M., Trippetta, S., and Macchiato, M. (2009a). Analysis of dynamics in Cd, Fe, and Pb in particulate matter by using the Fisher-Shannon method. *Water Air Soil Pollut*. 201, 33–41. doi: 10.1007/s11270-008-9924-4

Telesca, L., Chamoli, A., Lovallo, M., and Stabile, T. A. (2015a). Investigating the tsunamigenic potential of earthquakes from analysis of the informational and multifractal properties of seismograms. *Pure Appl. Geophys*. 172, 1933–1943. doi: 10.1007/s00024-014-0862-3

Telesca, L., and Lovallo, M. (2011). Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher-Shannon information plane. *J. Stat. Mech. Theory Exp*. 2011:P07001. doi: 10.1088/1742-5468/2011/07/P07001

Telesca, L., and Lovallo, M. (2013). Fisher-Shannon analysis of wind records. *Int. J. Energy Stat*. 01, 281–290. doi: 10.1142/S2335680413500208

Telesca, L., and Lovallo, M. (2017). On the performance of Fisher information measure and Shannon entropy estimators. *Phys. A Stat. Mech. Appl*. 484, 569–576. doi: 10.1016/j.physa.2017.04.184

Telesca, L., Lovallo, M., Alcaz, V., and Ilies, I. (2015b). Site-dependent organization structure of seismic microtremors. *Phys. A Stat. Mech. Appl*. 421, 541–547. doi: 10.1016/j.physa.2014.11.061

Telesca, L., Lovallo, M., and Carniel, R. (2010). Time-dependent Fisher information measure of volcanic tremor before the 5 April 2003 paroxysm at Stromboli volcano, Italy. *J. Volcanol. Geothermal Res*. 195, 78–82. doi: 10.1016/j.jvolgeores.2010.06.010

Telesca, L., Lovallo, M., Chamoli, A., Dimri, V., and Srivastava, K. (2013). Fisher-Shannon analysis of seismograms of tsunamigenic and non-tsunamigenic earthquakes. *Phys. A Stat. Mech. Appl*. 392, 3424–3429. doi: 10.1016/j.physa.2013.03.049

Telesca, L., Lovallo, M., Hsu, H. L., and Chen, C. C. (2011). Analysis of dynamics in magnetotelluric data by using the Fisher-Shannon method. *Phys. A Stat. Mech. Appl*. 390, 1350–1355. doi: 10.1016/j.physa.2010.12.005

Telesca, L., Lovallo, M., Ramirez-Rojas, A., and Angulo-Brown, F. (2009b). A nonlinear strategy to reveal seismic precursory signatures in earthquake-related self-potential signals. *Phys. A Stat. Mech. Appl*. 388, 2036–2040. doi: 10.1016/j.physa.2009.01.035

Vignat, C., and Bercher, J. F. (2003). Analysis of signals in the Fisher-Shannon information plane. *Phys. Lett. A* 312, 27–33. doi: 10.1016/S0375-9601(03)00570-X

Wand, M., and Jones, M. (1994). *Kernel Smoothing*. Boca Raton: Chapman & Hall; Taylor & Francis. doi: 10.1201/b14876

Zamir, R. (1998). A proof of the Fisher information inequality via a data processing argument. *IEEE Trans. Inform. Theory* 44, 1246–1250. doi: 10.1109/18.669301

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.

7. APPENDIX

The differential entropy H_X for Gamma, Weibull and log-normal distributions can be found in (Lazo and Rathie, 1978) and (Cover and Thomas, 2006). The SEP is simply a non-linear transform of H_X.

Proof of proposition 1: Computing the second derivative of $\log f(x)$, one has

$$\frac{\partial^2}{\partial x^2}\log f(x) = -\frac{k-1}{x^2},$$

and then, using (4), the variable change $x = \theta y$ and the properties of the Gamma function,

$$\begin{aligned}
I_X &= (k-1)\mathbb{E}[X^{-2}] \\
&= \frac{k-1}{\theta^k\Gamma(k)}\int_0^\infty x^{k-3}e^{-\frac{x}{\theta}}\,dx \\
&= \frac{k-1}{\theta^2\Gamma(k)}\int_0^\infty y^{k-3}e^{-y}\,dy \\
&= \frac{(k-1)\Gamma(k-2)}{\theta^2(k-1)(k-2)\Gamma(k-2)},
\end{aligned}$$

yielding the FIM for the Gamma distribution. The FSC is directly obtained by multiplying the SEP and the FIM. □

Proof of proposition 2: Starting from the Weibull PDF, one has

$$\frac{\partial^2}{\partial x^2}\log f(x) = -\frac{k-1}{(x-\mu)^2} - \frac{k(k-1)}{\lambda^k}(x-\mu)^{k-2},$$

and with the variable change $y = (\frac{x-\mu}{\lambda})^k$,

$$\begin{aligned}
I_X &= (k-1)\mathbb{E}\left[(X-\mu)^{-2}\right] + \frac{k(k-1)}{\lambda^k}\mathbb{E}\left[(X-\mu)^{k-2}\right] \\
&= \frac{k(k-1)}{\lambda^3}\left[\int_0^\infty \left(\frac{x-\mu}{\lambda}\right)^{k-3} e^{-(\frac{x-\mu}{\lambda})^k}\,dx \right.\\
&\quad \left. + k\int_0^\infty \left(\frac{x-\mu}{\lambda}\right)^{2k-3} e^{-(\frac{x-\mu}{\lambda})^k}\,dx\right] \\
&= \frac{k-1}{\lambda^2}\left[\int_0^\infty y^{-\frac{2}{k}}e^{-y}\,dy + k\int_0^\infty y^{1-\frac{2}{k}}e^{-y}\,dy\right] \\
&= \frac{k-1}{\lambda^2}\left[\Gamma\left(1-\frac{2}{k}\right) + k\Gamma\left(2-\frac{2}{k}\right)\right] \\
&= \frac{k-1}{\lambda^2}\left[1 + k\left(1-\frac{2}{k}\right)\right]\Gamma\left(1-\frac{2}{k}\right) \\
&= \frac{(k-1)^2}{\lambda^2}\Gamma\left(1-\frac{2}{k}\right).
\end{aligned}$$

□

Proof of proposition 3: The second derivative of $\log f(x)$ is

$$\frac{\partial^2}{\partial x^2}\log f(x) = \frac{\log x - \mu + \sigma^2 - 1}{\sigma^2 x^2},$$

and using the variable change $y = \log x - \mu$, one have

$$\begin{aligned}
I_X &= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \frac{1-\sigma^2-(\log x-\mu)}{\sigma^2 x^3} e^{-\frac{(\log x-\mu)^2}{2\sigma^2}}\,dx \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^\infty \frac{1-\sigma^2-y}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}-2y-2\mu}\,dy.
\end{aligned}$$

Note that

$$-\frac{y^2}{2\sigma^2} - 2y - 2\mu = -\frac{(y+2\sigma^2)^2}{2\sigma^2} + 2(\sigma^2-\mu).$$

Using this and the definition of a Gaussian distribution $\mathcal{N}(-2\sigma^2, \sigma)$,

$$\begin{aligned}
I_X &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^\infty \left(\frac{1-\sigma^2}{\sigma^2} - \frac{y}{\sigma^2}\right) e^{-\frac{(y+2\sigma^2)^2}{2\sigma^2}+2(\sigma^2-\mu)}\,dy \\
&= \frac{e^{2(\sigma^2-\mu)}}{\sigma^2}\left[\frac{1-\sigma^2}{\sigma\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{(y+2\sigma^2)^2}{2\sigma^2}}\,dy \right.\\
&\quad \left. - \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^\infty y e^{-\frac{(y+2\sigma^2)^2}{2\sigma^2}}\,dy\right] \\
&= \frac{e^{2(\sigma^2-\mu)}}{\sigma^2}\left[1-\sigma^2+2\sigma^2\right] \\
&= \frac{1+\sigma^2}{\sigma^2}e^{2(\sigma^2-\mu)},
\end{aligned}$$

and the FIM is obtained. The FSC is

$$C_X = (1+\sigma^2)e^{2\mu+2(\sigma^2-\mu)} = (1+\sigma^2)e^{2\sigma^2}.$$

□

frontiers in Earth Science

TECHNOLOGY AND CODE
published: 02 September 2020
doi: 10.3389/feart.2020.00296

Razorback, an Open Source Python Library for Robust Processing of Magnetotelluric Data

*Farid Smaï and Pierre Wawrzyniak**

BRGM, French Geological Survey, Orléans, France

Magnetotellurics (MT) is a geophysical method that investigates the relationships among the different components of the natural electromagnetic field related to the geoelectric structure of the subsurface. Data can be contaminated by anthropic noise sources and suffer from transient noise to signal variations. Since the 80s, robust processing methods have been introduced to minimize the impact of noise on sounding quality. This paper presents Razorback, an open source Python library, implemented to handle, manipulate, and combine time series of synchronous data. This modular library allows users to plug in data prefilters and includes both M-estimator and bounded influence techniques, as well as a two-stage multiple remote reference. Validation of this library is performed on a real data set by comparing the results with those of an existing code. In contrast to standalone codes, the developed library allows for the design of complex and specific processing procedures. As examples, Razorback is used to perform (i) continuous time lapse processing and (ii) processing of one site in a peri-urban context. In the latter case, we have tested all possible combinations of remote reference stations in an MT array. Our phase tensor analysis shows that the bounded influence outperforms the M-estimator in reducing the impacts of man-made electromagnetic noise on magnetotelluric soundings. The Razorback library is available at https://github.com/BRGM/razorback. Jupyter notebooks for data handling and MT robust processing are available at https://github.com/BRGM/razorback/blob/doc/docs/source/tutorials/.

Keywords: magnetotellurics, time-series analysis, Fourier analysis, robust methods, Python, M-estimator, bounded influence, remote reference

Edited by:
Reik Donner,
Hochschule Magdeburg-Stendal,
Germany

Reviewed by:
Leonardo Guimarães Miquelutti,
Universidade Federal Fluminense,
Brazil
Pavel Pushkarev,
Lomonosov Moscow State University,
Russia
Jared Peacock,
United States Geological Survey
(USGS), United States

****Correspondence:***
Pierre Wawrzyniak
p.wawrzyniak@brgm.fr

Specialty section:
This article was submitted to
Solid Earth Geophysics,
a section of the journal
Frontiers in Earth Science

Received: *05 February 2020*
Accepted: *25 June 2020*
Published: *02 September 2020*

Citation:
Smaï F and Wawrzyniak P (2020)
Razorback, an Open Source Python
Library for Robust Processing of
Magnetotelluric Data.
Front. Earth Sci. 8:296.
doi: 10.3389/feart.2020.00296

1. INTRODUCTION

The magnetotelluric (MT) method studies the relationships in the frequency domain among components of the natural electromagnetic (EM) field (Vozoff, 1972). MT fields are generated (i) by external geomagnetic sources (ionospheric currents) at frequencies under 1 Hz and (ii) by atmospheric lightnings propagating through the earth-ionosphere waveguide at frequencies above 1 Hz. Recorded at the ground surface, MT fields are supposed to be plane waves. As stated by Ward (1967), noise in EM fields can be either instrumental noise, "geological" noise or disturbance field EM noise. The latter is caused by fluctuations of the natural sources (mainly related to solar activity above 1 Hz) and artificial/man-made sources. In urban and industrialized areas, man-made EM sources contaminate MT fields and cause divergence from the plane wave model.

The first step in MT analysis is estimating transfer functions (TFs) between horizontal components of the electric and magnetic fields (i.e., the MT impedance tensor) and between the vertical and horizontal components of the magnetic field (the so-called tipper) (Sims et al., 1971; Vozoff, 1972). Such TF estimates were originally performed using the classical least-squares approach (Sims et al., 1971) but generally provide biased results (Goubau et al., 1978). To reduce the impact of noise on MT TFs, Gamble et al. (1979) introduced the remote reference method, where synchronous measurements of MT fields are performed at a second site to remove bias from the local TF estimates. This method still produces biased TFs if correlated noise contaminates both local and remote sites.

In the 80s, robust estimation techniques were introduced (Egbert and Booker, 1986; Chave et al., 1987) to handle a reasonable proportion of outliers in MT data sets. M-estimator (Chave and Thomson, 1989) and later bounded influence estimator (Chave and Thomson, 2004) are considered to be the most effective techniques for TF estimation and yield unbiased MT estimates if at least one of the remote sites is uncontaminated by correlated noise. Note that robust estimation techniques are also now used in controlled-source EM data processing to estimate TFs between transmitter and receiver data (Streich et al., 2013). An alternative procedure, the robust multivariate errors-in-variables (RMEV) approach, was proposed by Egbert (1997). This procedure aims to identify by principal component analysis the different sources present in the acquired data and separate correlated noise from plane wave MT data. However, despite successful examples of coherent noise removal (Di Giuseppe et al., 2018), no automatic RMEV approach is currently available (Chave and Jones, 2012).

When considering "the future of magnetotellurics," Chave and Jones (2012) noted the significant improvement resulting from robust estimation techniques but also stressed that a major challenge remains, namely, data contamination by man-made correlated noise sources, which can be permanent or intermittent (Szarka, 1988; Junge, 1996). Wavelet transform of MT fields can be used to select geomagnetic events in the time-frequency plane (Zhang and Paulson, 1997) and to identify intermittent or permanent correlated noise sources in the data (Trad and Travassos, 2000; Escalas et al., 2013; Carbonari et al., 2017). Once data filtering is performed, classical weighted least squares or a robust TF estimator are still used on a reduced filtered subset of the Fourier or wavelet transforms (Larnier et al., 2016).

Existing standalone robust codes have specific input formats (time series of electric and magnetic fields, for example) that may be incompatible with Fourier data pre-filtering and that are controlled at a higher level by parameter files. Once the code is launched, tracking the influence of processing parameters in the framework of the code can be difficult. At present, MT processing faces the need for integration and requires an approach to conveniently combine the available techniques as the steps of a full process. We suggest that an easy-to-use open source library would greatly help integrate and experiment with the well-established, the new and the forthcoming techniques. Thus, we implemented the library in Python, a modern programming language that allows modular codes and that is being increasingly adopted in scientific computing. A similar approach has already be taken by the MTPY library (Krieger and Peacock, 2014; Kirkby et al., 2019) that assists with MT data processing, analysis, modeling, visualization and interpretation while the focus of the present library is the actual processing of MT data. The Razorback library is constructed over elementary components that, when combined, allows for a simple modular implementation of classical weighted least squares and robust TF techniques, such as the M-estimator and bounded influence estimator for MT. The library aims to allow the user to be in control of data transformation at each step of the processing. In MT robust processing, it implies accessing robust Fourier transforms of electric and magnetic fields in the time-frequency representation (i.e., view of electric and magnetic signals, initially taken to be a functions of time represented over both time and frequency) to track which ones are rejected/down-weighted during pre-filtering and during the robust TF estimation procedure.

In this paper, we first recall the basics of MT TF theory. Next, we define the features, functions and objects provided by the Razorback library. Then, MT data processing examples are shown for validation. Finally, we show two uses of the standard package with higher level functions: (a) performing processing with all possible combinations of synchronous remote reference stations, and (b) performing time lapse processing by subdividing continuous data in consecutive and overlapping window portions. The Razorback library is licensed under GPL v3.0 and is available at https://github.com/BRGM/razorback. Jupyter notebooks for data handling and MT robust processing are available at https://github.com/BRGM/razorback/blob/doc/docs/source/tutorials/.

2. MT TRANSFER FUNCTIONS

MT is a natural source geophysical method introduced by Tikhonov (1950) and Cagniard (1953), linking surface electric and magnetic field measurements to subsurface electrical properties. In MT theory, surface EM fields are assumed to be plane waves.

The common practice is to perform a Fourier transform of the MT time series, leading to a set of N complex coefficients for each MT field channel and for each frequency. For a fixed frequency, one horizontal component of the electric field $e_i \in \mathbb{C}^N$, indexed by $i \in \{x, y\}$, is linked to the horizontal magnetic field $b \in \mathbb{C}^{N\times 2}$ through the relation

$$e_i = bz_i + \epsilon \tag{1}$$

where $z_i \in \mathbb{C}^2$ denotes the MT impedance associated with direction i of the electric field, expressed in millivolts per km per nanotesla [mV/(km× nT)], and $\epsilon \in \mathbb{C}^N$ is the error. The impedance z can be regarded as a TF to be estimated. Hereafter, $\mathbb{O}$ denotes an operator yielding such estimations, namely, $z_i \simeq \widehat{z_i} = \mathbb{O}(e_i, b)$. Residual $r_i \in \mathbb{C}^N$ is classically computed from

$$r_i = e_i - b\,\mathbb{O}(e_i, b) \tag{2}$$

Similarly, the vertical magnetic field $b_z \in \mathbb{R}^N$ is related to the horizontal components of the magnetic field:

$$b_z = bt + \epsilon \tag{3}$$

where $t \in \mathbb{C}^2$ is the dimensionless tipper representing the tipping of the magnetic field out of the horizontal plane (Vozoff, 1972). The same approach can be used to estimate the tipper t with $t \simeq \widehat{t} = \mathbb{O}(b_z, b)$.

In the presence of noisy data, the definition of the operator $\mathbb{O}$ is not unique. The following sections present several ways to define and compute this operator.

2.1. Single Site Least-Squares Estimate

Under the Gauss–Markov conditions on the noise (zero mean, homoscedasticity, i.e., all the data don't have the same finite variance, and no correlation), the ordinary least squares (OLS) operator $\mathbb{O}_{OLS}$ can be applied to impedance estimation for a single site (SS) dataset:

$$\widehat{z}_i = \left(b^H b\right)^{-1} \left(b^H e_i\right). \tag{4}$$

The accuracy of the OLS estimates can be compromised because (i) data rarely satisfy the homoscedastic assumption due to transient noise sources and (ii) the electromagnetic field data deviate from the plane wave assumption. When a small portion of the dataset is corrupted by (i) and (ii), the OLS method can still provide a reliable estimate. However, correlated EM noise that may exists between the local electric and magnetic fields can lead to strongly biased estimates.

2.2. Remote Reference Least-Squares Estimate

The remote reference (RR) method introduces an additional magnetic field (b_r) from a distant second site to minimize the impact of noise. The RR method requires uncorrelated noise between the RR site and the local station. The RR-OLS impedance estimate is given by

$$\widehat{z}_i = \left(b_r^H b\right)^{-1} \left(b_r^H e_i\right). \tag{5}$$

MT data can still be affected by correlated noise between the RR and the local station, whose impact can be reduced with several RR stations.

2.3. Two-Stage Remote Reference Estimate

Chave and Thomson (2004) proposed a generalization of the RR method to multiple RR data sets. The principle is to use a set of q RR horizontal magnetic fields, collected in the array $Q \in \mathbb{C}^{N\times q}$. The local magnetic field is linked to Q by

$$b = QW + \epsilon, \tag{6}$$

where $W \in \mathbb{C}^{q\times 2}$ is a TF between local and remote magnetic field data. The predicted local magnetic field $\widehat{b}$ is expressed as

$$\widehat{b} = Q\widehat{W}, \quad \text{with} \quad \widehat{W} = \mathbb{O}(b, Q), \tag{7}$$

where $\mathbb{O}$ denotes the operator used to estimate the TF W. The impedance can then be estimated by using the predicted local magnetic field $\widehat{b}$ rather than the local magnetic field b:

$$\widehat{z}_i^q = \mathbb{O}(e_i, \widehat{b}) \tag{8}$$

This formulation for multiple RR impedance estimates (Equations 7 and 8) will be referred as the two-stage RR estimate. Chave and Jones (2012) demonstrated that when only one RR ($q = 2$) and the OLS are used ($\mathbb{O} = \mathbb{O}_{OLS}$), the two-stage RR estimate is equivalent to the classical RR OLS estimate.

MT data suffer from time and frequency-dependent signal-to-noise ratio variations, and the Gauss–Markov conditions to use the OLS method are violated even if several remote stations are involved. To overcome those problems, robust methods were introduced in the 80s.

2.4. M-estimator

The M-estimator (Egbert and Booker, 1986) is a robust TF estimator designed to minimize the influence of data associated with large residuals (Equation 2) in the regression.

The M-estimate of the MT impedance TF, $\widehat{z}$, is obtained through the non-linear weighted least squares operator denoted $\mathbb{O}_{ME}$ and defined by the implicit relation

$$\widehat{z}_i = \mathbb{O}_{ME}(e_i, b) = \left(b^H \nu(\widehat{z}_i) b\right)^{-1} \left(b \nu(\widehat{z}_i) e_i\right) \tag{9}$$

where $\nu(\widehat{z}) \in \mathbb{R}^{N\times N}$ is the weighting diagonal matrix depending on the residual and given by

$$\nu_{j,j}(\widehat{z}_i) = \nu(x_j(\widehat{z}_i)) \quad , \quad x(\widehat{z}) = \frac{1}{d}(e_i - b\widehat{z}_i). \tag{10}$$

Here, d is an estimate of the scale of the residual population. The weighting function, $\nu(x)$, must be designed to reduce the influence of large residuals.

In practice, Equation (9) is solved iteratively, starting with an initial value $\widehat{z}_i^0$ and defining the sequence $(\widehat{z}_i^k)_{k\geq 0}$ by

$$\widehat{z}_i^{k+1} = \left(b^H \nu(\widehat{z}_i^k) b\right)^{-1} \left(b \nu(\widehat{z}_i^k) e_i\right) \quad \text{for } k \geq 0. \tag{11}$$

The convergence of the above sequence is not guaranteed in general. However, for some weighting functions, the procedure (Equation 11) is stable in the sense that the sequence converges to the same limit independently from $\widehat{z}_i^0$. In practice, the procedure is stopped when the variation of the weighted residuals sum of squares, $r_i^H r_i$, is lower than a user-defined tolerance *tol* (typically 1%).

The two key parameters of the M-estimator method are the scale of the residual population d and the weighting function ν of equations (10) (Chave et al., 1987). The scale d is approximated by the median absolute deviation (MAD) from the median computed from the first iteration (with $\widehat{z}_i^{k=0}$) given by

$$d = \frac{1}{0.44845} \text{median}\left(|r_i^{k=0} - \text{median}(r_i^{k=0})|\right).$$

Here, the scale is computed with $k = 0$ and is fixed for $k \geq 1$. Two weighting functions are proposed, namely, the Huber weighting function,

$$\nu_{\text{Huber}}(x) = \left| \begin{array}{ll} 1 & \text{if } |x| \leq \alpha \\ \alpha/|x| & \text{if } |x| > \alpha \end{array} \right. , \quad \alpha = 1.5 \tag{12}$$

and the Thomson weighting function,

$$\nu_{\text{Thomson}}(x) = \exp\left(\exp(-\xi^2)\right) \exp\left(-\exp(\xi((|x| - \xi)))\right), \quad \xi = \sqrt{2\log(2N)} \tag{13}$$

where ξ is the N-th quantile of the Rayleigh distribution and N is the size of the sample. The choice of the weighting function ν depends on the robustness and the stability of the procedure defined by (11). Robustness means lowering the influence of large residuals, and stability corresponds to the existence and uniqueness of $\widehat{z}$. The Thomson function is more robust, but unlike the Huber function, it does not ensure stability. The non-stability encountered with the Thomson function can be overcome in practice by using a good initial value for $\widehat{z}$ when solving (9). A procedure to compute the impedance M-estimate with the Thomson weighting function can be implemented with the following steps:

```
Input: $e, b$
    1. Compute $\widehat{z}_{\text{OLS}}$ defined by Equation (4).
    2. Compute $\widehat{z}_{\text{Huber}}$ defined by Equation (11) with
       $\nu_{\text{Huber}}$ (12) as weighting function
       and $\widehat{z}_{\text{OLS}}$ as initial value.
    3. Compute $\widehat{z}_{\text{Thomson}}$ defined by Equation (11) with
       $\nu_{\text{Thomson}}$ (13) as weighting function
       and $\widehat{z}_{\text{Huber}}$ as initial value.
Output: $\widehat{z} = \widehat{z}_{\text{Thomson}}$
```

(14)

The M-estimator TF provides reliable protection against strong data residuals but remains highly sensitive to extreme values of the magnetic field, known as *leverage points*.

2.5. Bounded Influence Estimator

To prevent the effect of leverage points, Chave and Thomson (2004) proposed using the diagonal part of the hat matrix for improving the weighting matrix of the M-estimator method. The hat matrix H depends on the diagonal weighting matrix ν and is defined as

$$H(\nu) = \sqrt{\nu} b \left(b^H t \nu b\right)^{-1} b^H \sqrt{\nu} \tag{15}$$

Chave and Thomson (2003) showed that the hat matrix diagonal follows the beta distribution, $\beta(h, p, M - p)$, where h is a diagonal element, p is the number of independent sources ($p = 2$ for MT), and M is the diagonal size. In practice, $M \gg p$, and the distribution of $y = hM$ tends toward the gamma distribution with $f(y) = \frac{1}{(p-1)!} y^{p-1} e^{-y}$ as the probability density function. The corresponding cumulative distribution function is the regularized lower incomplete gamma function $\gamma(p, y)$. Note that the expected mean of hM/p is 1 for both beta and gamma distributions. If we take a probability of rejection p_{rej} of 2×0.05, we expect that 90% of the values of hM/p associated with non-leverage points lie in the interval $[\chi_{lower}; \chi_{upper}]$, with $\chi_{lower} = \frac{1}{p}\gamma^{-1}(p, 0.05)$ and $\chi_{upper} = \frac{1}{p}\gamma^{-1}(p, 0.95)$. For $p = 2$, the interval is approximately $[0.178; 2.372]$ and increases for smaller probabilities of rejection. In contrast, the hat matrix values associated with leverage points do not follow the beta distribution and are more likely to lie outside the interval.

This information can be integrated in the weights by replacing the definition (11) of the sequence $(\widehat{z_i}^k)_k$ by

$$\widehat{z_i}^{k+1} = \left(b^H w^k \nu(\widehat{z_i}^k) b\right)^{-1} \left(b w^k \nu(\widehat{z_i}^k) e_i\right) \quad \text{for } k \geq 0 \tag{16}$$

Here, w^k is the diagonal matrix of the leverage weights, and its diagonal elements lie in $[0; 1]$ such that w_{ii}^k is close to 1 when $h_{ii}(\nu^{k-1})M/p \in [\chi_{lower}; \chi_{upper}]$ and close to 0 otherwise. Several definitions of w^k are possible. The simplest definition would be the indicator function of the interval $[\chi_{lower}; \chi_{upper}]$. One can also use a smooth approximation of the indicator function, such as

$$f(y) = \exp\Big[e^{-\chi_{upper}^2} - e^{\chi_{upper}(y_i - \chi_{upper})} + e^{-\log(\chi_{lower})^2} - e^{\log(\chi_{lower})(\log(y_i) - \log(\chi_{lower}))}\Big] \tag{17}$$

from which we can define w^k by $w_{ii}^k = f(h_{ii}(\nu^{k-1})M/p)$. To avoid oscillations preventing the convergence of $(\widehat{z_i}^k)_k$, Chave and Thomson (2003) proposed computing w^k as $w_{ii}^k = w_{ii}^{k-1} f(h_{ii}(\nu^{k-1})M/p)$. This ensures that excluding an extreme value at one step of the iterations is permanent.

As with the M-estimate computation, the most robust weighting is not stable; thus, a sequence of estimates are computed beginning with the most stable weighting and ending with the most robust. We define a sequence of N_{BI} leverage intervals $[\chi_{lower}^i; \chi_{upper}^i]$, where

$$\chi_{lower}^i = \frac{1}{2^{N_{BI}-i}} \chi_{lower} \quad \text{and} \quad \chi_{upper}^i = 2^{N_{BI}-i} \chi_{upper} \tag{18}$$

The impedance bounded influence estimate is then computed as follows:

$$
\left|
\begin{array}{rl}
1. & \text{Compute } \widehat{z}_{\mathrm{OLS}} \text{ defined by Equation (4).}\\
2.1 & \text{Compute } \widehat{z}^{1}_{\mathrm{Huber}} \text{ defined by Equation (16) with}\\
 & \nu_{\mathrm{Huber}} \text{ (12) as weighting function,}\\
 & \widehat{z}_{\mathrm{OLS}} \text{ as initial value}\\
 & \text{and } [\chi^{1}_{lower};\, \chi^{1}_{upper}] \text{ as leverage weight interval.}\\
2.2 & \text{Compute } \widehat{z}^{2}_{\mathrm{Huber}} \text{ defined by Equation (16) with}\\
 & \nu_{\mathrm{Huber}} \text{ (12) as weighting function,}\\
 & \widehat{z}^{1}_{\mathrm{Huber}} \text{ as initial value}\\
 & \text{and } [\chi^{2}_{lower};\, \chi^{2}_{upper}] \text{ as leverage weight interval.}\\
\ldots & \\
2.N_{BI} & \text{Compute } \widehat{z}^{N_{BI}}_{\mathrm{Huber}} \text{ defined by Equation (16) with}\\
 & \nu_{\mathrm{Huber}} \text{ (12) as weighting function,}\\
 & \widehat{z}^{N_{BI}-1}_{\mathrm{Huber}} \text{ as initial value}\\
 & \text{and } [\chi^{N_{BI}}_{lower};\, \chi^{N_{BI}}_{upper}] \text{ as leverage weight interval.}\\
3. & \text{Compute } \widehat{z}_{\mathrm{Thomson}} \text{ defined by equation (16) with}\\
 & \nu_{\mathrm{Thomson}} \text{ (13) as weighting function,}\\
 & \widehat{z}^{N_{BI}}_{\mathrm{Huber}} \text{ as initial value}\\
 & \text{and } [\chi^{N_{BI}}_{lower};\, \chi^{N_{BI}}_{upper}] \text{ as leverage weight interval.}
\end{array}
\right.
\tag{19}
$$

2.6. Discrete Fourier Transform Computation and Prefilters

Given magnetic and electric time-series data of duration D (in seconds), it can be split into N_{stat} time windows of the same size and evenly spaced. For each window and for each time signal, one Discrete Fourier Transform (DFT) coefficient will be computed for the target frequency f_k. In practice, the window length is adapted to the target frequency f_k by fixing the number period, N_{per}, thus the duration of one window is $\frac{N_{per}}{f_k}$. Further more, the shifting between two consecutive windows can be controlled to increase N_{stat} or on the contrary to decrease the correlation between windows. This shifting is expressed by the overlapping ratio, $c_{overlap}$. All these parameters are linked together by the relationship:

$$D = c_{overlap} N_{stat} \frac{N_{per}}{f_k}$$

that is used to determined the number of windows as follow

$$N_{stat} = \frac{D f_k}{c_{overlap} N_{per}}.$$

Once time series are divided into N_{stat} time portions, DFTs are computed using Slepian data taper windows (also known as discrete prolate spheroidal Slepian sequences), using a time bandwidth τ (τ=1, 2, 3 or 4). No prewhitening is performed on the data.

After DFT computation, MT data coefficients can be pre-filtered using thresholds on the coefficient of determination. One can use both a lower and an upper threshold value. The filtering eliminates some of the DFT pairs and keeps the others unchanged. The filter works on blocks of consecutive DFT pairs (EB or BBr). It computes the coefficient of determination for the block (a default value of 10 data by block is used). The block is eliminated if this coefficient is outside the thresholds.

Available thresholds are (i) Th_{EB} when pairs of local electric and magnetic fields are considered and (ii) Th_{BB^r} when pairs of local magnetic and remote magnetic fields are considered.

3. MAIN FEATURES OF THE RAZORBACK LIBRARY

In this section, we describe three core features of Razorback: handling time-series data, computing Fourier coefficients, and estimating the response function. We show how they are designed to preserve modularity and flexibility. The presentation relies on the theoretical background presented in section 2 and is accompanied by code examples. The full documentation is available at http://razorback.readthedocs.io.

3.1. Handling Time-Series Data

Time-series data are the primary material of the processing. Handling these data does not involve sophisticated algorithms; however, such algorithms should be easy to work with, particularly when one aims to consider several combinations among many data. This is what Razorback proposes by introducing dedicated structures with specific behaviors. **Appendix 1** presents an extended example of time data manipulation with the library.

3.1.1. `SyncSignal`

The most elementary time data that we consider are sampled signals. A sampled signal is fully described by the sequence of the values taken by the signal, the sampling rate and the starting time. However, some measuring instruments provide a raw signal that differs from the signal of interest. The filter to be applied depends on the instrument, and the corresponding TF is called the calibration function. Rather than forcing an early conversion from a raw signal to a calibrated signal, Razorback proposes the possibility of attaching a calibration function to any raw signal.

Razorback adopts a strict definition for synchronous signals: two signals are synchronous when they have the same sampling rate, the same starting time and the same length. According to this definition, one can easily store a group of synchronous signals in an object gathering the common sampling rate, the common starting time, and, for each individual signal, the sequence of values and the calibration function. The corresponding class is named `SyncSignal`. In this simple structure, an individual signal can be retrieved from its index.

3.1.2. `SignalSet`

With our strict definition of synchronousness, there are many cases where one cannot store all the relevant data in one `SyncSignal` object. For instance, the starting and ending times of acquisition may differ for some signals, acquisition may be discontinuous and present time intervals without any data, or consecutive acquisitions may have been run with different sampling rates. These complications are overcome by the `SignalSet` object, which gathers distinct acquisition runs of the same channels. In practice, a `SignalSet` object is a group of `SyncSignal` objects with a labeling system. All the `SyncSignal` objects handled by a `SignalSet` object must have the same number of individual signals, and they cannot overlap in time. These restrictions ensure that at any given time, either all channels of a `SignalSet` have one and only one value or there is no value for any channel. The labeling system is a meaningful way to retrieve one channel, or a group of channels, from a `SignalSet` object. It associates some character strings to groups of one or more indices. In this way, one can refer to channels, or groups of channels, by names rather than by indices.

To gather the relevant set of data to process in one `SignalSet` object from among all the available data, it is possible to group or split existing `SignalSet` objects to produce new ones. These operations do not duplicate the data values, which allows building any combination without memory cost. `SignalSet` objects can be grouped in two ways: joining different acquisition runs of the same channels or merging distinct channels. They can also be split in several ways: selecting some channels and runs or narrowing the time range.

3.1.3. `Inventory`

A `SignalSet` object is able to gather all the time data required to compute the TF between different channels. Building it requires exploring among all the available data to select and combine the relevant ones; this process can be performed using the `SignalSet` features. Razorback also provides the `Inventory` object to ease this task. An `Inventory` object is a container that gathers several `SignalSet` objects without any constraints. The `Inventory` object offers two types of behavior. On the one hand, it can build a new `Inventory` containing less data by selecting some channels and runs or by narrowing the time range. On the other hand, it can produce a `SignalSet` that contains as many signals as possible. In practice, we start by constructing an `Inventory` object that gathers all the data acquired for the survey. Then, for each site, we extract a sub `Inventory`, produce the adapted `SignalSet` and estimate the TF.

The data volume of an entire survey exceeding the available computer memory is not rare. This would prevent fully using the `Inventory` object. Razorback overcomes this limitation thanks to the Dask library (Dask Development Team, 2016). By storing the time data in a `dask.array` object rather than a classical (Numpy) array, the memory cost of the `Inventory` of an entire survey becomes negligible. The time data files are only loaded when needed during the computation of the Fourier coefficients, and then they are unloaded. When we use the `dask.array`, the memory footprint of the processing no longer depends on the size of the survey.

A jupyter notebook dedicated to data handling is provided by the authors and available at https://github.com/BRGM/razorback/blob/doc/docs/source/tutorials/signalset.ipynb.

3.2. Computing Fourier Coefficients

The `SyncSignal` and `SignalSet` objects provide the method `fourier_coefficients(freq, Nper, overlap, window)` to compute the Fourier coefficients of time data. Here, `freq` is the frequency of interest f_k, `Nper` is the number of $1/f_k$ periods, `overlap` is the overlapping coefficient $0 < c < 1$, and `window` is an object coding the data taper window. The `window` argument must be a function that takes the size of the window (a positive integer) and returns a discrete data taper (an array of the given size). When a calibration has been provided to the `SyncSignal` or `SignalSet` object, the calibration value at the given frequency is integrated in the resulting Fourier coefficients. More precisely, the returned value is the Fourier coefficient computed from the raw time value divided by the calibration value. Razorback provides the `slepian_window(half_bandwidth)` function that builds the `window` argument corresponding to the Slepian data taper with a given half bandwidth. The following code shows how to compute the Fourier coefficients from a given `signal`:

```
>>> from razorback.fourier_transform import
    slepian_window
>>> window = slepian_window(4)
>>> coeffs, winfo = signal.fourier_coefficients (freq
    , Kper, overlap, window)
>>> Nw, Lw, shift = winfo
```

As shown, two values are returned, `coeffs` and `winfo`. `coeffs` contains the Fourier coefficients of the different channels of `signal`. `winfo` summarizes information on the sliding window: the number of windows, the size of the discrete window and the index shift between consecutive windows. Note that this code works for `SyncSignal` and `SignalSet` objects. However, a `SignalSet` object can hold multiple runs, possibly at different sampling rates. In that case, `coeffs` contains the Fourier coefficients of all the runs, and `winfo` describes the sliding window on each run.

3.3. Estimating Response Function

Algorithms (14) and (19) show a similar structure. They compute a sequence of TFs, where each step of the sequence uses a different weighting strategy and is initiated with the result of the previous step. Each step solves a non-linear weighted least squares problem; see equations (11) and (16). Razorback provides a single function that implements the common logic of these algorithms: `transfer_function(outputs, inputs, weights=(None,), init=None, invalid_idx=None)`. Here, `outputs` is the list of arrays of the Fourier coefficients of the output (electric field components in MT), `inputs` is the list of arrays of the Fourier coefficients of the input (magnetic field components in MT), `weights` is the list of functions implementing the weighting strategy at each step, `init` is the initial estimate of the TF, and `invalid_idx` is the array of indices of the initially rejected coefficients. The important argument here is `weights`. Depending on its value, the function `transfer_function` can perform algorithm (14) or (19). It also allows performing other algorithms with different weighting strategies. When the special value `None` is placed in the list `weights`, it indicates that the corresponding step is the least squares estimation (see equation (4)). Thus, `None` is often the first element of `weights`. In Razorback, a weight function is a function that returns a one-dimensional array of the size of the data that contains the computed weight values. The function signature must be `func(it, residual, outputs, inputs, invalid_idx)`, where `it` is the iteration number of the inner loop (see Equation 11 or 16), `residual` is an array containing the residual values (see Equation 2), `outputs` and `inputs` are the same as for `transfer_function`, and `invalid_idx` is the array of indices of the rejected coefficients at the current step.

In the following, we show how the function `transfer_function` is used to perform the least squares estimator (4), the M-estimator (14) and the bounded influence estimator (19). We consider two arrays of complex values and size $2 \times N_k$, `E` and `B`, that contain the Fourier coefficients of the two electric field components and those of the magnetic field components. Computing the least squares impedance estimate (4) with the function `transfer_function` is simply done by using the default parameter values:

```
from razorback.mestimator import transfer_function
Z_lsq, rixd = transfer_function(E, B)
```

For the M-estimate, we need to specify the `weights` argument according to (14). Razorback provides implementations of the Huber weighting function (12) and the Thomson weighting function (13). The code is written as follows:

```
from razorback.mestimator import transfer_function
from razorback.weights import Huber, Thomson
weights_mest = [None, Huber(), Thomson()]
Z_mest, rixd = transfer_function(E, B, weights_mest)
```

The weighting function sequence for the bounded influence method (see Algorithm 19) is more elaborate. Preparing the `weights` argument can be performed as follows:

```
from razorback.mestimator import transfer_function
from razorback.weights import Huber, Thomson,
    BoundedInfluenceStep
from scipy.special import gammaincinv
p = 2  # for MT
n_bi_steps = 3  # number of intermediate BI step
reject_prob = 0.1
lower = gammaincinv(p, 0.5*reject_prob)
upper = gammaincinv(p, 1-0.5*reject_prob)
weights_bi = ([None]
    + [BoundedInfluenceStep(Huber(), lower*2**-i,
        upper*2**i)
       for i in range(n_bi_steps)[::-1]]
    + [BoundedInfluenceStep(Thomson(), lower, upper)
        ])
Z_bi, rixd = transfer_function(E, B, weights_bi)
```

To ease the use of the M-estimator and bounded influence methods, Razorback provides helpers for the corresponding weighting function sequences. Here is a shorter way to compute the same quantities `Z_mest` and `Z_bi` as above:

```
from razorback.mestimator import transfer_function
from razorback.weights import mest_weights,
    bi_weights
Z_mest, rixd = transfer_function(E, B, mest_weights)
Z_bi, rixd = transfer_function(E, B, bi_weights(
    reject_prob, n_bi_steps, p){})}
```

The use of a pre-filter on the Fourier coefficients is achieved thanks to the `invalid_idx` argument of `transfer_function`. Razorback provides an implementation of the filter described in section 2.6. Using the M-estimator with this filter can be performed as follows:

```
from razorback.mestimator import transfer_function
from razorback.weights import mest_weights
from razorbck.prefilter import cod_filter
filtered_idx = [cod_filter(ei, B) for ei in E]
Z_mest, rixd = transfer_function(E, B, mest_weights,
    invalid_idx=filtered_idx)
```

The two-stage RR estimate defined in Equation (8) can also easily be computed by using the `transfer_function` function. For instance, we show how it can be computed in combination with the M-estimator. Here, `Q` is an array of complex values of size $q \times N_k$ that contains the Fourier coefficients of all the components of the RR magnetic fields:

```
from razorback.mestimator import transfer_function,
    merge_invalid_indices
from razorback.weights import mest_weights
T, rr_ridx = transfer_function(B, Q, mest_weights)
predicted_B = T.dot(B)
rr_ridx = len(E) * [merge_invalid_indices(rr_idx)]
Z_rr, ridx = transfer_function(E, predicted_B,
    mest_weights, invalid_idx=rr_ridx)
```

3.4. Helper Function for Impedance Estimate

Section 3.3 shows how to use the low-level function `transfer_function` to compute an impedance estimate for given data Fourier coefficients with different algorithms. At a higher level, Razorback provides the helper function `impedance` that integrates the different algorithms, the use of a pre-filter and the computation of the Fourier coefficients. It aims to provide a simple yet fully controlled use of the methods described in section 2. The function signature is `impedance(data, l_freq, weights=(None,), prefilter=None, fourier_opts=None, remote=None, remote_weights=None, remote_prefilter=None, tag_elec=’E’, tag_mag=’B’)`. The first two arguments are the `SignalSet` object containing all the relevant channels and the list of the frequencies to investigate. The other arguments are optional and are used for customizing the processing. The `weights` and `prefilter` arguments correspond to those described in section 3.3. The `fourier_opts` argument corresponds to the options to pass to the `fourier_coefficients` method described in section 3.2. Its default value is `dict(Nper=8, overlap=.71, window=slepian_window(4))`. The `remote`, `remote_weights` and `remote_prefilter` arguments are relative to the two-stage RR method defined in equation (8) and illustrated in section 3.3. To activate the RR method, one must provide the tag gathering the remote channels in `data` to the `remote` argument. The `remote_weights` and `remote_prefilter` arguments are used to customize the first stage of the method, and the default value `None` indicates to use the corresponding settings for the second stage (`weights` and `prefilter`). The `tag_elec` and `tag_mag` arguments are the tags gathering the electric and magnetic channels, and their default values are `’E’` and `’B’`.

The following code shows the computation of impedance for 5 frequencies using the RR method with the bounded influence estimator:

```
>>> from razorback.utils import impedance
>>> from razorback.weights import bi_weights
>>> l_freq = [0.125, 1, 8, 64, 512]
>>> result = impedance(data, l_freq, bi_weights(0.1),
     remote=’Bremote’)
>>> result.impedance  # computed impedance at each
    frequency
>>> result.error  # estimated error on impedance at
    each frequency
>>> result.invalid_time  # rejected times at each
    frequency
```

The value returned by the `impedance` function contains three values: the estimated impedance tensor at each frequency, the estimated error and the times of the sliding windows excluded by the robust estimator. Using the `impedance` function is simple; it mainly requires the preparation of the `SignalSet` object (the `data` argument) as described in section 3.1.

4. VALIDATION

We perform data processing using single site (SS) and two-stage RR configurations with both the M-estimator (ME) operator $\mathbb{O}_{ME}$ and bounded influence (BI) operator $\mathbb{O}_{BI}$. We compare the ME and BI results obtained with Razorback and the BIRRP code (bounded influence remote reference processing, Chave and Thomson, 2004) using the same processing parameters.

The dataset consists of two synchronous and permanent MT stations installed on the La Fournaise volcano (Réunion Island, France): the BAV station is located 8.2 km northwest of the summit of the volcano, while the CSV station is on the western base of the caldera (see **Figure 1**). Both stations recorded the horizontal NS and EW components of the electromagnetic field with a 50 *mHz* sampling rate during the year 1997. A detailed description and analysis of the dataset can be found in Wawrzyniak et al. (2017).

The CSV data are processed using SS configuration, and RR configuration with BAV magnetic data. Sixteen target frequencies $(f_k)_k$ are defined ranging from 1.56 to 12.5 mHz. K_{per} is fixed to 128, and $c_{overlap}$ is 0.71. This yields a number N_{stat} of 10,000 data Fourier coefficients for the highest frequency and 1,000 for the lowest. The time bandwidth factor τ (parameter TBW in the code) is 4, and the lower thresholds Th_{BB^r} and Th_{EB} are set to 0. The BI regression is controlled by the probability of rejection p_{rej}, which is 0.05, and the number of BI iterations N_{BI} is set to 3.

The ME and BI results in SS configuration are shown in **Figures 2** and **3**, respectively. First, we observe SS apparent resistivity relative difference and phase difference between BIRRP and Razorback. The ME apparent resistivity differs by less than 10% on the xy, yx, and yy components (**Figure 2**). The phase difference is less than 2°. The relative differences on the xx component are higher, but the absolute value ρ_{xx} is two orders of magnitude lower than ρ_{xy}, ρ_{yx}, and ρ_{yy}.

The BI apparent resistivity differs from less than 12% on the *xy* and *yx* components to less than 20 % on the yy component (**Figure 3**). The phase difference is less than 3° on xy and yx and 4° on yy. Although the principle is the same, our implementation of the bounded influence algorithm differs in the computation of leverage weights from the hat matrix and in the definition of the increment of intermediate steps of the BI algorithm.

The two-stage RR ME (**Figure 4**) shows the apparent resistivity relative difference reaching 20 % on the xy component, less than 3 % on the yx component and less than 30 % on yy. The *xy* and *yy* components involve the H_y local and remote magnetic fields. The remote magnetic field component H_y has been diagnosed as biased in Wawrzyniak et al. (2017).

Thus, the impact of introducing a noisy RR station in the two-stage RR method leads to a moderate discrepancy between the BIRRP and Razorback estimates.

The two-stage RR BI (**Figure 5**) shows the apparent resistivity relative difference reaching 10 % on the *xy* and *yx* component and reaching several tens of percent on *xx* and *yy*. Absolute phase difference is still below 3° on*xy* and *yx*.

Apart from the data loading [the `load_data()` function], the impedance estimates presented in **Figures 5**, **3** for Razorback are obtained with the following code:

```
>>> data = load_data(’path/to/data/file’)
>>> print(data)
SignalSet: 6 channels, 1 run
tags: {’B’: (2, 3), ’Hremote’: (4, 5), ’E’: (0, 1)}
----------  -------------------  -------------------
  sampling                start                 stop
      0.05  1997-01-01 00:00:00  1997-12-31 23:59:40
----------  -------------------  -------------------

>>> # frequency parameters
>>> freq_min = 0.0015625
>>> freq_max = 0.02187
>>> Nfreq = 16

>>> # weighting parameters
>>> prej = 0.05
>>> n_bi_steps = 3

>>> # Fourier parameters
>>> TBW = 4
>>> c_overlap = 0.71
>>> Nper = 120

>>> l_freq = np.logspace(np.log10(freq_min), np.log10
    (freq_max), Nfreq)

>>> z_ss, l_ivt_ss, l_err_ss = impedance(
...     data, l_freq,
...     weights=bi_weights(prej, n_bi_steps, 2),
...     fourier_opts=dict(Nper=Nper, overlap=
    c_overlap, window=slepian_window(TBW))
...)

>>> z_rr, l_ivt_rr, l_err_rr = impedance(
...     data, l_freq,
...     weights=bi_weights(prej, n_bi_steps, 2),
...     fourier_opts=dict(Nper=Nper, overlap=
    c_overlap, window=slepian_window(TBW)),
...     remote=’Hremote’
...)
```

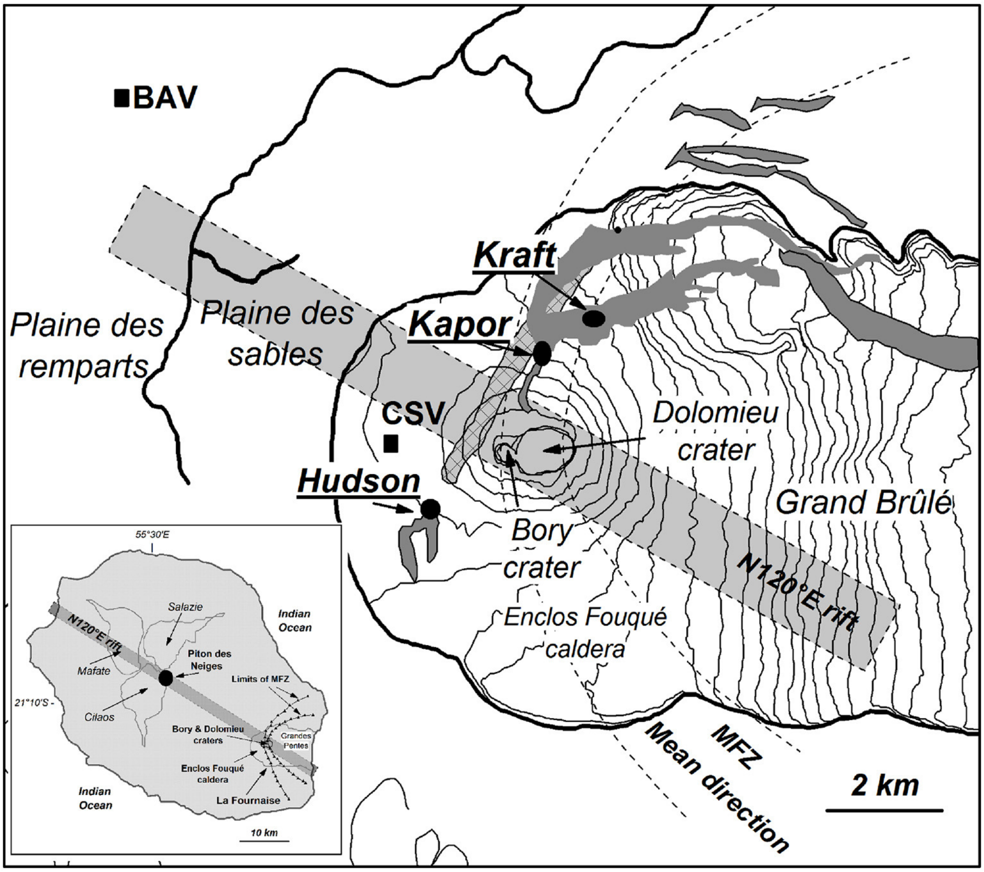


FIGURE 1 | Sketch of La Fournaise volcano, from Wawrzyniak et al. (2017). BAV and CSV are the electromagnetic stations. Lava flows emitted by Krafft, Kapor, and Hudson cones are in dark gray color. Dashed lines represent the main fracture zone along which most fissure eruptions occur. The gray rectangle illustrates the regional N120°E volcanic and fissural axis. Gray cross pattern corresponds to the trace of main earthquakes associated with the March 9, 1998, crisis.

5. ADVANCED USES

5.1. Testing All Possible Remote Reference Combinations

In the following, we show how the Razorback library allows the user to run one processing method, such as classical M-estimator or bounded influence regression, on a given signal set for any combination of RR stations. We assess the efficiencies of both ME and BI methods on a noisy peri-urban dataset.

Our experimental dataset is a synchronous array of MT stations from a CSEM/MT survey realized in the framework of the European FP7 project IMAGE. The survey area is a 10 by 10 km square located on the western side of the city of Strasbourg (**Figure 6**). The main geothermal targets in the Upper Rhine Graben are fractured zones within the basement or at the transition zone between the basement and the sedimentary cover. Unfortunately, in such a peri-urban context, the exploration of sedimentary basins using MT is challenging due to the presence of anthropogenic sources (DC railway, factories, power lines and so forth), leading to biased MT soundings.

We use 6 synchronous MT stations (ADU07 acquisition system, Metronix). Four "local" stations are located in the area of interest (stations 2, 4, 6, and 9) and have a 128 Hz sampling rate. Two "distant" remote reference stations are also used: one is installed in Schwabwiller (5 channel MT station, 30 km North), and the other is the Welschbruch geomagnetic observatory (recording horizontal magnetic field only, located at the Welschbruch pass, Le Howald, 30 km South).

We process station 4 using sites 2, 6, and 9 as "local" remote stations and Schwabwiller (RR99) and Welschbruch (RR100) stations as "distant" remote stations. This allows 31 possible combinations of remote stations in addition to the single site processing. We perform two-stage M-estimate (ME) and bounded influence (BI) robust processing without pre-filtering for both the first and second stages. Thirty-two output frequencies, ranging from 1 mHz to 32 Hz, are targeted. At the lowest frequencies, for some combinations of RRs, the computation does not converge. This is due to the weak signal-to-noise ratio and the small amount of Fourier data.

Quality assessment is performed on the MT soundings obtained from combinations of remote channels. First, components of apparent resistivity $\rho^{a}_{xy,yx}$ and impedance phase $\phi_{xy,yx}$ are displayed in **Figure 7** for ME and BI. ME

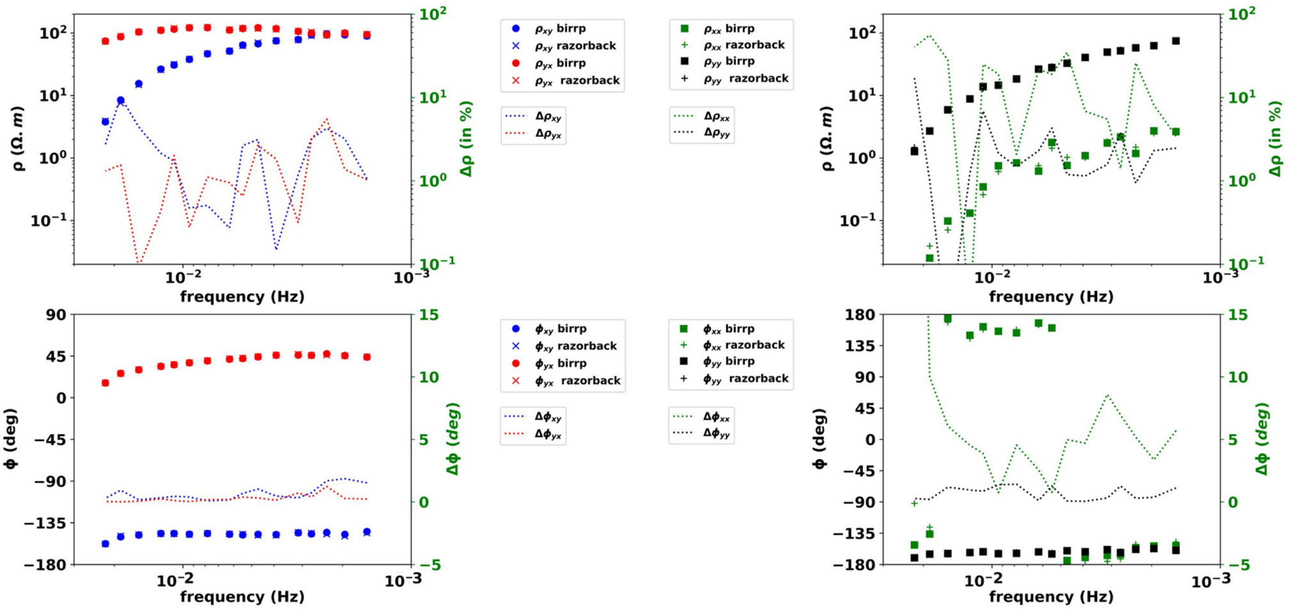


FIGURE 2 | M-estimator results in SS configuration. Upper left: apparent resistivity xy in blue and yx in red from birrp (circles) and razorback (crosses) with scales on the left axis, relative differences between birrp and razorback results (dotted lines) for xy (blue) and yx (red) with scales on the right axis. Lower left: same as upper left but for phase but with absolute differences (dotted lines). Upper Right: apparent resistivity xx in green and yy in black from birrp (squares) and razorback (plus) with scales on the left axis, relative differences between birrp and razorback results (dotted lines) for xx (green) and yy (black) with scales on the right axis. Lower Right: same as upper left but for phase but with absolute differences (dotted lines).

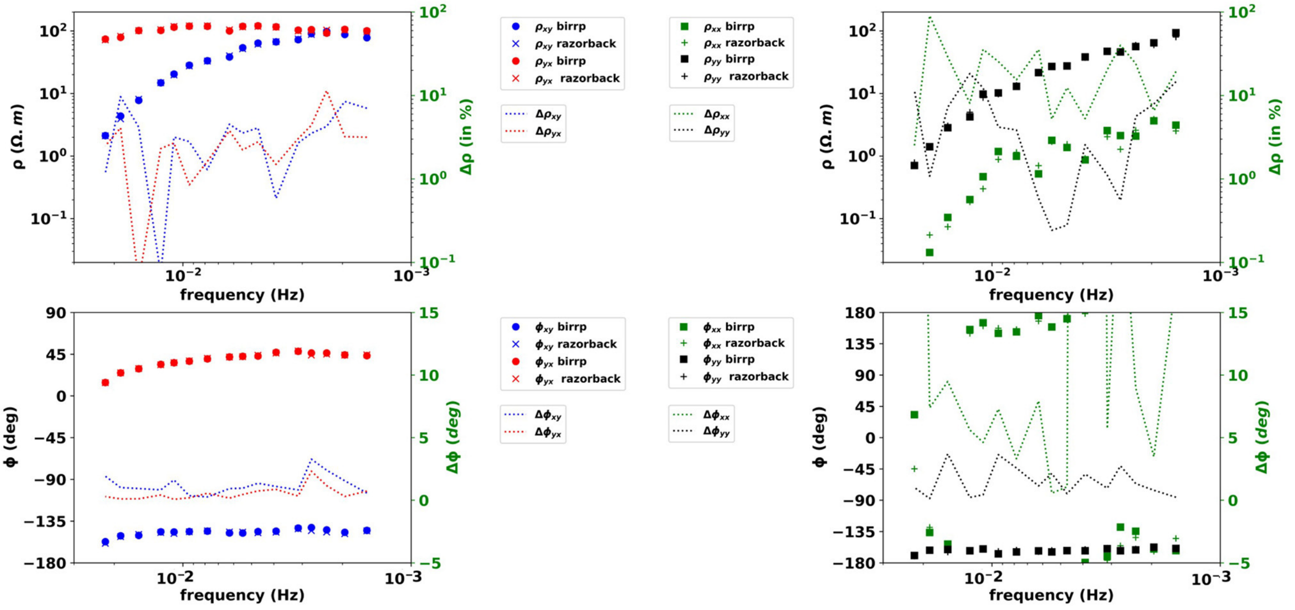


FIGURE 3 | Bounded Influence results in SS configuration. Same legend as **Figure 2** but for BI results.

soundings (**Figure 7**, upper part) exhibit significant variability in the [20 mHz–1 Hz] band. Some combinations show non-physical resistivity and phase variations (i.e., artifacts up to one order of magnitude on the ρ_{yx} component and 40° on ϕ_{yx}). BI soundings (**Figure 7**, lower part) show less variability in the same frequency band as soon as distant RRs are used. Similar to the ME results, ϕ_{yx} exhibits a 35° shift, centered on the 0.1 Hz frequency, but with a narrower frequency band imprint. There are more

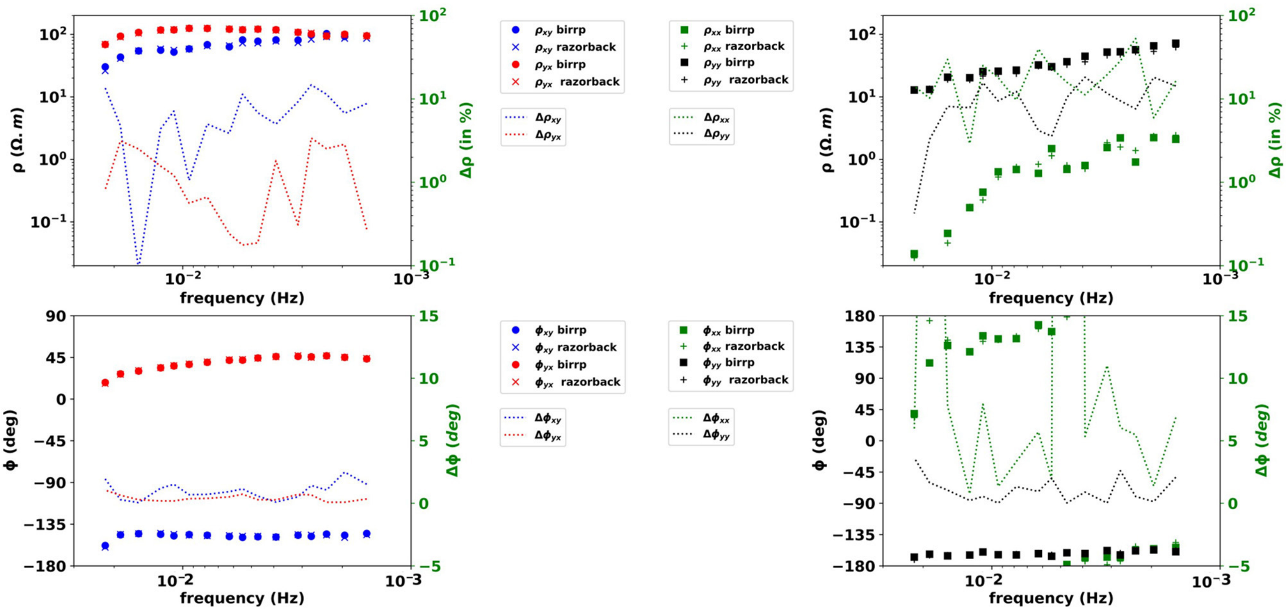


FIGURE 4 | M-estimator results in RR configuration. Same legend as **Figure 2** but for RR results.

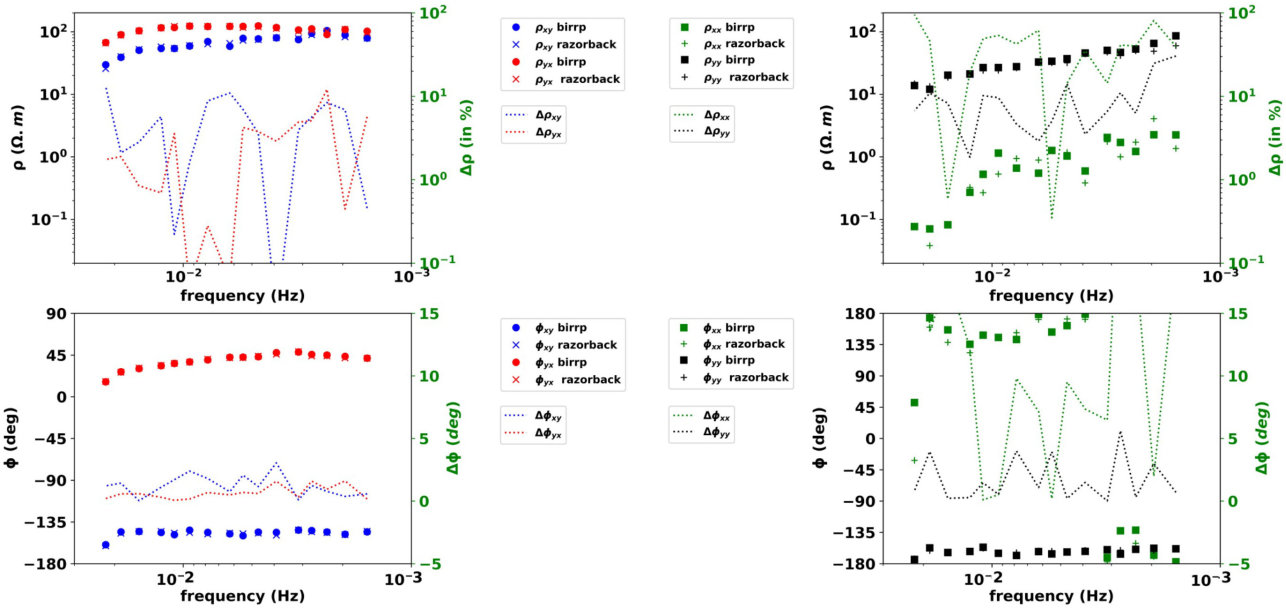


FIGURE 5 | Bounded Influence results in RR configuration. Same legend as **Figure 2** but for RR results.

artifacts with wider frequency bands and larger amplitudes in the ME results than in the BI results.

The sounding quality assessment is completed with phase tensor (PT) analysis (Caldwell et al., 2004). Booker (2014) suggests that "smooth variation of the phase tensor with period and position is a strong indicator of data consistency." Some main features of the PT are the orientation of its principal axis $\alpha - \beta$; the length of its principal axis Φ_{max}; its ellipticity λ that ranges from 0 to 1, and its skew angle β which are indicators of the dimensionality of the data. Low values of ellipticity indicate a 1D medium (Bibby et al., 2005), whereas absolute values of β angles below 10° indicate a 2D medium (Booker, 2014). A 1D medium is characterized by a low λ value and associated with $|\beta| < 10°$. A 2D medium is characterized by larger values of λ and $|\beta| < 10°$. A 3D medium is characterized by $|\beta| > 10°$. Consequently, the normalized phase tensor, i.e., the PT with the

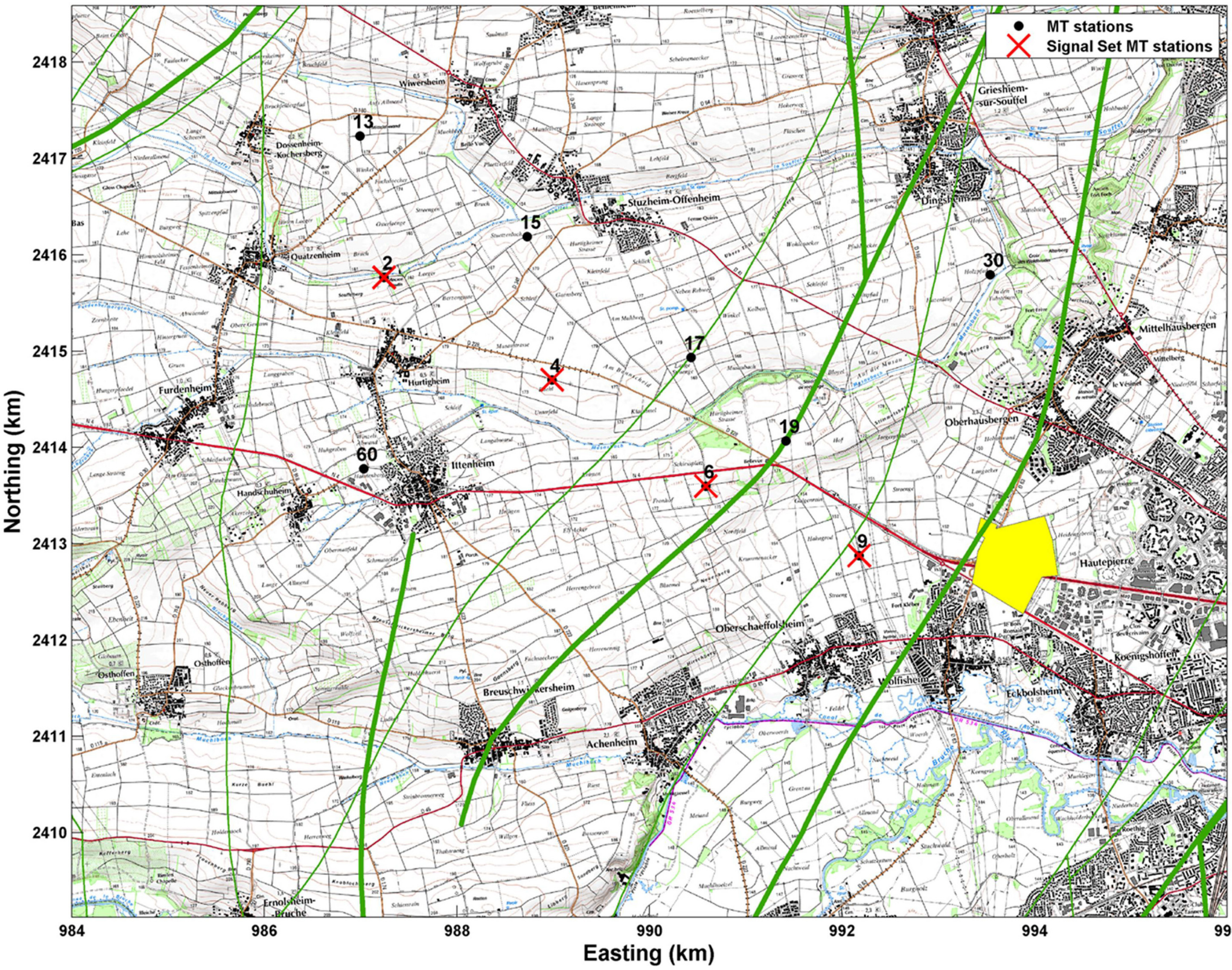


FIGURE 6 | Map of the MT stations deployed on the western side of Strasbourg. Black dots: complete station set. Red crosses: "local" station inserted in the `SignalSet` object. Green lines: major geological faults. Red line: seismic profile. Yellow area: geothermal plant area.

longer axis Φ_{max} normalized to 1, is displayed for all frequencies and RR combinations. Ellipses are filled with a color bar indexed either on their ellipticity value (left panel in **Figure 8**) or their β angle (right panel, same figure).

Comparing upper and lower part of **Figure 8** highlights the superiority of BI processing in the [0.2–2 Hz] band. In this band, the M-estimator leads to high ellipticity values, which would be associated with a 3D medium. In contrast, the BI results indicate a 1D medium with low ellipticity values. In addition, the BI PT curves exhibit smoother frequency variations.

An important observation can be made from the M-estimator results: the combination of a maximum of RR leads to discontinuous PT behavior. Smoother behavior is obtained for combinations of sites 02, 99 and 100. When sites 06 and 09 are added as RRs, discontinuous PTs are observed. This can have a high impact for any MT operator working on a peri-urban context and using ME two-stage processing: introducing additional noisy RR can degrade the sounding quality.

However, ME MT soundings associated with a selection of PT curves (RRs 99, 100, 99+100, 6+99+100, 6+9+99+100, and 2+6+9+99+100) are shown in **Figure 9**, upper part. The selected results are scattered, and phase and apparent resistivity artifacts still persist in the [0.02–1 Hz] band. Similarly, the BI soundings associated with the same combinations of RRs in **Figure 9**, lower part. Since 4 remotes are used, the MT soundings are similar. However, both phases ϕ_{xy} and ϕ_{yx} show sharp variations in the [0.5–1 Hz] band that can be attributed to persistent noise contamination of the data. Multiple RR two-stage BI processing helps reduce noise contamination of the dataset but cannot eliminate it in this case.

A jupyter notebook dedicated to the robust processing and handling of this data set is provided by the authors and available at https://github.com/BRGM/razorback/blob/doc/docs/source/tutorials/survey-study.ipynb.

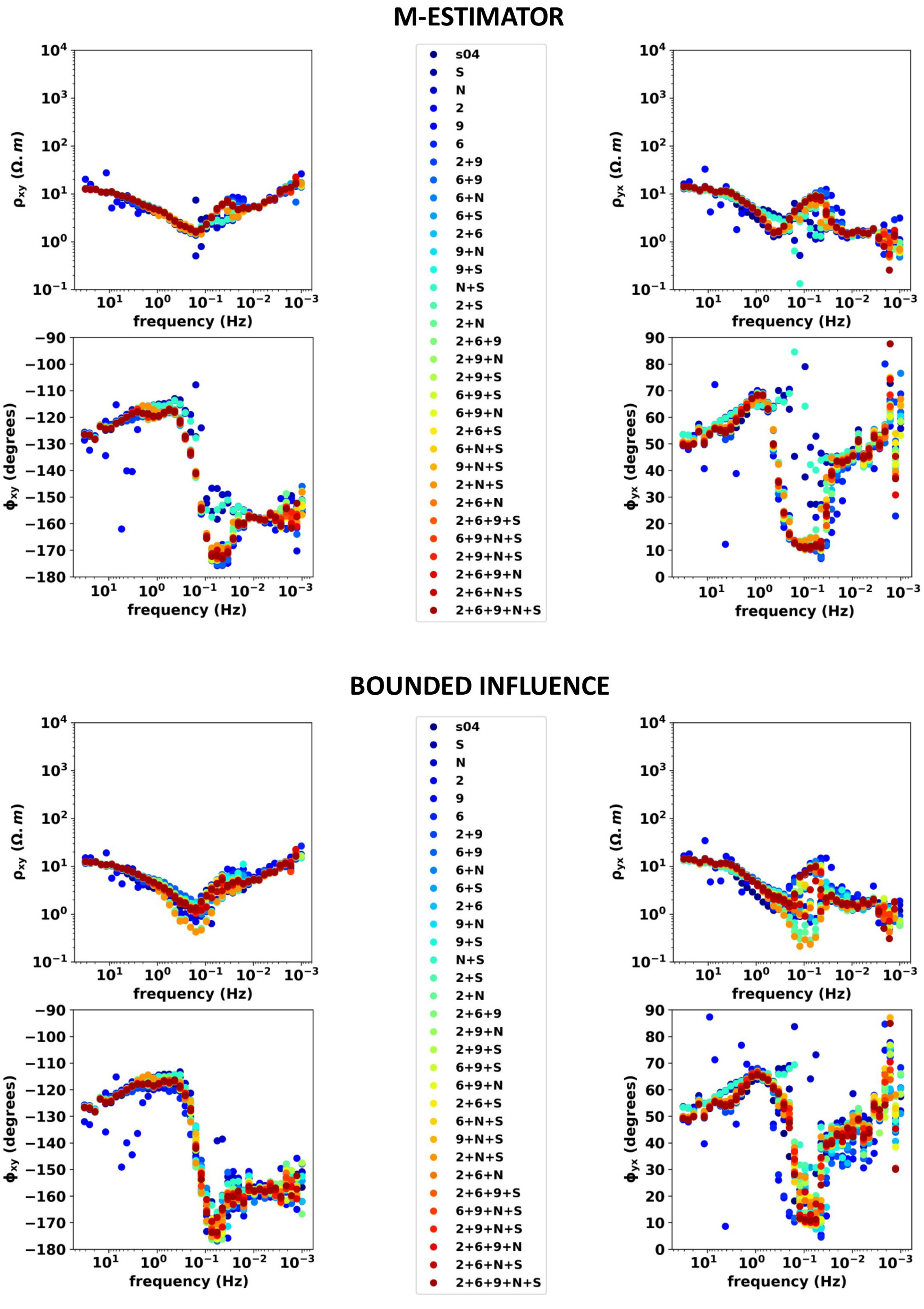


FIGURE 7 | Upper section: M-estimator results. SITE 04 with local RRs 2, 6, and 9 and distant RRs Welschbruch (RR100) and Schwabwiller (R99). No error bars. TF estimates for all possible combinations of RR stations. Apparent resistivity curves ρ_{xy} (top left) and ρ_{yx} (top right) and phases ϕ_{xy} (bottom left) and ϕ_{yx} (bottom right) are shown in dots, with a color code corresponding to the associated combination of RR stations(the associated legend is displayed on the center). Lower section: bounded Influence results. Same legend as above.

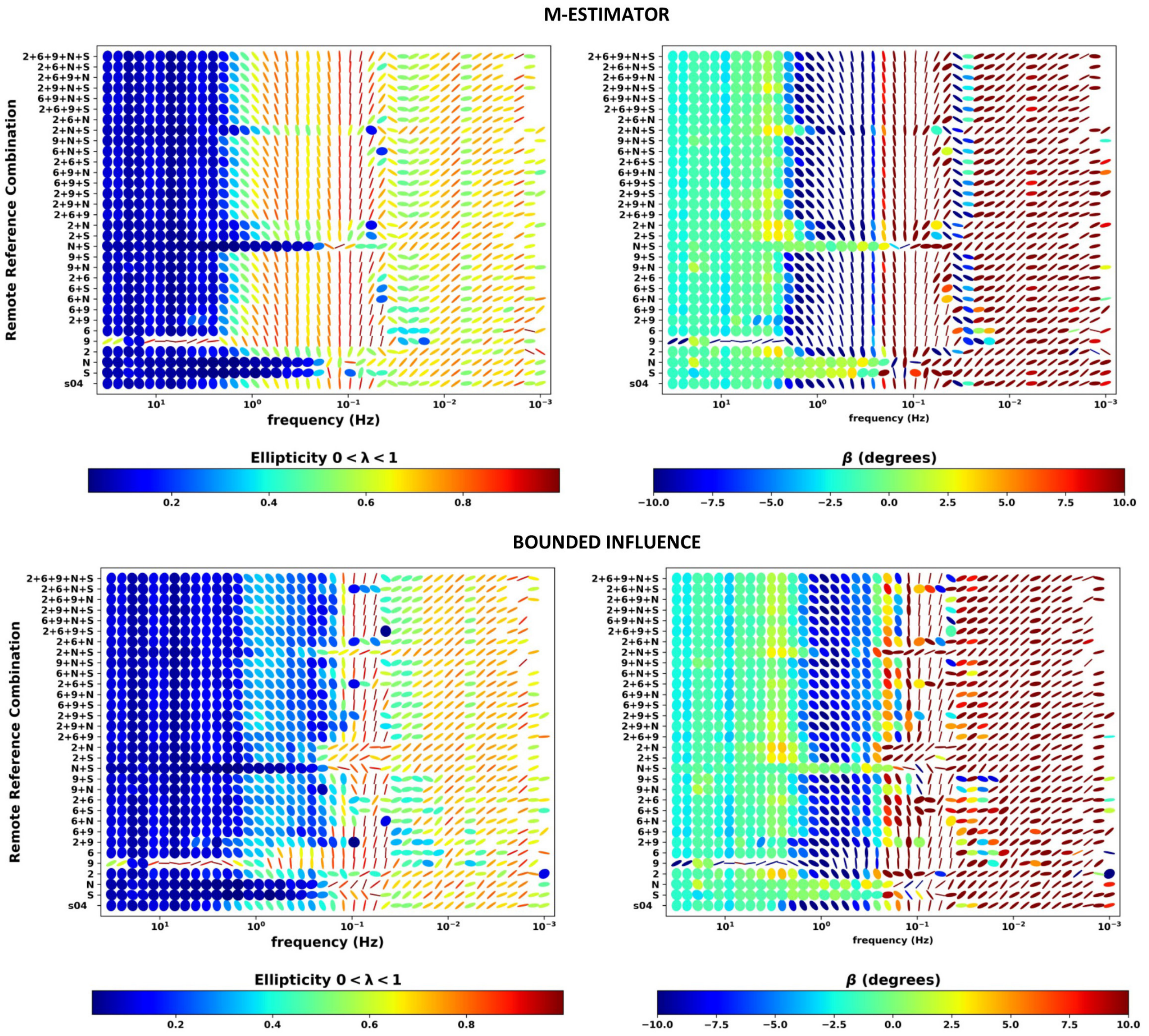


FIGURE 8 | Upper section: M-estimator results. Site 4 with local RRs 2, 6, and 9 and distant RRs Welschbruch (rr100) and Schwabwiller (rr99). Left: normalized phase tensor for all combinations of RR as a function frequency filled with their ellipticity λ value (1D indicator). Right: normalized phase tensors filled with *β* angle (2D indicator) value; the color bar is limited to [-10° 10°] range. Lower section: bounded influence results. Same legend as above.

5.2. Time-Lapse Magnetotellurics

The library also allows performing continuous time-lapse processing. In Wawrzyniak et al. (2017), time-lapse MT estimates were computed using bounded influence robust processing in both single site and RR configurations in the framework of volcano monitoring. The time resolution between consecutive estimates is of 48 h.

The dataset is the same as in section 4. Horizontal components of the electric and magnetic fields were sampled every 20 s. Continuous time series were available from 1996 to 1999 at CSV and from 1997 to March 20, 1998, at BAV. In March 1998, a major eruption occurred and lasted for 6 months, during which 60 10^6 m^3 of lava was expelled.

RR (not shown here) and single site processing at a period of 80 s show apparent resistivity determinant variations synchronous with the eruption (**Figure 10**). At CSV, the resistivity shows a continuous two order of magnitude decrease, reaching several local minima until April, and then slowly recovers its pre-eruptive values when the volcanic crisis reduced in activity. BAV shows a short one and a half order of magnitude decrease at the beginning of the eruptive crisis. Further data analyses are provided in Wawrzyniak et al. (2017).

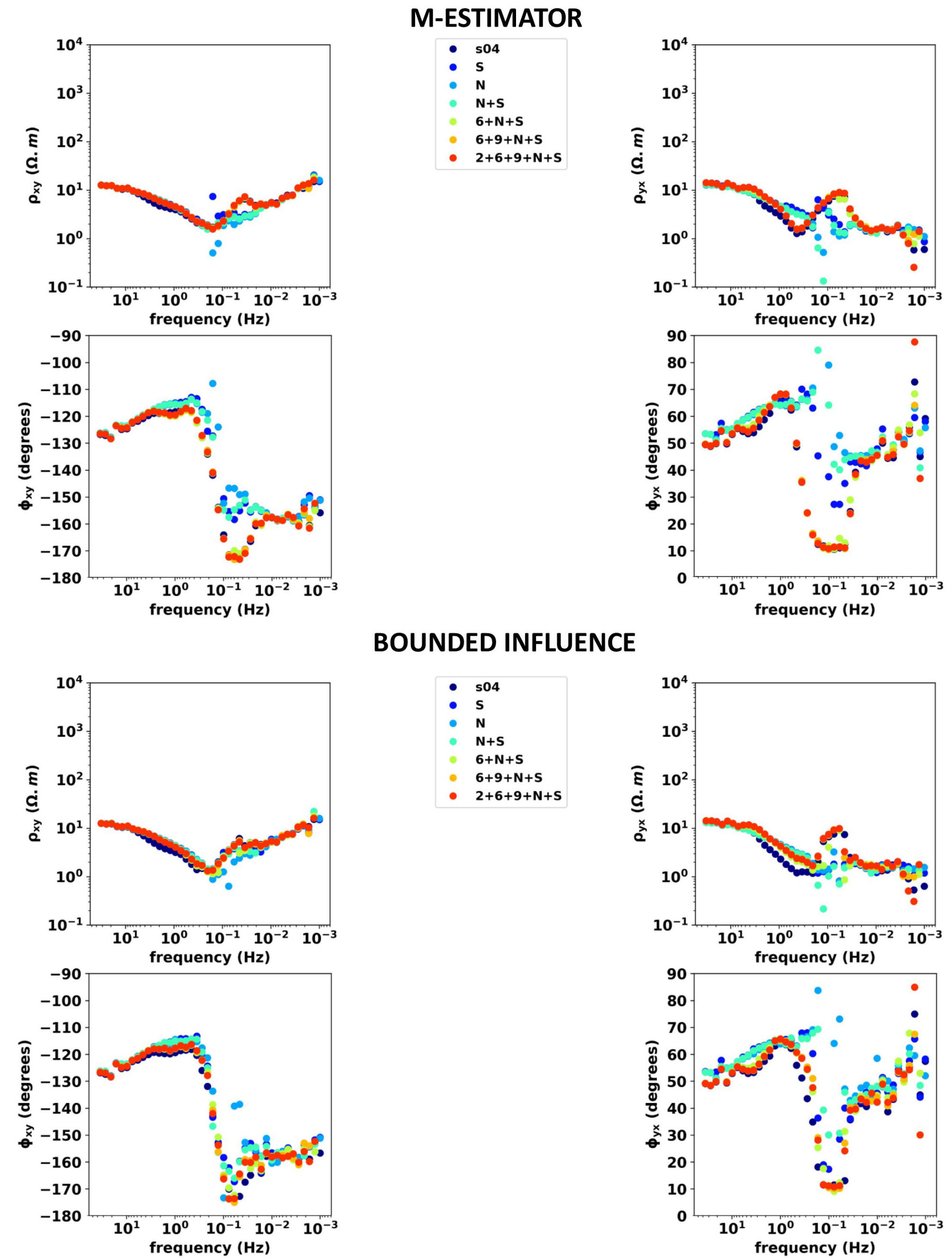


FIGURE 9 | Upper section: selected M-estimator results. SITE 04 with local RRs 2, 6, and 9 and distant RRs Welschbruch (RR100) and Schwabwiller (R99). No error bars. TF estimates for a selection of combinations of RR stations. Apparent resistivity curves ρ_{xy} (top left) and ρ_{yx} (top right) and phases ϕ_{xy} (bottom left) and ϕ_{yx} (bottom right) are shown in dots, with a color code corresponding to the associated combination of RR stations(the associated legend is displayed on the center). Lower section: selected Bounded Influence results. Same legend as above.

6. CONCLUSIONS

This paper shows the advantages of using a modular library for robust processing of MT array data. Razorback is designed to handle and process the large datasets commonly encountered in MT exploration surveys, with minimal memory footprint. After validation of TF estimates by comparison with existing codes, two kinds of study have been performed. First, we explored combinations of RRs for different robust procedures. This results in a large amount of estimates of one TF. The phase tensor analysis is used to compare the quality of the estimates. Moreover, the ability of the different robust methods to reduce the impact of

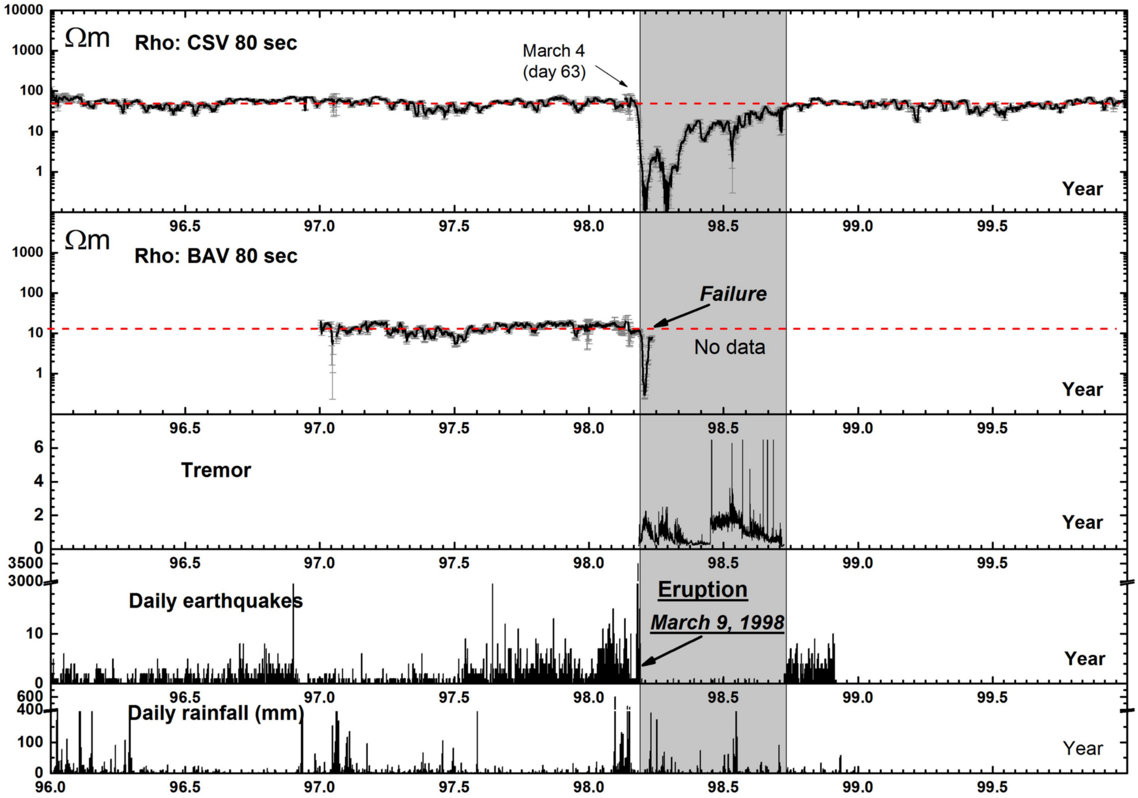


FIGURE 10 | Time-Lapse bounded influence processing, in single site configuration, from Wawrzyniak et al. (2017). From top to bottom: resistivity of the determinant of the impedance tensor values at CSV and BAV for the 80 s period computed by single MT method between 1996 and the end of 1999, tremor activity and daily number of earthquakes, and daily rainfall.

noise on soundings has been investigated. This is of particular interest for geophysicists processing a full MT survey dataset in anthropogenic environment. In addition, continuous time-lapse MT processing has been performed and shows promising results for subsurface monitoring of volcanoes or geothermal reservoirs.

The MT processing workflow mainly consists of (i) data analysis and transformation, (ii) TF estimation, and (iii) quality check of estimated quantities. We propose the open source Razorback library as a collaborative tool to perform these different tasks. For the first step, the library offers elaborated time series manipulations, state of the art DFT computation, and coefficient of determination pre-filtering. Alternate types of pre-filtering exists and can be included. Provided a detection method, the current features can eliminate the identified corrupted time segments. Concerning the second step, a category of standard robust procedures has been implemented. A well-tested set of weighting function sequences is available and can be easily enriched. Alternative categories of TF estimation procedures (e.g., the RMEV approach proposed by Egbert, 1997) could be included and would benefit from the other library features. Regarding the third step, a range of quality check methods exists and could be integrated in the library. Using the modular Razorback library, the MT practitioner fully controls the above workflow.

DATA AVAILABILITY STATEMENT

The Razorback library is available at https://github.com/BRGM/razorback. The data sets analyzed in this article will be made available by the authors, without undue reservation, to any qualified researcher upon reasonable request.

AUTHOR CONTRIBUTIONS

FS did most of the python language programming. PW designed the MT transfer function tools and validated extensively the library.

FUNDING

The Razorback library has been developed in the framework of the BRGM-funded projects EDGARR, OPTMT, and Benchmark MT. The MT dataset from Alsace has been acquired and processed in the framework of the IMAGE FP7 project, which has received funding from the European Union's Seventh Program for research, technological development and demonstration under grant agreement No: 608553. Part of the dataset provided as test-data were acquired with instruments of the French national pool of portable magnetotelluric instruments EMMOB (INSU-CNRS) and within activities of the Labex G-Eau-Thermie profonde program. MT data from La Fournaise volcano were used courtesy of Jacques Zlotnicki.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.00296/full#supplementary-material

REFERENCES

Bibby, H., Caldwell, T., and Brown, C. (2005). Determinable and non-determinable parameters of galvanic distortion in magnetotellurics. *Geophys. J. Int*. 163, 915–930. doi: 10.1111/j.1365-246X.2005.02779.x

Booker, J. R. (2014). The magnetotelluric phase tensor: a critical review. *Surv. Geophys*. 35, 7–40. doi: 10.1007/s10712-013-9234-2

Cagniard, L. (1953). Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics* 18, 605–635. doi: 10.1190/1.1437915

Caldwell, T. G., Bibby, H. M., and Brown, C. (2004). The magnetotelluric phase tensor. *Geophys. J. Int*. 158, 457–469. doi: 10.1111/j.1365-246X.2004.02281.x

Carbonari, R., D'Auria, L., Di Maio, R., and Petrillo, Z. (2017). Denoising of magnetotelluric signals by polarization analysis in the discrete wavelet domain. *Comput. Geosci*. 100, 135–141. doi: 10.1016/j.cageo.2016.12.011

Chave, A. D., and Jones, A. G. (2012). *The Magnetotelluric Method: Theory and Practice*. Cambridge University Press. doi: 10.1017/CBO9781139020138

Chave, A. D., and Thomson, D. J. (1989). Some comments on magnetotelluric response function estimation. *J. Geophys. Res. Solid Earth* 94, 14215–14225. doi: 10.1029/JB094iB10p14215

Chave, A. D., and Thomson, D. J. (2003). A bounded influence regression estimator based on the statistics of the hat matrix. *J. R. Stat. Soc. Ser. C Appl. Stat*. 52, 307–322. doi: 10.1111/1467-9876.00406

Chave, A. D., and Thomson, D. J. (2004). Bounded influence magnetotelluric response function estimation. *Geophys. J. Int*. 157, 988–1006. doi: 10.1111/j.1365-246X.2004.02203.x

Chave, A. D., Thomson, D. J., and Ander, M. E. (1987). On the robust estimation of power spectra, coherences, and transfer functions. *J. Geophys. Res. Solid Earth* 92, 633–648. doi: 10.1029/JB092iB01p00633

Dask Development Team (2016). *Dask: Library for Dynamic Task Scheduling*.

Di Giuseppe, M. G., Troiano, A., and Patella, D. (2018). Separation of plain wave and near field contributions in magnetotelluric time series: a useful criterion emerged during the Campi Flegrei (Italy) prospecting. *J. Appl. Geophys*. 156, 55–66. doi: 10.1016/j.jappgeo.2017.03.019

Egbert, G. D. (1997). Robust multiple-station magnetotelluric data processing. *Geophys. J. Int*. 130, 475–496. doi: 10.1111/j.1365-246X.1997.tb05663.x

Egbert, G. D., and Booker, J. R. (1986). Robust estimation of geomagnetic transfer functions. *Geophys. J. Int*. 87, 173–194. doi: 10.1111/j.1365-246X.1986.tb04552.x

Escalas, M., Queralt, P., Ledo, J., and Marcuello, A. (2013). Polarisation analysis of magnetotelluric time series using a wavelet-based scheme: a method for detection and characterisation of cultural noise sources. *Phys. Earth Planet. Interiors* 218, 31–50. doi: 10.1016/j.pepi.2013.02.006

Gamble, T. D., Goubau, W. M., and Clarke, J. (1979). Magnetotellurics with a remote magnetic reference. *Geophysics* 44, 53–68. doi: 10.1190/1.1440923

Goubau, W. M., Gamble, T. D., and Clarke, J. (1978). Magnetotelluric data analysis: removal of bias. *Geophysics* 43, 1157–1166. doi: 10.1190/1.1440885

Junge, A. (1996). Characterization of and correction for cultural noise. *Surv. Geophys*. 17, 361–391. doi: 10.1007/BF01901639

Kirkby, A., Zhang, F., Peacock, J., Hassan, R., and Duan, J. (2019). The MTPY software package for magnetotelluric data analysis and visualisation. *J. Open Sour. Softw*. 4:1358. doi: 10.21105/joss.01358

Krieger, L., and Peacock, J. R. (2014). MTPY: a python toolbox for magnetotellurics. *Comput. Geosci*. 72, 167–175. doi: 10.1016/j.cageo.2014.07.013

Larnier, H., Sailhac, P., and Chambodut, A. (2016). New application of wavelets in magnetotelluric data processing: reducing impedance bias. *Earth Planets Space* 68:70. doi: 10.1186/s40623-016-0446-9

Sims, W., Bostick Jr, F., and Smith, H. (1971). The estimation of magnetotelluric impedance tensor elements from measured data. *Geophysics* 36, 938–942. doi: 10.1190/1.1440225

Streich, R., Becken, M., and Ritter, O. (2013). Robust processing of noisy land-based controlled-source electromagnetic data. *Geophysics* 78, E237-E247. doi: 10.1190/geo2013-0026.1

Szarka, L. (1988). Geophysical aspects of man-made electromagnetic noise in the Earth-a review. *Surv. Geophys*. 9, 287–318. doi: 10.1007/BF01901627

Tikhonov, A. (1950). On determining electric characteristics of the deep layers of the Earth's crust. *Dolk. Acad. Nauk. SSSR* 73, 295–297.

Trad, D. O., and Travassos, J. M. (2000). Wavelet filtering of magnetotelluric data. *Geophysics* 65, 482–491. doi: 10.1190/1.1444742

Vozoff, K. (1972). The magnetotelluric method in the exploration of sedimentary basins. *Geophysics* 37, 98–141. doi: 10.1190/1.1440255

Ward, S. H. (1967). "Part A. Electromagnetic theory for geophysical applications," in *Mining Geophysics, Volume II, Theory* (Society of Exploration Geophysicists), 13–196. doi: 10.1190/1.9781560802716.ch2a

Wawrzyniak, P., Zlotnicki, J., Sailhac, P., and Marquis, G. (2017). Resistivity variations related to the large March 9, 1998 eruption at la Fournaise volcano inferred by continuous MT monitoring. *J. Volcanol. Geotherm. Res*. 347, 185–206. doi: 10.1016/j.jvolgeores.2017.09.011

Zhang, Y., and Paulson, K. (1997). Enhancement of signal-to-noise ratio in natural-source transient magnetotelluric data with wavelet transform. *Pure Appl. Geophys*. 149, 405–419. doi: 10.1007/s000240050033

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.

frontiers in Earth Science

ORIGINAL RESEARCH
published: 30 October 2020
doi: 10.3389/feart.2020.561148

Multi-Decadal Variability in the Snow-Cover Reconstruction at Parma Observatory (Northern Italy, 1681–2018 CE)

Nazzareno Diodato[1], Chiara Bertolin[2] and Gianni Bellocchi[1,3]*

[1]Met European Research Observatory – International Affiliates Program of the University Corporation for Atmospheric Research, Benevento, Italy, [2]Department of Mechanical and Industrial Engineering, Faculty of Engineering, Norwegian University of Science and Technology, Trondheim, Norway, [3]Université Clermont Auvergne, INRAE, Clermont-Ferrand, France

Edited by:
Reik Donner,
Hochschule Magdeburg-Stendal,
Germany

Reviewed by:
Jiri Miksovsky,
Charles University, Czechia
Dörthe Handorf,
Alfred Wegener Institute Helmholtz
Centre for Polar and Marine Research
(AWI), Germany

***Correspondence:**
Chiara Bertolin
chiara.bertolin@ntnu.no

Specialty section:
This article was submitted to
Hydrosphere,
a section of the journal
Frontiers in Earth Science

Received: *11 May 2020*
Accepted: *22 September 2020*
Published: *30 October 2020*

Citation:
Diodato N, Bertolin C and Bellocchi G
(2020) Multi-Decadal Variability in the
Snow-Cover Reconstruction at Parma
Observatory (Northern Italy,
1681–2018 CE).
Front. Earth Sci. 8:561148.
doi: 10.3389/feart.2020.561148

Emerging negative trends in snow depth and cover days highlight the challenges posed by changing snow patterns around the world. They suggest that snow-dependent regions in southern Europe could be affected by these changes because the number of days with snow on the ground (DSG) determines soil processes and water-flow in rivers, streams, lakes and reservoirs. We present here the first homogeneous, annually-resolved (from October to April), multi-centennial (1681–2018 CE) DSG time-series for the Parma meteorological observatory (OBS), in northern Italy, which to date is also the longest DSG series reconstructed in the world. DSG data are in fact still poorly documented and misunderstood due to the limited and fragmentary data measurements of the past. DSG recording only began in 1938 at Parma OBS. To generate the long-term annual DSG time-series at the study site, we develop a model consistent with calibration (1938–1990 CE) and validation (1991–2018 CE) samples of observed data. We show that the variability of DSG depends on winter precipitation and air temperature, as well as on winter-spring temperature variability, suggesting that long sequences of DSG are dominated by cold air masses in years with cold weather and high variability. Modeled DSG data show a downward trend from the 19th century, in the transition period from the cold of the Little Ice Age to the warmth of modern times, followed by a more rapid decline in the five most recent decades. The DSG at Parma OBS appear to have followed over the last century trends similar to those observed throughout Eurasia and across the Northern Hemisphere, where a marked decline of snow-cover duration has been reported in the transition seasons (spring and autumn).

Keywords: snow cover, snow on the ground, negative trend, Southern Europe, Parma, multi-centennial, reconstruction, historical dataset

INTRODUCTION

Escon gli armenti; e non appar su i campi
Erba, o fronda su gli alberi, ma sotto
Monti di neve desolata giace
La terra intorno, e d'aspro gel coperta,
Che alto a più braccia sovra lei s'indura. (…)
Cade la neve a larghi fiocchi intorno.
Ne muor la greggia intirizzita, e oppressi
Vi rimangono i buoi; ristretti in branco
Giacciono i cervi, e torpidi e sepolti.

Get the herds; and not in the fields
Grass, or branch on the trees, but below
Mountains of desolate snow lies
The earth around, and rough ice covered,
That high in more arms above her hardens. (…)
Snow falls broadly around.
The flock dies, numb, and oppressed
The oxen remain there; restricted in herds
The deer lie, and sluggish and buried.

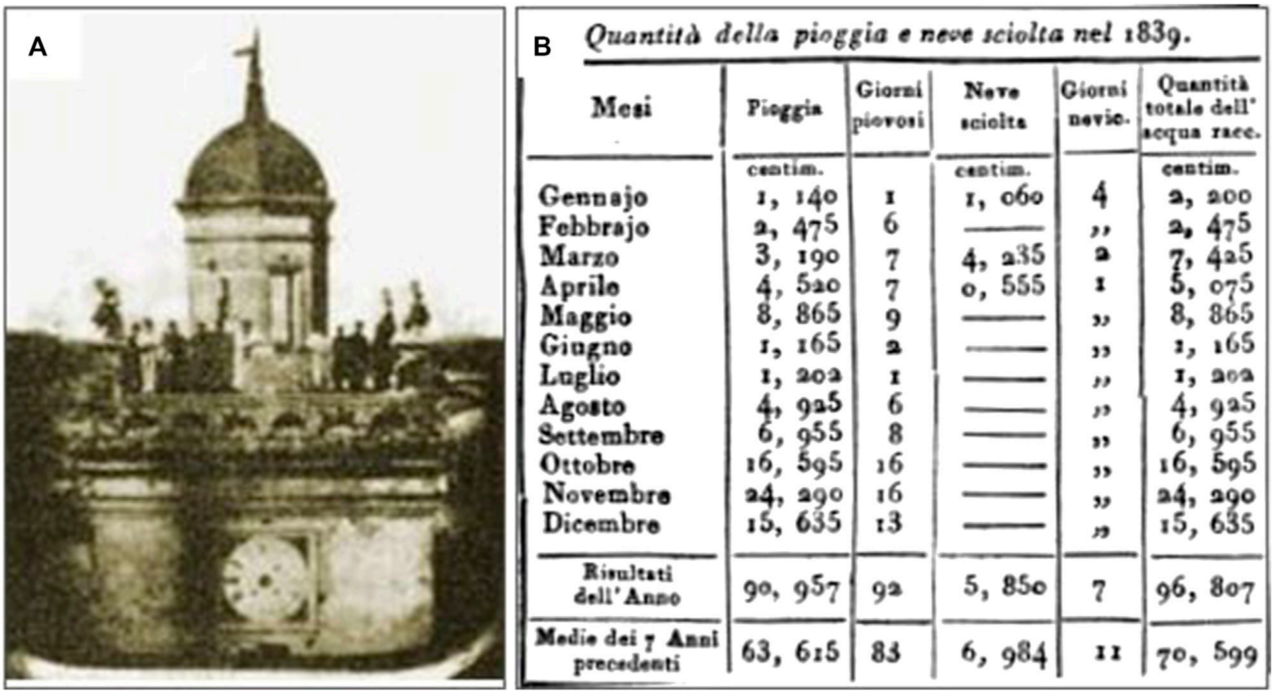
> Publio Virgilio Marone – Georgiche, I century B.C., Libro III (Italian translation from Latin: Clemente Bondi, 1801; our translation to English).

Every snowfall deposits its own layer on the soil surface. The timing, duration and persistence of the snowfall, which depend on climatic conditions, influence the amount of water stored in soils and its use for agriculture and other sectors such as industry, urban and rural domestic water supply (e.g., Lute et al., 2015; Barnhart et al., 2016; Wang et al., 2018a; Jennings et al., 2018). Large-scale snow-cover anomalies also cause important changes in the diabatic heating of the Earth's surface by enhancing the fraction of solar radiation reflected away by the surface (i.e., the surface albedo), thus becoming an essential component of the terrestrial radiation balance (e.g., Cohen and Rind, 1991; Goward, 2005; Sandells and Flocco, 2014). Snow cover affects the timing and magnitude of flow peaks generated by snow melting (Wang et al., 2006). By delaying the transfer of precipitation to surface runoff and infiltration in catchment areas, it also influences flow reduction and the extent of the flow network during summer baseflow (Yarnell et al., 2010; Godsey et al., 2014).

Knowledge of past climate is an important key to understanding long-term snow-cover variability. For instance, capturing anomalously cold and snowy winters can be consistent with, and help explain, a persistence of snow cover on the surface in late spring (after Jungclaus, 2009; Handmer et al., 2012; Enzi et al., 2014). In particular, the interannual variability of the number of days of snow on the ground (DSG) is an important part of the climate signal to detect potential changes of snowfall intensity and spatial distribution of snowpack variables that may have important impacts on both the environment and the society (Changnon and Changnon, 2006; O'Gorman, 2014). Our understanding of the DSG characteristics in several regions of the world is limited because of the short records available and a poor knowledge of the complex weather and climate patterns that occur locally. In Europe, the snowfall dynamics depend on the latitude and atmospheric circulation patterns, modulated by local orographic situations (Croce et al., 2018; Kretschmer et al., 2018). Polar maritime air masses originating from over the Atlantic collide here with the polar continental air masses connected with the Asian high pressure (Bednorz, 2004). The temperature control is dominant in the transient snow regions where the mean winter temperature is slightly below, and often crosses, the melting point (Diodato and Bellocchi, 2020). The situation is different in regions where the mean winter temperature is well below 0°C, as is the case of the Alps, where increases in DSG are controlled mainly by precipitation inputs (Clark et al., 1999). At lower-elevations, DSG are instead affected principally by air temperature and only secondarily by precipitation conditions.

Data on snowfall and the persistence of snow cover on the surface are becoming increasingly important due to the high variability of snowfall rates worldwide (e.g., Davi et al., 2012), and the impact that snow cover can have on society, agriculture and water resources. A quantitative assessment of the long-term and interannual DSG variability is noticeable for the identification and framing of signals of climate change, the validation of climate models, and a better understanding of the interactions among the different spheres of the Earth system, including the geosphere, biosphere, atmosphere, hydrosphere and cryosphere (Brown, 2014). Documentary sources, usually in manuscripts or annotations in different formats, provide evidence and useful information to study the variability of the climate over historical periods (Glaser and Riemann, 2009; Dobrovolný et al., 2010). However, the quality and quantity of sources are often unevenly distributed in space and time (Bradley and Jones, 1992; Mann et al., 2000; Brázdil et al., 2005; Adamson, 2015). This is especially true for snowfall, considering that *in situ* continuous snow monitoring is an arduous task (De Walle and Rango, 2008) and satellite remote sensing data capturing the snow-cover evolution are limited to recent decades (e.g., Pimentel et al., 2017). Also for historical times, snow characteristics data remain still poorly documented and understood due to limited and fragmented measurements (Kunkel et al., 2016; **Figure 1**). One of the longest-running snowfall records is the series of Parma observatory (Parma OBS), in the Padan plain (northern Italy), where however recording of DSG only started in the winter season 1937–1938. This coincides with the setup of new measuring meteorological facilities at ground level and no longer on the tower of the OBS, the suburbs and the city center.

Centrally located in the Po River Basin (PRB), the Parma OBS started weather observations on 1777 thanks to the local Jesuit community with the support of the University of Parma, but only in the 20th century regular and continuous daily snowfall measurements were recorded. Diodato et al. (2020a) provided



Quantità della pioggia e neve sciolta nel 1839.

Mesi	Pioggia	Giorni piovosi	Neve sciolta	Giorni nevic.	Quantità totale dell' acqua racc.
	centim.		centim.		centim.
Gennajo	1, 140	1	1, 060	4	2, 200
Febbrajo	2, 475	6	———	,,	2, 475
Marzo	3, 190	7	4, 235	2	7, 425
Aprile	4, 520	7	0, 555	1	5, 075
Maggio	8, 865	9	———	,,	8, 865
Giugno	1, 165	2	———	,,	1, 165
Luglio	1, 202	1	———	,,	1, 202
Agosto	4, 925	6	———	,,	4, 925
Settembre	6, 955	8	———	,,	6, 955
Ottobre	16, 595	16	———	,,	16, 595
Novembre	24, 290	16	———	,,	24, 290
Dicembre	15, 635	13	———	,,	15, 635
Risultati dell'Anno	90, 957	92	5, 850	7	96, 807
Medie dei 7 Anni precedenti	63, 615	83	6, 984	11	70, 599

FIGURE 1 | (A) The Parma meteorological observatory from a historical picture; **(B)** monthly summary of precipitation records (rainfall and snowfall) for the year 1839 (note the absence of recording in days with snow on the ground). Historical pictures are: **(A)** made freely available from Fabrizio Bonoli (University of Bologna, Italy) through the web-based content store (in Italian) of Ravenna Planetarium (http://planet.racine.ra.it/testi/astr_it.htm) [Accessed September 10, 2020]; **(B)** an excerpt of a printed image from Colla. (1840).

the series of monthly snowfall reconstruction in Parma over the 1777–2018 period, with the detailed analysis of metadata. In this regard, Italian regions hold among the world's longest monthly snowfall time series (Enzi et al., 2014), going back to 1780 at Rome (Mangianti and Beltrano, 1993), 1788 at Turin (Leporati and Mercalli, 1994) and 1884 at Montevergine (Diodato, 1997). Nevertheless, Parma OBS possesses the longest continuous record, as it goes back to 1777, although we know (Camuffo and Bertolin, 2012) that meteorological observers were already active in Parma from 1654 to 1660, supported by the Grand Duke of Tuscany Ferdinand II de' Medici (1610–1670).

The current work presents the first annually-resolved reconstruction of DSG since 1681 (i.e., the snow winter season going from October 1680 to April 1681) for the Parma OBS. A day with snow cover is a day with snow of at least 0.01 m depth observed in over 50% of an open neighboring area (UNESCO/IASH/WMO, 1970; WMO, 2008). Winter maximum number of days with snow depth matching these criteria was used to characterize snow cover duration for the purpose of this study (e.g., Falarz, 2004; Petkova et al., 2004). We first acquired a comprehensive knowledge about potential drivers of DSG in the PRB (*Materials and Methods*), where snowfall occurs from October to April, and then we used these factors (amounts of precipitation and air temperature) to develop a tight or "parsimonious" model for reconstructing annual DSG data over the period 1681–2018 CE (*Results and Discussion*). This allowed us to capture a wide range of climate variability and identify patterns of snowfall changes reported and discussed in section.

MATERIALS AND METHODS

Environmental Setting and Data

Parma OBS (44° 48′ N; 10° 19′ E, 49 m a.s.l.) is located in the central part of the PRB, in the Italian administrative region Emilia-Romagna (**Figure 2A**). This area is surrounded by the pre-Alps to the north and by the Emiliano Apennine to the south (**Figure 2B**). The interaction between these morphology and weather characteristics governs the occurrence of DSG (whose measurements are available since July 1937) under different conditions. The Parma OBS (black dot in **Figure 2C**) outlines a separation between the snowy peaks over the northern and western pre-alpine and alpine chains and the Apennines to the south where the climate is less continental. Complex reliefs, different slope exposures, as well as various distances to sea (**Figure 2B**), exert highly contrasting effects on snow-cover depth, persistence and spatial extent. The colored bands in **Figure 2C** do explain the amplitude of sub-regional variations around the Parma OBS. For instance, south of Parma the average snow-cover duration rises up to ~80 days year^{-1} in the Apennines.

During the year, cyclonic Atlantic air masses and areas of anti-cyclonic airflows from central and eastern Europe alternate. Fields of variable winds are associated with wet air masses

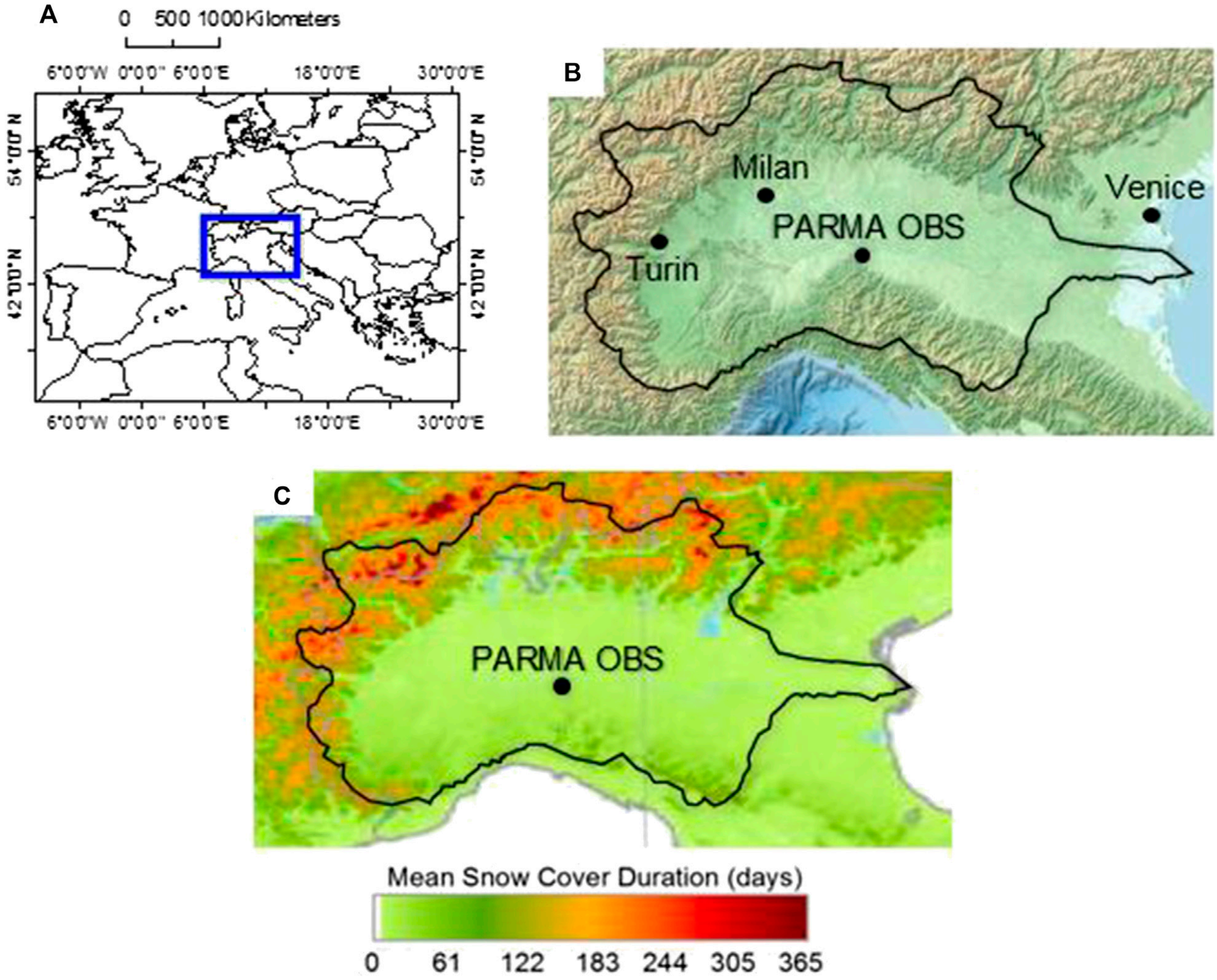


FIGURE 2 | Geographical setting: **(A)** North Italy (blue square); **(B)** the Po River Basin (black curve) with main locations and Parma OBS; **(C)** related winter (November-March, 2000–2011) snow-cover duration. Maps are authors' own elaboration from free, public domain images: **(A)** ESRI (http://www.esri.com) [Accessed September 10, 2020]; **(B)** Elastic Terrain Map (http://elasticterrain.xyz) [Accessed September 10, 2020]; **(C)** Dietz et al. (2012).

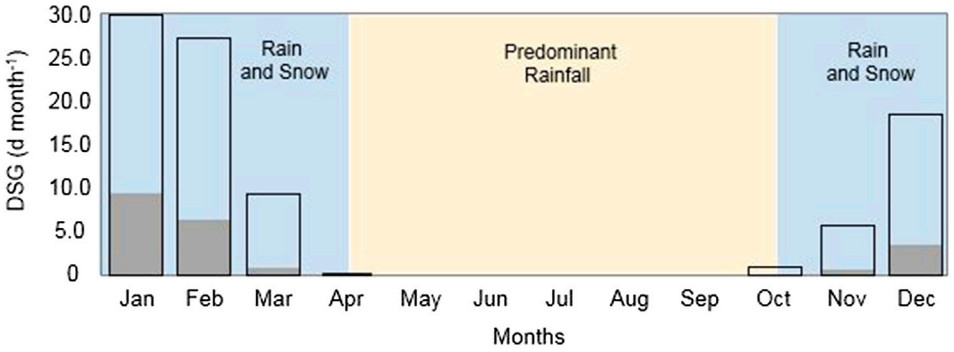


FIGURE 3 | Monthly mean distribution of days of snow on the ground (DSG) at Parma OBS (gray bars) with the related 98th percentile (empty bars), calculated over the period 1938–2018.

from southwest, and dry and continental winds from northeast. Snowfall increases with elevation until 2000 m a.s.l. **Figure 3** (gray bars) refers to observed monthly snow cover days. At Parma OBS, which can be assumed sufficiently representative of the PRB area, the average of DSG in the period 1938–2018 is about 20 days year^{-1}. Over this period of 81 years, the month of January had the greatest number of days with snow cover, with the soil covered by snow for about 10 days on average. On February, the number of days with snow on the ground roughly doubled the days of snow coverage in December. March and November had a similar number of days with snow cover, while the lowest number was registered in October. A similar distribution is noticeable for the 98th percentiles (**Figure 3**, empty bars).

The most abundant snowfall in the Alps and Apennines occurs during cold periods, in particular with the arrival of warmer and moist air from the south, with a rise in temperatures up to about 0°C. If plain areas are exposed to a degree of cold sufficiently intense, with temperature values below 0°C, snowfall and its persistence on the ground are also possible at very low altitudes (Bettini, 2016). This is what happens in the central-western sector of the PRB, when warm and humid air coming from the Mediterranean Sea slips on the cushion of cold air previously trapped in the low layers of mountain landscapes in the Alps and Apennines. There is a chance that the arrival of below-freezing temperatures can also bring snow along the PRB. In the presence of air masses close to saturation, very cold air suddenly entering the PRB raises near saturation leading wet air to condensation. This phenomenon, although rare, can be observed in the PRB when cold Siberian air penetrates westward into northern Italy, while very damp air or even a misty blanket stagnates in the plains. In these cases, snowfall is weak, of short duration and limited to the plains or the lowest slopes of the pre-Alps.

As covariates for the modeling of DSG, we referred to winter precipitation, and winter and spring air temperatures. Winter and spring are defined as periods from December to February, and from March to May, respectively. For the precipitation input, we used the seasonally-resolved data with 0.5° resolution, as arranged from Pauling et al. (2003). Likewise, we used seasonally-resolved winter and spring temperatures as arranged by Luterbacher et al. (2004) since 1500, and interpolated at the Parma OBS grid point. Both datasets were updated from the Climatic Research Unit global climate dataset (Jones and Harris, 2018). The database of observed DSG contains also the year 2019 (with 4 days of snow cover on the ground) but this year was left out of the analysis because the Climatic Research Unit data are updated until 2018 (checked on June 2020). For both winter precipitation and temperature, the analysis of the yearly data used (1681–2018 CE) shows that they are compliant with a normal distribution using the Kolmogorov-Smirnov test $p > 0.05$ (**Supplementary Appendix A**).

Development and Parameterization of the Statistical Model

The principle of parsimony as described in Mulligan and Wainwright (2013) states that a parsimonious model is the one with the greatest explanation or predictive power and least parameters or process complexity. Based on this principle, we developed a parsimonious model that could be easy parameterized and validated, and is reliable and applicable to the reconstruction of long-term historical DSG data. To increase predictability and minimize uncertainties in the modeling of annual DSG data, we built a nonlinear-multivariate regression with three input variables and three parameters considering the original observational 1938–2018 CE timeframe. The data resource of annual DSG data (81 years) was separated into two sub-sets to use roughly two-thirds (53 years) for calibration (1938–1990 CE) and one-third (28 years) for validation (1991–2018 CE). The calibration effort was thus anchored to the reality of snow-cover duration in the past to enable a reasonably accurate reconstruction of historical snow-cover days. A number of monthly climatic explanatory variables was considered during the input selection process. First, in order to reduce the number of inputs, we investigated the effects of single variables, or sets of variables, on DSG over some climatologically meaningful periods. Then, an iterative process (trial-and-error to decompose relevant drivers) enabled us to explain the dynamics of DSG in relatively simple terms. A stepwise selection logic was used to alternate between adding and removing terms. The following nonlinear multivariate regression model - *DSG(Parma)* - was thus derived to estimate DSG (day year^{-1}) at Parma OBS:

$$DSG(Parma) = A \cdot \left[ln(P_W) \cdot (\beta - T_w)^{\eta} \cdot VC\left(\frac{SD^{\eta}}{Mean}\right) \times \left(T(w \to s)_{t=-1}^{t=0}\right)\right] \qquad (1)$$

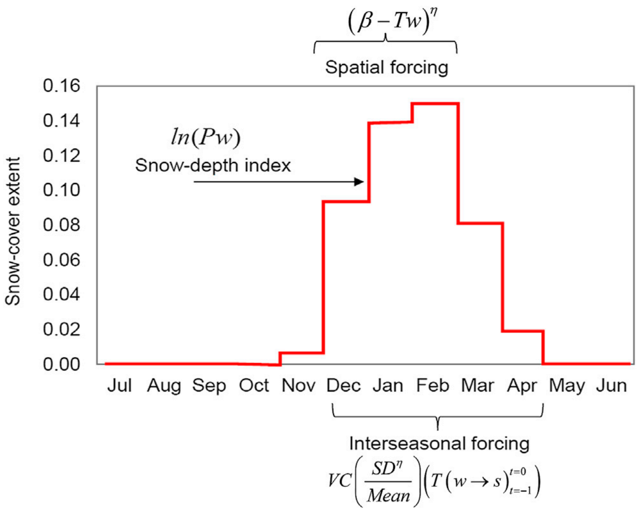


FIGURE 4 | Monthly snow-cover extent in Parma territory (100-km square grid cell), with the associated driving factors. Monthly mean values (red line) were calculated on data provided for the period 1966–2018 by Global Snow Lab - Rutgers University Climate Lab (https://climate.rutgers.edu/snowcover) via KNMI-Climate Explorer Climate Change Atlas (http://climexp.knmi.nl) [Accessed September 10, 2020]. The three main terms of **Eq. 1** are reported.

where: A (day year^{-1}) is a scale parameter corresponding to the number of DSG when the term between brackets is equal to unity; β (°C) and η are process parameters. Winter temperature, T_w (°C), is negatively related to DSG, as mediated by the shape in parameter η. Winter (w) and spring (s) temperature variability accounted for by the term VC is indicative of the instability of air masses near the Parma OBS. This approach discloses the association between the snowfall response variable *DSG (Parma)* and multiple predictors, such as the winter precipitation amount (*P_w*), winter mean air temperature (T_w), and winter-spring modified variation coefficient VC, that is, a standard deviation/mean ratio of winter and spring temperatures ($T(w \rightarrow s)$), calculated across the current ($t = 0$) and previous ($t = -1$) years. This variability term is a key factor to infer inter-seasonal temperature fluctuations, as it is larger in the presence of snow.

The concept of the model (**Figure 4**) summarizes the mechanisms mostly driving the common patterns of change of snow-cover extent and duration (e.g., Brown et al., 2010; Brown, 2019). It shows that ~10–15% of the Parma territory (central tendency) is covered by snow over wintertime. In this context, cool-season precipitation amounts are particularly important because the majority of annual precipitation occurs during the wintertime at medium-high latitudes, and preferentially as snow at high elevations through orographic enhancement (e.g., Dettinger et al., 1998; Selkowitz et al., 2002; Terzago et al., 2012; Luce et al., 2013). The logarithmic transformation of the total (solid and liquid) winter precipitation ($ln(P_w)$), which acts as snow-depth index in the model, was applied to underweight the amount of freshly fallen snow associated with broad-scale mixed precipitation over flat areas (e.g., Meir et al., 2016; Fehlmann et al., 2018). Snow-cover response to temperature includes the spatial extent (0.5°-grid resolution of winter temperatures) and the temporal scale (winter and spring months) over which the thermal forcing becomes important, by assuming that the object under study is a fractal (i.e., the site is a fractal element of the area). In this case, the scaling of the site can be reasonably modeled by a power-law process relationship, whose exponent η quantifies the spatial dependence (scaling nature) of the snow-cover days (Zhang et al., 2013). This is an issue of determining how the climate of a specific surface type differs from a grid-average climate, without formulating different model equations (Harvey, 2013). In particular, the η exponent does not only provide a parsimonious description of the process under study, but it is also a generic mechanism governing the process. In particular, higher exponent values obtained at a single site (compared to mean areal response) suggest that they can be regarded as a method to downscale areal (smoothed out) approximations to finer units (e.g., Spence and Mengistu, 2019).

We have also found that a strong correlation ($r = 0.84$) exists between snowfall frequency and snow-cover duration at Parma in the period 1938–2018 CE (**Supplementary Appendix B**). However, since snowfall frequency measurements (i.e., snow days per year) only started in 1777 at Parma OBS, we excluded this input from the model, which was built on the temperature- and precipitation-based proxies allowing to extend the reconstruction period back to 1681. With this long-term series, we can detect patterns of climate change explaining fluctuations and tendencies in DSG data. Going back about 100 years before the beginning of the snowfall-frequency observations at Parma OBS, we could in fact start our DSG series at the turning point of the 17th and 18th centuries, at the so-called Maunder Minimum, which started in 1645 and lasted until 1715, a period characterized by exceedingly rare sunspots and lower-than-average temperatures (Eddy, 1976).

Spreadsheet-based model development and assessment were performed with the analytical and graphical support of STATGRAPHICS (Nau, 2005) and WESSA (Wessa, 2009) statistical software. The Mean Absolute Error (*MAE*, day year^{-1}) was calculated to quantify the differences between actual and modeled DSG values, while the Kling-Gupta index ($-\infty < KGE \leq 1$; Kling et al., 2012) was used as efficiency measure (with $KGE > -0.41$ indicating that a model is better performing that the mean of observations as a benchmark predictor; Knoben et al., 2019). The Nash-Sutcliffe efficiency ($-\infty < EF \leq 1$, optimum; Nash and Sutcliffe, 1970) was also calculated as an uncertainty indicator of the model performance because greater values than 0.6 indicate limited model uncertainty, likely associated with narrow parameter uncertainty (Lim et al., 2006). With the determination coefficient ($0 \leq R^2 \leq 1$, optimum), or the correlation coefficient ($r = \sqrt{R^2}$), and the slope of the regression actual vs. modeled data ($b = 1$, optimum), we selected the set of important covariates for the parsimonious model for estimating DSG (after Mulligan and Wainwright, 2013). F-ratio p-values were used to present the statistical significance of the linear regression between actual and estimated data, and of the inputs' relationship to the dependent variable. The Durbin-Watson statistic (Durbin and Watson, 1950; Durbin and Watson, 1951) was used as a test for autocorrelation in the model residuals, considering that strong serial dependencies may induce spurious correlations (Granger et al., 2001). Exploratory and time-series analysis were carried out using AnClim (http://www.climahom.eu/software-solution/anclim). A MATLAB toolbox (https://noc.ac.uk/business/marine-data-products/cross-wavelet-wavelet-coherence-toolbox-matlab) was used for wavelet coherence analysis (Grinsted et al., 2004).

RESULTS AND DISCUSSION

Model Parameterization and Evaluation

For the reconstruction of a long-term annually-resolved DSG series, we first calibrated the *DSG(Parma)* model. The calibration work was performed through a trial-and-error process comparing the model estimates with observations until small *MAE* and large R^2 values were obtained. Then, for the final selection of the parameter values, a third criterion (*KGE* closer to 1) was additionally involved. The calibrated parameters are: A = 0.0404 day year^{-1}, β = 6.80°C, and η = 2.00. The value of the R^2 statistic of observed (y) vs. modeled (x) data (**Figure 5A**) indicates that the fitted model explains 72% of the observed variability. The regression line $y = 4.36\ (\pm 8.39) + 0.83\ (\pm 0.27) \cdot x$

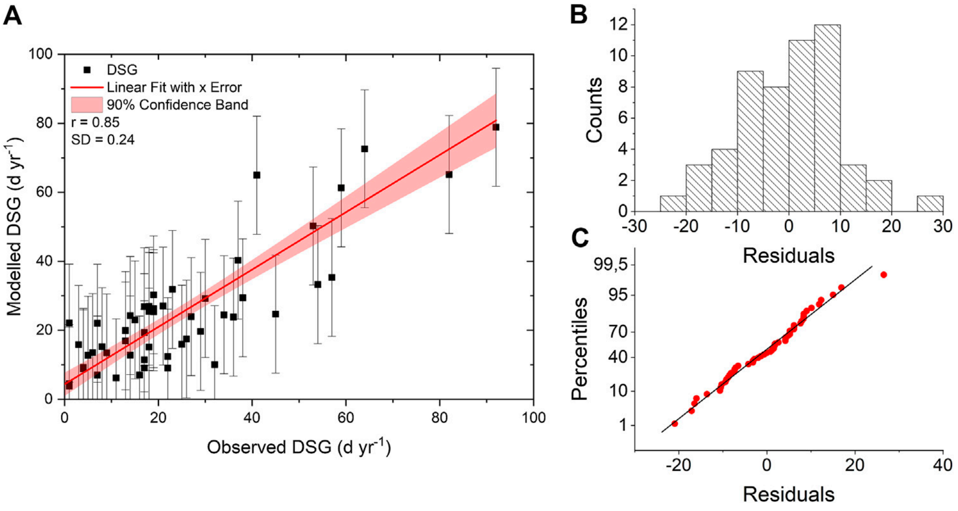


FIGURE 5 | (A) Scatterplot of observed and modeled (**Eq. 1**) DSG at Parma OBS for the calibration sub-set (1938–1990), with regression line (red line), the bounds showing 90% confidence limits (pink colored band); **(B)** histogram of residuals; **(C)** Q-Q plot (sample vs. theoretical quantile values).

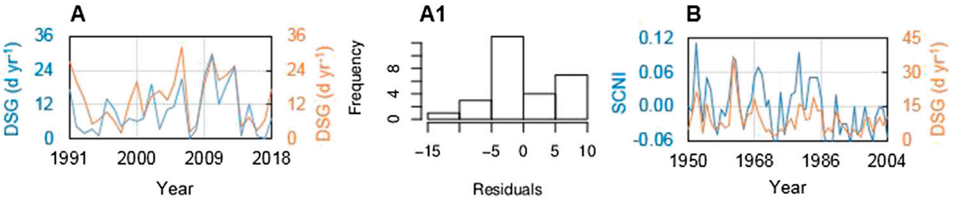


FIGURE 6 | (A) Timeline evolution of observed (blue curve) and modeled (**Eq. 1**, orange curve) DSG for the validation sub-set (1991–2018); **(A1)** histogram of residuals; **(B)** timeline coevolution of observed Snow-Cover Normalized Index (SCNI) for the Emilian Apennine mountain range (blue curve) and modeled DSG at Parma OBS (orange curve) for the period 1950–2004.

has intercept ($a = 4.36$) and slope ($b = 0.83$) with a relatively high standard error of the intercept (±8.39 days year^{-1}) and a root mean standard error of the fit equals to 0.24 day year^{-1}. This indicates the model's lesser predictive ability for near-zero DSG, there is a superior predictive ability of **Eq. 1** compared with a purely linear multivariate regression against the same independent variables, which produces several (unrealistic) negative values of DSG in both the calibration and validation periods (**Supplementary Appendix C**). The Nash-Sutcliffe efficiency value obtained in the calibration stage ($EF = 0.73$) indicates limited uncertainty in model estimates. Since the F-ratio p-value is less than 0.05, the linear regression between actual and estimated data is statistically significant. The mean absolute error (MAE), used to quantify the amount of error, was equal to 8.9 days year^{-1}, which is lower than the standard error of the estimates (10.9 days year^{-1}), while the Kling-Gupta index of 0.6 suggests a model of sufficient quality. The Durbin-Watson statistic ($DW = 1.72$, $p = 0.14$) reveals that there is no significant autocorrelation in the residuals.

The model residuals approximate a normal distribution (**Figure 5B**; normality test, $p > 0.05$, Jarque and Bera, 1981) and the Q-Q plot (**Figure 5C**) exhibits a distribution of sample-quantiles around the theoretical line, indicating only a few biased high DSG values.

At the validation stage, the value of the R^2 statistic indicates that 56% of the total variability in the observation is explained by the model, while MAE is equal to 4.4 days year^{-1}, which is lower than the standard error of the estimates (5.7 days year^{-1}), while the Kling-Gupta index is 0.6 (equal to the value obtained with the calibration set). Also in this case, the Durbin-Watson statistic ($DW = 1.60$, $p = 0.13$) indicates that there is no significant autocorrelation in the residuals. The timeline of DSG validation period shows the coevolution of observations and estimates (**Figure 6A**). The related distribution of model residuals (**Figure 6A1**) does not deviate significantly from normality ($p > 0.05$). In **Figure 6B**, the model estimates are compared to the decreasing trend (from 1950 to 2005) of the index obtained by De Bellis et al. (2010) for the Emilian Apennine, by averaging the snow-cover values of different mountain stations normalized with respect to the variability. We evaluated the relative performance of the DSG(Parma) model, without comparing the absolute estimates. Coevolution between the SCNI (snow-cover normalized index) and the modeled DSG illustrates a substantial agreement ($r = 0.74$).

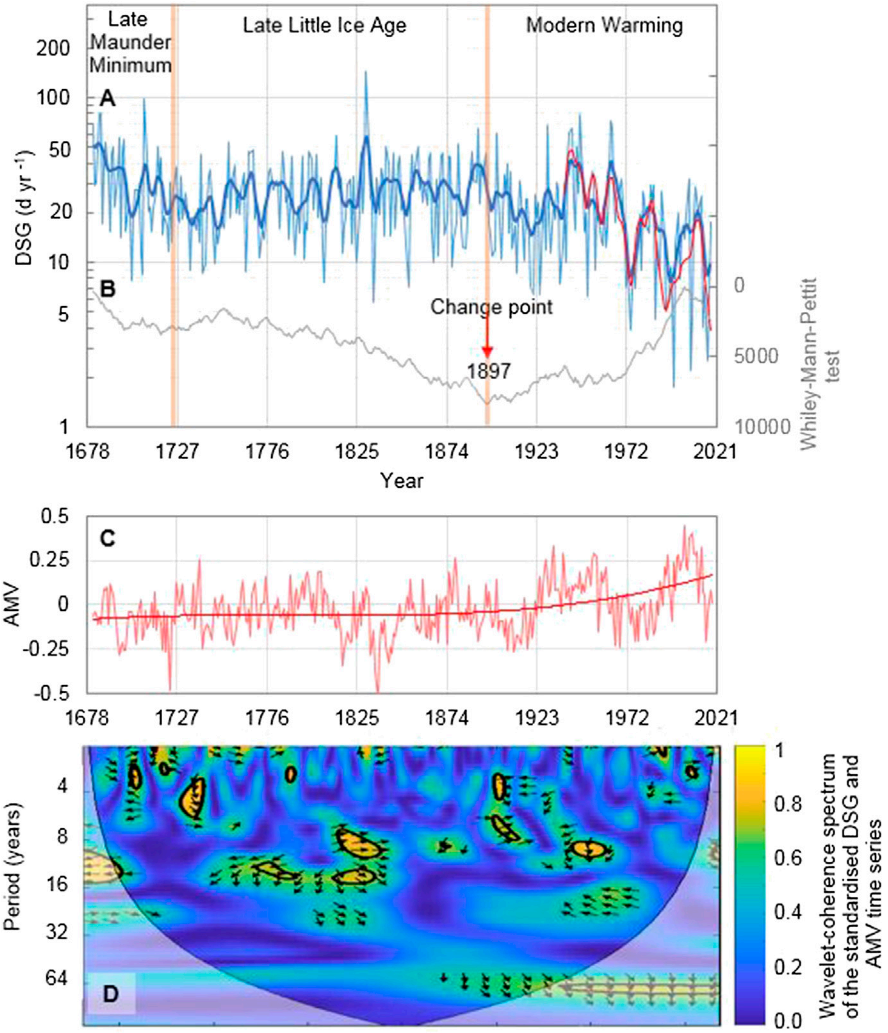


FIGURE 7 | Overview of days with snow on the ground (DSG) patterns at Parma OBS along the period 1681–2018 CE: **(A)** Timeline evolution of DSG data reconstructed by **Eq. 1** (light blue curve), with over-imposed 11-years Gaussian-filtered series for the whole period (bold blue curve) and the observational period 1938–2018 CE (red curve); **(B)** Mann-Whitney-Pettitt test statistic with the change point of 1897 (vertical red arrow); **(C)** time series of the Atlantic Multidecadal Variability (AMV; Wang et al., 2017) together with a polynomial (third-order) interpolation curve; **(D)** wavelet-coherence spectrum of the standardized DSG and AMV time series; bounded colors identify the 0.05 significance level areas; the bell-shaped, black contour marks the limit between the reliable region and the region below the contour where the edge effects occur (a.k.a. cone of influence); black arrows show the relative phase relationship, with in-phase pointing right, anti-phase pointing left, and AMV leading DSG by 90° pointing straight down.

We could get a slightly better performance by replacing P_w in **Eq. 1** with the square root of snow days per year during both the calibration ($R^2 = 0.76$, $KGE = 0.81$, $MAE = 8.6$ days year^{-1}) and validation ($R^2 = 0.61$, $KGE = 0.52$, $MAE = 4.6$ days year^{-1}) periods. We concluded that this improvement is not such as to justify the limitation of the reconstruction to the SDY observation period alone (1777–2018). **Equation 1** thus supports a broader reconstruction perspective. In determining whether the DSG (Parma) model can be simplified, we have fitted a multiple linear regression model to describe the relationship between DSG and the three independent variables P_w, T_w and VC in **Eq. 1**. The analysis of variance of the model gave $p < 0.05$ for each independent variable, the highest p-value being 0.04 for the term VC. This means that the model cannot be simplified further and its results correspond to criteria of stability, interpretability and usefulness (after Royston and Sauerbrei, 2008).

Time-Series Reconstruction of Snow-Cover Persistence on the Ground (1681–2018 CE)

Equation 1 was used to reconstruct the evolution of DSG over the period 1681–2018 CE (**Figure 7A**). The reconstructed time-series was then analyzed to find out possible climate patterns explaining variation in long-term trends of DSG, and to compare contemporary with historical DSG anomalies.

TABLE 1 | Descriptive statistics of the DSG (days with snow on the ground, day $year^{-1}$) time series for two climatic periods.

Climatic periods	DSG statistics			
	Mean	Median	75th percentile	95th percentile
Until the change point (1681–1897 CE)	28	25	37	53
After the change point (1898–2018 CE)	21	18	26	67

Overall, the temporal evolution of annual values presents a downward trend (Mann-Kendall M-K trend test $p < 0.01$; Kendall, 1975), which appears more prominent in recent times (**Figure 7A**, blue curve and related Gaussian fitting; **Figure 7B**, Mann-Whitney-Pettitt statistic). We have used different statistical methods to identify distinct climatic patterns in the long-term DSG series. In fact, different test statistics are variously sensitive to change points located at the beginning, in the middle, or at the end of a time series (Martínez et al., 2009; Toreti et al., 2011), and a combination of statistical methods is considered to be most effective to track down change points (e.g., Wijngaard et al., 2003). The application of test statistics by Pettitt (1979) and Buishand (1982) suggests the existence of a change point in the year 1897. Since the late 19th century, the Atlantic Multidecadal Variability (AMV) has experienced a significant upward trend (Si et al., 2020), coinciding with the onset of the warming period (**Figure 7C**). Based on terrestrial proxy records from the circum-North Atlantic region, the AMV reconstruction by Wang et al. (2017) exhibits pronounced variability on multidecadal time-scales. This multidecadal climate mode originates dynamically in the North Atlantic Ocean and propagates throughout the Northern Hemisphere via a suite of atmospheric and oceanic processes (Wyatt et al., 2012; Wang et al., 2018b). Sutton and Dong (2012) argued for the existence of a causal link between a positive phase of the AMV and drier conditions over the Mediterranean Basin. It is also known that winter precipitation in northern Italy is mainly caused by large-scale fronts of North Atlantic and Mediterranean synoptic low-pressure systems, which produce moderate but continuous precipitation (Hawcroft et al., 2012). Interannual to multidecadal variability of this kind of precipitation is mainly modulated by the North Atlantic Oscillation (NAO; Pinto and Raible, 2012; Gómara et al., 2014; Gómara et al., 2016), describing the fluctuations in the difference of atmospheric pressure at sea level between the Icelandic Low and the Azores High.

Other test statistics detected a change point in 1830 (SNHT-double shift; Alexandersson and Moberg, 1997) or 1967 (Worsley likelihood; Worsley, 1986; penalized maximal *t*-test; Wang et al., 2007; SNHT-single series, Alexandersson, 1986), which merely support the idea of a long transition period going from the final phase of the Little Ice Age (LIA; ~1300–1850 CE; Miller et al., 2012) to the most recent warming. These different statistically-relevant years provide a loose picture of climate-related DSG variations with changing climate patterns, where the cold conditions of the LIA are still dominating after the end of the Dalton minimum of reduced solar activity (~1790–1830; Wagner and Zorita, 2005) until toward the end of the 19th century, but in the process of evolving into an incipient warming that becomes noticeable later in the 20th century (mostly by about 1960s). We refer hereafter to the year 1897 as a relevant change-point year, which is also the starting point of more erratic weather conditions. The long-term mean value of DSG (**Table 1**) is equal to 28 days $year^{-1}$ (±16.8 days $year^{-1}$ standard deviation) until the change-point year (**Figure 7B**) and to 21 days $year^{-1}$ (±14.8 days $year^{-1}$) after that point for the continuous predicted time series (**Figure 7A**, blue curve), while it is equal to 20 days $year^{-1}$ (±11.9 days $year^{-1}$) for the joined modeled and observed series (**Figure 7A**, joined blue and ochre curves after the year 1897).

Considering that these two series, as well as the predicted and observed series for the period covered by observations (1938–2018 CE), are not statistically different (paired Student-*t* test, *p*-values >0.05), our analysis was based on the modeled values only. **Table 1** also shows that the median and the 75th percentile values were higher over the period 1681–1897 CE, while more extreme values (95th percentile) were higher after the change point. The trend toward anomalously warm conditions over the 20th century and the beginning of the 21st century indicates that increasing temperatures represent the main factor triggering the decline of DSG, in particular over the most recent five decades where DSG have been subject to further decrease. This is in agreement with the contraction of snow-cover extent recently observed across Siberia and North America (e.g., Musselman et al., 2017; Zhang and Ma, 2018). It is striking that snow covered the ground over 145 days in the year 1830 and only over 99 days in the year 1709 (**Figure 7A**). The latter proved to be the coldest in memory, with a particularly long and hard freeze in a winter passed to the historical chronicles because of its cold and snowstorms. In the central Mediterranean, the severity of this winter was accompanied by ice and frozen water bodies like lakes and rivers (Glaser et al., 1994; Grove and Conterio, 1994), while snowfall continued from late December to mid-February, after which rain and snowmelt caused several rivers to overflow (Diodato et al., 2019). However, the 1829–1830 winter was not only described as harsh in all parts of Europe, but also early, with long and heavy snowfall from mid-November to the end of March (Corradi, 1865–1890, vol. III, p. 403). In January

Il massimo freddo osservato fu nel gennajo in cui nel giorno due giunse a gradi 9 sotto lo zero, limite però inferiore a quello degli anni 1795, 1800, 1812, 1826, 1830, senza ricordar quello di epoca più remota.	The maximum cold observed was in January where, in the day two, it reached degrees 9 below zero, a limit however lower than that of the years 1795, 1800, 1812, 1826, 1830, without recalling the most remote epoch.

Poichè di nevi ebbe tanta abbondanza, che nemmeno i più vecchi ricordavano, non che l'eguale; per questo la terra ne restò coperta fin dopo Pasqua, ed il freddo fu così crudo che il Po e l'Arno ghiacciarono; tanto freddo fu universale: sul Tamigi passavano i carri	For the snow was so abundant that not even the elders remembered it, not that there had been any equal; so the Earth was covered with snow until after Easter, and the cold was so raw that the Po and the Arno froze; so the cold was universal: the chariots passed over the Thames.

1830, even the press was suspended because the roads, due to the abundance of snow, had not allowed the arrival of newspapers. The year 1830 was rather the end of a period that extended over important parts of the LIA, during which Europe experienced predominant cooling (Xoplaki et al., 2001). In fact, winter 1835–1836 was also remarkable for Milan and its region but not quite as severe as others (Stella, 1836, pp. 350–351):

The year 1684, at the beginning of our series, is also distinguished by a high number of days with snow on the ground (88 days). In this regard, Corradi Annals' (1865–1890, vol. II, p. 257) define 1684 as the year in which snow covered the ground until after Easter:

After Xoplaki et al. (2001), the severe conditions during the winter 1683–1684 can be attributed to the presence of an extended high-pressure system over central and western Europe, with its center over north Iberia. This distribution of atmospheric pressure centers in Europe led to persistent northwesterly and/or northeasterly circulation over the Balkans and the eastern Mediterranean. These cold air outbreaks caused low or extremely low temperature conditions connected with precipitation events (rainfall and/or snowfall). Also noteworthy are the years 1947 and 1963, with snow on the ground over 79 and 73 days, respectively. January 1947 was recognized as historical for Italy, with very strong waves of frost and snow all over the country. The coldest winter in the post-World War II period was, however, 1962–1963: started on December, it continued with an escalation of continuous Siberian cold spell, alternating with Atlantic phases, until the month of March, which was equally rigid. In Britain, although the Thames had not frozen as in previous centuries, that winter turned out to be not only the coldest of the century but among the coldest ever. After that winter, the DSG median dropped further (only 15 days per year), and with it also the 95th percentile reached the lowest values ever (30 days per year). These trends are confirmed by the decrease in snow-cover extent in spring, as observed in Eurasia. Although the extent of the autumn snowpack has been limited in recent decades, and the winter trend has remained unchanged in Europe and Asia, a decrease in the snowpack extent and a greater variability in the transition seasons (spring and autumn) have been documented at the hemispheric scale since the 1920s (Krasting et al., 2013). Such snow-cover reduction was particularly significant in the mid-latitudes (40°–60°) of the Northern Hemisphere (Brown, 2000). Accordingly, Déry and Brown (2007) noted that the mean monthly snow-cover extent strongly anti-correlates with the mean air temperature in the Northern Hemisphere. This occurs particularly in April–June, when extended snow-cover duration is accompanied by intense incoming solar radiation.

We used the wavelet-coherence analysis for highlighting the time and frequency intervals when DSG and AMV have a significant interaction (Grinsted et al., 2004). The wavelet coherence method offers a bivariate extension of the wavelet analysis to identify regions with large common power in the time-frequency domain of the two time series, and reveals information about their phase relationship (which may be suggestive of dependency; Maraun and Kurths, 2004). The wavelet-coherence spectrum of the two time series (**Figure 7D**) displays sporadic high-frequency periodicities around 11 years. The quasi-11-year sunspot cycle is a main feature of solar activity (as identified by Wolf, 1852; Wolf, 1853). However, the 5% significance level is a not a reliable indication of dependency for erratic periods in the long-term behavior although Italian Alps snow-cover trends have been observed to oscillate with 11-year periods (Valt and Cianfarra, 2010). The region of the spectrum at lower-frequency periodicity of ~60 years is quite extensive and suggestive of a causal AMV-DSG relationship emerging at the onset of the warm period (across the change-point year of 1897 with AMV leading in time), which appears to be linked to North Atlantic internal ocean-atmosphere variability (e.g., Knudsen et al., 2011; Mazzarella and Scafetta, 2012). However, the significant area falls outside of the cone of influence where edge effects become important and prevents this analysis from obtaining a robust interpretation in this sense.

Influence of Atmospheric Circulation Patterns of Snow-Cover Periods

The occurrence of snow events in Parma is to some extent related to the presence of a low-pressure system over the central Mediterranean area, which is a favourable configuration for several snow days and persistent snow. In order to discover if different climatic periods are somewhat the result of changes occurred in the continental atmospheric circulation, we created composite plots of sea-level pressure for the winters of the late Maunder Minimum (**Figure 8A**), the late LIA until the change-point year (**Figure 8B**) and the modern warming (**Figure 8C**). These climatic sub-periods reflect many dominant types of large-scale atmospheric circulation patterns, which are responsible for wintertime conditions over the Mediterranean region. During the first period, 1681–1715 (**Figure 8A**), the atmospheric pattern shows that the central Mediterranean area is characterized by a winter depression, which attracts cold arctic air masses from the north/northeast of Europe. In the period 1716–1897 (**Figure 8B**), a similar atmospheric circulation pattern continued to favor the transition of cold air masses from northern Europe though with a shallower depression center lessening snowfall and snow-cover duration. There were certainly less frequent and severe cold spells during the late LIA than in the Maunder Minimum.

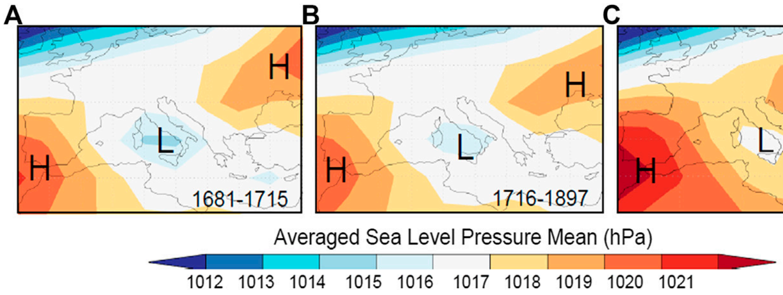


FIGURE 8 | Mean reconstructed sea-level pressure (SLP) maps over southern Europe over the winters (December to March) of three periods: **(A)** late Maunder Minimum (1681–1715 CE); Late Little Ice Age (1716–1897 CE); **(C)** Modern Warming (1897–2008). L = low-pressure system (cyclone), H = high-pressure system (anticyclone). Maps are authors' own elaboration using Climate Explorer (https://climexp.knmi.nl/start.cgi) [Accessed September 10, 2020] on data from: **(A,B)** Luterbacher et al. (2002); **(C)** from Küttel et al. (2010), which stops at year 2008.

However, the late LIA was also characterized by some severe winters. To give an example, taken from Corradi (1865–1890, vol. II, p. 623), in 1782 the whole of Italy was besieged by a deadly bad weather:

> Il freddo acuto e straordinario del mese di Febbraio è avvertito altresì dei Diarj Manoscritti di Montecassino, i quali pure lamentano la mortalità degli animali: secondo il Calandrelli nella campagna romana perirono nel Marzo, a cagione dei geli del mese precedente, da 102000 bestie, e pecore sopratutto.

> The sharp and extraordinary cold of the month of February is felt also by the Diarj Manuscripts of Montecassino, who also complain about the mortality of the animals: according to Calandrelli, in the Roman countryside by 102000 beasts perished in March, and especially sheep, due to the frosts of the previous month.

The exit of the LIA generally offered less inclement winter weather. The map of the third period, 1898–2008 (**Figure 8C**), shows the squeezing and reabsorption of the depression center by two anticyclones, which is indicative of snow weather (both mean and extreme snowfall) becoming less common. The two high-pressure systems (Azores high to the west and Russian high to the east) prevented the Mediterranean from prolonged snow events of certain importance, as compared to the previous periods.

CONCLUSION

The annually-resolved DSG (days with snow on the ground) time-series reconstruction documents variations at Parma OBS since 1681. With this reconstruction, unprecedented in time length, we have deepened our understanding of the characteristics of snowfall in the Po valley (northern Italy) and reported the consistency of our model-based reconstruction of DSG with historical data. That consistency suggests that the reconstructed historical signal may be representative of real multi-decadal variations. In fact, our 339-year long time series of DSG shows a particularly marked downward trend in recent decades, after the change point detected in 1897, and suggests that important shifts in mean DSG values and their variability may be resolved by extended reconstruction. Interannual and inter-decadal variations are evident in the reconstruction. While the causes of the decrease observed in DSG are manifold and varied, two factors are recognisable: an increase in seasonal temperatures (Dobrovolný et al., 2010; Serquet et al., 2013) and a contraction in the variability of snow precipitations from 1 year to the next. With respect to the latter, we highlight a general shortening of snow duration due to a delay in the start date of snow cover in autumn and an advancement of the end date in spring, associated with a retreat of glaciers, mainly due to a decrease in winter precipitation (Vincent et al., 2005; Huss et al., 2008). However, other factors besides the increase in temperature may have caused a decrease in the snow-cover duration. Already in 1881, on the occasion of the royal meeting of the Accademia dei Lincei ("Lincean Academy") in Rome (Italy), the Italian Catholic priest and geologist Antonio Stoppani (1824–1891) in a speech *Sull'attuale regresso dei ghiacciai sulle Alpi* ("On the current regression of glaciers in the Alps") provided as a main cause of the regression of the glaciers, not the variations of temperature (even if they were of common knowledge), but the reduction of snowiness (e.g., Scaramellini and Bonardi, 2001). In fact, in that period the number of snow days had been reduced to roughly a third in just about 50 years, a worrying phenomenon, especially because at that time there were no dams and artificial reservoirs, and glaciers were among the few large reserves of water (e.g., Stoppani, 1876). Again, from 1797 to 1806 the days of snow in Milan (Italy) had been 243 (i.e., 26 on average per year), but the situation had substantially changed from 1857 to 1876 with 166 snow days (i.e., eight snow days per year). Recently, Diodato et al. (2020b) provided evidence for Switzerland that this decrease is continuing in the Alpine range.

In our study, we report the results without providing conclusive evidence regarding the causes of the observed changes in the reconstructed DSG. Our model results offer robustness against anthropogenic disturbance in urban environments. The latter is relevant for the Parma territory, where Zanella (1976) showed an average difference of 1.4°C between urban and rural temperature data in the period 1959–1973, and reported that the difference varied seasonally (especially in spring and summer). Heat islands may thus have been a contributing factor of the snow-cover decline (e.g., Musco, 2016) observed in this urban site (from about 80 days year^{-1} in the 1940s to <20 days year^{-1} in the most recent years), which appears to be sufficiently captured by the modeled time series. We have corroborated our findings suggesting that change points in snowfall series in northern Italy can be linked to large-scale changes in the modes of climate (e.g., the Atlantic Multidecadal Variability), identifying in this way the need for future research on the subject. Though the long-term trend of the DSG series may be represented by a combination of intermittent periodicities, due to the brevity of the time interval the presence of these periodicities and their relationship with snowfall mechanisms in northern Italy should be regarded as tentative, and in need of confirmation by additional studies on extended spatial scales.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

ND conceived the study, performed the analysis and drafted the manuscript with GB, who wrote the final manuscript. CB prepared and organized the snowfall data used in the study, and contributed to scientific discussion of the article.

ACKNOWLEDGMENTS

We thank the NTNU publishing fund support to cover article-processing charges. The authors would like to thank Iñigo Gómara (Universidad Complutense de Madrid, Spain) for his support on wavelet-coherence analysis (wavelet software was provided by A. Grinsted).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.561148/full#supplementary-material

REFERENCES

Adamson, G. C. D. (2015). Private diaries as information sources in climate research. *WIREs Clim. Change.* 6, 599–611. doi:10.1002/wcc.365

Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *J. Climatol.* 6, 661–675. doi:10.1002/joc.3370060607

Alexandersson, H., and Moberg, A. (1997). Homogenization of Swedish temperature data. Part I: homogeneity test for linear trends. *Int. J. Climatol.* 17, 25–34. doi:10.1002/(sici)1097-0088(199701)17:1<25::aid-joc103>3.0.co;2-j

Barnhart, T. B., Molotch, N. P., Livneh, B., Harpold, A. A., Knowles, J. F., and Schneider, D. (2016). Snowmelt rate dictates streamflow. *Geophys. Res. Lett.* 43, 8006–8016. doi:10.1002/2016GL069690

Bednorz, E. (2004). Snow cover in eastern Europe in relation to temperature, precipitation and circulation. *Int. J. Climatol.* 24, 591–601. doi:10.1002/joc.1014

Bettini, E. (2016). *Climatologica dinamica delle precipitazioni sulla località di Milano*. Environmental and Land Planning Engineering Msthesis. Milan (Italy): Polytechnic University of Milan [in Italian].

Bradley, R. S., and Jones, P. D. (1992). "Records of explosive volcanic eruptions over the last 500 years," in *Climate since A.D. 1500*. Editors R. S. Bradley and P. D. Jones (London, UK: Routledge), 606–622.

Brázdil, R., Pfister, C., Wanner, H., Von Storch, H., and Luterbacher, J. (2005). Historical climatology in Europe - the state of the art. *Clim. Change.* 70, 363–430. doi:10.1007/s10584-005-5924-1

Brown, A. (2014). Mean and extreme snowfall. *Nat. Clim. Change.* 4,860. doi:10.1038/nclimate2404

Brown, I. (2019). Snow cover duration and extent for Great Britain in a changing climate: altitudinal variations and synoptic-scale influences. *Int. J. Climatol.* 39, 4611–4626. doi:10.1002/joc.6090

Brown, R. D. (2000). Northern Hemisphere snow cover variability and change, 1915–97. *J. Climate.* 13, 2339–2355. doi:10.1175/1520-0442(2000)013<2339:NHSCVA>2.0.CO;2

Brown, R., Derksen, C., and Wang, L. (2010). A multi-data set analysis of variability and change in Arctic spring snow cover extent, 1967-2008. *J. Geophys. Res.* 115, D16111. doi:10.1029/2010JD013975

Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *J. Hydrol.* 58, 11–27. doi:10.1016/0022-1694(82)90066-X

Camuffo, D., and Bertolin, C. (2012). The earliest temperature observations in the world: the Medici Network (1654-1670). *Clim. Change.* 111, 335–363. doi:10.1007/s10584-011-0142-5

Changnon, S. A., and Changnon, D. (2006). A spatial and temporal analysis of damaging snowstorms in the United States. *Nat. Hazards* 37 (3), 373–389.

Clark, M. P., Serreze, M. C., and Robinson, D. A. (1999). Atmospheric controls on Eurasian snow extent. *Int. J. Climatol.* 19, 27–40. doi:10.1002/(SICI)1097-0088(199901)19:1<27::AID-JOC346>3.0.CO;2-N

Cohen, J., and Rind, D. (1991). The effect of snow cover on the climate. *J. Climate.* 4, 689–706. doi:10.1175/1520-0442(1991)004<0689:TEOSCO>2.0.CO;2

Colla, A. (1840). *Giornale Meteorologico dell'anno 1839. Parma: registro manoscritto delle osservazioni fatte nella specola dell'Università*. Parma: University of Parma [in Italian].

Corradi, A. (1865–1890). *Annali delle epidemie occorse in Italia dalle prime memorie fino al 1850*. Bologna, Italy: Arnaldo Forni, Vol. 5 (reprint in 1972). [in Italian].

Croce, P., Formichi, P., Landi, F., Mercogliano, P., Bucchignani, E., Dosio, A., et al. (2018). The snow load in Europe and the climate change. *Clim. Risk Manage.* 20, 138–154. doi:10.1016/j.crm.2018.03.001

Davi, N. K., Jacoby, G. C., D'Arrigo, R. D., Baatarbileg, N., Jinbao, L., and Curtis, A. E. (2012). A tree-ring-based drought index reconstruction for far-western Mongolia: 1565-2004. *Int. J. Climatol.* 29, 1508–1514.

De Bellis, A., Pavan, V., and Levizzani, V. (2010). Bologna: Technical Report ARPA-SIMC – 19. Climatologia e variabilità interannuale della neve sull'Appennino Emiliano-Romagnolo. Available at: https://www.arpae.it/cms3/documenti/_cerca_doc/meteo/quaderni/quaderno_neve_2010.pdf (Accessed September 10, 2020).

De Walle, D., and Rango, A. (2008). *Principles of snow hydrology*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511535673

Déry, S. J., and Brown, R. D. (2007). Recent Northern Hemisphere snow cover extent trends and implications for the snow-albedo feedback. *Geophys. Res. Lett.* 34, L22504. doi:10.1029/2007GL031474

Dettinger, M. D., Cayan, D. R., Diaz, H. F., and Meko, D. M. (1998). North-South precipitation patterns in western North America on interannual-to-decadal timescales. *J. Clim.* 11, 3095–3111. doi:10.1175/1520-0442(1998)011<3095:NSPPI10.1175/1520-0442(1998)011<3095:nsppiw>2.0.co;2

Dietz, A. J., Wohner, C., and Kuenzer, C. (2012). European snow cover characteristics between 2000 and 2011 derived from improved modis daily snow cover products. *Rem. Sens.* 4, 2432–2454. doi:10.3390/rs4082432

Diodato, N. (1997). *Paesaggi d'inverno: aspetti naturalistici e climatologici delle nevicate sulla Campania interna. Benevento: La Provincia Sannita*. 17 [in Italian].

Diodato, N., and Bellocchi, G. (2020). Climate control on snowfall days in peninsular Italy. *Theor. Appl. Climatol.* 140, 951–961. doi:10.1007/s00704-020-03136-0

Diodato, N., Büntgen, U., and Bellocchi, G. (2019). Mediterranean winter snowfall variability over the past Millennium. *Int. J. Climatol.* 39, 384–394. doi:10.1002/joc.5814

Diodato, N., Bertolin, C., Bellocchi, G., Ferri, L., and Fantini, P. (2020a). New insights into the world's longest series of monthly snowfall (Parma, Northern Italy, 1777-2018). *Int. J. Climatol.* 6766, 1–17 [accepted]. doi:10.1002/joc.6766

Diodato, N., Fratianni, S., and Bellocchi, G. (2020b). Reconstruction of snow days based on monthly climate indicators in the Swiss pre-alpine region. *Reg. Environ. Change.* 20, 55. doi:10.1007/s10113-020-01639-0

Dobrovolný, P., Moberg, A., Brázdil, R., Pfister, C., Glaser, R., Wilson, R., et al. (2010). Monthly, seasonal and annual temperature reconstructions for Central Europe derived from documentary evidence and instrumental records since AD 1500. *Clim. Change.* 101, 69–107. doi:10.1007/s10584-009-9724-x

Durbin, J., and Watson, G. S. (1950). Testing for serial correlation in least squares regression. *I. Biometrika* 37 (3–4), 409–428. doi:10.1093/biomet/37.3-4.409. JSTOR 2332391

Durbin, J., and Watson, G. S. (1951). Testing for serial correlation in least squares regression. *II. Biometrika* 38 (1–2), 159–179. doi:10.1093/biomet/38.1-2.159. JSTOR 2332325

Eddy, J. A. (1976). The maunder minimum. *Science.* 192, 1189–1202. doi:10.1126/science.192.4245.118910.1126/science.192.4245.1189

Enzi, S., Bertolin, C., and Diodato, N. (2014). Snowfall time-series reconstruction in Italy over the last 300 years. *Holocene.* 24, 346–356. doi:10.1177/0959683613518590

Falarz, M. (2004). Variability and trends in the duration and depth of snow cover in Poland in the 20th century. *Int. J. Climatol.* 24, 1713–1727. doi:10.1002/joc.1093

Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M., and Stoffel, M. (2018). Estimating the snowfall limit in alpine and pre-alpine valleys: a local evaluation of operational approaches. *Atmos. Res.* 204, 136–148. doi:10.1016/j.atmosres.2018.01.016

Glaser, R., and Riemann, D. (2009). A thousand-year record of temperature variations for Germany and Central Europe based on documentary data. *J. Quat. Sci.* 24, 437–449. doi:10.1002/jqs.1302

Glaser, R., Militzer, S., and Busche, D. (1994). "A contribution to the reconstruction of climate in Germany during the Late Maunder Minimum (1675 to 1715)," in *Climatic trends and anomalies in Europe 1675–1715*. Editors B. Frenzel, C. Pfister, and B. Gläser (Stuttgart, Germany: Gustav Fischer), 173–190.

Godsey, S. E., Kirchner, J. W., and Tague, C. L. (2014). Effects of changes in winter snowpacks on summer low flows: case studies in the Sierra Nevada, California, USA. *Hydrol. Process.* 28, 5048–5064. doi:10.1002/hyp.9943

Gómara, I., Rodríguez-Fonseca, B., Zurita-Gotor, P., and Pinto, J. G. (2014). On the relation between explosive cyclones affecting Europe and the North Atlantic Oscillation. *Geophys. Res. Lett.* 41, 2182–2190. doi:10.1002/2014GL059647

Gómara, I., Rodríguez-Fonseca, B., Zurita-Gotor, P., Ulbrich, S., and Pinto, J. G. (2016). Abrupt transitions in the NAO control of explosive North Atlantic cyclone development. *Clim. Dynam.* 47, 3091–3111. doi:10.1007/s00382-016-3015-9

Goward, S. N. (2005). "Albedo and reflectivity," in *Encyclopedia of world climatology*. Editor J. E. Oliver (*Dordrecht, Neitherlands: encyclopedia of earth sciences series.*, Springer), 32–35.

Granger, C. W. J., Hyung, N., and Jeon, Y. (2001). Spurious regressions with stationary series. *Appl. Econ*, 33 (7), 899–904.

Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* 11, 561–566. doi:10.5194/npg-11-561-2004

Grove, J. M., and Conterio, A. L. (1994). "Climate in the eastern mediterranean, 1675 to 1715," in *Climatic trends and anomalies in Europe 1675-1715*. Editors B. Frenzel, C. Pfister, and B. Gläser (Stuttgart, Germany: Gustav Fischer), 275–285.

Handmer, J., Honda, Y., Kundzewicz, Z., Arnell, N., Benito, G., Hatfield, J., et al. (2012). "Changes in impacts of climate extremes: human systems and ecosystems," in Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. Editors C. Field, V. Barros, T. Stocker, and Q. Dahe (Cambridge: Cambridge University Press),. 231–290. doi:10.1017/CBO9781139177245.007

Harvey, L. D. D. (2013). "Climate and climate-system modelling," in *Environmental modelling: finding simplicity in complexity*. Editors J. Wainwright and M. Mulligan (Chichester, UK: John Wiley & Sons), 151–164.

Hawcroft, M. K., Shaffrey, L. C., Hodges, K. I., and Dacre, H. F. (2012). How much Northern Hemisphere precipitation is associated with extratropical cyclones? *Geophys. Res. Lett.* 39, L24809. doi:10.1029/2012GL053866

Huss, M., Bauder, A., Funk, M., and Hock, R. (2008). Determination of the seasonal mass balance of four Alpine glaciers since 1865. *J. Geophys. Res.* 113, F01015. doi:10.1029/2007JF000803

Jarque, C. M., and Bera, A. K. (1981). An efficient large-sample test for normality of observations and regression residuals. Working Papers in Economics and Econometrics 40. Canberra, Australia: The Australian National University.

Jennings, K. S., Winchell, T. S., Livneh, B., and Molotch, N. P. (2018). Spatial variation of the rain-snow temperature threshold across the Northern Hemisphere. *Nat. Commun.* 9, 1148. doi:10.1038/s41467-018-03629-7

Jones, P. D., and Harris, I. C. (2018). Climatic Research Unit (CRU): time-series (TS) datasets of variations in climate with variations in other phenomena v3. Leeds: NCAS British Atmospheric Data Centre. Available at: http://catalogue.ceda.ac.uk/uuid/3f8944800cc48e1cbc29a5ee12d8542d (Accessed September 10, 2020). doi:10.4324/9781315268590

Jungclaus, J. H. (2009). Lessons from the past millennium. *Nat. Geosci*. 2 (7), 468–470.

Kendall, M. (1975). *Rank correlation measures*. London, UK: Charles Griffin.

Knudsen, M. F., Seidenkrantz, M.-S., Jacobsen, B. H., and Kuijpers, A. (2011). Tracking the atlantic multidecadal oscillation through the last 8,000 years. *Nat. Commun.* 2, 178. doi:10.1038/ncomms1186

Knoben, W. J. M., Freer, J. E., and Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrol. Earth Syst. Sci.* 23 (10), 4323–4331.

Kling, H., Fuchs, M., and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J. Hydrol.* 424–425, 264–277. doi:10.1016/j.jhydrol.2012.01.011

Krasting, J. P., Broccoli, A. J., Dixon, K. W., and Lanzante, J. R. (2013). Future changes in Northern Hemisphere snowfall. *J. Clim.* 26, 7813–7828. doi:10.1175/JCLI-D-12-00832.1

Kretschmer, M., Cohen, J., Matthias, V., Runge, J., and Coumou, D. (2018). The different stratospheric influence on cold-extremes in Eurasia and North America. *NPJ Clim. Atmos. Sci.* 1, 44. doi:10.1038/s41612-018-0054-4

Kunkel, K. E., Robinson, D. A., Champion, S., Yin, X., Estilow, T., and Frankson, R. M. (2016). Trends and extremes in Northern Hemisphere snow characteristics. *Curr. Clim. Change Rep.* 2, 65–73. doi:10.1007/s40641-016-0036-8

Küttel, M., Xoplaki, E., Gallego, D., Luterbacher, J., García-Herrera, R., Allan, R., et al. (2010). The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Clim. Dynam.* 34, 1115–1128. doi:10.1007/s00382-009-0577-9

Leporati, E., and Mercalli, L. (1994). Snowfall series of Turin, 1784-1992: climatological analysis and action on structures. *Ann. Glaciol.* 19, 77–84. doi:10.3189/S0260305500011010

Lim, K. J., Engel, B. A., Tang, Z., Muthukrishnan, S., Choi, J., and Kim, K. (2006). Effects of calibration on L-THIA GIS runoff and pollutant estimation. *J. Environ. Manag.* 78, 35–43. doi:10.1016/j.jenvman.2005.03.014

Luce, C. H., Abatzoglou, J. T., and Holden, Z. A. (2013). The missing mountain water: slower westerlies decrease orographic enhancement in the Pacific Northwest USA. *Science.* 342, 1360–1364. doi:10.1126/science.1242335

Lute, A. C., Abatzoglou, J. T., and Hegewisch, K. C. (2015). Projected changes in snowfall extremes and interannual variability of snowfall in the Western United States. *Water Resour. Res.* 51, 960–972. doi:10.1002/2014WR016267

Luterbacher, J., Xoplaki, E., Dietrich, D., Rickli, R., Jacobeit, J., Beck, C., et al. (2002). Reconstruction of sea level pressure fields over the eastern North Atlantic and Europe back to 1500. *Clim. Dynam.* 18, 545–561. doi:10.1007/s00382-001-0196-6

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science.* 303, 1499–1503. doi:10.1126/science.1093877

Mangianti, F., and Beltrano, M. C. (1993). *La neve a Roma dal 1741 al 1990*. Roma: Central Office of Agricultural Ecology of the Italian Ministry of Agriculture. [in Italian].

Mann, M. E., Gille, E., Bradley, R. S., Hughes, M. K., Overpeck, J., Keimig, F. T., et al. (2000). Global temperature patterns in past centuries: an interactive presentation. *Earth Interact.* 4, 1–29. doi:10.1175/1087-3562(2000)004<0001:GTPIPC>2.3.CO;2

Maraun, D., and Kurths, J. (2004). Cross wavelet analysis: significance testing and pitfalls. *Nonlinear Process. Geophys.* 11, 505–514. doi:10.5194/npg-11-505-2004

Martínez, M. D., Serra, C., Burgueño, A., and Lana, X. (2009). Time trends of daily maximum and minimum temperatures in Catalonia (ne Spain) for the period 1975-2004. *Int. J. Climatol.* 30, 267–290. doi:10.1002/joc.1884

Mazzarella, A., and Scafetta, N. (2012). Evidences for a quasi 60-year North Atlantic Oscillation since 1700 and its meaning for global climate change. *Theor. Appl. Climatol.* 107, 599–609. doi:10.1007/s00704-011-0499-4

Meir, P., Cox, P., and Grace, J. (2006). The influence of terrestrial ecosystems on climate. *Trends Ecol. Evol.* 21, 254–260.

Miller, G. H., Geirsdóttir, Á., Zhong, Y., Larsen, D. J., Otto-Bliesner, B. L., Holland, M. M., et al. (2012). Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophys. Res. Lett.* 39, L02708. doi:10.1029/2011GL050168

Mulligan, M., and Wainwright, J. (2013). "Modelling and model building," in *Environmental modelling: finding simplicity in complexity*. Editors J. Wainwright and M. Mulligan (Chichester: John Wiley & Sons), 5–29.

Musco, F. (2016). Counteracting Urban Heat Island Effects in a Global Climate Change Scenario. Switzerland: Springer International Publishing AG.

Musselman, K. N., Clark, M. P., Liu, C., Ikeda, K., and Rasmussen, R. (2017). Slower snowmelt in a warmer world. *Nat. Clim. Change.* 7, 214–219. doi:10.1038/nclimate3225

Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - a discussion of principles. *J. Hydrol.* 10, 282–290. doi:10.1016/0022-1694(70)90255-6

Nau, R. (2005). STATGRAPHICS V.5: Overview & tutorial guide, http://www.duke.edu/~rnau/sgwin5.pdf.

O'Gorman, E. J. (2014). Integrating comparative functional response experiments into global change research. *J. Anim. Ecol.* 83 (3), 525–527.

Pauling, A., Luterbacher, J., and Wanner, H. (2003). Evaluation of proxies for European and North Atlantic temperature field reconstructions. *Geophys. Res. Lett.* 30, 1787. doi:10.1029/2003GL017589

Petkova, N., Koleva, E., and Alexandrov, V. (2004). Snow cover variability and change in mountainous regions of Bulgaria, 19312000. *Meteorol. Z.* 13, 19–23. doi:10.1127/0941-2948/2004/0013-0019

Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Appl. Stat.* 28, 126–135. doi:10.2307/2346729

Pimentel, R., Herrero, J., and Polo, M. (2017). Quantifying snow cover distribution in semiarid regions combining satellite and terrestrial imagery. *Rem. Sens.* 9, 995. doi:10.3390/rs9100995

Pinto, J. G., and Raible, C. C. (2012). Past and recent changes in the North Atlantic oscillation. WIREs Clim. Change, 7, 79–90. doi:10.1002/wcc.150

Royston, P., and Sauerbrei, W. (2008). *Multivariate model-building*. Chichester: John Wiley & Sons.

Royston, P., and Sauerbrei, W. (2008). *Multivariate model-building*. Chichester: John Wiley & Sons.

Sandells, M., and Flocco, D. (2014). "Snow," in *Introduction to the physics of the cryosphere*. Editors M. Sandells and D. Flocco (Reading, UK: Morgan & Claypool Publishers), Vol. 3, 3.1–3.17.

Scaramellini, G., and Bonardi, L. (2001). La géographie italienne et les Alpes de la fin du XIXe siècle à la Seconde Guerre mondiale. *Revue de Géographie Alpine.* 89, 133–158. [in French]. doi:10.3406/rga.2001.3062

Selkowitz, D. J., Fagre, D. B., and Reardon, B. A. (2002). Interannual variations in snowpack in the crown of the continent ecosystem. *Hydrol. Process.* 16, 3651–3665. doi:10.1002/hyp.1234

Serquet, G., Marty, C., and Rebetez, M. (2013). Monthly trends and the corresponding altitudinal shift in the snowfall/precipitation day ratio. *Theor. Appl. Climatol.* 114, 437–444. doi:10.1007/s00704-013-0847-7

Si, D., Jiang, D., and Wang, H. (2020). Intensification of the atlantic multidecadal variability since 1870: implications and possible causes. *J. Geophys. Res. Atmos.* 125, e2019JD030977. doi:10.1029/2019JD030977

Spence, C., and Mengistu, S. G. (2019). On the relationship between flood and contributing area. *Hydrol. Process.* 33, 1980–1992. doi:10.1002/hyp.13467

Stella, A. F. (1836). *Biblioteca italiana ossia giornale di letteratura scienze ed arti compilato da una societa di letterati*. Tomo, LXXXI. Milano, Italy: Direzione del Giornale di Letteratura, Scienze ed Arti [in Italian].

Stoppani, A. (1876). *Il Bel Paese. Conversazioni sulle bellezze naturali, la geologia, e la geografia fisica d'Italia*. Milan: Tipografia e Libreria Editrice Ditta Giacomo Agnelli. [in Italian].

Sutton, R. T., and Dong, B. (2012). Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat. Geosci.* 5, 788–792. doi:10.1038/ngeo1595

Terzago, S., Fratianni, S., and Cremonini, R. (2012). Winter precipitation in Western Italian Alps (1926-2010). *Meteorol. Atmos. Phys.* 119, 125–136. doi:10.1007/s00703-012-0231-7

Toreti, A., Kuglitsch, F. G., Xoplaki, E., Della-Marta, P. M., Aguilar, E., Prohom, M., et al. (2011). A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series. *Int. J. Climatol.* 31, 630–632. doi:10.1002/joc.2088

UNESCO/IASH/WMO (1970). *Seasonal snow cover*. Paris: United Nations Educational, Scientific and Cultural Organization.

Valt, M., and Cianfarra, P. (2010). Recent snow cover variability in the Italian Alps. *Cold Reg. Sci. Technol.* 64, 146–157. doi:10.1016/j.coldregions.2010.08.008

Vincent, C., Le Meur, E., and Six, D. (2005). Solving the paradox of the end of the Little Ice Age in the Alps. *Geophys. Res. Lett.* 32, L09706. doi:10.1029/2005GL022552

Wagner, S., and Zorita, E. (2005). The influence of volcanic, solar and CO2 forcing on the temperatures in the Dalton Minimum (1790-1830): a model study. *Clim. Dynam.* 25, 205–218. doi:10.1007/s00382-005-0029-0

Wang, J., Liang, Z., Wang, D., Liu, T., and Yang, J. (2016). Impact of climate change on hydrologic extremes in the Upper Basin of the Yellow River Basin of China. *Adv. Meteorol.* 2016, 1–13. doi:10.1155/2016/1404290

Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R., et al. (2017). Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years. *Nat. Geosci.* 10, 512–517. doi:10.1038/ngeo2962

Wang, X. L., Wen, Q. H., and Wu, Y. (2007). Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteorolog. Clim.* 46, 916–931. doi:10.1175/JAM2504.1

Wang, Y., Yang, J., Chen, Y., Wang, A., and De Maeyer, P. (2018a). The spatiotemporal response of soil moisture to precipitation and temperature changes in an arid region, China. *Rem. Sens.* 10, 468. doi:10.3390/rs10030468.

Wang, Y., Yang, B., Osborn, T. J., Ljungqvist, F. C., Zhang, H., and Luterbacher, J. (2018b). Causes of East Asian temperature multidecadal variability since 850 CE. *Geophys. Res. Lett.* 45, 13485–13494. doi:10.1029/2018GL080725

Wessa, P. (2012). Free statistics software, version 1.1.23-r7. Office for Research Development and Education, http://www.wessa.net.

Wijngaard, J. B., Klein Tank, A. M. G., and Können, G. P. (2003). Homogeneity of 20th century European daily temperature and precipitation series. *Int. J. Climatol.* 23, 679–692. doi:10.1002/joc.906

WMO (2008). *Annual bulletin on the climate in WMO region VI - Europe and Middle East - 2008*. Offenbach am Main: Deutscher Wetterdienst.

Wolf, R. (1852). Neue Untersuchungen über die Periode der Sonnenflecken und ihre Bedeutung. *Mitt. Naturforsch. Ges. Bern.* 255, 249–270. [in German].

Wolf, R. (1853). Ueber den Zusammenhang magnetischer Erscheinungen mit dem Zustande der Sonne. *Astr. Nachr. AN.* 35, 59–60. [in German]. doi:10.1002/asna.18530350407

Worsley, K. J. (1986). Confidence regions and tests for a change-point in a sequence of exponential family random variables. *Biometrika.* 73, 91–104. doi:10.1093/biomet/73.1.91

Wyatt, M. G., Kravtsov, S., and Tsonis, A. A. (2012). Atlantic multidecadal oscillation and Northern Hemisphere's climate variability. *Clim. Dynam.* 38, 929–949. doi:10.1007/s00382-011-1071-8

Xoplaki, E., Maheras, P., and Luterbacher, J. (2001). Variability of climate in Meridional Balkans during the periods 1675-1715 and 1780-1830 and its impact on human life. *Clim. Change.* 48, 581–615. doi:10.1023/A:1005616424463

Yarnell, S. M., Viers, J. H., and Mount, J. F. (2010). Ecology and management of the spring snowmelt recession. *Bioscience.* 60, 114–127. doi:10.1525/bio.2010.60.2.6

Zanella, G. (1976). Il clima urbano di Parma Rivista Meteorologia Aeronautica, 36, 125–146.

Zhang, X., Drake, N. A., and Wainwright, J. (2013). "Spatial modelling and scaling issues," in *Environmental modelling: finding simplicity in complexity.* Editors J. Wainwright and M. Mulligan (Chichester: John Wiley & Sons), 69–90.

Zhang, Y., and Ma, N. (2018). Spatiotemporal variability of snow cover and snow water equivalent in the last three decades over Eurasia. *J. Hydrol.* 559, 238–251. doi:10.1016/j.jhydrol.2018.02.031

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.

frontiers
in Earth Science

METHODS
published: 07 December 2020
doi: 10.3389/feart.2020.537028

Instrument Bias Correction With Machine Learning Algorithms: Application to Field-Portable Mass Spectrometry

B. Loose[1]*, **R. T. Short**[2] **and S. Toler**[2]

[1]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, United States, [2]Advanced Technology and Systems Division, SRI International, St. Petersburg, FL, United States

In situ sensors for environmental chemistry promise more thorough observations, which are necessary for high confidence predictions in earth systems science. However, these can be a challenge to interpret because the sensors are strongly influenced by temperature, humidity, pressure, or other secondary environmental conditions that are not of direct interest. We present a comparison of two statistical learning methods—a generalized additive model and a long short-term memory neural network model for bias correction of *in situ* sensor data. We discuss their performance and tradeoffs when the two bias correction methods are applied to data from submersible and shipboard mass spectrometers. Both instruments measure the most abundant gases dissolved in water and can be used to reconstruct biochemical metabolisms, including those that regulate atmospheric carbon dioxide. Both models demonstrate a high degree of skill at correcting for instrument bias using correlated environmental measurements; the difference in their respective performance is less than 1% in terms of root mean squared error. Overall, the long short-term memory bias correction produced an error of 5% for O_2 and 8.5% for CO_2 when compared against independent membrane DO and laser spectrometer instruments. This represents a predictive accuracy of 92–95% for both gases. It is apparent that the most important factor in a skillful bias correction is the measurement of the secondary environmental conditions that are likely to correlate with the instrument bias. These statistical learning methods are extremely flexible and permit the inclusion of nearly an infinite number of correlates in finding the best bias correction solution.

Keywords: neural network, long short-term memory, mass spectrometry, generalized additive model, bias, ocean carbon, ocean oxygen

Edited by:
Flavio Cannavo',
National Institute of Geophysics and Volcanology, Italy

Reviewed by:
Adil Rasheed,
Norwegian University of Science and Technology, Norway
Lingxin Chen,
Chinese Academy of Sciences, China

****Correspondence:***
B. Loose
bloose@uri.edu

Specialty section:
This article was submitted to Environmental Informatics and Remote Sensing, a section of the journal Frontiers in Earth Science

Received: *21 February 2020*
Accepted: *26 October 2020*
Published: *07 December 2020*

Citation:
Loose B, Short RT and Toler S (2020) Instrument Bias Correction With Machine Learning Algorithms: Application to Field-Portable Mass Spectrometry. Front. Earth Sci. 8:537028. doi: 10.3389/feart.2020.537028

INTRODUCTION

The uncalibrated signal (s) produced by an environmental sensor contains the superposition of multiple influences. These include the instrument response to an environmental property of interest: $y(\vec{x}, t)$, but it also includes some instrument responses (β) that are not of interest, as well as some uncorrelated or random error (ε). The undesirable influences in β can be represented if the environmental influences or correlates, X, are separately measured. An example of $\beta(X)$ would be changes in the internal resistance of a circuit board as the room temperature varies. We refer to $\beta(X)$ as instrumental bias, and their influence on s can be treated as additive,

$$s = y\left(\vec{x}, t\right) + \beta(X) + \varepsilon \tag{1}$$

and therefore separable from $y(\vec{x}, t)$, the desired environmental response.

Experimental chemistry has been slow to consider bias and systematic error, in part because the end goal of many studies was the demonstration of a corollary relationship rather than a process model (Newman, 1993). However, when the same relationships are used in a predictive capacity, the uncorrected bias can lead to erroneous results. Recently, bias has been given more explicit treatment through applications such as air quality for human health (Delle Monache et al., 2006) and charge state in electric vehicles (Sun et al., 2016). These and other applications demand accurate forecasts, thereby renewing focus on elimination of bias from the process model.

Within the geosciences, the problem of chronic undersampling in diffusive environments, such as air and water (Pimentel, 1975), has created a strong incentive to take instruments out of the lab to increase sample density and better characterize the tracer field. If samples are analyzed in a discrete fashion, instrumental drift that leads to bias can be accounted for with pre/postcalibration to constrain the instrument drift. This was the approach adopted by, e.g., Guegen and Tortell (2008) to measure dimethyl sulfide (DMS) and carbon dioxide—two climatically important gases—during a shipboard expedition in the Southern Ocean. However, the continuous sampling that takes place with *in situ* or underway chemical sensors requires a slightly different approach to account for instrument drift as a source of bias. One clever solution has been to switch to reference compound(s) at regular intervals as part of the measurement protocol. This has the effect of chopping up the time series and introducing data gaps, but these gaps are often small (minutes) in comparison to the averaging interval (tens of minutes to hours) that is utilized for final data presentation. Takahashi et al. (2002) and Takahashi et al. (2009) have used the approach of reference compounds at intervals to create very precise coverages of ocean surface carbon dioxide concentration for several decades. Cassar et al. (2009) showed that mass spectrometer drift, while measuring oxygen and argon, could be characterized by switching regularly to measure atmospheric air. Saltzman et al. (2009) describe a detailed method for continuous measurement of DMS using a chemical ionization mass spectrometer. Their approach, which uses DMS isotope dilution, also uses switching at intervals to characterize several bias corrections and account for internal sources of DMS, as well as sensitivity of the instrument to changes in seawater temperature and other environmental factors. These biases are reported at less than 1% of the overall DMS signal.

The approach of regular switching to a reference compound is a proven means to correct for drift in continuous instruments. However, the instrumental conditions that we confront in this study differ in two significant ways from the previously described continuous measurement methods. The first difference has to do with the magnitude of the bias, compared to the signal of interest. Previous underway studies have confronted bias corrections of a few to 10% of the overall instrumental signal, while the instrumental bias that we face can vary by 100% or more. The magnitude of this bias renders the true environmental signal unrecognizable until the correction has been applied. The second major difference is that previous studies have identified the most likely sources of bias, but they have not quantified those sources to implement the bias correction. When the instrumental bias masks the true environmental signal, the bias must be treated as a continuously varying function, and therefore, a simple linear correction to baseline drift is not adequate. This bias correction problem lends itself to time series and multivariate regression techniques, including partial least squares, ridge regression, generalized linear, and generalized additive models (Hastie et al., 2001).

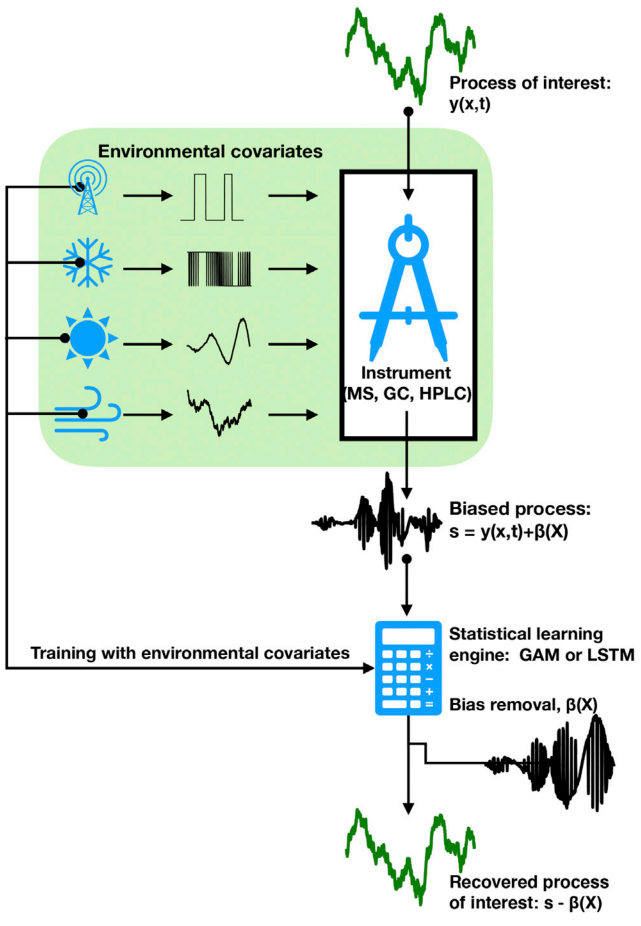


FIGURE 1 | A schematic depiction of the effect of environmental factors on the introduction of bias into *in situ* chemical instrumentation and the subsequent identification and removal of bias using environmental covariates to train a statistical learning engine.

Multivariate time series predictions have undergone a period of rapid development and availability thanks to the popularity of another member of the statistical learning family, neural networks, which have proven facile at, e.g., image and speech recognition. Neural networks are also suited for time series applications including forecasting or prediction (Brownlee, 2019a). Specifically, the long short-term memory (LSTM) algorithm combines the learning power of neural networks with a capacity to down-weight or "forget" information that does not prove relevant, leading to the overall stability of the network optimization (Hochreiter and Schmidhuber, 1997).

In this application, we apply and compare a generalized additive model (GAM) and a LSTM neural network model to observe their performance in baseline correction to mass spectrometer data. A schematic depiction of the bias correction workflow can be observed in **Figure 1**. Both the GAM and LSTM models use the statistical learning approach to optimize their calibration and weight coefficients. However, there is a fundamental difference in approach and user control. The LSTM weights and tradeoffs are largely abstracted from the user; one has to trust the algorithm without being able to interrogate the details of the solution. The consolation is the tremendous skill that the LSTM models exhibit in preserving the information that is necessary to discriminate or predict while avoiding the spurious oscillations that can characterize simpler, stiffer models. Unlike the LSTM, the GAM represents a linear combination of regression models (Wood, 2017) between each environmental correlate (X_i) and the instrument signal (s). This allows the user to observe and evaluate the partial dependence of the GAM solution on each X_i and to alter the functional form (e.g., linear, polynomial, and cubic spline) that is fit between s and each X_i. The effect is to give the user greater control over the functional form and the partial influence of each correlate on the total solution.

The signals of interest to this study are measurements of gases dissolved in water and seawater using field-portable quadrupole mass spectrometers (QMS). We present examples of the GAM and LSTM applied to data from a submersible wet inlet mass spectrometer (SWIMS) that was used to measure dissolved oxygen in the top 150 m of the Sargasso Sea and Gulf Stream, in the subtropical Atlantic Ocean. We present a second example of signals collected with a similar mass spectrometer aboard a ship that was used to measure dissolved carbon dioxide at the ocean surface, within the sea ice-covered Ross Sea, Antarctica.

Throughout this text, we make references to Python modules that were used to implement the individual solutions. The implementation of the GAM backfit algorithm, as well as example scripts for applying these methods to SWIMS data, can be found in the **Supplementary Material** and in the Acknowledgments.

METHODS

The bias correction models were each applied to ocean measurements of gases dissolved in seawater. These measurements were made using a QMS. The QMS is an ideal tool for ocean measurements because it is compact, and it can scan over a large range of atomic masses. In this study, we refer to the mass-to-charge ratio (m/z), where m represents the atomic mass of the molecule of interest, and z represents the positive charge state. For example, water vapor is measured in the QMS at m/z = 18, and molecular oxygen (O_2) is measured at m/z = 32. In this study, z = 1 in every instance. The QMS can be connected to a variety of gas inlet configurations. Further detail on the principles of quadrupole mass spectrometry can be found in Dawson and Herzog, 1995, but they are not needed to follow the methods presented here.

Ocean Data Used to Evaluate the Bias Correction Models

Submersible Wet Inlet Mass Spectrometer Tow

The first ocean dataset was collected in July 2017 along a dynamic section of the subtropical Atlantic between 35° and 40° N latitude (**Figure 2**). The QMS was incorporated into a submersible wet inlet mass spectrometer (SWIMS), which is capable of *in situ* gas analysis to a water depth of 2000 m; in this application, we towed the mass spectrometer through water depths from 0 to 150 m aboard a Triaxus tow vehicle, corresponding to a region where sunlight penetrates the surface ocean. The triaxus tow vehicle was also equipped with a CTD to measure water column properties. The SWIMS position can be visualized by the gray saw-tooth pattern in panel **b** of **Figure 3**. Calibration of the SWIMS instrument is described below in *In Situ Calibration of the Submersible Wet Inlet Mass Spectrometer*.

This ocean section began in the North Atlantic subtropical gyre, a circulation feature that is known to be highly depleted of nutrients with low biomass (Jenkins, 1982). In summer, surface waters in the Gulf Stream and gyre can exceed 30°C, and nearly 1% of temperature measurements in this section fell between 30° and 35°C (**Figure 3**). North of the Gulf stream, waters cool and become significantly fresher, reflecting river inputs and the influence of the southward-flowing Labrador Current (Chapman and Beardsley, 1989). We chose to test the bias correction models in this region because the environment is highly changeable on a small horizontal and vertical scale; so, the SWIMS is subjected to a wide range of environmental conditions, including temperature, salinity, and dissolved organic matter—all of which can cause the dissolved gas burden of the seawater to vary.

The SWIMS was being used to measure oxygen, argon, carbon dioxide, nitrogen, and methane in the surface ocean. Each of these dissolved gases has significance for biology and geochemistry of the ocean. Our *in situ* calibration system included reference gases for each of these compounds, allowing the SWIMS to reproduce realistic concentration distributions for each analyte. Here, we will restrict analysis of the bias correction to the SWIMS signal at m/z = 32, corresponding to dissolved oxygen. By developing the bias correction at m/z = 32, we are able to take advantage of independent measures of dissolved oxygen using a membrane oxygen sensor, the Seabird model SBE 43, which allows for a detailed reference time series, throughout the vehicle tow. Ultimately, we use the root mean square error between the SBE43 and the SWIMS to establish a truly independent measure of the bias correction algorithm.

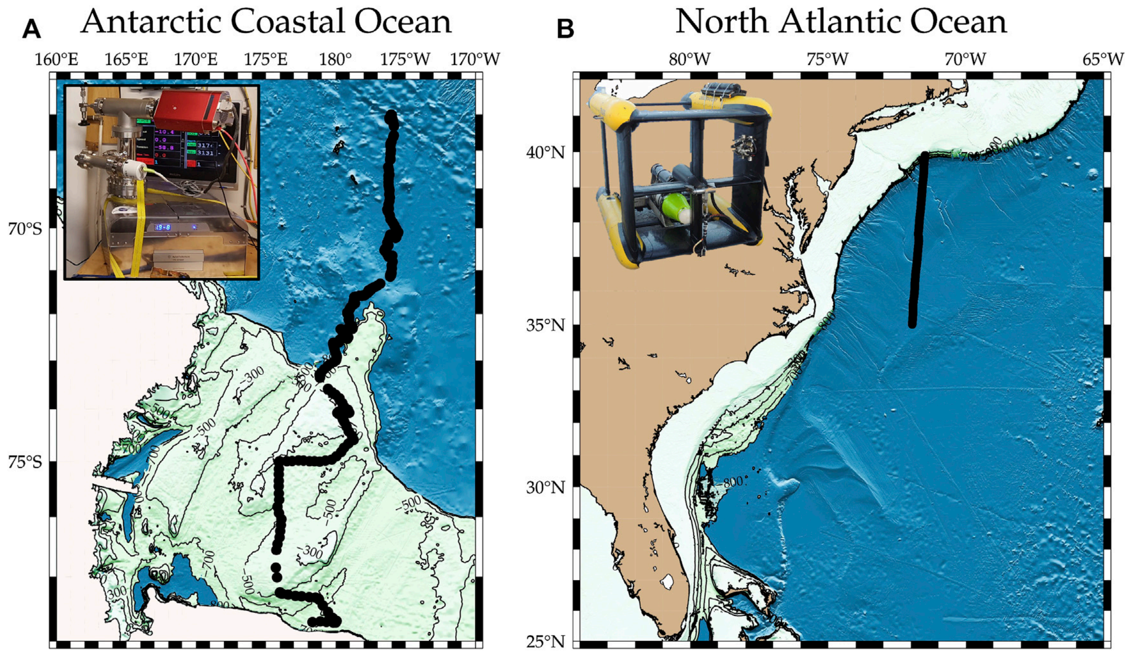


FIGURE 2 | Maps showing the ocean regions where QMS data were collected as part of oceanographic surveys. Panel **(A)** is a region in the Ross Sea, along the coast of Antarctica in the Atlantic sector; panel **(B)** shows a region of the North Atlantic, including the Gulf Stream and Labrador water. The measurements collected at these locations are the subjects of the additive model and neural network bias correction algorithms that we compare in this study.

Shipboard Quadrupole Mass Spectrometers

The bias correction models were also tested on data from a shipboard QMS that continuously sampled dissolved gases in the Ross Sea sector of the South Atlantic, south of 75°S. These measurements were collected between May 16 and June 4, 2017. The partial pressure of carbon dioxide (pCO_2) was measured by connecting the QMS directly to a turbulent air-water equilibrator of the type described by Takahashi (1961). The same equilibrator was used to measure pCO_2 by infrared absorption spectroscopy (Takahashi et al., 2002; Takahashi et al., 2009); again, this provided an independent measurement to compare against the bias correction. The QMS was connected to the equilibrator with a 2 m × 50 μm (len × dia) capillary, which served to throttle the gas flow into the QMS and thereby maintain a vacuum below 10^{-5} torr.

Carbon dioxide was measured with the QMS by scanning at the atomic mass $m/z = 44$. The reconstruction of pCO_2 was carried out with a daily 3-point linear calibration with reference gases of $pCO_2 = 0\%$, 0.4%, and 0.1%. These signals can be seen in the expanded scale on the right side of **Figure 4**. Unlike the SWIMS tows, these calibrations were not long enough in duration to record the bias while sampling from a stable gas concentration. Therefore, we apply the bias correction to a time series of CO_2 partial pressure (pCO_2) measured at atomic mass $m/z = 44$. Instead, the GAM and LSTM models were trained on relatively stable ion current signals measured during a four-day period between May 27 and June 1.

This late autumn period in the Southern Hemisphere was cold and windy with continual disaggregated ice formation in the surface ocean. The principal source of bias appeared from the thermal cycling in the room where the QMS and equilibrator were operating (**Figure 4**). The heating system in that room would cause the temperature in the room to increase and decrease by 2-3°C every 30 min. Additionally, the seawater intake was periodically clogged with ice crystals, causing the equilibrator flow rate to vary.

In Situ Calibration of the SWIMS Instrument

The SWIMS passes seawater directly over a gas-permeable silicone membrane under conditions that approach a constant flow rate while maintaining constant water temperature using a resistive heater and aluminum block (Short et al., 2001; Wenner et al., 2004). The wet membrane inlet is a simple and elegant design that allows for a submersible instrument, but it is subjected to a number of confounding environmental influences that complicate interpretation of the SWIMS ion current. The most significant of these is a change in membrane permeability as it is compressed under the increasing water pressure (Bell et al., 2007). The permeability behavior is made more complex by hysteresis between the compression and decompression cycles (Futó and Degn, 1994; Lee et al., 2016). Over progressive cycles, the silicon membrane can become tempered and eventually exhibits less compressibility (Futó and Degn, 1994; Lee et al., 2016), which indicates that any bias correction should include multiple

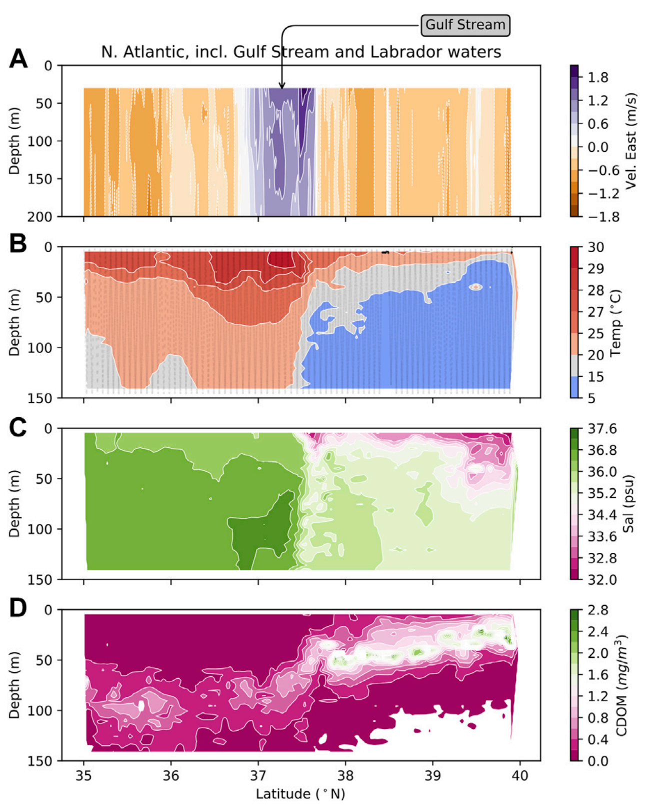


FIGURE 3 | Ocean properties during the SWIMS tow in the N. Atlantic, across the Gulf Stream and into coastal waters influenced by the Labrador Current. The track lines of the tow are shown in panel **(B)**.

compression-decompression cycles to capture the longer term transients. To capture this and other sources of bias, we designed an *in situ* calibration method that involves connecting the SWIMS to a 1 L Tedlar bag that contains seawater, equilibrated with a reference gas mixture. The sample in the compressible Tedlar bag is subjected to the same pressure variations as the water column sample, but gas concentrations remain constant because there is no gas headspace in the bag. Using a 3-way solenoid switching valve, the SWIMS can change states from sampling the environment to sampling the constant reference gas. Because the gas concentration is invariant, any trends in ion current that are observed must be due to instrumental bias. An example of this instrumental bias can be observed in **Figure 5**, which shows the environmental correlates measured during approximately 4.8 tow cycles while measuring from the *in situ* calibration reference. These signal variations are what we seek to correct.

Calibration after Bias Removal

To discover the instrument response, it is necessary to remove $\beta(X)$, the instrumental bias, and rearrange **Eq. 1** as follows:

$$y\left(\vec{x}, t\right) = f\left(s - \beta(X)\right). \tag{2}$$

Here forward, we drop the explicit reference to uncorrelated error (ε), which means that this error source is still a part of s. After bias correction, it is still necessary to estimate the uncertainty on y that is caused by ε, but that topic is extensively covered by other studies, so it will not be addressed here.

Therefore, the steps to obtain y are to the first model $\beta(X)$, so that it can be removed, and then to calibrate to obtain the empirical dependency, $f()$, between y and the bias-corrected signal. To make this procedure less abstract, we focus on measuring the oxygen concentration in seawater $y = [O_2]$ using the ion current measured at m/z = 32. The raw ion current (s) in amps at m/z = 32 responds directly not only to the amount of O_2 dissolved in the water but also to other environmental correlates, X. The values of X must be measured as a time series, coincident with the instrument's deployment. Other properties that we might include in X are, for example, the duty cycle of a heater or chiller, the atmospheric pressure, the temperature of a chemically reactive solute (e.g., pH-sensitive dye), or the electrical conductivity of a water solution. The environmental correlates used to model $\beta(X)$ in the SWIMS are shown in **Figure 5**, and the correlates used to model $\beta(X)$ in the shipboard QMS are shown in **Figure 4**.

After bias removal, s reflects only the environmental signal of interest and some component of random error; $s_{-\beta}$ denotes the ion current after bias removal, and this term is calibrated against the reference gas concentrations using a linear equation,

$$\begin{aligned} f\left(s_{-\beta}\right) &= m\left(s_{-\beta}\right) + b \quad \text{or} \\ y\left(\vec{x}, t\right) &= m\left(s_{-\beta}\right) + s^0. \end{aligned} \tag{3}$$

Here, the terms m and s^0 are the slope and intercept, and these terms are estimated as described in *Ocean Data Used to Evaluate the Bias Correction Models*. Practically, we estimate m as

$$m = \frac{y - y^0}{s_{-\beta} - s^0_{-\beta}}. \tag{4}$$

At the limit of $y^0 = [O_2] = 0$, the ion current does not reach zero because of electronic noise, and the potential for "virtual leaks" as gas is desorbed from the walls of the QMS under vacuum. In other words, y^0 is always zero, but in practice, $s^0_{-\beta}$ in **Eq. 3** reflects the nonzero ion current at undetectable gas concentrations leaving the following linear calibration:

$$y\left(\vec{x}, t\right) = m\left(s_{-\beta} - s^0_{-\beta}\right). \tag{5}$$

The technique for determining m and $s^0_{-\beta}$ for the shipboard QMS was determined by measuring m/z = 44 at $pCO_2 = 0$ in ultrapure N_2 gas, as described in *Ocean Data Used to Evaluate the Bias Correction Models*. The SWIMS determination of m occurred during *in situ* calibration, which took place ca. every other day; however, we did not determine $s^0_{-\beta}$, so it became necessary to account for baseline drift in the SWIMS using an external reference. We implemented a reference to the equilibrium oxygen solubility, based on seawater temperature and salinity. We also used the SBE43 as a daily reference.

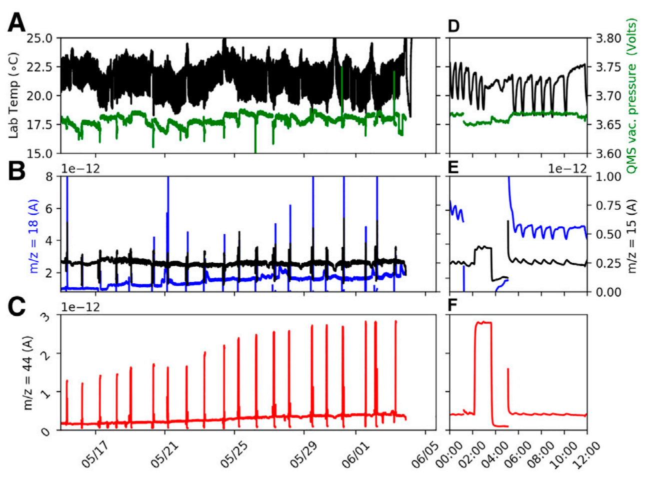


FIGURE 4 | Time series of environmental correlates used to bias correct the pCO_2 signal measured by shipboard QMS at m/z = 44 (panels **(C)** and **(F)**). Panels **(A)** and **(D)** show the lab temperature and QMS vacuum pressure; panels **(B)** and **(E)** show water vapor (m/z = 18) and m/z = 15. Panels **(D), (E),** and **(F)** show the time series for one 12 h period on June 1, 2017.

General Approach of Statistical Learning

The bias corrections that we evaluate here belong to a family of statistics called supervised learning. These corrections compare correlating inputs with corresponding outputs to develop a predictor that can be applied to any set of inputs. To develop the prediction, a sufficiently large dataset is divided into subsets—often referred to as "train" and "test" subsets (Ahmed et al., 2010). Separating in this manner allows the learning algorithm to develop a fit using the "train" dataset and evaluate the quality of that fit by predicting the data in the "test" dataset. The Scikit-learn module library in Python has been designed around the test-train convention and allows the user to subset using a number of different methods (Pedregosa et al., 2011). Last, the "test" dataset is used to estimate general error between the bias corrector and the actual data (Hastie et al., 2001).

Statistical learning models are exceedingly flexible and conform to almost any feature at any scale within a time series. This can result in "overfitting," a condition where the learning algorithm attempts to reproduce small scale noise or other shapes in the data that do not improve the prediction or bias correction. Overfitting results because of the imperfect separation between the bias and the random error. This imperfect separation between β and ε, called the bias-variance tradeoff (Wood, 2017), results in a degradation of the fit as greater degrees of freedom are introduced to the model. Statistical learning algorithms included penalty parameters that can be adjusted to iteratively reduce the degrees of freedom. When this is done iteratively, one can probe the range of model-data misfit and determine the point where improved fitting becomes overfitting and then choose penalties accordingly in a process called regularization (Hastie et al., 2001). We describe the application of penalty regularization to the GAM in *Implementation of the Generalized Additive Model and Backfit Algorithm* and to the LSTM in *Implementation of the Long Short-Term Memory Algorithm.*

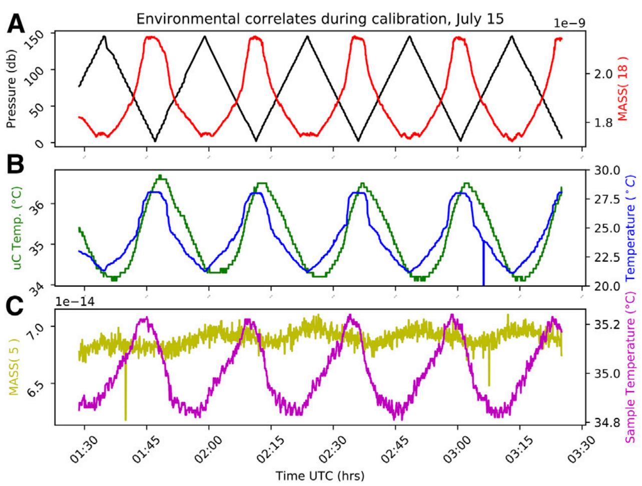


FIGURE 5 | The six environmental correlates that were measured by the SWIMS instrument or CTD to capture variations in the environment that are likely to influence the signal response (s) of the SWIMS for m/z = 32 and other dissolved gases. Panel **(A)** shows instrument water depth as hydrostatic pressure and the ion current for water vapor (mass 18). Panel **(B)** shows the temperature of the electronics (uC temp) and the water temperature. Panel **(C)** shows the sample temperature or temperature of the heater block where gas is introduced across the membrane, and it shows ion current at mass 5, the electronic noise baseline.

Implementation of the Generalized Additive Model and Backfit Algorithm

A GAM achieves smooth fitting by using the sum of fitting functions that individually represent the covariance between an individual input ($X = p_i, q_i, r_i$) and the response (y_i) data,

$$y_i = y_0 + f_1(p_i) + f_2(q_i) + f_3(r_i) + \varepsilon_i. \tag{6}$$

The choice for fitting functions (f_j) is flexible, although a typical choice is a natural cubic spline. Natural cubic splines are a collection of polynomials, with second derivative equal to zero at the endpoints or knots. By specifying more knots, the splines can represent a higher frequency fluctuations. The fit between y and $f_1(p)$ can be generated through any penalized linear least-squares algorithm,

$$\|\boldsymbol{y} - \mathrm{f_j}(\boldsymbol{x})\| + \lambda \int_0^1 \left[\mathrm{f_j}''(\boldsymbol{x})\right]^2 dx = 0. \tag{7}$$

The fit penalization, λ, is the primary means by which the solution is tuned. The fit between y and the sum of f_j's means that the influence of each f_j on the global solution can be observed, plotted, and evaluated. As mentioned, this is one of the principal strengths of the GAM, and it permits a more interactive and nuanced approach to determining the significance of each input variable and the behavior of each f_j.

We implemented the penalized least squares using the ridge regression algorithm in the Scikit-learn library with a specified value for penalization and normalization of all input variables;

>> model = Ridge(alpha = λ, normalize = true).

The natural cubic spline matrix with k = 9 knots was implemented using the Patsy module.

>> basis = dmatrix("cr(train, df=10)-1", {"train": X[j]}).

We incorporated this penalized regression into the global fit using the backfit algorithm (Wood, 2017), which permits an iterative approach to fitting where each environmental correlate, j, is fit against the partial residuals (e_p), or the difference between the signal response (s) and the spline fit to all inputs except X_j,

$$e_p^j = \hat{s} - \sum_{k \neq j} f_k(X_k). \tag{8}$$

Here, s has already been standardized or normalized to have zero mean. The backfit algorithm described by Wood (2017) has been reproduced here for clarity. The Python code can be found in the **Supplementary Material**.

(1) Standardize or remove the mean from s: $\hat{s} = s - \bar{s}$
(2) Set the initial spline functions to zero: $\mathbf{f}_j = 0$
(3) Use linear regression to fit f_j to e_p: basis = model.fit(basis,e_p)
(4) Estimate y from f_j: $\hat{s} = \sum_j \hat{f}_j$, news = basis.predict(dmatrix ("cr (valid, df = 10)-1", {"valid": X[j]})))
(5) Recompute e_p: $e_p^j = \hat{s} - \sum_{k \neq j} \hat{f}_k(x_k)$
(6) Repeat steps 3 thru 5 until e_p stops changing

More complex examples, involving other link functions between y and f_j and the imposition of different probability distributions on y_i (e.g., Gamma, Poisson or exponential), are all treated in more detail in Hastie et al. (2001).

To determine the optimal fit, we iteratively apply the backfit algorithm to the training data subset and then compute the generalized cross validation (GCV), as it varies with λ, the penalization parameter,

$$V(\lambda) = \frac{\frac{1}{n}\|I - \boldsymbol{A}(\lambda)y\|^2}{\left[\frac{1}{n} tr(I - \boldsymbol{A}(\lambda))\right]^2}. \tag{9}$$

In **Eq. 9**, n is the number of records of instrument signal response, I is the identity matrix, and A is the "influence" matrix, reflecting the penalized linear least-squares solution that can be applied as a step during the GAM fit (Golub et al., 1979),

$$\boldsymbol{A}(\lambda) = \boldsymbol{X}\left(\boldsymbol{X}^T\boldsymbol{X} + n\lambda I\right)^{-1}\boldsymbol{X}^T. \tag{10}$$

The GCV approach is to look for the minimum in $V(\lambda)$ to determine the most appropriate regularization penalty and strike the best balance between fit complexity and overfit. The GCV metric is better suited for this task than seeking the minimum residual sum of squares because that value decreases continuously with n and with the magnitude of λ.

The GCV score can be computed directly using **Eq. 9**. It is also computed and can be output by the Scikit-learn regression() toolbox. We used ridge regression, and the GCV score is output as

>> model = Ridge (alphas = λ, store_cv_values = True).fit(X_train, s_train).

>> gcv = model.score (X_test, s_test).

Because the components of the GAM model are separable, it is also possible to determine which environmental correlates contribute most to the best-fit solution. This avoids the inclusion of correlates that make no contribution or may even degrade the GAM solution. The Bayesian information criterion (BIC) considers the model fit quality but also penalizes for models of increasing complexity (Burnham and Anderson, 2004), providing a measure for each correlate's contribution to the GAM solution,

$$BIC = n\log_e(RSS/n) + k\log_e(n). \tag{11}$$

This version of the BIC applies when using a maximum likelihood estimator (such as ridge regression). The term k is the number of parameters included in the model. In this case, k is equivalent to the number of environmental correlates. The absolute value of BIC is not important; rather, the goal is to seek a minimum in BIC, which indicates the model best fit with the fewest parameters. For this task,

$$\Delta BIC_i = BIC_i - \min(\text{BIC}) \tag{12}$$

will achieve a value of zero when the best set of environmental correlates have been used. The ΔBIC is further useful as it allows the user to determine if certain environmental correlates degrade the overall solution or make no contribution (**Figure 6**).

Implementation of the Long Short-Term Memory Algorithm

Recurrent neural networks (RNN) can be used to interpret sequential data, like time series, where each data record may be related to the records that preceded it. The neural network uses functional dependencies along a network of nodes, and the influence of these dependencies is weighted based on their relative importance. The RNN keeps track of these network weights as a means to archive predictive information as memory (Brownlee, 2019b). Since their development, RNNs sometimes have difficulty converging to a solution when attempting to optimize weights at all the nodes. This problem was solved by the LSTM algorithm (Hochreiter and Schmidhuber, 1997) that discards or "forgets" weight information that is not pertinent to the solution. The documentation of RNN theory, concepts, and implementation is very extensive, rapidly evolving, and available in the public domain, so we will move straight to a discussion of the implementation for instrument bias correction. We used the Keras API (Chollet, 2018) which serves as an interface to the TensorFlow toolbox to develop, train, and implement the LSTM network.

Taxonomy of the Time Series Forecast

Because there are so many types of problems that can be solved using neural networks, it is helpful to list out the characteristics of this particular time series solution because this affects the structure of the neural network (Brownlee, 2019b). In our case, we are determining a single output from multivariate inputs; the neural network is a regression, rather than

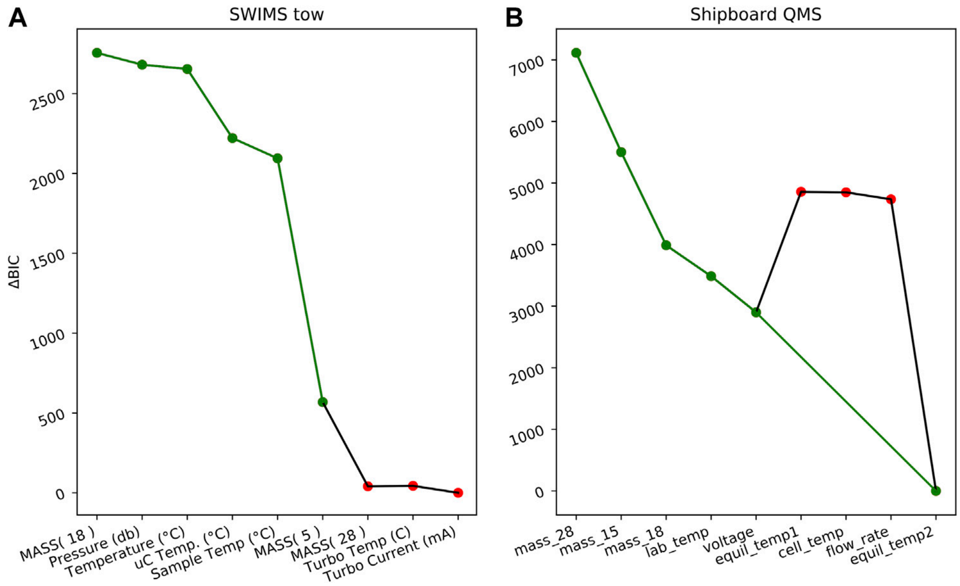


FIGURE 6 | The normalized Bayesian information criterion (ΔBIC) which was used to determine the set of the environmental correlates that best reproduce the bias in the GAM model. The black line with red dots indicates environmental correlates that did not measurably improve the ΔBIC score. These were left out of the final set of environmental correlates, which are indicated by the green line and green dots.

classification; we seek a multitime step output to be able to predict over an unspecified range of time, and the current solution is static because it has been trained using *in situ* calibration data and does not update the solution over time. The exogenous inputs are water temperature and water hydrostatic pressure that the SWIMS experiences. The endogenous inputs, which are coinfluenced by the environment are water vapor inside the SWIMS detector, measured at m/z = 18, the sample temperature, the circuit board temperature, and the mass spectrometer background noise measured at m/z = 5 (**Figure 4**).

Instrument bias correction can be thought of as time series prediction. Even though our approach is to use a multivariate set of inputs to help develop the bias prediction, the potential for long-term transients in the instrument signal encourage the interpretation of bias correction as a sequential and time-dependent statistical problem. Examples of instrumental memory can include, e.g., the silicon membrane stiffening (*In Situ Calibration of the Submersible Wet Inlet Mass Spectrometer*) or the thermal inertia, a pressure casing that may dampen the heat transfer between the environment and electronics inside the housing.

We use the Keras sequential() model. The 2D environmental array X of n data records through time by k input parameters (e.g., temperature and pressure) must be reshaped into a 3D array or tensor. The n data records in time are decomposed into p sequences of t time steps: $n = p \times t$ (Stevens and Antiga, 2019). Tensor creation provides the RNN with multiple time series realizations against which to train and develop network weights. The fundamental choice for the user is to decide how many t time steps to include in each sequence. If data are periodic, it may be instructive to break the data into lengths that roughly capture an interval of the period. For example, two years of solar radiation data or sea level data measured every 10 min may be naturally broken into $t = 144$ or $t = 36$ time steps corresponding to the 1-day or one half tidal period. However, this choice is rarely carried out a priori and must be determined iteratively.

After the $p \times t \times k$ tensor dimensions have been established, the user must choose a functional relationship or "activation function" between input and response at each network node, the number of iterations or "epochs" over which the RNN algorithm will train, and the number of "neurons" and the "optimizer" or metric that is used to evaluate the goodness of fit. As with the time steps, the settings for these parameters cannot be determined a priori, so we establish appropriate values through iteration (Brownlee, 2019b).

Keras allows a user to take control of when the RNN weights are updated; this is known as controlling the model state or "stateful = True." By default, Keras updates the LSTM state after a "batch" is processed. A batch is a collection of sample sequences, where each sample sequence has t timesteps, as we defined above. A batch size of one causes the model weights to be updated after each sample, but the penalty in processing speed and computation often requires a large batch size. Ideally, the batch size is a factor of p, the number of sample sequences; otherwise, a set of left over sequences are processed in an additional step (Brownlee, 2019b).

Determining Fit Quality

During tuning and iteration of the GAM model, we used GCV(λ) to test for overfitting and the root mean square error (RMSE), which is a measure of the deviation between modeled bias $\widehat{\beta}(X)$ and instrument bias, using the train datasets. We also evaluated the neural network LSTM model using the RMSE between $\widehat{\beta}(X)$ and the instrument bias, measured during *in situ* calibrations.

To evaluate the overall fit quality, we measured the RMSE between the independent O_2 and CO_2 instruments (y_{ind}), and the bias-corrected signal from the QMS and SWIMS instruments as defined by **Eq. 5**, ($y(\vec{x},t)$):

$$RMSE = \sqrt{\frac{1}{n}\left(y\left(\vec{x},t\right) - y_{ind}\right)^2}. \tag{13}$$

RESULTS AND DISCUSSION

The bias correction workflow is depicted in **Figures 1 and 7**; the calibrated GAM solution is shown in **Figure 7** panels **b** through **d**, but the steps are essentially the same for the LSTM solution. In this section, we present the details of the GAM and LSTM fits and contrast the two bias correction models.

Generalized Additive Model Fit

The ability to choose a functional form for each X_j environmental correlate was an attractive feature of the GAM because early tests revealed that oxygen (m/z = 32) strongly correlated with water vapor (m/z = 18), and signal from the SWIMS showed m/z = 18 ion currents outside the range observed during *in situ* calibration. Consequently, it appeared necessary to have a linear or proportional correction to m/z = 18. Water is present in solution at nearly 1 mol/mol; so, its concentration far exceeds the other analytes. Somewhat counterintuitively, m/z = 18 correlated positively with m/z = 32, perhaps suggesting a similar response to membrane permeability rather than competition for ionization inside the SWIMS source (**Supplementary Figure S1**).

All the environmental correlates (**Figure 5**) negatively covaried with the water depth. More subtle features, such as lag between the circuit board temperature (uC temp) and the sample temperature can also be observed in the SWIMS electronics temperature (**Figure 5**, panel **b**). Using the flexibility of the GAM, we tested both linear and quadratic fits between m/z = 18 and the target output variable $[O_2]$ or m/z = 32. While these parameterizations showed a stiffer, more proportional response to the large-scale variations in m/z = 18; ultimately, the natural cubic spline produced the best RMSE solution.

Having chosen a cubic spline functional form for *f()* for each X_i, there remain only two additional parameters that can be used to tune the solution—the number of knots in each spline and the value of the penalty function, λ (**Eq. 9**). We tested the fit to *in situ* calibration data for a range from 3 to 30 knots and observed no significant change in fit quality above 10 knots, so all cubic spline fits used a total of 10 knots. The term GCV(λ) was computed iteratively over a range from $\lambda = 10^{-10}$ to $\lambda = 10^{10}$ (**Figure 8**); the minimum GCV(λ) suggests the region where fit complexity and minimization of bias are optimal (Wood, 2017). We found GCV(λ) was not sensitive to the penalty, outside the range $10^{-2} < \lambda < 10^5$, with a minimum near $\lambda = 10^5$, so this value of the penalty was implemented in the solution.

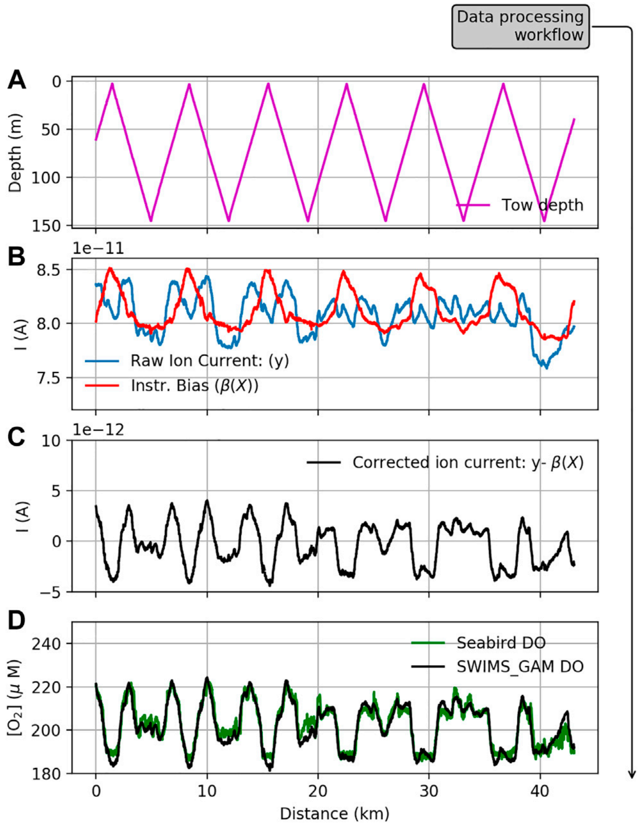


FIGURE 7 | Sequence showing the SWIMS tow and bias correction using the generalized additive model (GAM). Panel **(A)** reveals the depth recorded by the Triaxus CTD during vertical tows; panel **(B)** shows the raw signal recorded for oxygen at m/z = 32 and the estimate of instrumental bias. Panel **(C)** shows the corrected ion current, and panel **(D)** shows the calibrated ion current in O_2 concentration units, alongside the Seabird DO sensor.

Long Short-Term Memory Fit

As noted, the Keras LSTM algorithm requires iteration to choose appropriate values for the *t* time steps in each sample, the batch size, and epochs, as well as the choice for how often to update the weights of the RNN or statefulness. We chose to optimize based on the RMSE and used a hyperbolic tangent activation function. We found the LSTM solution was most sensitive to batch size and the number of epochs, especially as they related to overfitting. To mitigate overfitting, we implemented node dropout regularization using the Keras dropout() attribute. The approach is to assign a dropout likelihood between 0 and 1, wherein the model will randomly remove some nodes during training,

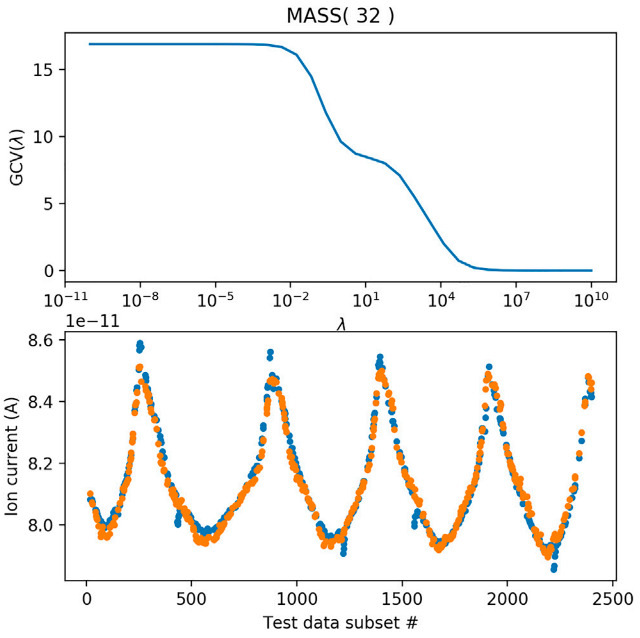


FIGURE 8 | Test of the GAM solution using a range of values for the penalty term λ in **Eq. 9**. Orange circles represent the GAM reconstruction and blue circles represent the raw ion current during in-situ calibration.

thereby reducing codependence and overweighting of certain nodes (Srivastava et al., 2014).

Because the choice of batch_size, epoch number, and dropout regularization cannot be determined a priori, but have a preponderant influence on overfitting, we objectively determined the optimal values for these three hyper parameters using the GridSearchCV() algorithm in Keras. The approach tries all permutations of the hyper parameters and measures the fit quality using the RMSE and a *k*-fold cross validation (with k = 5). The *k*-fold cross validation randomly samples the training data to produce test data subsets, which are then used to measure fit quality *k* times. We tested batch_sizes ranging from 20 to 80, epoch numbers ranging from 5 to 30, and dropout likelihood ranging from 0.1 to 0.8. The smallest *k*-fold RMSE value was found at with a batch_size = 80, epochs = 20, and dropouts = 0.4. The residual error between the training data and the LSTM solution, "train RMSE" in **Figure 9**, reveals a continual reduction in both test and train RMSE through epoch = 20. Beyond epoch = 20, the test RMSE increases, suggesting an overfit (**Figure 9**).

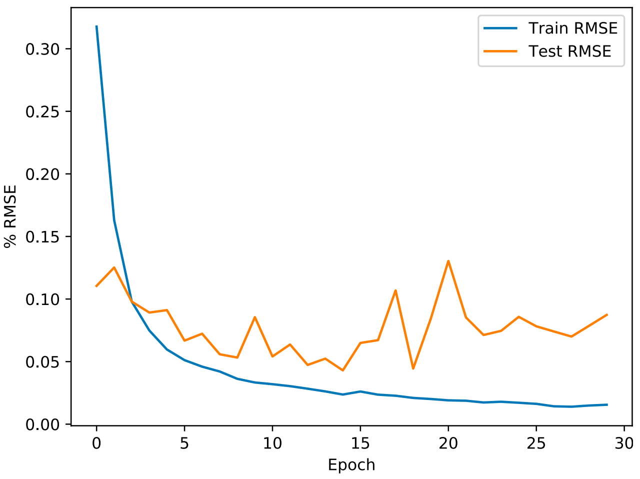


FIGURE 9 | Root mean squared deviation between the train and test subsets during successive training epochs. While the training RMSE continually decreases, suggesting improvement, the test RMSE begins to increase after 20 epochs, suggesting that the solution is being overfit.

Finally, the choice of *t* timesteps in each sample can be an important consideration. Because time series may have quasiperiodic correlations, it is desirable to have *t* be large enough to capture the full period, in order to make future predictions based on past time series behavior (Brownlee, 2019a). The Triaxus tow vehicle was programmed to ascend and descend at 0.2 m/s, so a full tow from surface to 150 m and back to the surface took approximately 25 min or *t* = 750 time steps at data reported every 2 s, which is the scan rate of the SWIMS. Initially, we anticipated that a sample size of *t* > 750 would provide the best fit. However, splitting the *in situ* calibration data into a test and train subset did not permit the inclusion of sample sizes of *t* = 750 because we felt it is necessary to validate against a test dataset that was at least 2 *t* in length. In practice, we tested values of *t* = 50, 100, 200, and 300. The test RMSE actually improved significantly as *t* was reduced. Eventually, *t* = 100 provided both computational efficiency and low RMSE, even though this number of time steps does not encompass the full profile tow. The tow profile may not be as necessary, suggesting that the information used to reconstruct the bias comes from the environmental correlates that are available at the prediction timestep, rather than from the learned temporal dependence.

GAM vs. LSTM Bias Correction, SWIMS Tow

Normally, the procedure to evaluate a statistical learning algorithm involves validating the solution against the test data (*General Approach of Statistical Learning*), which was set aside before the training stage. However, the independent measurement of oxygen by the SBE43 (*Ocean Data Used to Evaluate the Bias Correction Models*) provides an opportunity to quantify the bias correction against an entirely unique measure of oxygen. It should be noted that the SBE43 probe can also be subject to its own sources of bias, some of which may not be accounted for, but this instrument has a long performance history in oceanography (e.g., Helm et al., 2011) that supports the choice to use it as a reference instrument.

The final list of environmental correlates was determined using the ΔBIC metric (**Eq. 12**). In addition to water vapor, we tested for environmental covariation in the water pressure, seawater temperature, the sample temperature inside the SWIMS heater block, the circuit board temperature, the temperature of the turbo pump, current draw of the turbo pump, and the duty cycle of the membrane heater. Using ΔBIC, it was determined that these last parameters did not add any meaningful additional constraints beyond what the first six environmental correlates. That is, ΔBIC achieved a minimum after including water vapor, water pressure, circuit board temperature, sample temperature, and instrument noise at m/z = 5 (**Figure 6**). The remaining correlates were eliminated from the GAM solution.

The SWIMS tow between 35° and 40° N recorded a total of N = 49,181 individual measurements of dissolved O_2. A contour plot of dissolved O_2 reveals the tracer field in **Figure 10**. The RMSE between SBE43 and bias-corrected SWIMS data using the GAM was 11.2 µM (micromoles per liter of seawater); the units of RMSE are the same as the concentration data itself. The mean $[O_2]$ in this section was 196 µM, suggesting a 5.7% deviation between the two instruments. Within the same section, the neural network LSTM bias correction yielded RMSE = 9.8 µM or 5.0% deviation overall. Both GAM and LSTM bias corrections tended to fit some regions better than others; however, the fit quality of the GAM and fit quality of the LSTM did not degrade in the same places, suggesting some differences in how the two models respond to the environmental correlates (**Figure 5**).

It should be noted that we are focusing on interpretation of the relative RMSE between the GAM and LSTM solutions. The absolute value of the RMSE is less meaningful because the calibration intercept (s^0) was not measured on the SWIMS *in-situ* calibrations. This term, s^0, represents the instrument baseline drift, and so, we determined s^0 by optimal fit to the SBE43. The same baseline drift can be determined by fitting to another independent reference, such as the equilibrium oxygen solubility (Garcia and Gordon, 1992). When we use equilibrium solubility, the shape or trend in the daily estimates of s^0 remains the same, but the magnitude of s^0 shifts, causing a larger misfit between the SBE43 and the SWIMS. During future *in situ* calibrations, we think it is possible to implement a workable measure of s^0 by shutting off the water pump, causing all the gas around the silicon membrane to be depleted and achieving a practical value of zero concentration for all gases except water vapor.

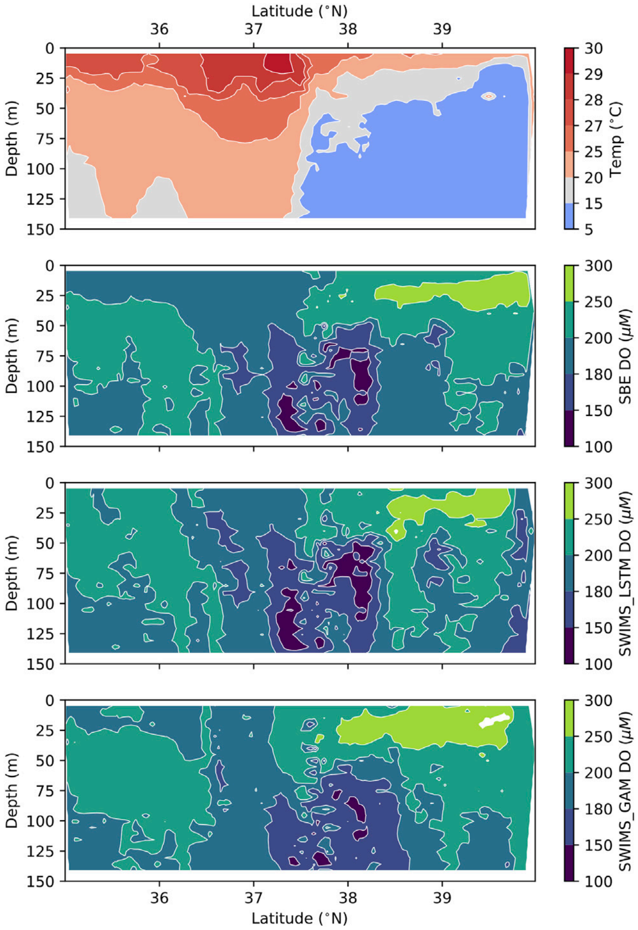


FIGURE 10 | North Atlantic section, including Gulf Stream and Labrador waters showing temperature **(top)** and oxygen from the Seabird SBE 43 membrane and the SWIMS with bias corrections using GAMs and the neural network LSTM model.

GAM vs. LSTM Bias Correction, Shipboard Quadrupole Mass Spectrometers

The bias corrections in the shipboard QMS were fit using training data over a four-day period of the surface ocean equilibrator time series from May 27 to June 1. The RMSE between the GAM solution and training data subset was 3.5%, and the LSTM misfit was 1.8%. Unlike the SWIMS tows, it was not possible to evaluate $\beta(X)$ independent of the environmental signal $y(\vec{x}, t)$. The daily calibrations with reference gases did not take place for long enough to properly observe and decompose the time series aliasing. Instead, it was necessary to train the LSTM and GAM models on a section of the real-time series. This approach can lead to muddling the separation between $y(\vec{x}, t)$ and $\beta(X)$, potentially correcting away some of the environmental signals in pCO_2 during the bias correction. However, the ambient changes in pCO_2 should reflect the biology and chemistry which in turn are only partly dependent on the exogenous environmental correlates. The endogenous environmental correlates reflect instrument behavior, which should have zero correlation with environmental pCO_2. The environmental correlates used to develop the bias correction model included, 1) temperature of the lab where the QMS was installed, 2) the total gas pressure in the QMS measured as voltage, 3) the seawater flow rate through the turbulent equilibrator, 4) water vapor measured at m/z = 18, and 5) m/z = 15. Similar to the SWIMS tow, we found that three environmental correlates caused an increase (no decrease) in ΔBIC metric, signaling that they contributed no meaningful constraint. Consequently, the IR pCO_2 cell temperature, the water wall flow rate, and the second equilibrator temperature reading were eliminated from the bias correction solutions (**Figure 6**).

After bias correction, the raw ion current was calibrated to CO_2 partial pressure, using the three-point calibration of reference standards that were measured daily. There are additional corrections to gas measurements that are made using a turbulent equilibrator, and these are described by Takahashi et al. (2009). These corrections have not been implemented here; while their implementation might improve the overall misfit between the two measurements of pCO_2, they would drop out of the comparison between GAM and LSTM bias corrections; so, these additional data corrections are not material to this evaluation.

In this case, the GAM model was better at removing the periodic oscillation in the QMS ion current at m/z = 44 (). However, a level of noise persists even after the bias correction, suggesting that the environmental correlates may be missing some components of the bias. In total, the 18-day time series contains 5043 unique measurements of pCO_2 by infrared absorption spectroscopy and by QMS. The RMSE between the IR pCO_2 and GAM-corrected pCO_2 was 31.3 μatm; the average pCO_2 was 411 μatm, revealing an overall misfit of 7.5% (**Figure 11**). The LSTM RMSE was 35.2 μatm or 8.5% of the mean pCO_2. In this case, it appears the LSTM (not pictured) may have slightly overfit the training data, resulting in a degraded fit to the overall time series. Nevertheless, the difference in RMSE between GAM and LSTM was less than 1%, which suggests that both methods produce very similar overall bias correction outcomes.

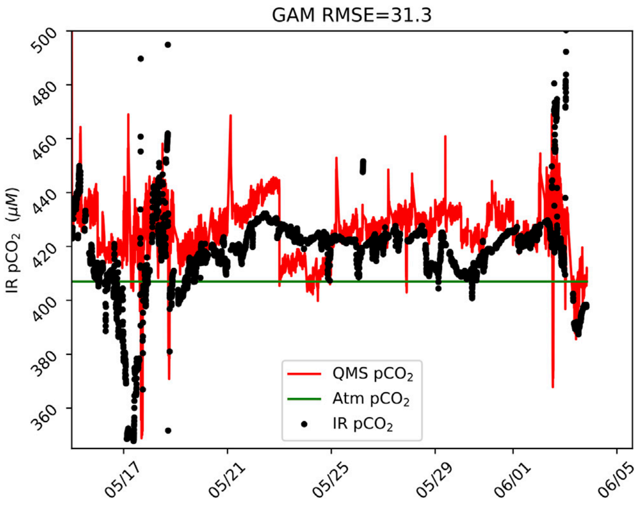


FIGURE 11 | Bias-corrected and calibrated pCO_2 from shipboard QMS alongside measurements of pCO_2 by infrared absorption spectroscopy (IR pCO_2) in the Ross Sea, 2017.

SUMMARY

This study presents two models for instrument bias correction, a GAM and a LSTM neural network model. The two models represent philosophically different approaches to the multivariate prediction; the GAM allows the user to investigate the intermediate model fit products and choose the functional form f() for optimal regression between the results and the individual environmental correlates in *X*. This advantage was particularly useful when interrogating which environmental influences to include as correlates in the model solution, using the *BIC* criterion. This calculation is straightforward and can be determined offline without iteration of the GAM model, precisely because the solution is separable. The procedure eliminated three environmental correlates from both Shipboard and SWIMS ocean datasets (**Figure 6**). The six remaining correlates were also used to fit the LSTM solution.

The LSTM RNN model gives the user fewer intermediate diagnostics, which produces an initial lack of confidence in the robustness of the solution because it can be challenging to understand or visualize the nature of the solution. Nevertheless, there is an emerging recognition that, compared to the human brain, computers are much more capable instruments at assigning appropriate weights to an *n*-dimensional set of variates in pursuit of a solution. By accepting these models, we implicitly acknowledge that the multivariate weights in the solution are beyond our capacity to evaluate simultaneously, thus rendering the "black box" criticism somewhat moot. However, the procedures for implementing RNNs, including the grid search or random search (B. Nakisa et al., 2018; Bergstra and Bengio, 2012), provide a systematic approach to determining the optimal tuning of hyperparameters (e.g., times steps, batch size, epochs, and hidden nodes), and the eventual robustness of the solution has held up under rigorous testing and comparison. In the SWIMS dataset with *in situ* calibration, the LSTM solution proved more effective at removing bias in the high gradient oceanic region, with tows across the Gulf Stream. However, the GAM exhibited better fit quality in the Antarctic shipboard QMS dataset, as compared to the LSTM.

The difference between GAM and LSTM RMSE was 1% or less for both ocean sections, suggesting that both models performed similarly well. The RMSE for both methods were better than 6% for O_2 and less than 9% for CO_2, demonstrating a predictive accuracy of better than 91% for both dissolved gases. The quality of the bias removal solution was significantly more dependent on the availability of coincidently sampled environmental correlates as inputs. We further found that the *in situ* calibration for SWIMS data was a significant factor in producing a high fidelity bias correction. Several attempts were made to produce the same bias correction using just SWIMS tow data (without the *in situ* calibration) as training data, and the solution was significantly diminished with an RMSE for the LSTM model of 17% as compared to 5% with the *in situ* calibration. These results demonstrate that the bias corrections are most effective when they can be tuned using the *in situ* calibration with an invariant reference gas to reveal the instrument bias.

The overall performance of the GAM and LSTM models was highly comparable, making it difficult to declare a clear winner in this case. The primary advantage conferred by the GAM model is the ability to evaluate the fit to each individual correlate, separately. This is a big advantage when it is necessary to better understand an instruments behavior and might even lead to engineering solutions that eliminate the biggest source of bias. In comparison, the skill that an LSTM RNN brings to time series prediction can potentially serve to model longer-term transients in the signal, which could lead to a better bias model when few or no environmental correlates been measured.

DATA AVAILABILITY STATEMENT

The entire workflow including code, writeup, and source data with BCO-DMO DOI can be found at https://github.com/bloose/bias_correction_by_ML/.

AUTHOR CONTRIBUTIONS

BL developed and tested the bias correction methods described here, and lead the field data collection program. RT Short co-developed the SWIMS instrument and in-situ calibration system, and participated in the field data collection. ST participated in the field data collection.

FUNDING

This work was supported by a grant from the National Science Foundation, Award # 1429940.

ACKNOWLEDGMENTS

This research was supported by an award from the National Science Foundation Chemical and Biological Oceanography Program #1429940. We thank two anonymous reviewers for the comments and suggestions that have improved this manuscript. The GAM backfit algorithm is available at https://github.com/bloose/Python_GAM_Backfit. The supplemental contains annotated Python scripts and SWIMS example data to demonstrate application of the GAM and LSTM to bias correction.

SUPPLEMENTARY MATERIALS

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.537028/full#supplementary-material

REFERENCES

Ahmed, N. K., Atiya, A. F., Gayar, N. E., and El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econom. Rev.* 29, 594–621. doi:10.1080/07474938.2010.481556

Bell, R. J., Short, R. T., van Amerom, F. H. W., and Byrne, R. H. (2007). Calibration of an in situ membrane inlet mass spectrometer for measurements of dissolved gases and volatile organics in seawater. *Environ. Sci. Technol.* 41, 8123–8128. doi:10.1021/es070905d

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305.

Brownlee, J. (2019a). Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python (Jason Brownlee).

Brownlee, J. (2019b). *Long short-term memory networks with Python*. Australia: Jason Brownlee.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference. *Socio. Methods Res.* 33, 261–304. doi:10.1177/0049124104268644

Cassar, N., Barnett, B. A., Bender, M. L., Kaiser, J., Hamme, R. C., and Tilbrook, B. (2009). Continuous high-frequency dissolved O2/Ar measurements by equilibrator inlet mass spectrometry. *Anal. Chem.* 81, 1855–1864. doi:10.1021/ac802300u

Chapman, D. C., and Beardsley, R. C. (1989). On the origin of shelf water in the middle atlantic bight. *J. Phys. Oceanogr.* 19, 384–391. doi:10.1175/1520-0485(1989)019<0384:otoosw>2.0.co;2

Chollet, F. (2018). *Deep learning with Python*. Shelter Island, NY: Manning Publications.

Dawson, P. H., and Herzog, R. F. (1995). *Quadrupole mass spectrometry and its applications*. New York, NY: American Institute of Physics.

Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., and Stull, R. (2006). Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. *J. Geophys. Res. Atmospheres* 111, D05308 doi:10.1029/2005jd006311

Futó, I., and Degn, H. (1994). Effect of sample pressure on membrane inlet mass spectrometry. *Anal. Chim. Acta* 294, 177–184. doi:10.1016/0003-2670(94)80192-4

Garcia, H. E., and Gordon, L. I. (1992). Oxygen solubility in seawater: better fitting equations. *Limnol. Oceanogr.* 37, 1307–1312. doi:10.4319/lo.1992.37.6.1307

Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* 21, 215–223. doi:10.1080/00401706.1979.10489751

Guegen, C., and Tortell, P. D. (2008). High-resolution measurement of Southern Ocean CO2 and O2/Ar by membrane inlet mass spectrometry. *Mar. Chem.* 108, 184–194. doi:10.1016%2Fj.marchem.2007.11.007

Hastie, T., Tibshirani, T., and Friedman, J. (2001). *The elements of statistical learning: data mining, Inference, and prediction*. Berlin: Springer.

Helm, K. P., Bindoff, N. L., and Church, J. A. (2011). Observed decreases in oxygen content of the global ocean. *Geophys. Res. Lett.* 38, 23602. doi:10.1029/2011gl049513

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735

Jenkins, W. J. (1982). Oxygen utilization rates in North Atlantic subtropical gyre and primary production in oligotrophic systems. *Nature* 300, 246–248. doi:10.1038/300246a0

Lee, W. S., Yeo, K. S., Andriyana, A., Shee, Y. G., and Mahamd Adikan, F. R. (2016). Effect of cyclic compression and curing agent concentration on the stabilization of mechanical properties of PDMS elastomer. *Mater. Des.* 96, 470–475. doi:10.1016/j.matdes.2016.02.049

Nakisa, B., Rastgoo, M. N., Rakotonirainy, A., Maire, F., and Chandran, V. (2018). Long Short term memory hyperparameter optimization for a neural network based emotion recognition framework. *IEEE Access* 6, 49325–49338. doi:10.1109/access.2018.2868361

Newman, M. C. (1993). Regression analysis of log-transformed data: statistical bias and its correction. *Environ. Toxicol. Chem.* 12, 1129–1133. doi:10.1002/etc.5620120618

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pimentel, K. D. (1975). *Toward a mathematical theory of environmental monitoring: the infrequent sampling problem*. Dissertation. CA, United States: California University.

Saltzman, E. S., De Bruyn, W. J., Lawler, M. J., Marandino, C. A., and McCormick, C. A. (2009). A chemical ionization mass spectrometer for continuous underway shipboard analysis of dimethylsulfide in near-surface seawater. *Ocean Sci.* 5, 537–546. doi:10.5194/os-5-537-2009

Short, R. T., Fries, D. P., Kerr, M. L., Lembke, C. E., Toler, S. K., Wenner, P. G., et al. (2001). Underwater mass spectrometers for *in situ* chemical analysis of the hydrosphere. *J. Am. Soc. Mass Spectrom.* 12, 676–682. doi:10.1016/s1044-0305(01)00246-x

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.

Stevens, E., and Antiga, L. (2019). *Deep learning with PyTorch*. Shelter Island, NY: Manning Publications.

Sun, F., Xiong, R., and He, H. (2016). A systematic state-of-charge estimation framework for multi-cell battery pack in electric vehicles using bias correction technique. *Appl. Energy* 162, 1399–1409. doi:10.1016/j.apenergy.2014.12.021

Takahashi, T. (1961). Carbon dioxide in the atmosphere and in Atlantic Ocean water. *J. Geophys. Res.* 66, 477–494. doi:10.1029/jz066i002p00477

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., et al. (2002). Global sea-air CO2 flux based on climatological surface ocean pCO2, and seasonal biological and temperature effects. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 49, 1601–1622. doi:10.1016/s0967-0645(02)00003-6

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO2, and net sea-air CO2 flux over the global oceans. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 56, 554–577. doi:10.1016/j.dsr2.2008.12.009

Wenner, P. G., Bell, R. J., van Amerom, F. H. W., Toler, S. K., Edkins, J. E., Hall, M. L., et al. (2004). Environmental chemical mapping using an underwater mass spectrometer. *TrAC. Trends Anal. Chem.* 23, 288–295. doi:10.1016/s0165-9936(04)00404-2

Wood, S. (2017). *Generalized additive models: an introduction with R*. United States: CRC Press.

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.

frontiers
in Earth Science

ORIGINAL RESEARCH
published: 21 January 2021
doi: 10.3389/feart.2020.577588

Time-Scales of Inter-Eruptive Volcano Uplift Signals: Three Sisters Volcanic Center, Oregon (United States)

Sara Rodríguez-Molina[1], *Pablo J. González*[2,3]*, *María Charco*[1], *Ana M. Negredo*[1,4] *and David A. Schmidt*[5]

[1]Institute of Geosciences (IGEO, CSIC, UCM), Madrid, Spain, [2]COMET, Department of Earth, Ocean and Ecological Sciences, School of Environmental Sciences, University of Liverpool, Liverpool, United Kingdom, [3]Volcanology Research Group, Department of Life and Earth Sciences, Institute of Natural Products and Agrobiology (IPNA-CSIC), La Laguna, Spain, [4]Department of Earth Physics and Astrophysics, Faculty of Physical Sciences, Complutense University of Madrid, Madrid, Spain, [5]Department of Earth and Space Sciences, University of Washington, Seattle, WA, United States

A classical inflation-eruption-deflation cycle of a volcano is useful to conceptualize the time-evolving deformation of volcanic systems. Such a model predicts accelerated uplift during pre-eruptive periods, followed by subsidence during the co-eruptive stage. Some volcanoes show puzzling persistent uplift signals with minor or no other geophysical or geochemical variations, which are difficult to interpret. Such temporal behaviors are usually observed in large calderas (e.g., Yellowstone, Long Valley, Campi Flegrei, Rabaul), but less commonly for stratovolcanoes. Volcano deformation needs to be better understood during inter-eruptive stages, to assess its value as a tool for forecasting eruptions and to understand the processes governing the unrest behavior. Here, we analyze inter-eruptive uplift signals at Three Sisters, a complex stratovolcano in Oregon (United States), which in recent decades shows persistent inter-eruptive uplift signals without associated eruptive activity. Using a Bayesian inversion method, we re-assessed the source characteristics (magmatic system geometry and location) and its uncertainties. Furthermore, we evaluate the most recent evolution of the surface deformation signals combining both GPS and InSAR data through a new non-subjective linear regularization inversion procedure to estimate the 26 years-long time-series. Our results constrain the onset of the Three Sisters volcano inflation to be between October 1998 and August 1999. In the absence of new magmatic inputs, we estimate a continuous uplift signal, at diminishing but detectable rates, to last for few decades. The observed uplift decay observed at Three Sisters is consistent with a viscoelastic response of the crust, with viscosity of ~10^{18} Pa s around a magmatic source with a pressure change which increases in finite time to a constant value. Finally, we compare Three Sisters volcano time series with historical uplift at different volcanic systems. Proper modeling of scaled inflation time series indicates a unique and well-defined exponential decay in temporal behavior. Such evidence supports that this common temporal evolution of uplift rates could be a potential indicator of a rather reduced set of physical processes behind inter-eruptive uplift signals.

Keywords: inter-eruptive deformation, characteristic relaxation time, continuous GPS, interferometric synthetic aperture radar, geodetic time series, Three Sisters volcano

Edited by:
Flavio Cannavo',
National Institute of Geophysics and Volcanology,
Section of Catania, Italy

Reviewed by:
Maurizio Battaglia,
United States Geological Survey (USGS), United States
Gilda Maria Currenti,
Istituto Nazionale di Geofisica e Vulcanologia (INGV), Italy

***Correspondence:**
Pablo J. González
pjgonzal@liverpool.ac.uk

Specialty section:
This article was submitted to Volcanology,
a section of the journal Frontiers in Earth Science

Received: *29 June 2020*
Accepted: *23 November 2020*
Published: *21 January 2021*

Citation:
Rodríguez-Molina S, González PJ, Charco M, Negredo AM and Schmidt DA (2021) Time-Scales of Inter-Eruptive Volcano Uplift Signals: Three Sisters Volcanic Center, Oregon (United States). Front. Earth Sci. 8:577588. doi: 10.3389/feart.2020.577588

1 INTRODUCTION

Many volcanoes follow a common deformation pattern consisting of uplift during inter-eruptive periods and subsidence in co-eruptive stages, occasionally interrupted by periods of quiescence or subsidence. Some other volcanoes do not however exhibit this simple behavior (Biggs and Pritchard, 2017). Part of them show puzzling non-steady persistent uplift signals that can last from days to years with minor or no other geophysical or geochemical variations, which are difficult to interpret. Therefore, uplift during inter-eruptive episodes cannot be only interpreted as a pre-eruptive precursory indicator. Such temporal behavior is usually observed in large calderas (e.g., Yellowstone, Long Valley, Campi Flegrei, Rabaul), but less commonly for stratovolcanoes.

An important goal in volcano eruption forecasting is to find how the deformation time-series can distinguish among physical processes, especially during inter-eruptive periods leading to a pre-eruptive scenario. The latter are characterized by new injections of magma/increment of volatiles, viscoelastic relaxation of the media, or a mixing of different coeval processes. Therefore, we must constrain what controls usually long-lived or persistent uplift at volcanic centers. Le Mével et al. (2015) show that the temporal evolution of deformation surprisingly follows the same pattern for different volcanic systems at specific analyzed periods (Yellowstone, Long Valley, Laguna de Maule and Three Sisters). This is consistent with the hypothesis that similar processes may be at work in similar volcanoes. After this stage, these volcanoes presented an eventual pause and/or change to subsidence (related to seismic events and/or hydrothermal changes), but did not produce an eruption.

Stratovolcano behavior contrasts with better documented restless calderas (Acocella et al., 2015; Galetto et al., 2017). Calderas usually show long repose periods between large eruptions, but without quiescence (e.g., resurgent volcanism, subsidence, multiple pulses of uplift) (Biggs and Pritchard, 2017). The Three Sisters complex volcano is a good example of a system with long lasting monotonous inter-eruptive uplift without associated eruptive activity or significant seismicity and might correspond with different magmatic system behaviors.

In this work, we studied the decadal deformation time-series of the Three Sisters volcano in order to understand the processes governing its unrest behavior and to find out whether it is still inflating. For this purpose, we needed to perform a consistent analysis of the volume change time-series underlying the Three Sisters uplift signal, which is a challenging task. We analyze available continuous GPS data since 2001 and multiple satellite interferometric data spanning the 1993–2020 period. We proposed combining multiple geodetic data-sets using an improved linear regularization method based on Truncated Singular Value Decomposition (TSVD) (González et al., 2013) to find an optimal regularization criteria for Three Sister data-set combination. We obtained a seamless, continuous time series of volume change (and its uncertainties) with which to rigorously assess changes over the 26-years studied period. Finally, we compared the Three Sisters temporal behavior to other well-known examples of uplifting volcanoes to understand 1) whether a variety or not of physical mechanism are at work behind deformation and, if so, 2) if uplift time-scales are informative of whether a certain volcano is on a late or early stage of the inter-eruptive period.

2 BACKGROUND

The Three Sisters Volcanic Complex forms a N-S chain of stratovolcanoes located at 44.1°N in the central Oregon Cascades (**Figure 1**), an active volcanic arc produced by the subduction of the Juan de Fuca and Gorda plates beneath the North American plate. In addition to this convergent motion, there is an oblique relative plate motion and northward push of the Sierra Nevada-Great Basin microplate, favoring a N-S stress orientation of the vents within the Oregon Cascades (Mccaffrey et al., 2007). South Sisters is near the propagating tip of a crustal melting anomaly westward across Oregon, progressing since the mid-Miocene, going through the Cascades in the Quaternary (Fierstein et al., 2011). All these circumstances influence on the eruptive history of Three Sisters.

The Three Sister area includes shields, composite volcanoes and cinder cones, with basaltic to rhyolitic volcanism. The three eponymous volcanoes are progressively younger from north to south and exhibit little family resemblance (Hildreth et al., 2012). North Sister is a monotonously mafic edifice created 120 ka ago, formed by long-lived effusive volcanism (Schmidt and Grunder, 2011); Middle Sister is an andesite-basalt-dacite cone constructed between 48 ka and 14 ka and South Sister is a bimodal rhyolitic-intermediate edifice built between 50 ka and 2 ka, both with histories of explosive volcanism (Scott et al., 2001). However, most of the volcanic activity is identified with mafic shields and cones around the major composite volcanoes (Hildreth, 2007). Geochemical anomalies suggest that episodes of intrusion may be more frequent in the Three Sisters area than the age of eruptive vents would involve (Evans et al., 2004). The most recent eruptions were rhyolitic close to South Sisters, 2000 years ago (Hildreth et al., 2012). There is a potential volcanic hazards threat if future eruptions are similar to South Sister's recent past. Tephra fallout might accumulate to 1–2 cm thick in the Bend area, and small-volume lahars and pyroclastic flows could pose a hazard to nearby areas (Sherrod et al., 2004).

Before 2001, Three Sisters was considered a dormant volcano. Nonetheless, ERS-1/2 satellite interferometric synthetic aperture data (InSAR) analysis from 1992 to 2000 revealed active uplifting located 6 km west of South Sister (Wicks, 2002). Geodetic information for Three Sisters volcanic center has been accumulating since this discovery. Nowadays, deformation is continuously monitored through Leveling surveys since 2002, Continuous GPS (CGPS), GPS campaign since 2001 and Semi Permanent GPS (SPGPS) since 2009 (Dzurisin et al., 2009; Dzurisin et al., 2017). The uplift at Three Sisters has been aseismic except for a swarm of ~300 small earthquakes ($M_{\max} = 1.9$) in March 2004 (Moran, 2004; Dzurisin et al., 2009). Previous studies (Wicks, 2002; Dzurisin et al., 2006; Dzurisin et al., 2009; Riddick and Schmidt, 2011) show evidence that observed uplift

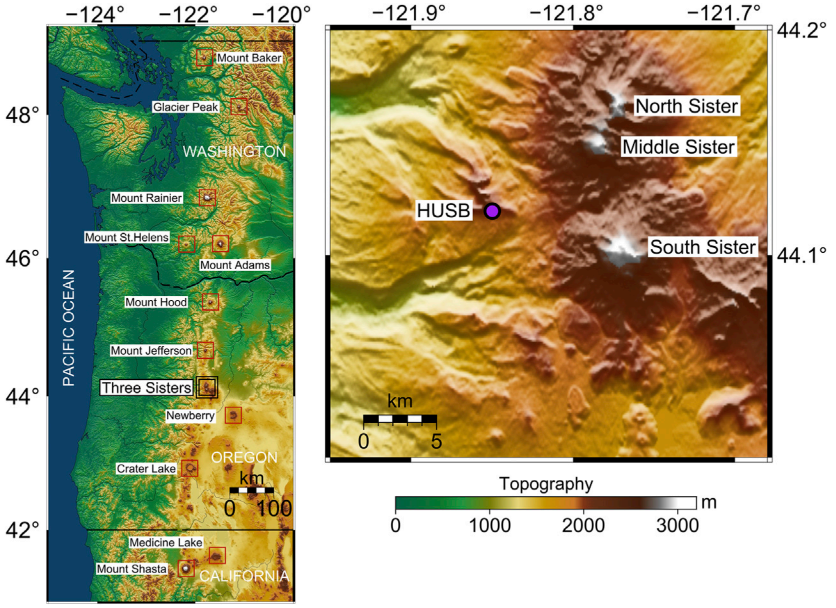


FIGURE 1 | (A) Shaded relief map for the Cascade range, with locations of representative Cascade volcanoes. Three sisters volcano is highlighted with black squares. **(B)** Regional map of Three Sisters volcano complex. The continuous GPS station HUSB, located at ~5 km west of South Sister volcano, is shown as a purple circle with black outline.

can be described by a spherical point source within an homogeneous isotropic elastic half-space. Nevertheless, deformation source geometry is non-unique and sources as horizontal crack, vertical prolate spheroid, and sill-like have been proposed at Three Sisters to fit geodetic data. Interpretation of the temporal evolution of InSAR, leveling and GPS data suggests the beginning of deformation in late 1997 or 1998, with a maximum uplifting rate of 3-5 cm/year during 1998–2001. Microgravity data collected between 2002 and 2009 show no significant change in the mass flux across the deforming area (Zurek et al., 2012). No studies have been published about the uplift evolution over the last decade. The uplift process was still on-going in January 2020, when this manuscript was prepared.

3 GEODETIC DATA PROCESSING

We aim to extend the detailed uplift history at Three Sisters already mentioned above to the present (2020) by using the available CGPS and InSAR data-sets.

3.1 Continuous GPS

In May 2001, the U.S Geological Survey (USGS), in collaboration with the U.S. Forest Service, installed a continuous GPS station (HUSB) near the actively deforming area. It was strategically installed at a location approximately ~2 km northwest of the detected uplift center. HUSB is part of the USGS Pacific Northwest Network, so it is automatically processed to obtain daily coordinates. No other regional and local continuous GPS station falls within the deformation area. Hence, the HUSB time-series is particularly important to understand the surface deformation time-scales at Three Sisters.

Daily GPS data (coordinates and their uncertainties) are analyzed by the USGS using GIPSY/OASIS II software. Common-mode daily biases are estimated and removed using QOCA (Dzurisin et al., 2009). Three Sisters is located near the actively deforming Cascadia margin, so any geodetic data and coordinates must consider the wider regional deformation patterns. The motion of a background steady rigid-body with a rotation pole situated near the eastern limit of Oregon must therefore be removed from the time-series, and a correction for predicted horizontal tectonic motion should be applied. Here, we remove a linear trend of 4.29 mm/yr for the North component and 1.50 mm/yr for the East component. This model prediction represents an update and improved version of the horizontal displacements at HUSB (Dzurisin et al., 2006; Dzurisin et al., 2009; M. Lisowsky personal communication; Cascades Volcano Observatory, 2017).

Figure 2 shows the resulting GPS displacements between July 2001 and January 2020. CGPS data reveals several gaps that occurred due to snowfall in the winter seasons. Furthermore, CGPS shows a gap and a posterior data offset during August 2017-August 2018. USGS data site reports some readjustment of HUSB permanent station during this period and these could explain some of the gaps and offsets in the time series.

3.2 Interferometric Synthetic Aperture Radar (InSAR)

Our InSAR data set includes 85 interferometric pairs, with temporal baselines from 35 to 2,894 days, from four satellite

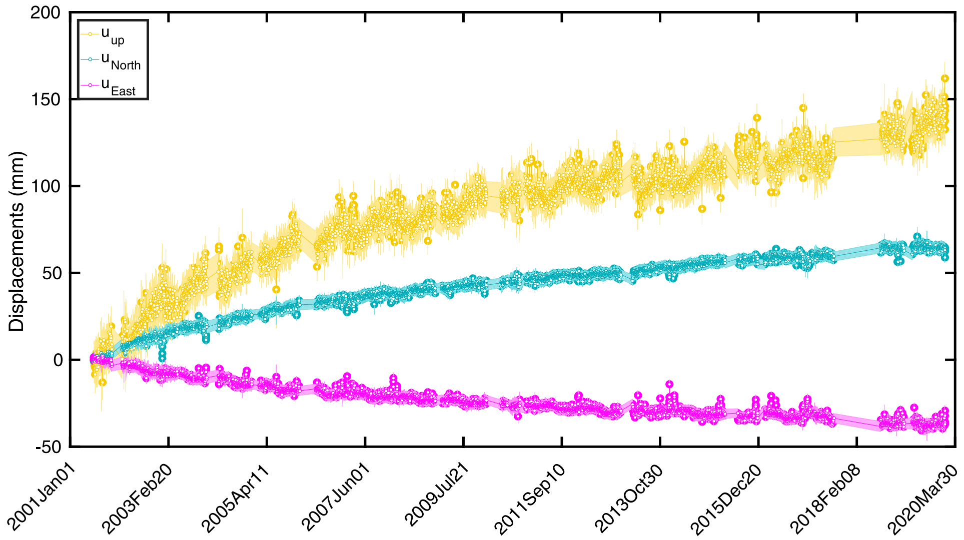


FIGURE 2 | North (U_{North}), east (U_{East}) and vertical (U_{up}) components of the continuous GPS displacements at HUSB.

missions (ERS, ENVISAT, ALOS-1 and Sentinel-1). ERS and ENVISAT SAR images were acquired during summer and fall between 1993 and 2010 (descending orbits, tracks 113 and 385; ERS and ENVISAT look angles 20.2° and 19.8°, respectively). We used 51 interferograms processed with the ROI PAC software (Rosen et al., 2004) and unwrapped using SNAPHU (Chen and Zebker, 2002), with perpendicular baselines up to 500 m, as explained by Riddick and Schmidt (2011).

To improve the temporal coverage of InSAR observations, we also analyzed data from the ALOS-1 and Sentinel-1 SAR data missions. The mean line-of-sight velocity of ALOS-1 data (path 219, ascending orbit, look angle 38.7°) was obtained during January 2007- March 2011. Most individual interferograms in the Cascades range show poor coherence because of vegetation and seasonal snow coverage, hence we also processed 4 Sentinel-1 summer-to-summer and summer-to-late spring interferometric pairs, between September 2015 and May 2018, for descending (path 115, look angle 39.8°) and ascending (path 137, look angle 38.8°) orbits. To provide deformation data during the GPS gap mentioned above, two Sentinel-1 interferometric pairs cover that period. We used JPL InSAR Scientific Computing Environment (ISCE) software (Rosen et al., 2012), processing Level-0 raw ERS-1 and ALOS-1, and SLC-level Sentinel-1 radar data. All interferograms were corrected for orbital and topographic contributions using precise orbit information and the SRTM digital elevation model (Farr et al., 2007). We also reprocessed a highly coherent interferometric pair for ERS-1 track 365 (descending orbit, corresponding to August 1997 - September 2000). This interferogram was essential to further re-evaluate the magmatic source location and constrain its uncertainties.

L-band data (wavelength ~24 cm) from ALOS-1 were very useful to avoid decorrelation owing to the vegetation encompassing the Three Sisters area. Although the LOS deformation rate from 2006 to 2010 is small (about 6–8 mm/yr)

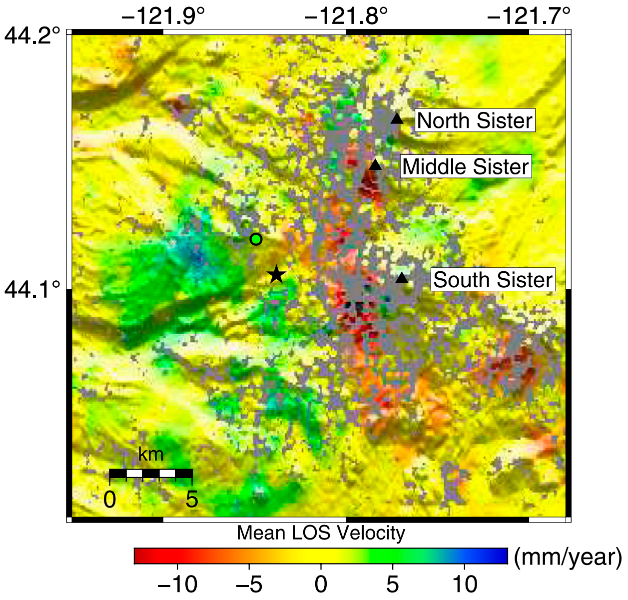


FIGURE 3 | The linear surface deformation LOS velocity (mm/year) obtained for the ascending Path 219 ALOS-1, using the small baselines method implemented in StaMPS (Hooper et al., 2012) for the period January 2007 to March 2011. Positive LOS velocity values corresponds to displacements toward the satellite, i. e., uplifting. Black triangles and star represent the Three Sister complex volcano system and the approximate center of the uplifting area. StaMPS LOS velocity results were noisy and we post-processed to reduce undesirable oscillations of non-volcanic origin. We applied a band pass filter to retain spatial deformation signals between 10 and 0.8 km using a median filter (GMT blockmedian). Results indicate a 6 km circular uplift pattern west of South Sister with a mean LOS velocity of approximately 5–10 mm/year, consistent with a value obtained for the Husband CGPS station during the same period (5.21 mm/year), shown as circle with a black outline.

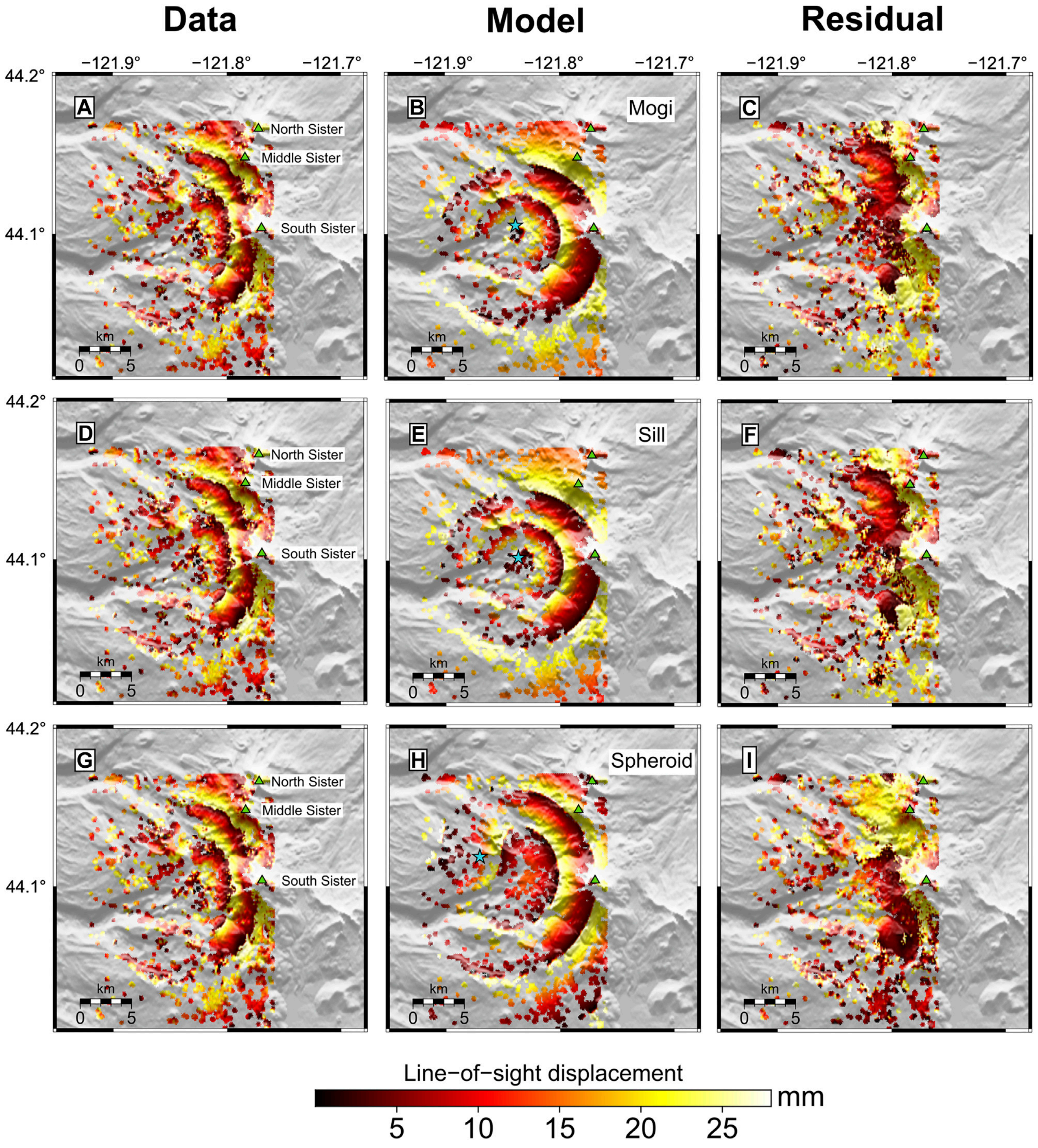


FIGURE 4 | Wrapped InSAR data and model results for Track 385 of the ERS satellite. **(A, D, G)** Line-of-sight (LOS) deformation observed in a nearly 3 years period from August 24, 1997 and September 17, 2000, considering only pixels with coherence >0.2. Green triangles represent the Three Sister complex volcano system. **(B, E, H)** Bayesian model and horizontal location for the median a posteriori probability solution (blue star) for a predicted Mogi, sill-like and prolate spheroid source, respectively. **(C, F, I)** Residual maps for Mogi, sill-like and prolate spheroid source. The model parameters results are presented in **Figure 5** and **Tables 1**, **2**.

making it difficult for a single L-band interferogram to detect the deformation signal (Riddick and Schmidt, 2011), a cumulative LOS deformation time-series can detect such changes in rate. The corresponding time-series was processed by StaMPS Version 3.3.b1 to study the surface deformation, applying the Small Baselines method for 30 interferometric pairs (Hooper et al., 2012; Bekaert et al., 2017). Small Baselines method minimizes decorrelation in natural terrains. So, it is an appropriate method for the Three Sisters area, which lacks man-made structures and hence offers few dominant persistent scatterers.

Due to the small deformation rate (5.8 mm/year for the period June 2015 - January 2020) and low signal-noise ratio in the Cascades, geodetic data must be analyzed carefully (Poland et al., 2017). Following this recommendation, we consider a 1 cm standard deviation for neighboring pixels in all interferograms. Only four good quality interferometric pairs were used for the Sentinel-1 observation period. Adding more interferograms did not significantly improve the analysis of the volume change time-series. Moreover, the analysis of a 6-years long Sentinel-1 dataset in Turkey indicates that surface displacement rate uncertainties are mostly dominated by length of observation, rather than larger numbers of available interferograms (Weiss et al., 2020). Hence, we consider Sentinel-1 summer-to-summer and summer-to-late spring InSAR data to avoid decorrelation due to snow coverage, and to fill a noticeable GPS time-series gap. However, the deformation rates could be reexamined in future, using longer Sentinel-1 datasets.

Figure 3 represents the mean LOS velocity (mm/year) for the ascending path 219 ALOS-1 from January 2007 to March 2011. Due to the high signal-to-noise ratio, the StaMPS LOS velocity results were noisy and we post-processed them to reduce undesirable oscillations of non-volcanic origin. We applied a band pass filter to retain spatial deformation signals between 10 and 0.8 km, using a median filter (GMT blockmedian). Although close to the signal-to-noise ratio value, results indicate a 6 km circular uplift pattern west of South Sister with a mean LOS velocity of ~5–10 mm/year. This mean velocity is consistent with a value obtained for the HUSB CGPS station during the same period (5.2 mm/year).

4 METHODS

Here, we introduce a mathematically rigorous strategy for the joint inversion of time-dependent InSAR (different look angles and sensors, high spatial resolution) and continuous GPS (daily sampling) data to achieve a complete timeline of volcanic activity and quantify a single time series of volume flux rates. The strategy captures the benefits of each system avoiding the time evolution determination on a point-by-point basis. It is based on the two-step approach proposed by González et al. (2013) that produces time series of volume from sets of different look angles and satellite sensors once an active source is determined for an inflation period. In this section, we provide a description of the González et al. (2013) algorithm and extent it 1) to include continuous GPS data and therefore to combine different components and/or look vectors using a unique source model and 2) to afford a defined method of truncation of the TSVD technique used to regularize the inverse problem, with the goal of finding the time evolution of volume and therefore to improve the accuracy on the estimation of volume time series. First, we show how the active source location (horizontal position and depth) and geometry is determined using a Bayesian inversion approach. Subsequently, we solve for temporal evolution of volume.

4.1 Source Characterization

First of all, we infer the active magmatic source through a Bayesian inversion, using InSAR data spanning the period of maximum displacement. The horizontal location, depth and geometry of the inflation source at Three Sisters were computed using the MATLAB®-based software package GBIS (Geodetic Bayesian Inversion Software) (Bagnardi and Hooper, 2018), which estimates source parameters through a Markov chain Monte Carlo method and uses, among others, analytical forward models from dMODELS software package (Battaglia et al., 2013). It obtains the posterior probability distributions (PDFs) for all model parameters by taking into account uncertainties in the data (e.g., data errors). To achieve this, considering the pattern of surface deformation, we employ simple elastic models such as point source (Mogi, 1958), prolate spheroid (Yang et al., 1988) and sill-like (Fialko et al., 2001) models. An elastic, homogeneous and isotropic half-space is assumed in all the approaches with Poisson's ratio 0.25. We assumed, as previous studies (Dzurisin et al., 2006, 2009; Riddick and Schmidt, 2011), a stationary source and we used an interferogram spanning August 1997 - September 2000 to look for source parameters. This interferogram fulfills two important conditions to determine the best-fitting static displacements: 1) it spans the shortest time during the period of maximum deformation; 2) it shows acceptable signal-to-noise ratio. InSAR spatially correlated error (caused mainly by the "wet" tropospheric delay) is estimated by modeling experimental semivariograms in deformation-free regions (Bagnardi and Hooper, 2018). InSAR data are subsampled with an adaptative quadtree method (Decriem et al., 2010) to reduce the computational cost of the Bayesian inversion, particularly for sill-like and prolate spheroid sources. The inversion computes 2×10^6 iterations for spherical point source, 5×10^6 iterations for sill-like source and 8×10^6 iterations for prolate spheroid, which stabilizes the inversion procedure.

4.2 Temporal Evolution of the Source Volume Changes

Once the magmatic source is fixed, we perform the quasi-dynamic time-dependent model using a linear inversion scheme to look for the volume changes at each interferogram's period and the cumulative volume changes since the first GPS observation. Both volume changes at each interferogram and cumulative volume changes from the GPS data are used to solve for the time evolution of volume using TSVD.

4.2.1 First Step: Piecewise Volume Changes Over Temporal Data Periods

Once the location and geometry of the inflation source are fixed, we determine the volume changes over the corresponding time intervals (increments of volume changes, ΔV) for both, InSAR and CGPS data sets, which are assumed to be uncorrelated. In this way, observations from several interferograms and GPS sites can be combined to estimate increments of volume changes assuming a unique source model. Each volume change, ΔV_{ij}, records 1) the incremental volume change between two acquisition dates, t_i and t_j from an interferogram or 2) the cumulative increment of volume since the first observation, i.e., $t_j = t_0$, being t_0 the starting date of CGPS.

A linear inversion scheme using Weighted Least Squares (WLS, Menke, 1989) is applied. The inversion is constrained by 55 interferometric pairs (ERS, ENVISAT, Sentinel-1), one ALOS-1 interferogram (cumulative LOS deformation time-series) and a 3-component GPS time series. The forward problem is defined by $\mathbf{d} = \mathbf{Gm}$, where $\mathbf{d}$ is the data vector (InSAR or GPS), $\mathbf{m}$ is the model parameter ($\mathbf{\Delta V}$) vector, and $\mathbf{G}$ is the Green's function matrix representing the impulse response for the specific elastic source, projected into the three components of GPS or the satellite line-of-sight. Therefore, a total of 5,064 independent linear inversions were performed to find the increments of volume changes, ΔV_{ij}, given the set of interferograms and 5,008 cumulative GPS displacements.

The least square estimator of each inversion, $\widehat{\mathbf{m}}$, is given by:

$$\widehat{\mathbf{m}} = \left[\mathbf{G}^T\mathbf{C}_\mathbf{d}^{-1}\mathbf{G}\right]^{-1}\mathbf{G}^T\mathbf{C}_\mathbf{d}^{-1}\mathbf{d}, \tag{1}$$

with the cofactor matrix $\mathbf{C}_{\widehat{\mathbf{m}}} = \left[\mathbf{G}^T\mathbf{C}_\mathbf{d}^{-1}\mathbf{G}\right]^{-1}$. We considered a diagonal variance-covariance matrix, $\mathbf{C}_\mathrm{d}$, assuming that all data are independent, which significantly reduces the computation time of the inversions. Hence, we ignore the possible spatial and temporal correlation noise in InSAR data (e.g., pixel correlation due to atmospheric artifacts, topography structures, repeated acquisitions) and between GPS components (Lohman and Simons, 2005; Biggs et al., 2010).

4.2.2 Second Step: Volume Changes Time-series

We want to solve for the temporal evolution of volume change for each observed epoch t_k from ΔV_{ij} obtained on first step considering both InSAR and CGPS data. Instead of volume change itself, the rate of volume changes is inverted as a function of time by applying the Short Baseline Subset Approach (SBAS, Berardino et al., 2002). This prevents the presence of large discontinuities in the final solution.

Let $\mathbf{\Delta V}$ be the data vector of volume changes over the corresponding time intervals ($N \times 1$), and $\dot{\mathbf{V}}$ the unknown vector of volume change rates ($M \times 1$) between adjacent epochs, t_j where the overdot means differentiation over time. Then,

$$\dot{\mathbf{V}} = \begin{bmatrix} \dot{v_1} = \dfrac{v_1}{t_1 - t_0} \\ \vdots \\ \dot{v_M} = \dfrac{v_M - v_{M-1}}{t_M - t_{M-1}} \end{bmatrix}. \tag{2}$$

The usual method of converting the observations $\mathbf{\Delta V}$ on volume change rates is:

$$\mathbf{B}\dot{\mathbf{V}} = \mathbf{\Delta V}, \tag{3}$$

where $\mathbf{B}$ is the design matrix ($N \times M$). To determine the components of $\mathbf{B}$, we define a ($M \times 1$) vector $\mathbf{E}$, containing the single epochs t_j present in all time intervals and sorted in chronological order, for $j = 1, ..., M$; and a ($N \times 2$) vector $\mathbf{F}$, whose columns are the slave (t_{slave_i}) and master (t_{master_i}) epochs of each i time interval, for $i = 1, ..., N$. Therefore, the (i, j) component of the design matrix is $B_{ij} = (t_{j+1} - t_j)$ for $t_{\mathrm{slave}_i} \le t_j < t_{\mathrm{master}_i}$, and zero elsewhere. In the case of cumulative $\mathbf{\Delta V}$ (i.e., continuous GPS), $\mathbf{B}$ presents lower triangular matrix blocks. For example, if volume changes are obtained over different time intervals, i.e., if t_{AB}, t_{BC}, and t_{AC} (from InSAR data), and t_{CD}, t_{CE}, t_{CF}, (from CGPS data) are Δv_{AB}, Δv_{BC}, Δv_{AC}, Δv_{CD}, Δv_{CE} and Δv_{CF}, the design matrix is given by:

$$\begin{pmatrix} (t_B - t_A) & 0 & 0 & 0 & 0 \\ 0 & (t_C - t_B) & 0 & 0 & 0 \\ (t_B - t_A) & (t_C - t_B) & 0 & 0 & 0 \\ 0 & 0 & (t_D - t_C) & 0 & 0 \\ 0 & 0 & (t_D - t_C) & (t_E - t_D) & 0 \\ 0 & 0 & (t_D - t_C) & (t_E - t_D) & (t_F - t_E) \end{pmatrix} \begin{pmatrix} \dot{v}_{AB} \\ \dot{v}_{BC} \\ \dot{v}_{CD} \\ \dot{v}_{DE} \\ \dot{v}_{EF} \end{pmatrix} = \begin{pmatrix} \Delta v_{AB} \\ \Delta v_{BC} \\ \Delta v_{AC} \\ \Delta v_{CD} \\ \Delta v_{CE} \\ \Delta v_{CF} \end{pmatrix}. \tag{4}$$

To illustrate the simple example in **Eq. 4**, let all dates, t_A, t_B, t_C, t_D, t_E, and t_F be equally spaced at time intervals of 1 year, i.e., t_{AB} = 1 year and so on, and $\mathbf{\Delta V} = [2, 1, 3, 1, 2, 3] \times 10^6\ \mathrm{m}^3$. In this case, standard least squares can be applied, given $\dot{\mathbf{V}} = [2, 1, 1, 1, 1] \times 10^6\ \mathrm{m}^3/\mathrm{year}$. The cumulative volume times series is then $\mathbf{V} = [2, 3, 4, 5, 6] \times 10^6\ \mathrm{m}^3$, meaning a linear inflation rate of $2 \times 10^6\ \mathrm{m}^3/\mathrm{year}$ in time interval t_{AB} and a posterior linear inflation rate of $1 \times 10^6\ \mathrm{m}^3/\mathrm{year}$.

However, the set of ΔV_{ij} forms, in general, an unconnected set of observations with at least one time step not directly related to data, making **Eq. 4** an ill-posed problem without solution even in the least square sense. Thus, given $\mathbf{P}$ a definite positive matrix, a least-square solution, $\hat{\dot{\mathbf{V}}} = (\mathbf{B}^T\mathbf{PB})^{-1}\mathbf{B}^T\mathbf{\Delta V}$, is not possible since **Eq. 4** constitutes an ill-posed unstable model, with one or more eigenvalues of the normal matrix $\mathbf{B}^T\mathbf{PB}$ close to zero. This fact is responsible for large uncertainty on the estimated volume change rates, $\hat{\dot{\mathbf{V}}}$. Alternatives to the least square method can be proposed for an improved estimate of $\hat{\dot{\mathbf{V}}}$: Tikhonov regularization (Tikhonov and Arsenin, 1977), Bayesian and stochastic inferences (Backus, 1988), Truncated Singular Value Decomposition (TSVD) or Principal Components (Lawless and Wang, 1976; Hansen, 1990; Hansen, 1992). Here, we consider TSVD as proposed by González et al. (2013).

4.2.3 Regularized Linear Joint Inversion

A key difficulty in applying the TSVD method is how to set up proper criteria to truncate eigenvalues due to the lack of a theoretically solid foundation to discard small nonzero eigenvalues. We developed a strategy to circumvent this difficulty based on Picard condition and L-curve methodology. In such way, we are assured a good balance, filtering out enough noise without losing too much information in the computed solution (Hansen and O'Leary, 1993). Furthermore, we included some estimations of the error of data ($\mathbf{\Delta V}$) to establish some uncertainty in the $\hat{\dot{\mathbf{V}}}$ estimator. To estimate the uncertainties, the Weighted Generalized Inverse method (Menke, 1989) permitted the use of the “a priori” information from the data $\mathbf{C}_{\Delta\mathbf{V}}$ (and optional model, $\mathbf{C}_{\dot{\mathbf{V}}} = \mathbf{A}^T\mathbf{P}\mathbf{A}$) covariance matrix. Such matrices can be decomposed as:

$$\begin{aligned}\mathbf{C}_{\Delta\mathbf{V}} &= \left[\mathbf{D}\mathbf{D}^T\right]^{-1}\\ \mathbf{C}_{\dot{\mathbf{V}}} &= \left[\mathbf{S}\mathbf{S}^T\right]^{-1}\end{aligned} \tag{5}$$

where the $\mathbf{D}$ ($N \times N$) and $\mathbf{S}$ ($M \times M$) matrices are determined from the eigenvalue problem of each covariance matrix. In our case, no model covariance information is used, so $\mathbf{C}_{\dot{\mathbf{V}}} = \mathbf{I}.\mathbf{C}_{\Delta\mathbf{V}}$ is obtained by error data propagation through $\widehat{\Delta\mathbf{V}}$ estimator. The utility of $\mathbf{D}$ and $\mathbf{S}$ is the introduction of a transformed coordinate system where data (and optional model) parameters each have uncorrelated errors and unit variance. Therefore, $\Delta\mathbf{V}_{\mathbf{new}} = \mathbf{D}\Delta\mathbf{V}$, $\dot{\mathbf{V}}_{\mathrm{new}} = \mathbf{S}\dot{\mathbf{V}}$ and $\mathbf{B}_{\mathrm{new}} = \mathbf{D}\mathbf{B}\mathbf{S}^{-\mathbf{1}}$ give the transformation of data, model parameters and forward operator in the new system of coordinates. TSVD is applied to $\mathbf{B}_{\mathrm{new}}$ with a specific regularization method to find the Principal Components of the observation set ($\Delta\mathbf{V}$). Then, the problem is back to the original coordinates to achieve the solution and finally, the volume change time-series is obtained by integrating the volume change rate in time:

$$V(t) = \sum \dot{V}\delta t \tag{6}$$

4.2.4 Regularization: Techniques Used for Truncation of Small Eigenvalues

Some workable criteria for truncation in interdisciplinary problems include L-curve, Discrete Picard condition and Generalized cross validation (GCV) (Hansen, 1992; Hansen, 2007; Hansen and O’Leary, 1993). Methods like GCV sometimes fail to find the appropriate regularization parameter (flat local minima), whereas the L-curve gives a robust estimation (Hansen and O’Leary, 1993) and the appropriate smoothing solution, which is very attractive from a mathematical point of view. We thus designed a strategy based on a L-curve to set up proper criteria to truncate the eigenvalues.

First, we considered the Discrete Picard condition to explain the instability of the transformed linear inverse problem (**Eq. 3**) and disregarded the smallest singular values (Hossainali et al., 2010; Hansen, 1990; Hansen, 2007):

$$\left\|\dot{\mathbf{V}} - \dot{\mathbf{V}}_{T_1}\right\|_2 \le P^{\frac{1}{2}} \max_{1\le i\le M-P}\left\{\frac{\left|u_i^T\Delta V_i\right|}{s_i}\right\} \tag{7}$$

where $\left\|\dot{\mathbf{V}} - \dot{\mathbf{V}}_{T_1}\right\|_2$ is the regularization error, $\dot{\mathbf{V}}$ and $\dot{\mathbf{V}}_{T_1}$ being the exact and truncated SVD solutions; P is the regularization parameter value, s_i are the singular values, and $\left|u_i^T\Delta V_i\right|$ are called Fourier coefficients (ΔV_i are data and u_i the corresponding eigenvectors of the data space). The Discrete Piccard condition is satisfied if, for all singular values larger than P, the corresponding Fourier coefficients decay faster on average than s_i.

The L-curve method is applied to $\dot{\mathbf{V}}_{T_1}$ resulting in turn from applying (**Eq. 7**) through a log-log plot of the norm of a regularized solution $\left\|\dot{\mathbf{V}}_{T_2}\right\|_2$ vs. the norm of the corresponding residual norm $\left\|\mathbf{B}\dot{\mathbf{V}}_{T_2} - \Delta\mathbf{V}\right\|_2$. As recommended by Hansen (1992) we fit the log-log plot of discrete points with some curves, choosing a 2D spline curve and then search for the truncation parameter by computing the L-corner (maximum curvature point). This corner of the spline curve is approximated to the closest discrete point. The L-corner is located exactly where the solution changes in nature from being dominated by regularization errors to being dominated by the residual size. This regularization filters out the contribution of small singular values and noisy data.

5 RESULTS

5.1 Re-Evaluation of Source Location and Geometry

We performed the Bayesian inversion for point source, prolate spheroid and sill-like sources, with similar results. The results are represented in **Figures 4**, **5** and **Tables 1**, **2** report the prior information and the PDFs with the 95% credible intervals for all model parameters, respectively. The inversion reveals that the surface displacements can be explained by a spherical point source with depth (4500–6000)m and ΔV (7–13)$\times 10^6$ m^3, by a sill-like source with depth (5600–7200)m, radius (220–400)m and dimensionless excess pressure 0.05–0.30 and by a prolate spheroid source with depth (5300–7400)m, major semi-axis (240–720)m, aspect ratio (0.22–0.37)m, dimensionless excess pressure 0.38–8.61, strike angle (49–102)° and plunge (78–85)°. The descending ERS Track 385 wrapped interferogram reveals a near axisymmetric deformation pattern, with a maximum LOS surface displacement of ~5 cm recorded at the center of the uplifting pattern (**Figures 4A, D** and **G**). **Figures 4B, E** and **H** present the predicted spherical point, sill-like and prolate spheroid forward models, using the median value of the PDF of the model parameters. As expected from the deformation pattern, spherical point and sill-like models are very similar, suggesting that the geometry of the source is far from unique. The extra modeling parameters of the prolate spheroid do not improve the misfit. Therefore, we favor the simplest spherical point source model over a sill-like and prolate spheroid source, to fit deformation pattern displayed in **Figures 4A, D** and **G**. Blue stars represent the horizontal location of spherical point, sill-like and prolate spheroid estimated sources (**Table 2**). **Figures 4C, F** and **I** show the residuals between observed LOS displacement, and spherical point, sill-like and prolate spheroid model predicted displacements. The residual is larger close to the Three Sisters complex volcano (green triangles), due to orbital and topographic contributions, and also in the western half of the uplift pattern, where data are less dense.

Figure 5 displays the histograms of marginal PDF for the four spherical point, five sill-like and eight prolate spheroid source parameters. Black solid lines show the optimal values for the corresponding model parameters. For the sill-like source, the radius and dimensionless excess pressure PDFs exhibit bi-modality (slightly unstable inversion result). For the prolate spheroid source, the aspect ratio between semi-axes and the

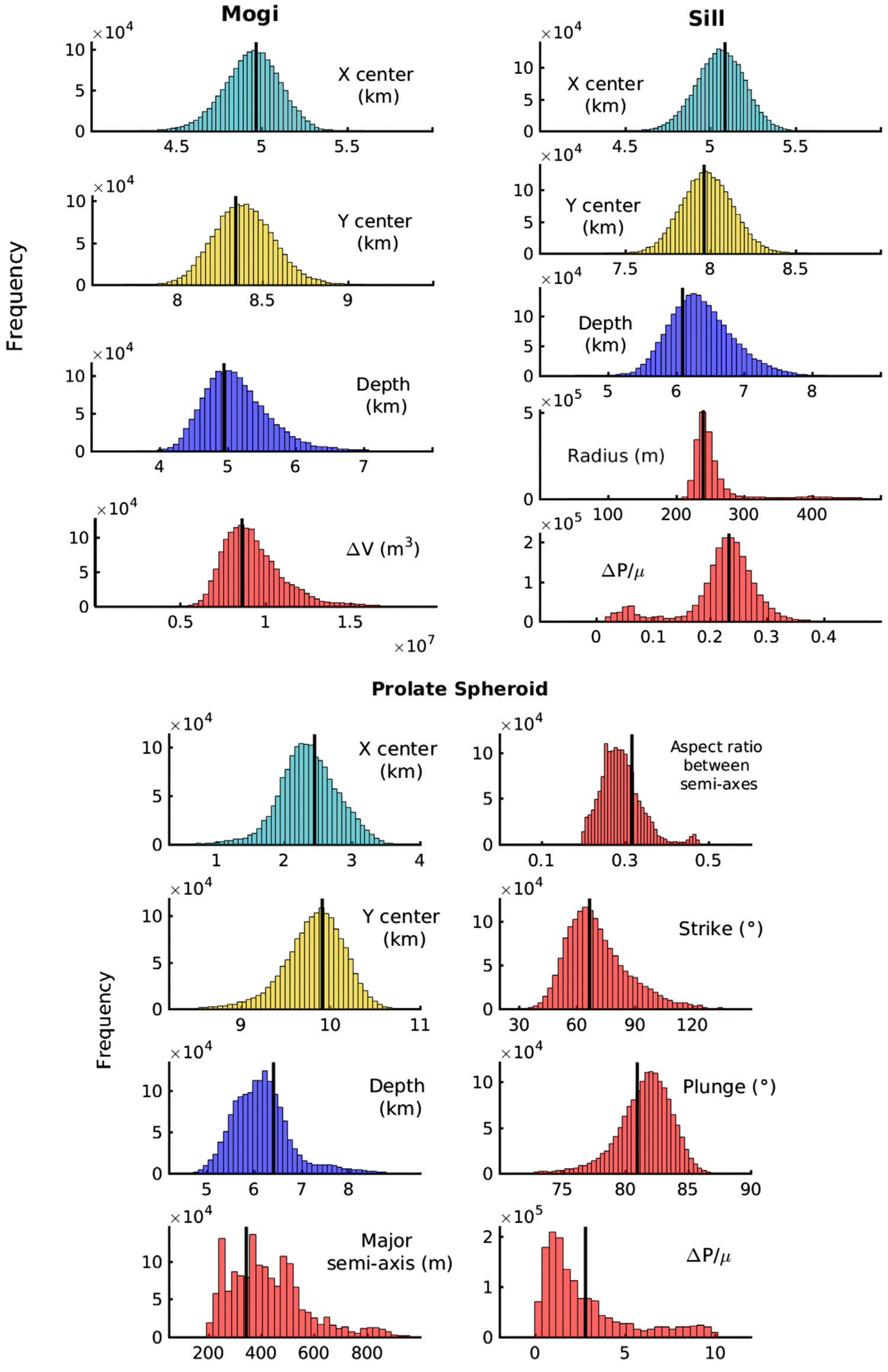


FIGURE 5 | Posterior probability distributions for the Mogi, sill-like and prolate spheroid source models. Black solid lines show the optimal value for the corresponding model parameter (**Tables 1**, **2**).

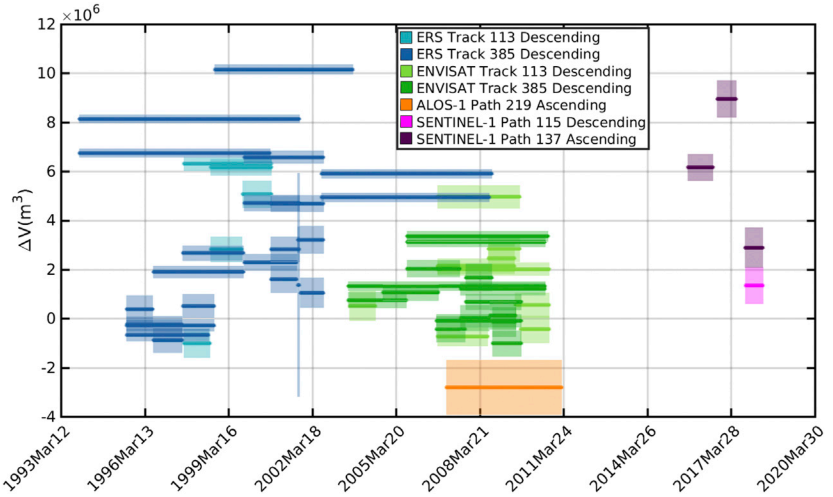


FIGURE 6 | Increments of volume change obtained for all interferograms (ERS, ENVISAT, ALOS-1 and SENTINEL-1), according with the median value for the source depth.

excess pressure PDFs display also a bi-modal shape. The major semi-axis PDF exhibits multi-modality.

5.2 Source Inflation Time-Series

We performed a CGPS and InSAR joint inversion to obtain the time-line of volume changes, considering the best fitting source location for the point source geometry to better characterize the time-dependent inflation of the magma source at Three Sisters.

To apply our two-step-approach (**section 3.2**), we use the median, and 5% and 95% percentile values of the PDF of depth estimated by the Bayesian inversion. The corresponding values are: $d_{\text{median}} = 5000$ m, $d_{5\%} = 4500$ m and $d_{95\%} = 6000$ m. The volume change time-series is determined using InSAR data from four satellite missions (ERS, ENVISAT, ALOS-1 and Sentinel-1), on five different tracks and look angles, and CGPS data from HUSB.

First, we obtain the increments of each volume change, ΔV_i, relating the Green's functions (representing the source impulse for a buried point source) to the LOS deformation data observed along each satellite track and the three component CGPS data. **Figures 6**, **7** respectively show results regarding estimation of the median value of source depth ($d_{\text{median}} = 5000$ m) for InSAR and CGPS data. The cumulative increments of volume changes detected at HUSB show gaps due to ice and snow accumulation during winter. By means of the daily GPS measurements, the corresponding increments of volume change, ΔV_i, are more uniform for the CGPS data sets (**Figure 7**), but more variable for the individual SAR data sets (**Figure 6**).

Finally, we applied the Picard Plot condition suitable to understand the conditions of the ill-posed problem (**Eq. 3**). **Figure 8A** shows how s_i only decays faster than the Fourier coefficients ($|u_i^T \Delta V_i|$) for the smallest nonzero singular values. Hence, the problem can be considered stable, discarding the last 10% of the singular values. Due to the stability of the problem, the Picard Plot provides no clues about the appropriate level of truncation (**Eq. 7**). Therefore, we use L-curve to determine the truncation level. L-curve criterion is fulfilled when L-corner = 1198, i.e., when only the first 24.1% nonzero singular values are used in the inversion (**Figure 8B**).

TABLE 1 | Prior information for the Elastic half-space Spherical Point Pressure (Mogi, 1958), Penny-shaped sill-like (Fialko et al., 2001) and Spheroid (Yang et al., 1988) deformation sources.

	Lower	Upper	Step	Start
Mogi				
X_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Y_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Depth (m)	5×10^2	2.00×10^4	50	1.00×10^3
ΔV ($\times 10^6$m³)	0.1	1.00×10^4	1.00×10^{-3}	0.1
Sill				
X_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Y_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Depth (m)	5×10^2	2.00×10^4	50	1.00×10^3
Radius (m)	100	4,000	50	100
$\Delta P/\mu$	1×10^{-5}	10	1×10^{-6}	1×10^{-2}
Spheroid				
X_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Y_{center} (m)	-1.00×10^3	1.50×10^4	25	2.00×10^3
Depth (m)	5×10^2	2.00×10^4	50	1.00×10^3
Major semi axis (m)	100	4,000	50	100
Aspect ratio	0.2	0.99	0.01	0.5
$\Delta P/\mu$	1×10^{-5}	10	1×10^{-6}	1×10^{-2}
Strike (°)	1	359	1	270
Plunge (°)	0	90	1	45

The analysis and results of the final inflation time-series are shown in **Figures 9** and **10**, and **Tables 3** and **4**. The inflation time-series associated with the median estimation of the source depth suggests a maximum volume change rate of $\sim 1.60 \times 10^6$ m³/yr during 1999–2001 and a subsequent rate as much as $\sim 0.75 \times 10^6$ m³/yr for the period 2015- January 2020. Data since

TABLE 2 | Bayesian inversion results, with the median a posterior probability solution and the 95% credible intervals, for the Elastic half-space Spherical Point Pressure (Mogi, 1958), Penny-shaped sill-like (Fialko et al., 2001) and Spheroid (Yang et al., 1988) deformation sources.

	median	5%	95%
Mogi			
Lon (°)	−121.8382	−121.8418	−121.8350
Lat (°)	44.1055	44.1030	44.1082
Depth (m)	5,000	4,500	6,000
ΔV ($\times 10^6 m^3$)	8.99	6.98	12.66
Sill			
Lon (°)	−121.8368	−121.8398	−121.8339
Lat (°)	44.1018	44.0997	44.1041
Depth (m)	6,300	5,600	7,200
Radius (m)	250	220	400
$\Delta P/\mu$	0.23	0.05	0.30
Spheroid			
Lon (°)	−121.8707	−121.8791	−121.8616
Lat (°)	44.1187	44.1130	44.1228
Depth (m)	6,100	5,300	7,400
Major semi-axis (m)	400	240	720
Aspect ratio	0.28	0.22	0.37
$\Delta P/\mu$	1.85	0.38	8.61
Strike (°)	68	49	102
Plunge (°)	82	78	85

Geo-reference point is [−121.9, 44.03]°.

2018 show a subtle, but significant change in the trend, instead of following asymptotic behavior.

6 DISCUSSION

6.1 Source Characterization

Studies at Three Sisters using InSAR interferometric pairs and stacks (Wicks, 2002; Riddick and Schmidt, 2011), GPS (campaign and continuous) and leveling (Dzurisin et al., 2009) assessed various source geometries such as spherical point, sill-like or crack and ellipsoidal. These different sources can all fit the data in a satisfactory way. Our results are consistent with previous findings (**Table 3**). However, volume change rates and depths vary slightly, possibly due to the fact that: 1) there may be a poorly resolved deeper magma source, 2) inversions were limited to purely kinematic models, 3) the source not being a stationary, pressurized cavity in an isotropic elastic half-space, thus producing bias due to spatial or temporal considerations, 4) diverse inversion techniques and related possible mathematical artifacts, 5) different types of data sets and 6) ambiguity of source geometries. We assumed a simple, stable, purely kinematic model as a valid approach, following the results of Dzurisin et al. (2009) and Riddick and Schmidt (2011), for estimating volume time series. Now, we focus on discussing the implications of 3), 4), 5), and 6).

We revisited the assumption of location stationarity of the inflation source, with focus on the most recent periods. Riddick and Schmidt (2011) already showed that the temporal evolution of the uplift signal can be represented by a stationary volcanic source geometry and location with a decreasing inflation rate at least from 1992 to 2010. The extent of deformation pattern remains constant over 1992–2010 time period providing a source depth range compatible with the uncertainties of inversion models, as it is expected given that the extent of deformation pattern mainly depend on the source depth and strength. The inversion for source parameters using Sentinel-1 interferograms (2014–2018) reveals very large uncertainties on the parameters. Such results could be due to the interferograms' low signal-to-noise ratio caused by slower uplift rates during this period. Nevertheless, the best fitting spherical sources are not able to predict the observed displacements in the HUSB CGPS displacement time series. Moreover, the lack of substantial changes in the trends of each component of the HUSB CGPS time series indicates that the source might not have changed

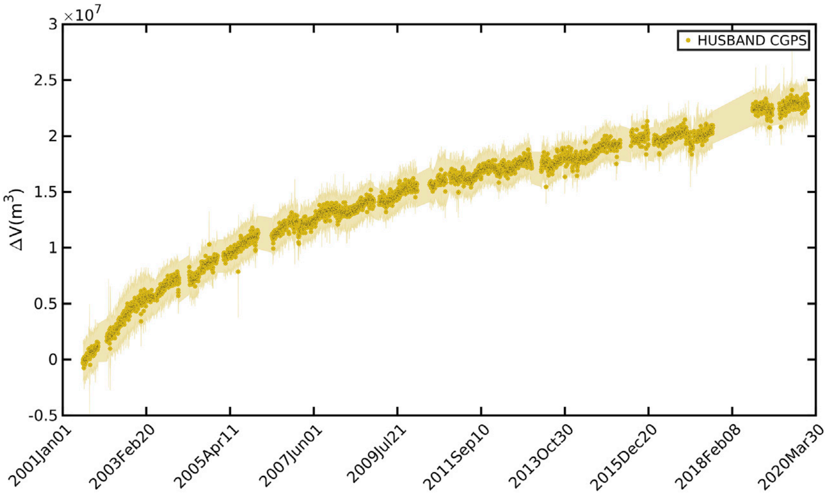


FIGURE 7 | Cumulative increments of volume changes for the CGPS station, Husband. The figure shows the results according with the median value for the source depth.

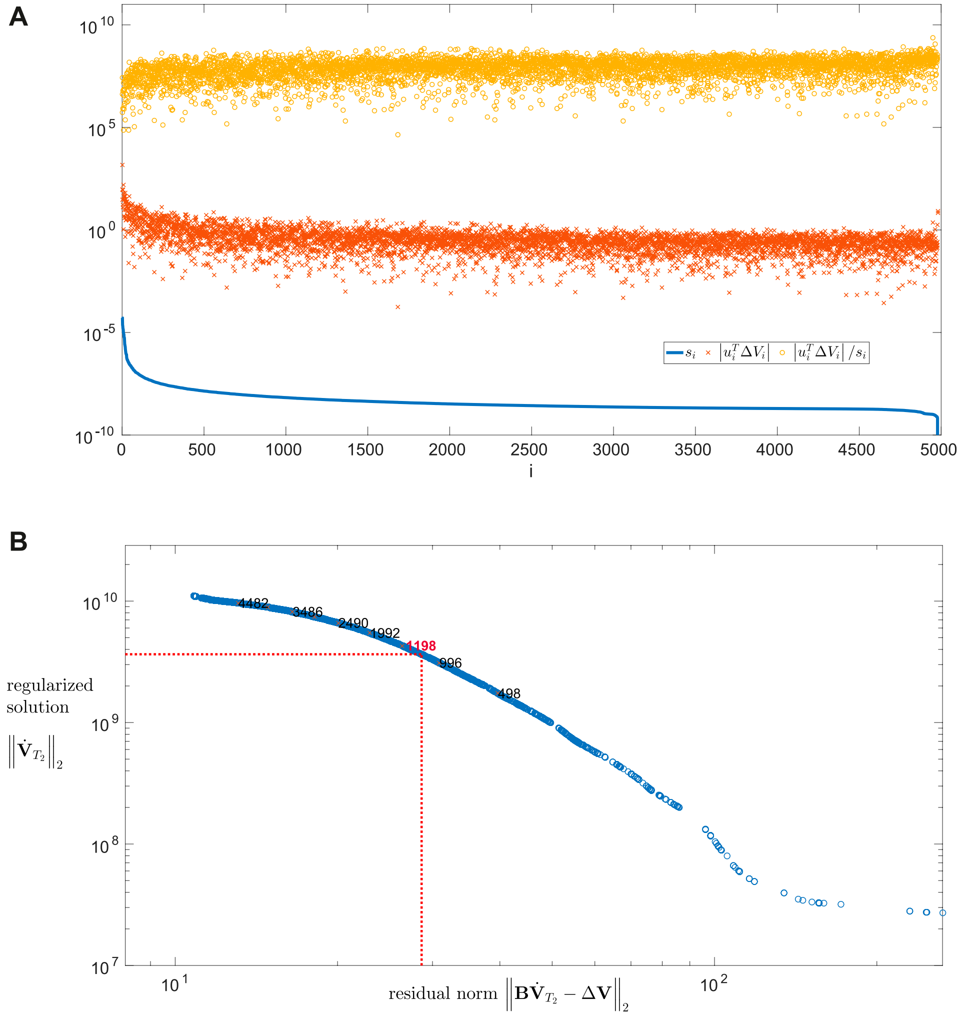


FIGURE 8 | (A) Discrete Picard Plot condition, suited for the analysis of ill-posed problems. The solution is stable when the Fourier coefficients, $|u_i^T \Delta V_i|$, on average decay to zero faster than the reciprocal singular values, s_i. In this case, the problem can be consider stable, discarding the last 10% singular values. **(B)** L-curve showing the trade-off between minimizing the residual norm ($\|\mathbf{B}\dot{\mathbf{V}}_{T_2} - \Delta\mathbf{V}\|_2$) and minimizing the regularized solution size ($\|\dot{\mathbf{V}}_{T_2}\|_2$). The L-corner (represented in red) is located exactly where the solution changes in nature from oversmoothing (i. e, dominated by regularization errors) to being dominated by residual size.

position. Therefore, we assumed the source has not changed significantly either in shape nor in location since the onset of deformation.

A range of common techniques to estimate source location has been used, like forward modeling, grid search by iteratively fixing of one parameter, arithmetic mean to obtain range values, or grid search. However, the Bayesian approach presents important advantages: 1) robust inversion for a single or more InSAR interferograms with an acceptable signal-to noise-ratio and/or GPS data, 2) rapid simultaneous inversion of all model parameters; 3) use of data uncertainty and prior model information; and 4) efficient sampling of posterior PDFs to estimate optimal model parameters and the associated range of error. Bearing in mind such advantages, to obtain a robust estimate of source geometry and location we only need geodetic data with high spatial coverage, spanning the most appropriate period (shortest time, high deformation). For this, we use the period undergoing maximum deformation, displaying as much as ~5 cm of line-of-sight deformation (**Figure 4**). Selection of the ERS-1 track 365 (descending orbit) interferometric pair spanning August 1997- September 2000 satisfies both criteria. To reduce the signal-to-noise ratio, the interferogram is filtered for a pixel coherence threshold of 0.2. Other ERS-1 InSAR data were also processed for similar periods of time, but not used due to the low signal-to-noise ratio. No GPS data were available until 2001 and cannot be used to study the maximum uplift rate period.

We favor the simplest source, a spherical point source, to infer volume time series at Three Sisters. Our inversion results for spherical, prolate spheroid and a sill-like sources showed quantitatively similar results, in terms of data misfit and surface displacement pattern. We noted that there is slightly elongated pattern of the InSAR data in the North-South direction (**Figure 4**). This pattern cannot be perfectly reproduced with an axisymmetric source geometry. The elongation could be also be

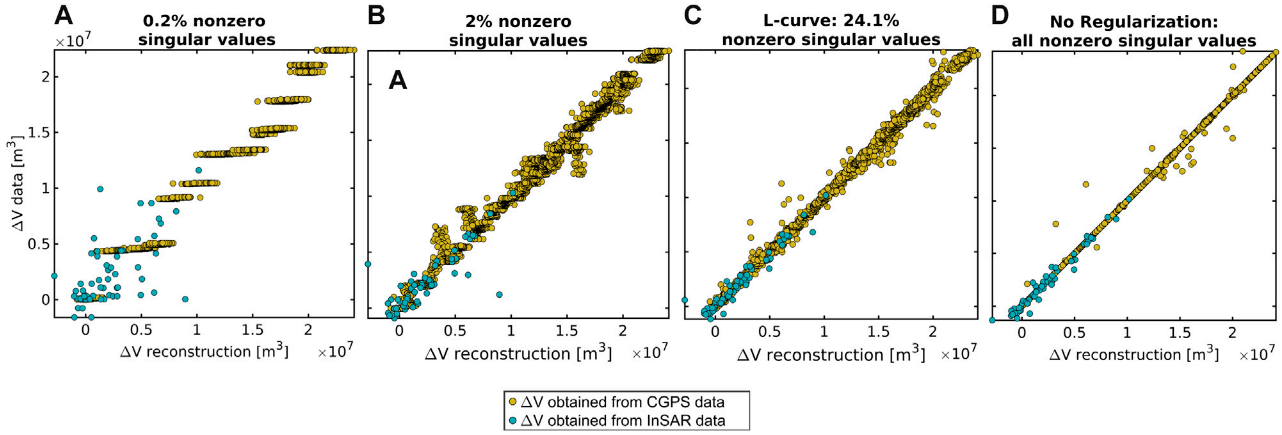


FIGURE 9 | ΔV_i data (increment of volume change obtained from InSAR and CGPS deformation data, **Eq. 1**), vs. Simulated observations ($\Delta V_{i_{\text{reconstruction}}} = V_{(i+1)} - V_i$, **Eq. 6**). **(A, B)** Inversion solutions with regularization, using only the first 0.2 and 2% nonzero singular values, respectively (oversmoothing solution). **(C)** Regularized solution using the L-curve criteria corresponding with 24.1% nonzero singular values (appropriate smoothing solution). **(D)** No regularization case (undersmoothing solution).

TABLE 3 | Source location comparison for Three Sisters (assuming a Mogi source) from Previous Studies and the bayesian inversion carried out in this Study.

Inversion	Depth (m)	ΔV (10^6 m^3)	ΔV$_{rate}$ (10^6 m^3/yr)
1996–2000, InSAR[a]	6500 ± 400	23.00 ± 3.00	~5.75 ± 0.75
2001–2008, Geodetic ground base data[b]	5,800	22.20	~3.14
1993–2008, InSAR[c]	5200 ± 100	57.00 [+1.95, −3.60]	~3.80 [+0.13, −0.24]
1997–2000, InSAR [this study][d]	5000 [+1000, −500]	8.99 [+3.67, −2.01]	~3.00 [+1.22, −0.67]

[a] *ERS Descending (08/1996–10/2000) (Wicks, 2002).*
[b] *Campaign GPS, CGPS and Leveling (05/2001-late 2008) (Dzurisin et al., 2009).*
[c] *ERS Track 385 Descending (08/1993-08/2008) (Riddick and Schmidt, 2011).*
[d] *ERS Track 385 Descending (08/1997-09/2000).*

caused by the topography of Three Sisters area at the east side of the deforming area or tropospheric delays in the interferometric data. To robustly distinguish between different source geometries, we must have full three dimensional surface displacements (Dieterich and Decker, 1975). Therefore, in our case, the almost symmetric shape of the 2D deformation pattern implies that source geometry remained far from being uniquely resolved. Nevertheless, to accept more complex geometries than spherical, we should have obtained significantly lower data misfit values. Furthermore, the spheroid modeling residuals are not fully consistent with the observed pattern in the western area of deformation (**Figures 4I**). In this case, we also consider that the topography could have a minor effect because the highest topographic relief is concentrated on the far field area of deformation signal. Therefore, we assume that the noted asymmetry in the InSAR data should not affect the overall interpretation of the time series of volume changes.

The differences between inversion methods and data selection might explain that our optimal inversion results suggest a slightly shallower source with a corresponding smaller increment of volume change. Despite that, considering that the models fit the data well and yield similar misfit values, we conclude that it is reasonable to assume the spherical point source as the simplest kinematic model that explains the signal. Furthermore, the values for depth and increments of volume change lie within our 95% credible intervals (**Table 3**). Wicks (2002) processed three interferograms, obtaining an increment of volume change of $\Delta V = 23 \times 10^6\,\text{m}^3$ and depth of 6500 m for the one with the largest apparent signal-to-noise ratio. Ultimately, a deeper source will trade-off with a greater ΔV, to fit the same surface deformation. Although the time acquisition of the best interferogram of Wicks (2002) spans only 1 year more than our InSAR data interferometric pair, the important difference between our $\Delta V_{rate} = \sim 3.00 \times 10^6\,\text{m}^3/\text{yr}$ and their $\Delta V_{rate} = \sim 5.75 \times 10^6\,\text{m}^3/\text{yr}$ is mainly due to their depth estimation. Our InSAR interferogram matches one of the other two by Wicks (2002). For that interferogram, their model gives depths ~1 km shallower, being closer to our depth estimation.

6.2 Time Series of Volume Changes: Regularization Using the Truncated SVD

To assess the effect of the regularization, we compared the increments of volume change (ΔV_i, **Eq. 1**) and the corresponding simulated observations ($\Delta V_{i_{\text{reconstruction}}} =$

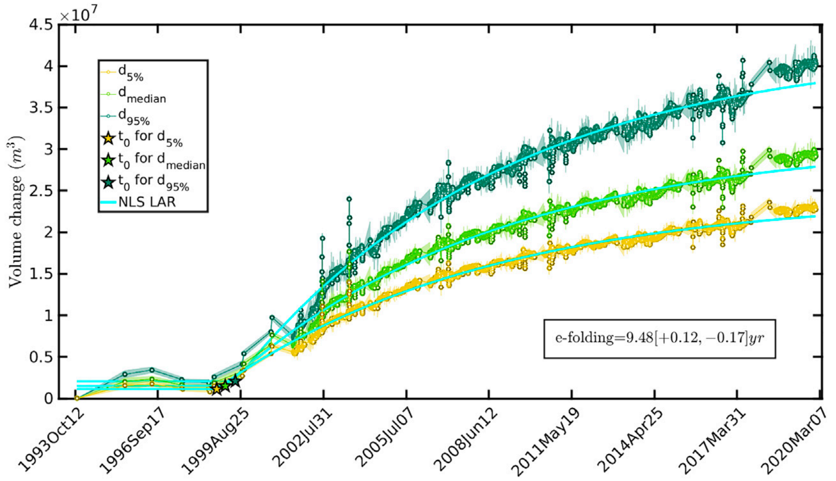


FIGURE 10 | Volume change time series for a Mogi source (constrained by both the GPS and InSAR data), at different depth estimations (**Table 1**): median value (light green), 95 and 5% percentile values (yellow and dark green). Cian solid lines represent predicted curves through Non Linear Least Squares (NLS), Least Absolute Residuals (LAR) method, following the e-folding characteristic shape function of **Eq. 8**. Stars show the location of the beginning of the exponential function: for data associated with median depth value estimation (light green), 95 and 5% percentile values (yellow and dark green).

TABLE 4 | e-folding parameters obtained by NonLinear Least Squares, using the Least Absolute Residual method (LAR).

Value	c (yr^{-1})	1/c(yr)	t_0 (yr)	R–square	RMSE
Optimal	0.1055	9.48	1999.09	0.989	0.60×10^6
Lower bound	0.1042	9.60	1998.75	0.998	0.18×10^6
Upper bound	0.1074	9.31	1999.59	0.988	0.87×10^6

Upper and lower bounds for the curve fitting parameters, encompassing the results for the volume changes series related to the 5%, median and 95% percentile values of source depth estimate.

$V_{(i+1)_{\text{inversion}}} - V_{i_{\text{inversion}}}$ from the results of **Eq. 6**) for different levels of regularization. **Figures 9A, B** show the solution for an extreme regularization, using only the first 0.2% and 2% nonzero singular values (oversmoothing solutions). For the 0.2% case, the values of ΔV_i associated with the InSAR data display a wide dispersion, and the ΔV_i associated with GPS data acquire discrete values, i.e., the same ΔV_i value is obtained for many different time intervals. **Figure 9D** represents the case of non-regularization (less filtering, maximum solution size and minimum misfit). The residual in **Figure 9D** is minimized because the solution reproduces even the seasonal perturbations of the GPS data (undersmoothing solution). There is a seasonal deformation of the crust associated with the surface load of the snow cover. It is possible that the magma reservoir's internal pressure also fluctuates seasonally in response to this effect. However, from CGPS data alone we cannot resolve the cause of these fluctuations. Therefore, a smoother solution is preferred to depict the long-term deformation time series. This is given by the combination of Picard condition and L-curve criteria; it corresponds to the appropriate smoothing solution, i.e, 24.1% nonzero singular values (**Figure 9C**).

6.3 Time-Scales of Inter-Eruptive Uplift Signals: Three Sisters and Other Volcanoes

The continuous and extended regularized time-series of volume change allows us to study the inflation process in detail. Riddick and Schmidt (2011) proposed a piece-wise linear parametrization with two changes in rate provides a good fit to uplift rates till 2009 explaining two different inflation processes beginning at 1998 and 2004 at Three Sisters. This model was supported by the detection of a seismic swarm in 2004. Our denser, longer time-series clearly shows a smooth and continuous function, which we interpret as a fast inflation followed by relaxation of the crust (**Figure 10**). We are specifically interested on the interval of decaying rates. Consequently, the time-series is divided into two main intervals separating increasing and decaying behavior of volume rates. An exponential function can reasonably reproduce the relaxation process. Therefore, we propose a piece-wise parametric of the form:

$$V(t) = \begin{cases} d, & t < t_0 \\ a - b\exp - (c \cdot (t - t_0)), & t \geq t_0 \end{cases} \tag{8}$$

where a, b and d are constants, $1/c = \epsilon$ is the characteristic relaxation time constant, here after named e-folding parameter, and t_0 is the start of the exponential trend. We solved the parameters of this model using a non-linear least-square fit. To minimize the influence of outliers, we used regression method: the Least Absolute Residuals (LAR) and Bisquare weights methods, considering also the data uncertainties (weighted). Four methods (Bisquare, Bisquare-Weighted, LAR, LAR-Weighted) show very similar fit, LAR performing the best (**Figure 10**; **Table 4**). Time-series from the median, and 5% and 95% percentiles of the PDF distribution for depth, along with the variance of the curve fitting, permit

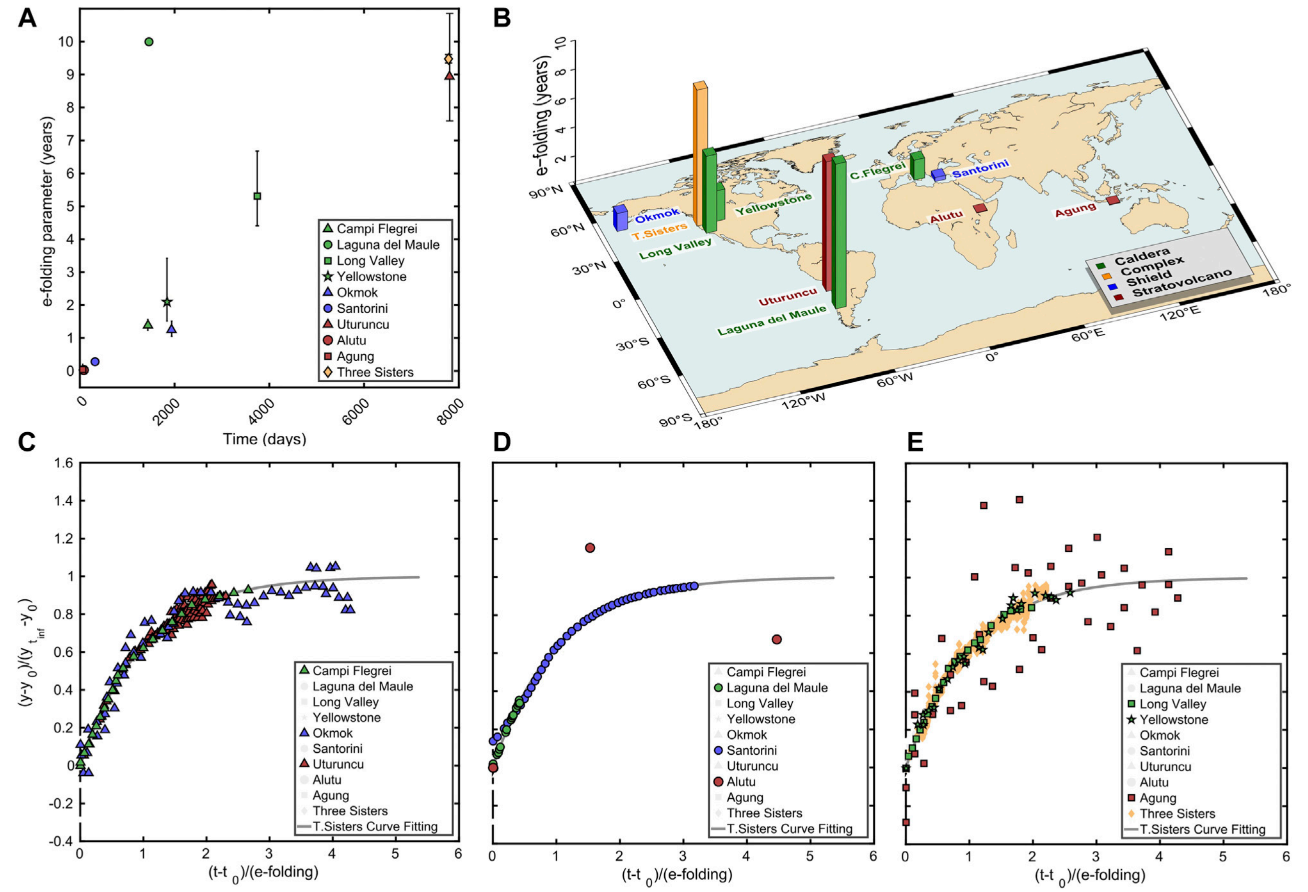


FIGURE 11 | (A) Estimated e-folding parameter of other available geodetic and volume time series from volcanoes with recent and well-studied unrest episodes: Okmok (Biggs et al., 2010), Long Valley (Hill et al., 2020), Uturuncu (Lau et al., 2018), Laguna del Maule (Le Mével et al., 2015), Yellowstone (Tizzani et al., 2015), Campi Flegrei (Troise et al., 2007), Santorini (Parks et al., 2015)), Alutu (Hutchison et al., 2016), Agung (Syahbana et al., 2019). **(B)** Geographic location and classification according to the type of Volcano. **(C, E, F)** Normalized uplift or volume change, *y*′ (**Eq. 9**) as function of normalized time, *t*′ (**Eq. 10**).

constraining the lower and upper limits of e-folding and t_0 parameters.

Our time-series spans 26 years and presents a characteristic time constant of 9.48 [+0.12, −0.17] years. This new result updates the value of 5.3 years obtained by Dzurisin et al. (2009) from an 8-years time-series of GPS displacements from 2001 to late 2008. Riddick and Schmidt (2011) hyphotesized that one injection of magma started between June 1996 and July 1997, given StaMPS results for T385 ERS. The updated volume time-series presented in this study shows a clear exponential decay trend. We estimate a start date for the exponential trend between October 1998 and August 1999 (**Table 4**). These results suggest that the continuous uplift signal will be detectable for a few decades, considering volume change rates as low as 0.1×10^6 m^3. As late as January 2020, our inflation time-series indicated that the cumulative volume of a spherical point source with $d_{\mathrm{median}} =$ 5000 m is 29.1×10^6 m^3. For $d_{5\%} = 4500$ m and $d_{5\%} = 6000$ m, the values are 22.7×10^6 and 39.9×10^6 m^3, respectively. This range of values is consistent with those predicted by Dzurisin et al. (2009) for a prolate spheroid model, 44.9 to 51.6×10^6 m^3 (where the uncertainty is 10–20% of those values). This estimation of the cumulative volume is 10–20 times less than the volume erupted from Mount St. Helens in May 1980 (Wicks, 2002).

To compare the characteristic relaxation time for other volcanoes with recent and well-studied unrest episodes, we compiled and modeled the available geodetic and volume time-series of the following volcanic systems: Okmok (Biggs et al., 2010), Long Valley (Hill et al., 2020), Uturuncu (Lau et al., 2018), Laguna del Maule (Le Mével et al., 2015), Yellowstone (Tizzani et al., 2015), Campi Flegrei (Troise et al., 2007), Santorini (Parks et al., 2015), Alutu (Hutchison et al., 2016), Agung (Syahbana et al., 2019). We have selected unrest periods showing deformation that deviates from the background trend and are characterised by an exponential decay. These unrest periods span from the beginning of the deformation series or from a new increment of uplift, until it occurs a new change in the trend, indicated by a departure from the flattened part of the exponential behavior. Therefore, the analyzed periods are based solely on the exponential shape of data regardless of the unrest outcome (non-eruptive unrest or pre-eruptive unrest) and of the duration of the entire inter-eruptive period. Only a posteriori can it be known whether a period of uplift with exponential decay trend can be taken as indicative or related to volcanic eruptions. Considering this, and to avoid possible bias, we analyze a heterogeneous group of volcanoes (different types, diverse geographical locations, distinct deformation/volume change time-series and different stages within each inter-eruptive period), in order to discern whether the e-folding parameter may be or not a useful parameter for understanding posterior volcanic activity. We estimated the e-folding parameter and t_0 following **Eq. 8** in order to compare different volcanic systems. Greater e-folding indicates longer characteristic relaxation times, as shown in **Figure 11A**. The estimated e-folding parameters vary between 0.033 and 10 years, for uplift and inflation episodes lasting between 60 days and 26 years. Our selected volcanoes, particularly those of the North and South American volcanic arcs, present the longest e-folding times (**Figure 11B**).

From the small selection of volcanoes, neither the type nor composition of the volcanoes seems to be decisive for the characteristic relaxation time associated with their inter-eruptive periods. However, those with a shorter e-folding display significant changes in their behavior. For instance, Campi Flegrei exhibited an increment in displacement from 1968–1972, with an e-folding of 1.38 years. Next, the subsidence rate was small until 1982, and the deformation trend changed due to a new uplift episode during 1982–1984, followed by subsidence during 1985–2004. Those features were related to an overpressure in the magmatic source and fracturing of the rock volume between the magmatic fluids and the aquifer (Troise et al., 2007). Alutu underwent two inflation pulses, the latest showing a short e-folding of 0.033 years, during the period October-December 2008, then a slow deflation took place. These short time-scale suggest the migration of hydrothermal or magmatic fluids or volatiles (Hutchison et al., 2016). Agung went through an uplift from August-October 2017, characterized by an e-folding of 0.038 years, then in late November, a phreatomagmatic eruption and stronger explosions took place (Syahbana et al., 2019). Santorini presented a source inflation process with an e-folding of 0.28 years for the period October 2011- August 2012 (Parks et al., 2015). Then, its subsidence rates increased in the post-unrest period 2012–2017, suggesting the superimposition of various deformation sources (Papageorgiou et al., 2019). The Okmok inflation episode during the period May 2002 - September 2007 had an e-folding of 1.24 years. Although there was a small amount of deflation (Biggs et al., 2010), the general trend can be modeled as an exponential decay. After this inter-eruptive episode, a phreatomagmatic eruption occurred in July–August 2008 (Larsen et al., 2015). Long Valley deformation series presents an e-folding of 5.31 years for the period 1978–1988. No significant seismicity was detected during the interval ~1986–1988. After the exponential decaying trend, there was a renewed unrest, characterized by recurring earthquake swarms and tumescence of the resurgent dome (Hill et al., 2020). Yellowstone went through an uplift during 2005–2010, with an estimated e-folding of 2.1 years. Since 2015, subsidence of Yellowstone caldera has occurred at an average rate of 2–3 cm per year, as reported by Yellowstone Volcano Observatory (USGS). Laguna del Maule is the only volcano that yields a high relaxation time value for a short inter-eruptive period (2010–2014), according to Le Mével et al. (2015). However, the fit of the data for this period could also be due to a linear inflation pulse superimposed on an exponentially decaying deformation rate. On the other hand, Uturuncu and Three Sisters present the longest e-folding (8.93 years and 9.48 years), without showing significant changes in their volcanic activity. Although the e-folding parameter seems to be an informative variable in the magnitude of the inter-eruptive period time scales, it does not provide any parametric information that differentiates pre-eruptive unrest (Agung, Okmok) from non-eruptive unrest (e.g. Alutu, Campi Flegrei).

We re-scaled the observed time-series to properly emphasize similarities on the exponential decay. We normalize displacement or volume change (y'), as a function of normalized time by means of the e-folding parameter (t'):

$$y'(t) = \left(y(t) - y_0\right) \big/ \left(y_{t_\infty} - y_0\right) \tag{9}$$

$$t' = (t - t_0)/\epsilon \tag{10}$$

where y_0 is the displacement or volume change at t_0, y_{t_∞} is the value after total relaxation (i.e., at $t = t_\infty$), and t_0 is the onset of the exponential function. It is worth noting that using the e-folding parameter accurately represents the characteristic relaxation times, and hence re-scales invariantly the observations.

Figures 11C,D,E show the resulting normalized time-series for each volcano, revealing a strikingly similar pattern. This behavior seems to be independent of the e-folding value or duration of the inter-eruptive episode. Accordingly, the temporal invariance could indicate that there is a limited set of physical scenarios underlying inter-eruptive inflation episodes. This evidence seems to support inter-eruptive physical processes with exponential time-dependent solution. Several physical models could explain deformation with this pattern (e.g., Lengliné et al., 2008; Reverso et al., 2014; Walwer et al., 2019). Dzurisin et al. (2009) put forward several mechanistic explanations for Three Sisters: 1) hydraulic pressurization or instantaneous response of the crust to continued intrusion at depth; 2) pressurization of the hydrothermal system; 3) viscoelastic response of the crust due to an intrusion emplaced at depth. Our analyses estimate the start of the exponential decay around 1998–1999, in agreement with previous studies (Wicks, 2002; Dzurisin et al., 2006; Dzurisin et al., 2009; Riddick and Schmidt, 2011). Although not unique on which physical mechanism is acting at Three Sisters, the observed uplift decay is consistent with a viscoelastic response of the crust.

The lack of hot springs and thermal pools rules out that the deformation could be due to an active hydrothermal system (Wicks, 2002). Recent rhyolitic eruptions close to South Sisters, 2000 years ago (Hildreth et al., 2012), and the evidence of a long-lived magmatic source at depth in Three Sisters, from spring geochemistry studies (Evans et al., 2004), are indicative of magma evolving during thousand of years before eruption. This is compatible with the formation of viscoelastic aureoles as a result of the alteration of the mineralogical and rheological properties of the surrounding rocks. Previously, Zurek et al. (2012) concluded, based on the lack of gravity change from 2002–2009, that deformation at Three Sisters reflects a viscoelastic response of the crust to an intrusion of magma. The estimated e-folding indicates viscosities of $\sim 10^{18}$ Pa *s* of a viscoelastic shell surrounding the magmatic source with a pressure change which increases in finite time from 0 to a constant value, considering theoretical estimations on the time constants and viscoelastic medium parameters (e.g., Newman et al., 2001; Bonafede and Ferrari, 2009; Del Negro et al., 2009; Segall, 2016).

7 CONCLUSION

The evolution of volume change time-series at active volcanoes can be studied by combining heterogeneous geodetic datasets. For Three Sisters volcano, we combined high spatial coverage from multiple InSAR satellite data and long term temporal information on the three-components of the only available continuous GPS. We improved a previous two-step approach to volume time-series reconstruction, by including a non-arbitrary truncation level. The cut-off criterion for truncation (i.e., type of filter) is necessary to obtain a solution without too much loss of resolution affecting the stability of the inversion. We proposed a method that combines the Discrete Picard Condition and the L-curve. Furthermore, our approach takes propagation errors into account in all inversion steps. The final time-series is determined considering volume change rates instead of increments of displacement, avoiding problems deriving from the amplification of uncorrelated noise between adjacent GPS data or propagation through the time-series of the uncertainty of the first acquisition.

The inflation time-series of Three Sisters since 2018 shows a noticeable change in the trend, which departs from the previous asymptotic trend toward a constant decay rate. This change can be explained by a fixed step in the position, such as that caused by a change in the instrumentation or monument stability. However, we cannot rule out a minor injection of magma or fluid pressurization beneath Three Sisters. Considering the wide range of erupted magma compositions and eruption styles, any changes in the Three Sisters background uplift behavior should be evaluated as an important indicator of future volcanic activity.

The Three Sisters volcano uplift is still on-going. The Bayesian inversion of source parameters gives 95% credible intervals, the depth for a spherical point source being between (4500–6000)m. Parametric modeling of the inflation time-series associated with the median, and 5% and 95% percentiles of source depth allows us to constrain the onset of the exponential trend to between October 1998 and August 1999 and the characteristic relaxation time to 9.48 $[+0.12, -0.17]$ years. Therefore, in the absence of different or new unrest signals, we estimate a continuous uplift signal, at diminishing but detectable rates, lasting for few decades (currently estimate to 2054 $[\pm 2 \text{ years}]$). The observed uplift decay is consistent with a viscoelastic response of the crust. Our future efforts will be focused on elucidating whether the deformation could be a viscoelastic response to a very rapid magma emplacement or to several years of active intrusion of magma.

This analysis is a step toward understanding the time-scale of inter-eruptive processes. Inter-eruptive uplift/volume change signals of analyzed volcanoes show rather simple and time-scale invariant behavior, after a proper scaling. We interpret this observation as pointing to a rather reduced set of physical mechanisms underlying inter-eruptive inflation episodes that are consistent with exponential decay (viscoelastic response and/or hydraulic pressurization). Furthermore, we suggest that the magnitude of the characteristic relaxation time can be indicative of significant changes of the background behavior on volcanoes. Temporally persistent, long-lasting and overlapping uplift signals are potential confounding indicators for the classic inflation-eruption-deflation cycle model. We highlight the importance of high-temporal and continuous surface deformation monitoring to identify any departures from background temporal behavior (potentially very complex), as an indicator of future eruption hazard in persistent uplifting volcanoes. In regards to eruption forecasting, the uplift/inflation itself cannot be used as a pre-eruptive precursor without knowing what controls it through the combination of petrological and/or geophysical data.

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 below: JAXA/METI ALOS PALSAR L1.0 was accessed through ASF DAAC, https://asf.alaska.edu [June 2018] doi:10.5067/J4JVCFDDPEW1. ERS-1, ESA [1997–2000] downloaded from ASF DAAC, https://asf.alaska.edu [June 2018]. Copernicus Sentinel data was retrieved from ASF DAAC, https://asf.alaska.edu [June 2018]. This work contains modified Copernicus Sentinel data [2014–2018], processed by ESA. HUSB CGPS data and errors were obtained from the Pacific Northwest network on the USGS website https://earthquake.usgs.gov/monitoring/gps/Pacific_Northwest/husb.

AUTHOR CONTRIBUTIONS

SR-M processed radar images, performed modeling, carried out the comparison among different volcanic systems, prepared the figures and co-wrote the manuscript. PG and DS processed radar images. MC and PG conceptualized the study, interpreted results and co-wrote the manuscript. AN co-supervised SR-M and contributed to the manuscript. DS contributed to the manuscript.

FUNDING

Part of this work was carried out during research visits of SR-M to the University of Liverpool and the U.S. Geological Survey (USGS) Cascades Volcano Observatory (CVO). PG was supported by the United Kingdom NERC project Centre for the Observation and Modeling of Earthquakes, Volcanoes and Tectonics (COMET, GA/13/M/031, comet.nerc.ac.uk). SR-M, MC, and AN research was supported by MINECO projects CGL2014-58821-C2-1-R and PGC2018-095154-B-I00. SR-M was also supported by MINECO fellowship BES-2015-074228.

ACKNOWLEDGMENTS

We would like to thank Mike Lisowsky for providing the predicted horizontal tectonic motion at HUSB. We would like to thank Yu Jiang for his support in InSAR processing and Xiao Wang for his help with the University of Liverpool, School of Environmental Sciences computing resources. This manuscript was edited by Guido Jones, currently funded by the Cabildo de Tenerife, under the TFinnova Program supported by MEDI and FDCAN funds. We also acknowledge Maurizio Battaglia, Gilda Gurrenti and Valerio Acocella their thoughtful comments and suggestions. This is a contribution of the CSIC Thematic Platform Volcanism and Society (www.ptivolcan.es).

REFERENCES

Acocella, V., Di Lorenzo, R., Newhall, C., and Scandone, R. (2015). An overview of recent (1988 to 2014) caldera unrest: knowledge and perspectives. *Rev. Geophys.* 53, 896–955. doi:10.1002/2015RG000492

Backus, G. E. (1988). Bayesian inference in geomagnetism. *Geophys. J. Int.* 92, 125–142. doi:10.1111/j.1365-246X.1988.tb01127.x

Bagnardi, M., and Hooper, A. (2018). Inversion of surface deformation data for rapid estimates of source parameters and uncertainties: a bayesian approach. *G-cubed* 19, 2194–2211. doi:10.1029/2018GC007585

Battaglia, M., Cervelli, P., and Murray, J. (2013). DMODELS: a MATLAB software package for modeling crustal deformation near active faults and volcanic centers. *J. Volcanol. Geoth. Res.* 254, 1–4. doi:10.1016/j.jvolgeores.2012.12.018

Bekaert, D. P. S., Hamlington, B. D., Buzzanga, B., and Jones, C. E. (2017). Spaceborne synthetic aperture radar survey of subsidence in Hampton roads, Virginia (United States). *Sci. Rep.* 7, 14752. doi:10.1038/s41598-017-15309-5

Berardino, P., Fornaro, G., Lanari, R., and Sansosti, E. (2002). A new algorithm for monitoring localized deformation phenomena based on small baseline differential SAR interferograms. *IEEE Internat. Geosci. Remote Sens. Symp.* 2, 2375–2383. doi:10.1109/IGARSS.2002.1025900

Biggs, J., and Pritchard, M. (2017). Global volcano monitoring: what does it mean when volcanoes deform? *Elements.* 13, 17–22. doi:10.2113/gselements.13.1.17

Biggs, J., Lu, Z., Fournier, T., and Freymueller, J. T. (2010). Magma flux at Okmok Volcano, Alaska, from a joint inversion of continuous GPS, campaign GPS, and interferometric synthetic aperture radar. *J. Geophys. Res.: Solid Earth* 115, 1–11. doi:10.1029/2010JB007577

Bonafede, M., and Ferrari, C. (2009). Analytical models of deformation and residual gravity changes due to a Mogi source in viscoelastic medium. *Tectonophysics* 471, 4–13. doi:10.1016/j.tecto.2008.10.006

Chen, C. W., and Zebker, H. A. (2002). Phase unwrapping for large SAR interferograms: statistical segmentation and generalized network models. *IEEE Trans. Geosci. Rem. Sens.* 40, 1709–1719. doi:10.1109/TGRS.2002.802453

Decriem, J., Árnadóttir, T., Hooper, A., Geirsson, H., Sigmundsson, F., Keiding, M., et al. (2010). The 2008 May 29 earthquake doublet in SW Iceland. *Geophys. J. Internat.—Geophys. J. Int.* 181, 1128–1146. doi:10.1111/j.1365-246X.2010.04565.x

Del Negro, C., Currenti, G., and Scandura, D. (2009). Temperature-dependent viscoelastic modeling of ground deformation: application to Etna volcano during the 1993–1997 inflation period. *Phys. Earth Planet. In.* 172, 299–309. doi:10.1016/j.pepi.2008.10.019

Dieterich, J. H., and Decker, R. W. (1975). Finite element modeling of surface deformation associated with volcanism. *J. Geophys. Res. (1896–1977).* 80, 4094–4102. doi:10.1029/JB080i029p04094

Dzurisin, D., Lisowski, M., and Wicks, C. W. (2009). Continuing inflation at Three Sisters volcanic center, central Oregon cascade range, United States, from GPS, leveling, and InSAR observations. *Bull. Volcanol.* 71, 1091–1110. doi:10.1007/s00445-009-0296-4

Dzurisin, D., Lisowski, M., Wicks, C. W., Poland, M. P., and Endo, E. T. (2006). Geodetic observations and modeling of magmatic inflation at the Three Sisters volcanic center, central Oregon cascade range, United States. *J. Volcanol. Geoth. Res.* 150, 35–54. doi:10.1016/j.jvolgeores.2005.07.011

Dzurisin, D., Lisowski, M., and Wicks, C. W. (2017). Semipermanent GPS (SPGPS) as a volcano monitoring tool: rationale, method, and applications. *J. Volcanol. Geoth. Res.* 344, 40–51. doi:10.1016/j.jvolgeores.2017.03.007

Evans, W., van Soest, M., Mariner, R., Hurwitz, S., Ingebritsen, S., Wicks, C., et al. (2004). Magmatic intrusion west of Three Sisters, central Oregon, United States: the perspective from spring geochemistry. *Geology* 32, 69–72. doi:10.1130/G19974.1

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar topography mission. *Rev. Geophys.* 45, RG2004. doi:10.1029/2005RG000183

Fialko, Y., Khazan, Y., and Simons, M. (2001). Deformation due to a pressurized horizontal circular crack in an elastic half-space, with applications to volcano geodesy. *Geophys. J. Int.* 146 (1), 181–190. doi:10.1046/j.1365-246X.2001.00452.x

Fierstein, J., Hildreth, W., and Calvert, A. T. (2011). Eruptive history of South sister, Oregon cascades. *J. Volcanol. Geoth. Res.* 207, 145–179. doi:10.1016/j.jvolgeores.2011.06.003

Galetto, F., Acocella, V., and Caricchi, L. (2017). Caldera resurgence driven by magma viscosity contrasts. *Nat. Commun.* 8, 1750. doi:10.1038/s41467-017-01632-y

González, P. J., Samsonov, S. V., Pepe, S., Tiampo, K. F., Tizzani, P., Casu, F., et al. (2013). Magma storage and migration associated with the 2011–2012 El Hierro eruption: implications for crustal magmatic systems at oceanic island volcanoes. *J. Geophys. Res.: Solid Earth* 118, 4361–4377. doi:10.1002/jgrb.50289

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* 34, 561–580. doi:10.1137/1034115

Hansen, P. C. (2007). Regularization tools version 4.0 for matlab 7.3. *Numer. Algorithm.* 46, 189–194. doi:10.1007/s11075-007-9136-9

Hansen, P. C. (1990). The discrete picard condition for discrete ill-posed problems. *Bit.* 30, 658–672. doi:10.1007/BF01933214

Hansen, P. C., and O'Leary, D. P. (1993). The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.* 14, 1487–1503. doi:10.1137/0914086

Hildreth, W. (2007). *Quaternary magmatism in the cascades-geologic perspectives.* USGS. Professional Paper Vol. 1744, 125.

Hildreth, W., Fierstein, J., and Calvert, A. T. (2012). *Geologic map of three sisters volcanic cluster, cascade range, Oregon*. US Geological Survey Scientific Investigations Map, Vol. 3186, pamphlet 107. p., 2 sheets, scale 1:24,000. Available at https://pubs.usgs.gov/sim/3186/.

Hill, D., Montgomery-Brown, E., Shelly, D., Flinders, A., and Prejean, S. (2020). Post-1978 tumescence at Long Valley caldera, California: a geophysical perspective. *J. Volcanol. Geoth. Res.* 400, 106900. doi:10.1016/j.jvolgeores.2020.106900

Hooper, A., Bekaert, D., Spaans, K., and Arikan, M. (2012). Recent advances in SAR interferometry time series analysis for measuring crustal deformation. *Tectonophysics* 514–517, 1–13. doi:10.1016/j.tecto.2011.10.013

Hossainali, M., Becker, M., and Groten, E. (2010). Comprhensive approach to the analysis of the 3D kinematics deformation with appliction to the kenai Peninsula. *J. Geodetic Sci.* 1, 59–73. doi:10.2478/v10156-010-0008-1

Hutchison, W., Biggs, J., Mather, T. A., Pyle, D. M., Lewi, E., Yirgu, G., et al. (2016). Causes of unrest at silicic calderas in the East African Rift: new constraints from InSAR and soil-gas chemistry at Aluto volcano, Ethiopia. *G-cubed.* 17, 3008–3030. doi:10.1002/2016GC006395

Larsen, J., Neal, C., Schaefer, J., Kaufman, A., and Lu, Z. (2015). The 2008 phreatomagmatIc eruption of Okmok volcano, AleutIan Islands, Alaska: chronology, deposits, and landform changes. *Bull. Volcanol.* 77, 69. doi:10.1007/s00445-015-0952-9

Lau, N., Tymofyeyeva, E., and Fialko, Y. (2018). Variations in the long-term uplift rate due to the Altiplano–Puna magma body observed with Sentinel-1 interferometry. *Earth Planet Sci. Lett.* 491, 43–47. doi:10.1016/j.epsl.2018.03.026

Lawless, J., and Wang, P. (1976). A simulation study of ridge and other regression estimators. *Commun. Statist. Theor. Meth*, A5, 307–323. doi:10.1080/03610927608827353

Le Mével, H., Feigl, K. L., Córdova, L., DeMets, C., and Lundgren, P. (2015). Evolution of unrest at Laguna del Maule volcanic field (Chile) from InSAR and GPS measurements, 2003 to 2014. *Geophys. Res. Lett.* 42, 6590–6598. doi:10.1002/2015GL064665

Lengliné, O., Marsan, D., Got, J.-L., Pinel, V., Ferrazzini, V., and Okubo, P. G. (2008). Seismicity and deformation induced by magma accumulation at three basaltic volcanoes. *J. Geophys. Res.: Solid Earth* 113. B12305. doi:10.1029/2008JB005937

Lohman, R. B., and Simons, M. (2005). Some thoughts on the use of InSAR data to constrain models of surface deformation: noise structure and data downsampling. *Geochem. Geophys. Geosyst.* 6, Q01007. doi:10.1029/2004GC000841

Mccaffrey, R., Qamar, A., King, R., Wells, R., Khazaradze, G., Williams, C., et al. (2007). Fault locking, block rotation and crustal deformation in the Pacific Northwest. *Geophys. J. Internat. – Geophys. J. Iint.* 169, 1315–1340. doi:10.1111/j.1365-246X.2007.03371.x

Menke, W. (1989). *Geophysiocal data analysis: discrete inverse theory*. San Diego, Calif: Academic.

Mogi, K. (1958). Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bull. Earthquake Res. Inst. Univ. Tokyo*. 36, 99–134.

Moran, S. C. (2004). *Seismic monitoring at cascade volcanic centers, 2004- status and recommendationsc*. US Geological Survey Scientific Investigations Report, 2004-5211, 28.

Newman, A., Dixon, T., Ofoegbu, G., and Dixon, J. (2001). Geodetic and seismic constraints on recent activity at Long Valley Caldera, California: evidence for viscoelastic rheology. *J. Volcanol. Geoth. Res.* 105, 183–206. doi:10.1016/S0377-0273(00)00255-9

Papageorgiou, E., Foumelis, M., Trasatti, E., Guido, V., Raucoules, D., and Mouratidis, A. (2019). Multi-sensor SAR geodetic imaging and modelling of Santorini volcano post-unrest response. *Rem. Sens.* 11, 259. doi:10.3390/rs11030259

Parks, M., Moore, J., Papanikolaou, X., Biggs, J., Mather, T., Pyle, D., et al. (2015). From quiescence to unrest: 20 years of satellite geodetic measurements at Santorini volcano, Greece. *J. Geophys. Res. B: Solid Earth*. 120, 1309–1328. doi:10.1002/2014JB011540

Poland, M. P., Lisowski, M., Dzurisin, D., Kramer, R., McLay, M., and Pauk, B. (2017). Volcano geodesy in the cascade arc, United States. *Bull. Volcanol.* 79. doi:10.1007/s00445-017-1140-x

Reverso, T., Vandemeulebrouck, J., Jouanne, F., Pinel, V., Villemin, T., Sturkell, E., et al. (2014). A two-magma chamber model as a source of deformation at Grímsvötn Volcano, Iceland. *J. Geophys. Res.: Solid Earth* 119, 4666–4683. doi:10.1002/2013JB010569

Riddick, S. N., and Schmidt, D. A. (2011). Time-dependent changes in volcanic inflation rate near Three Sisters, Oregon, revealed by InSAR. *G-cubed*. 12, Q12005. doi:10.1029/2011GC003826

Rosen, P. A., Gurrola, E., Sacco, G. F., and Zebker, H. (2012). "The InSAR scientific computing environment," in EUSAR 2012; 9th European conference on synthetic aperture radar, 23-26 April 2012, Nuremberg, Germany 730–733.

Rosen, P., Henley, S., Peltzer, G., and Simons, M. (2004). Updated repeat orbit interferometry package released. *Eos, Trans. American Geophys. Union*. 85, 47. doi:10.1029/2004EO050004

Schmidt, M., and Grunder (2011). Deep mafic roots to arc volcanoes: mafic recharge and differentiation of basaltic andesite at north Sister volcano, Oregon cascades. *J. Petrol.* 52, 603–641. doi:10.1093/petrology/egq094

Scott, W. E., Iverson, R., Schilling, S., and Fisher, B. (2001). *Volcano hazards in the three sisters region, Oregon*. US Geological Survey Open-File Report, Vol. 18, 99–437. doi:10.3133/ofr99437

Segall, P. (2016). Repressurization following eruption from a magma chamber with a viscoelastic aureole. *J. Geophys. Res.: Solid Earth* 121, 8501–8522. doi:10.1002/2016JB013597

Sherrod, D. R., Taylor, E. M., Ferns, M. L., Scott, W. E., Conrey, R. M., and Smith, G. A. (2004). *Geologic map of the Bend 30- × 60-minute quadrangle, central Oregon*. US Geological Survey Geological Investigations Series Map I-2683, 48.

Syahbana, D., Kasbani, K., Suantika, G., Prambada, O., Andreas, A., Saing, U., et al. (2019). The 2017–19 activity at Mount Agung in Bali (Indonesia): intense unrest, monitoring, crisis response, evacuation, and eruption. *Sci. Rep.* 9, 8848. doi:10.1038/s41598-019-45295-9

Tikhonov, A., and Arsenin, V. (1977). *Solutions of ill-posed problem*. New York: John Wiley & Sons.

Tizzani, P., Castaldo, R., Pepe, A., Zeni, G., Lanari, R., and Battaglia, M. (2015). Magma and fluid migration at Yellowstone Caldera in the last three decades inferred from InSAR, leveling and gravity measurements. *J. Geophys. Res.: Solid Earth* 120, 2627–2647. doi:10.1002/2014JB011502

Troise, C., De Natale, G., Pingue, F., Obrizzo, F., De Martino, P., Tammaro, U., et al. (2007). Renewed ground uplift at Campi Flegrei caldera (Italy): new insight on magmatic processes and forecast. *Geophys. Res. Lett.* 34, L03301. doi:10.1029/2006GL028545

Walwer, D., Ghil, M., and Calais, E. (2019). Oscillatory nature of the Okmok volcano's deformation. *Earth Planet Sci. Lett.* 506, 76–86. doi:10.1016/j.epsl.2018.10.033

Weiss, J., Walters, R., Morishita, Y., Wright, T., Lazecky, M., Wang, H., et al. (2020). High-resolution surface velocities and strain for Anatolia from Sentinel-1 InSAR and GNSS data. *Geophys. Res. Lett.* 47, e2020GL087376. doi:10.1029/2020GL087376

Wicks, C. W. (2002). Magmatic activity beneath the quiescent Three Sisters volcanic center, central Oregon cascade range, United States. *Geophys. Res. Lett.* 29, 26-1–26-4. doi:10.1029/2001GL014205

Yang, X.-M., Davis, P. M., and Dieterich, J. H. (1988). Deformation from inflation of a dipping finite prolate spheroid in an elastic half-space as a model for volcanic stressing. *J. Geophys. Res.: Solid Earth* 93, 4249–4257. doi:10.1029/JB093iB05p04249

Zurek, J., Williams-Jones, G., Johnson, D., and Eggers, A. (2012). Constraining volcanic inflation at Three Sisters volcanic field in Oregon, United States, through microgravity and deformation modeling. *G-cubed* 13, Q10013. doi:10.1029/2012GC004341

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.

frontiers in Earth Science

ORIGINAL RESEARCH
published: 28 January 2021
doi: 10.3389/feart.2020.581742

Clustering of Experimental Seismo-Acoustic Events Using Self-Organizing Map (SOM)

Flora Giudicepietro[1], Antonietta M. Esposito[1]*, Laura Spina[2], Andrea Cannata[3,4], Daniele Morgavi[5], Lukas Layer[1,6,7] and Giovanni Macedonio[1]

[1]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Napoli, Italy, [2]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Roma1, Roma, Italy, [3]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo-Sezione di Catania, Catania, Italy, [4]Università degli Studi di Catania, Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Catania, Italy, [5]Università degli Studi di Perugia, Dipartimento di Fisica e Geologia, Perugia, Italy, [6]Università degli Studi di Napoli "Federico II", Napoli, Italy, [7]Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Padova, Italy

The analogue experiments that produce seismo-acoustic events are relevant for understanding the degassing processes of a volcanic system. The aim of this work is to design an unsupervised neural network for clustering experimental seismo-acoustic events in order to investigate the possible cause-effect relationships between the obtained signals and the processes. We focused on two tasks: 1) identify an appropriate strategy for parameterizing experimental seismo-acoustic events recorded during analogue experiments devoted to the study of degassing behavior at basaltic volcanoes; 2) define the set up of the selected neural network, the Self-Organizing Map (SOM), suitable for clustering the features extracted from the experimental events. The seismo-acoustic events were generated using an ad hoc experimental setup under different physical conditions of the analogue magma (variable viscosity), injected gas flux (variable flux velocity) and conduit surface (variable surface roughness). We tested the SOMs ability to group the experimental seismo-acoustic events generated under controlled conditions and conduit geometry of the analogue volcanic system. We used 616 seismo-acoustic events characterized by different analogue magma viscosity (10, 100, 1000 Pa s), gas flux (5, 10, 30, 60, 90, 120, 150, 180 × 10^{-3} l/s) and conduit roughness (i.e. different fractal dimension corresponding to 2, 2.18, 2.99). We parameterized the seismo-acoustic events in the frequency domain by applying the Linear Predictive Coding to both accelerometric and acoustic signals generated by the dynamics of various degassing regimes, and in the time domain, applying a waveform function. Then we applied the SOM algorithm to cluster the feature vectors extracted from the seismo-acoustic data through the parameterization phase, and identified four main clusters. The results were consistent with the experimental findings on the role of viscosity, flux velocity and conduit roughness on the degassing regime. The neural network is capable to separate events generated under different experimental conditions. This suggests that the SOM is appropriate for clustering natural events such as the seismo-acoustic transients accompanying Strombolian explosions and that the adopted parameterization strategy may be suitable to extract the significant features of the seismo-acoustic (and/or infrasound) signals linked to the physical conditions of the volcanic system.

Keywords: self-organizing map, neural network, seismo-acoustic signals, experimental volcanology, clustering method

Edited by:
Nico Fournier,
GNS Science, New Zealand

Reviewed by:
Oliver D. Lamb,
University of North Carolina at Chapel Hill, United States
Arthur Jolly,
GNS Science, New Zealand
John J Lyons,
U.S. Geological Survey, United States

****Correspondence:***
Antonietta M. Esposito
antonietta.esposito@ingv.it

Specialty section:
This article was submitted to Volcanology,
a section of the journal
Frontiers in Earth Science

Received: *09 July 2020*
Accepted: *21 December 2020*
Published: *28 January 2021*

Citation:
Giudicepietro F, Esposito AM, Spina L, Cannata A, Morgavi D, Layer L and Macedonio G (2021).
Clustering of Experimental Seismo-Acoustic Events Using Self-Organizing Map (SOM).
Front. Earth Sci. 8:581742.
doi: 10.3389/feart.2020.581742

INTRODUCTION

In recent years, neural networks have been increasingly used thanks to the rapid progress of computer performances and the continuous growth of digital data worldwide, which are difficult to analyze with traditional search and classification methods. In the field of geophysics and volcanology, both supervised and unsupervised neural networks have been used in many applications, including tasks for the classification and recognition of seismic signals in volcanic (Carniel, 1996; Del Pezzo et al., 2003; Scarpetta et al., 2005; Ersoy et al., 2007; Langer et al., 2009; Messina and Langer, 2011; Esposito et al., 2013a; Carniel et al., 2013; Unglert et al., 2016; Burzynski et al., 2018; Ren et al., 2020; Watson, 2020) and tectonic (Köhler et al., 2009; Giudicepietro et al., 2017) environments, analysis of petrographic patterns (Esposito et al., 2020a; Esposito et al., 2020b), predictive analysis of seismicity (Esposito et al., 2014) and radon emissions (Ambrosino et al., 2020). In particular, among unsupervised neural networks, the Self-Organizing Map (SOM) is suitable for the discrimination of seismic signals generated by different sources in a composite seismic wavefield. For instance, several neural network based methods have been applied to study the seismicity of Stromboli (Esposito et al., 2006a, Esposito et al., 2008; Esposito et al., 2013b, Esposito et al., 2018), that is an example of seismo-acoustic wavefield dominated by signals produced by different sources linked to the degassing through a basaltic magma. More in general, the seismic and acoustic wavefield of an open conduit volcano might be originated from a wide spectrum of processes; unsupervised neural networks are fundamental for discriminating different sources of signals. Noteworthy, the investigation of seismo-acoustic transients, related to unsteady explosive activity, is known to provide fundamental information on the degassing dynamics also at other volcanoes such as Erebus (Rowe et al., 2000; Johnson et al., 2008) and Yasur (Spina et al., 2015; Capponi et al., 2016; Simons et al., 2020).

In parallel, analogue models of volcanic degassing have been developed to better understand the dynamics of open conduit volcanoes and define the elastic signature of magma-gas interaction with the volcanic conduit (e.g. James et al., 2004; James et al., 2006; Arciniega-Ceballous et al., 2014; Spina et al., 2018; Spina et al., 2019). Indeed, the laboratory approach offers the advantage of controlling physical properties and the thermodynamic state of the investigated systems. This is convenient for investigating physical processes whose controlling parameters are inaccessible to direct observation, as in a volcanic environment. The potentiality of constraining subsurface degassing dynamics by laboratory experiments on accurately scaled two-phase analogue systems has induced a flourishing number of literature studies focusing on the source of acoustic events (e.g. Vidal et al., 2006; Vidal et al., 2009; Divoux et al., 2008; Kobayashi et al., 2010; Lyons et al., 2013) or on the seismic signature of the ascent of volcanic gas slugs (e.g. James et al., 2004; James et al., 2006) and of the fragmentation processes (Arciniega-Ceballos et al., 2014; Arciniega-Ceballos et al., 2015). Moving from the evidence that conduit discontinuities are specific sites where pressure and momentum changes in the fluid are effectively coupled to the Earth (Chouet and Matoza, 2013 and references therein), Spina et al. (2019) assessed the effect of cross-sectional changes in the conduit geometry (in terms of irregularity of the conduit surface) and of physical properties on radiated seismic and acoustic signals. Noteworthy, cross-sectional changes of conduit shape are expected to result from the coupling between conduit shape and volcanic flow (Macedonio et al., 1994).

Innovative approaches to data analysis such as machine learning are important for studying the signals produced by analogue experiments. Actually, these data analysis techniques help to discover the relationships between experimental data and physical processes. The main objective of this work is to identify data preprocessing and clustering techniques capable to highlight a similarity among signals obtained from experiments featuring analogous starting conditions. In this way, the data similarity must reflect the activity of similar source processes, establishing a link between the signals that we can record and the physical processes that generate them, which in the case of natural systems are unknown. In this work, we apply a SOM based method for clustering a dataset of experimental events obtained through a novel experimental protocol described in Spina et al. (2019) suitable to investigate the seismo-acoustic signals generated at different experimental conditions (gas flux and liquid viscosity), taking into account analogue conduits with different extent of roughness of the internal wall. The dataset used in this work was produced by analogue experiments aimed at investigating the interplay among surface irregularity of the conduit, physical properties of the analogue magma in a range of viscosities and gas flux representative of basaltic systems (Spina et al., 2018, Spina et al., 2019).

Initially, we used two well established signal preprocessing approaches (Del Pezzo et al., 2003; Scarpetta et al., 2005) to extract the appropriate features to represent in a compressed form the significant characteristics of the signals such as the spectral content and the temporal evolution of the waveform. Then, we designed a SOM to group the experimental events. The results indicate that we are able to discriminate between analogue events generated under different experimental conditions, and thus this methodology may be applicable to volcanic datasets.

MATERIALS AND METHODS

The Experimental Setup

The experimental setup for generating seismo-acoustic events (**Figure 1**) is fully described in Spina et al. (2018), Spina et al. (2019) and consists of two main parts: 1) the analogue volcano that reproduces volcanic degassing phenomena, and 2) the sensor system, used to measure the seismo-acoustic signals, i.e. the elastic energy radiated by the degassing processes.

Part (1) is made up of a compressor system, connected to a set of flow-meters and injecting the air-gas into an epoxy conduit with a mean diameter of 3 cm and length of 80 cm. In order to investigate the role of conduit irregularity, a novel technique was developed to build epoxy conduits with different fractal

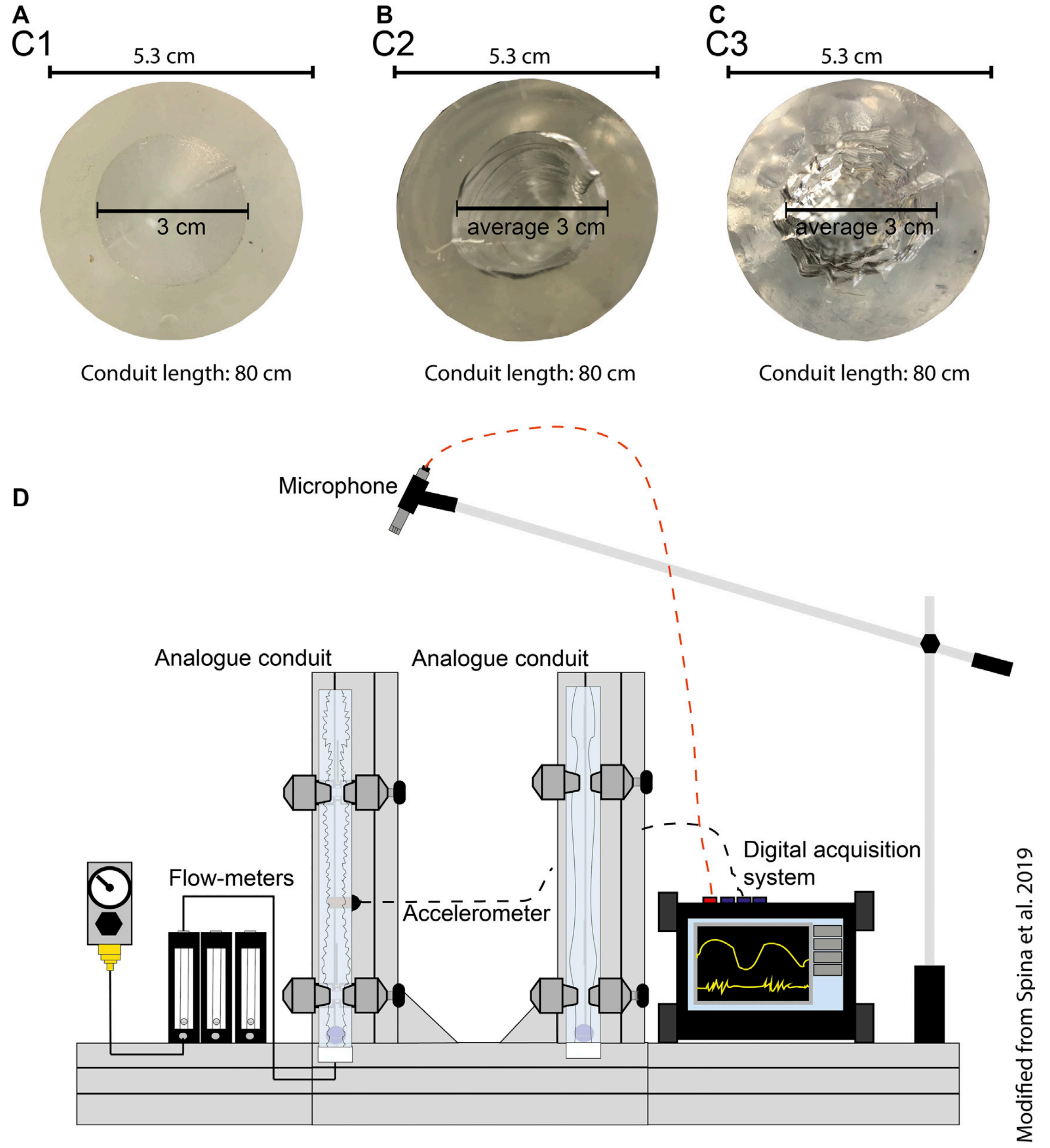


FIGURE 1 | (A–C) Top pictures of the analogue conduits SM (FD = 2), F1 (FD = 2.18) and F2 (FD = 2.99), respectively. **(D)** Simplified sketch of the experimental device (not to scale; modified from Spina et al., 2019). The tubes shown in **(D)** are 80 cm high. Modified from Spina et al. (2019).

dimension (FD) of the internal surface. Three conduits were used to carry out the experiments: a smooth cylinder (SM, FD = 2), and two cylinders with increasing extent of roughness of the internal wall (F1 and F2, FD = 2.18 and 2.99 respectively). The conduits were filled with silicone (Wacker©) oils as an analogue for magma, with viscosities in the range 10^1–10^3 Pa s and density of 970 kg/m^3.

The sensor system mainly consists of a ceramic shear accelerometer ICP J352C33 model (PCB Piezotronics) with a sensitivity of 0.1 V/g in the band 0.5–10,000 Hz, and a microphone ICP 378B02 model (PCB Piezotronics) with a sensitivity of 50 mV/Pa in the band 7–10,000 Hz (±1 dB). The digital acquisition system is a DAS50 (SEFRAM) with four channels and sampling rate up to 1 MHz. In our case signals are sampled at 50 kHz. Since the accelerometer and microphone have slightly different frequency responses (e.g. the lower frequency limit is different) both seismic and acoustic signals were band-pass filtered in the range 10–10,000 Hz, interval of common flat response of accelerometer and microphone. A video-camera (25 fps) allowed observation of the degassing regimes through the optically clear analogue magma and link them with the recorded seismo-acoustic signal.

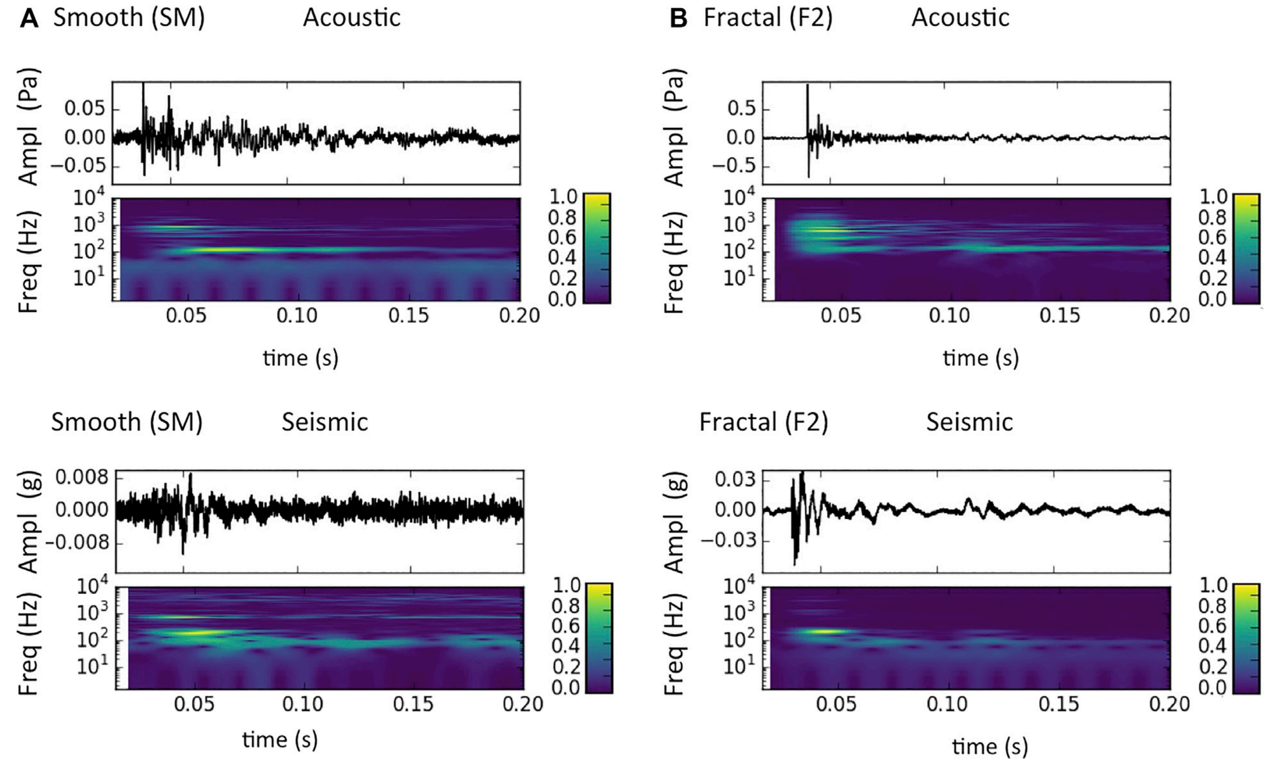


FIGURE 2 | Comparison between the raw signal and the spectrogram of a seismo-acoustic event generated with a smooth SM-type conduit, viscosity = 10 Pa s, flow velocity = 120 × 10^{-3} l/s **(A)**, and an event generated with a F2-type conduit, viscosity = 10 Pa s, flow velocity = 120 × 10^{-3} l/s **(B)**.

The Dataset

The dataset used in this work includes 616 experimental seismo-acoustic events generated under different experimental conditions. Each event is represented by a seismic recording and by an acoustic recording; therefore our dataset consists of 616 pairs of seismic and acoustic signal records (1232 files in total).

Three parameters were systematically changed to investigate the variability of the seismo-acoustic signals with different degassing states: the roughness of the epoxy conduits (SM, F1 and F2), the viscosity of the analogue magma (10, 100 and 1000 Pa s) and the air flow rate (5, 10, 30, 60, 90, 120, 150, 180 × 10^{-3} l/s). **Figure 2** shows two examples of seismo-acoustic signals. The first (**Figure 2A**) was generated using the smooth experimental conduit (SM), analogue magma with viscosity of 10 Pa s and a flow rate of 120 × 10^{-3} l/s. The second one (**Figure 2B**) was also generated with viscosity of 10 Pa s and a flow rate of 120 × 10^{-3} l/s, but using a F2-type experimental conduit.

The seismo-acoustic transients collected in the dataset were recorded during different degassing regimes, from bubbly flow (i.e. bubbles passively transported in the analogue magma) to slug regime (i.e. individual conduit-filling bubbles) to churn-annular flow (i.e. the gas phase is able to support the upward movement of the liquid along the conduit wall).

Feature Extraction Methods

The feature extraction (data preprocessing) is an important phase that precedes the application of the SOM neural network for the clustering. It provides an alternative data representation by removing redundant information and identifying the significant ones that uniquely describe them. In this way a compact and robust coding is obtained. There are several methods to do this, which depend on the particular application and its complexity degree. The crucial aspect is to choose the features that are relevant for the examined task. In our case, in order to select the most appropriate features, we took into account the information usually considered by analysts for the signal discrimination, i.e. the signal spectral content and waveform, in the frequency and time domain respectively. Then, to feature extraction techniques we relied on the good results obtained in Del Pezzo et al. (2003), Scarpetta et al. (2005), Esposito et al., (2006a), by using the Linear Predictive Coding (LPC) technique (Makhoul, 1975; Marple, 1980) for the frequency information and a waveform parametrization, for the amplitude-versus-time information, respectively. The advantage of using LPC over raw spectra is that LPC provides a more compressed representation of frequency content information. This aspect was discussed in Del Pezzo et al. (2003).

LPC is a widely used method in the speech recognition field to extract compact spectral features from audio signals, which are acoustic signals just like those considered in this work. However, it has also been successfully applied in the geophysical and volcanological fields (Esposito et al., 2013a; Giudicepietro et al., 2017; Esposito et al., 2018). In particular, the features extracted with the LPC were given in input both to unsupervised neural networks, such as the Self-Organizing Map (SOM) (Esposito et al., 2006a, Esposito et al., 2006b; Masiello et al., 2006), and supervised neural networks, such as the Multi-Layer Perceptron (MLP) (Del Pezzo et al., 2003; Scarpetta et al., 2005; Giudicepietro et al., 2017; Esposito et al., 2018). Therefore, we decided to use this technique following the good results provided in the previously mentioned works.

LPC predicts a signal sample through a linear combination of its previous samples:

$$\hat{x}(n) = \sum_{i=1}^{p} a_i x(n-i)$$

where $\hat{x}(n)$ is the predicted signal value for the sample n, $x(n-i)$ are the previously observed values, $p \le n$ is the model order and a_i are the prediction coefficients. The prediction coefficients are given by an optimization procedure that tries to reduce the error between the real signal and its LPC estimate. The number of prediction coefficients p is problem dependent and it must be determined via a trade-off between the information content and the representation compactness. We used the *librosa* python library (McFee et al., 2015) to obtain the LPC coefficients.

The waveform parametrization (WP) is performed by a function defined as the normalized difference between the maximum and minimum signal amplitude computed in an arbitrary duration window. In Scarpetta et al. (2005) it was shown that the addition of this feature as further information on the data led to an improvement in the performance of the neural network up to 5% (reaching 100% in some experiments aimed at the classification of seismic signals at Mt. Vesuvius) compared to the performance obtained using only the LPC features.

Waveform parametrization is expressed as:

$$f_m = \frac{(\max(A_{i,m}) - \min(A_{i,m}))*N}{\sum_{m=0}^{N} (\max(A_{i,m}) - \min(A_{i,m}))}$$

where $A_{i,m}$ are the amplitudes in the signal window and N is the total number of windows.

In general, the choice of the features used to represent the data can influence the results of the final clustering; for this reason this choice is critical for the good neural network performance. We chose the above described preprocessing methods because we are confident of their appropriateness for feature extraction on the basis of good results obtained in the previous cited works (Del Pezzo et al., 2003; Scarpetta et al., 2005; Esposito et al., 2006a; Esposito et al., 2018).

Self-Organizing Map (SOM)

SOM (Kohonen et al., 1996) is a special class of unsupervised artificial neural network (ANN) extensively used as a clustering and visualization tool in exploratory data analysis. In general, when a large data set is available and no information about which similarity measure is more appropriate to group them, then it may be useful to apply an unsupervised technique. Classic methods of cluster analysis often make assumptions regarding linearity, normal distribution or intrinsic clustering relationships in the data. For these reasons, they can fail for data with a complex structure.

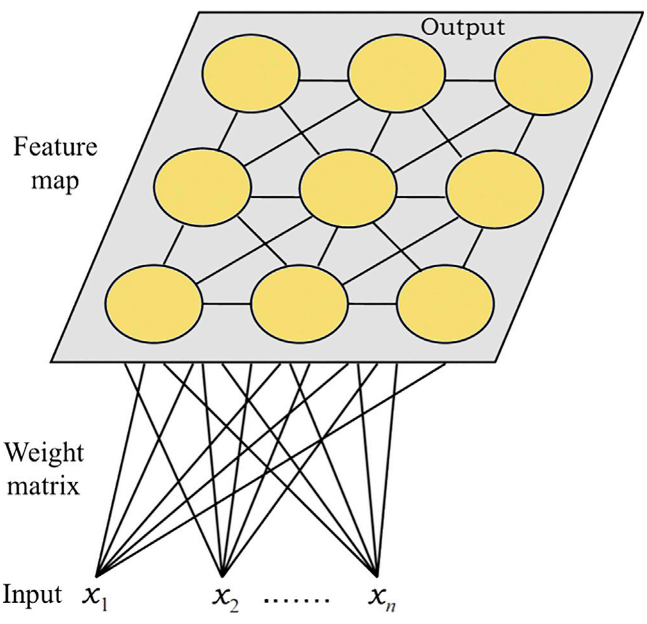


FIGURE 3 | Example of the architecture of a Self-Organizing Map. The network has two layers: the input and the output layer, with the nodes of the input layer directly connected to those of the output layer.

Contrary to classical methods, the main advantages of using the SOM technique are that it does not require any a-priori assumption on the data structure and it is able to manage large data sets with high-dimension inputs to detect patterns and isolated structures in the data. Moreover, it provides an easy visualization and interpretation of the results.

SOM performs a nonlinear mapping of the input space into a two-dimensional grid, called map, by preserving the topological and metric relationships of the data. This makes the SOM an effective method to visualize the similarity and the distances between the input vectors.

The SOM architecture has typically two layers, the input and the output one, with the nodes of the input layer directly connected to those of the output layer (**Figure 3**). Nodes in the output layer are arranged in a topological structure, usually a bi-dimensional grid, since it is generally agreed that a two-dimensional map provides a better representation of the clusters.

Usually, the network topology is divided into two factors: the local lattice structure and the global map shape. The first one can be rectangular or hexagonal. In the rectangular one each internal node has four neighbors, while in the hexagonal one six. Hence, the hexagonal structure is usually preferred since it displays greater variance in neighborhood size. The global map shape instead can be sheet, cylinder or toroid. In our experiment, we used a SOM map with a local hexagonal structure (**Figure 4A**)

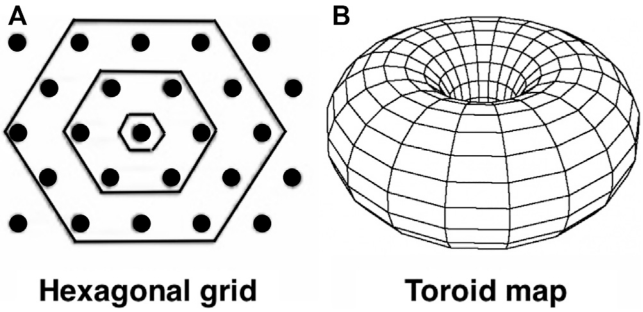


FIGURE 4 | Examples of a local hexagonal lattice structure **(A)**, and a global toroid map shape **(B)**. In this article the toroid shape, visualized as sheet, is used (see Results section).

and a global toroid shape (**Figure 4B**) visualized as a sheet to get a better interpretation of the obtained clusters.

The SOM algorithm is unsupervised: this means that the network tries to discover the hidden relationships in the data by applying a competitive learning rule where the output nodes compete among themselves for representing the vectors of the input space. In this competition only the node whose weight vector is most similar to the input vector will be the winner or the Best Matching Unit (BMU). To find the BMU, in each training step, the distances between an input vector, chosen randomly, and all the weight vectors of the SOM are computed by using the Euclidean distance method, as suggested by Kohonen et al. (1996). After finding the BMU, the prototype of the winning node is updated toward the input vector and also the prototypes of the neighborhood nodes are adjusted (cooperative aspect) in a way that depends on the neighborhood function. We adopted a Gaussian neighborhood function. The learning rate, which controls the intensity of the attraction of the input vector, and the neighborhood radius, which controls the number of vectors attracted other than the winning node, depend on the Gaussian function. The learning rate and the neighborhood radius are time-dependent functions and their values decrease during the training phase. At the end of the iterative process, there is the convergence phase in which the prototypes of the different nodes reach their final values and the final map is obtained.

In this work, the SOM parameter were chosen according to the SOM toolbox for Matlab (http://www.cis.hut.fi/somtoolbox/, last accessed 25 October, 2020) and Kohonen et al. (1996).

RESULTS

Data Preprocessing

Before extracting the features from the accelerometric and acoustic data produced during the experimental events, we cut the recordings using a standard Short Time Average/Long Time Average trigger algorithm (STA/LTA; e.g. Allen, 1978; Withers et al., 1998; Trnkoczy, 2012) to have a uniform criterion to generate the signal windows for preprocessing and analysis. Actually, the purpose of our study is not to develop an automatic system, but to discover the fingerprints of the degassing processes in seismo-acoustic signals; therefore the automatic procedure is used only to facilitate the picking of the event onsets.

We performed the automatic trigger of the events (**Figure 5**) using a classic STA/LTA algorithm (length of short time window = 0.01 s, length of long time window = 0.1 s, trigger on threshold = 2.5), included in the Python toolbox Obspy (Krischer et al., 2015). The STA/LTA was applied on the acoustic signals, and, each time an acoustic event was identified, 0.03 s windows of both acoustic and accelerometric signals (starting 0.002 s before each trigger time) were extracted. Although the microphone and accelerometer are located at different positions in the setup, their short distance (~60 cm) allows for a delay time on the order of 10^{-3} s, which is by far smaller than events recurrence time (Figure 11 in Spina et al., 2019) and of the selected signal windows. Accordingly, we are confident that triggered seismic and acoustic signals share the same source. Thus, we obtained a dataset of 616 events, each composed by seismic and acoustic recordings of 1500 samples (both signals being sampled at 50,000 samples per second). Finally, we parameterized the events in the frequency domain, through LPC, and in the time domain, through the waveform function described above (WP).

By using the *librosa* python library (McFee et al., 2015), we applied the LPC algorithm to extract the spectral content features. We made several experiments using a different number of windows and finally we found that the most appropriate solution to encode the LPC features of our data was to use a single 1024-sample length window for each signal, in order to correctly represent even the lower frequency events. So, we calculated the LPC on one signal window (first 1024 sample), extracting 46 LPC coefficients for each window. We chose the number of LPC coefficients as a trade-off between an adequate

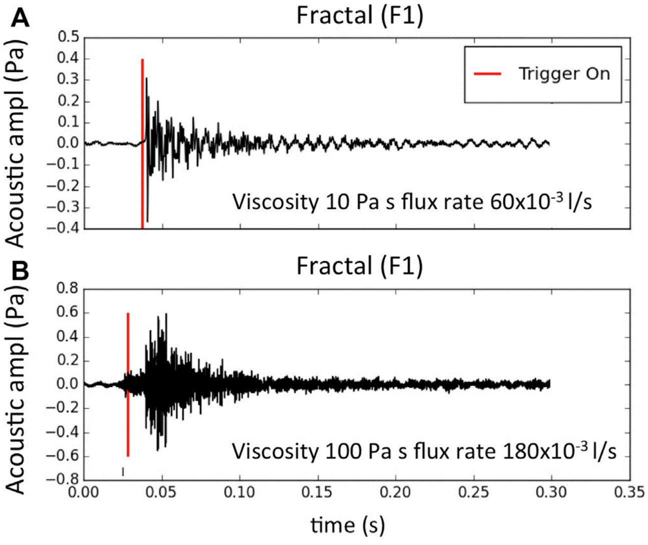


FIGURE 5 | Examples of automatic picking of two acoustic events by STA/LTA algorithm. Both events were generated with a F1 fractal roughness conduit but with different viscosity and flux rate conditions. For the event in plot **(A)** viscosity = 10 Pa s and flux rate = 60 × 10^{-3} l/s; for event in plot **(B)** viscosity = 100 Pa s and flux rate = 180 × 10^{-3} l/s.

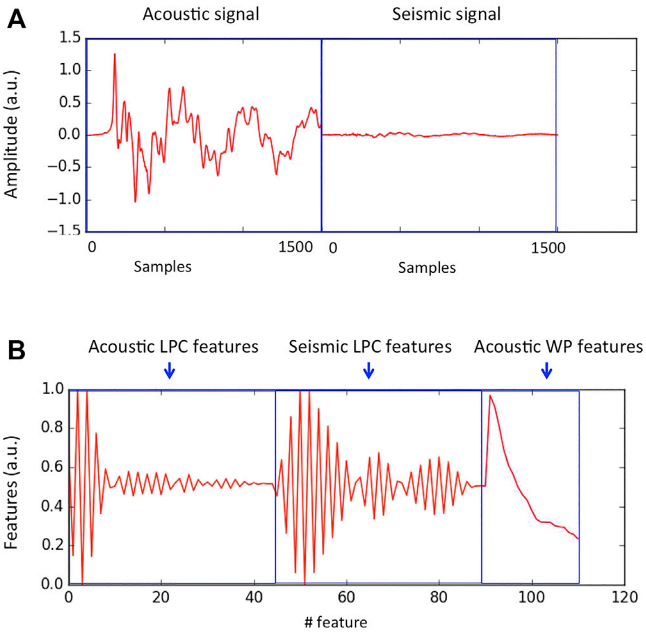


FIGURE 6 | (A) Acoustic and seismic signals of an experimental event; **(B)** input vector with the features extracted from the acoustic and seismic signals.

encoding of the spectral content and the compactness of the information. Thus, we encode the spectral content features of the experimental events with 46 features for each one of the seismic and acoustic signals.

For the waveform parameterization, we used only the acoustic signal features and we chose a 50 sample-long window, so obtaining 30 parameters from each 1500 sample-long recording. In order to avoid undesirable effects due to the automatic trigger of the signals that in different events can result in different pre-event signal segment, we sorted the waveform parameterization features in descending order. Finally, we took only the first 20 of these features, neglecting the queue which is less representative of the sorted signal. Therefore the waveform parameterization results in a 20-dimensional feature vector for each type of signal, acoustic and seismic. In this way the acoustic and seismic signals were encoded using for each of them 46 LPC features and 20 WP features. Finally, we combine the LPC features and the waveform parameterization (WP) of the acoustic signal obtaining the input vector of 46 + 46 + 20 = 112 features for each experimental event.

To scale the amplitude of the features extracted from acoustic and seismic signals, which are different in size and units of measurement (**Figure 6A**), we applied a feature scaling-normalization and obtained input vectors such as the one shown in **Figure 6B**.

SOM Clustering

In this study, we chose a SOM with a local hexagonal structure (**Figure 4A**) and a toroid shape (**Figure 4B**). The map size depends on the specific application. A "right number of nodes" does not exists, it depends on the detail one wants to have with the generated clusters. In our case we chose a map size of 14 × 8 nodes based on the number of events in the dataset. The map has 112 (14 × 8) nodes and the dataset contains 616 events, so we considered the ratio 616/112 = 5.5 that is a good compromise for data clustering.

Figure 7A shows the SOM map obtained on the examined dataset. The yellow hexagons indicate the non-empty nodes. Their size is proportional to the number, or data density, of input vectors, which fall in each of them. The numbers on the outside of the map (**Figure 7A**) indicate the order or numbering of the nodes on the map, that is, from top to bottom and from right to left. For some significant nodes the number is also reported inside the yellow hexagon.

The gray hexagons among the yellow ones are not real nodes and are used to represent the Euclidean distances between the nodes according to a gray scale. Also the empty nodes, i.e. those that do not contain any input vector, are filled according to the Euclidean distance gray scale. White or light gray indicates that neighborhood nodes are similar to each other and therefore may belong to the same cluster. Dark gray or black mark a clear separation between the nodes. In this way a visual qualitative measure of the cluster structure discovered by the SOM is possible, which allows identifying areas where the nodes are more similar to each other.

Following this qualitative approach and based on the analysis of the content of each node, on the SOM in **Figure 7A** we identify four main areas of nodes more similar to each other or four main clusters that we represent on the map with different colors. The clusters are mainly characterized by the degree of roughness of the conduit and by the viscosity of the analogue magma. The cyan dotted line in **Figure 7A** marks the Cluster 1 (130 events) that groups events that were generated mainly with smooth conduit (SM) and an analogue magma viscosity of 10 Pa s. Cluster 2 (63 events) is delimited by a blue dotted line and mainly includes events generated in experiments with a F1 conduit and an analogue magma viscosity of 10 Pa s. It also includes some nodes that are at the bottom of the map. This is because we used a toroidal map, so the upper and lower part and the right and left one are to be considered united. In **Figure 7A** we displayed the map as a two-dimensional sheet to allow an easy interpretation of the results. Cluster 3 (176 events), surrounded by a red dotted line, mainly groups events generated in experiments with a F2 conduit and an analogue magma viscosity of 10 Pa s. Finally, Cluster 4 (66 events), delimitated by a magenta dotted line, mainly contains events generated in experiments with analogue magma viscosity of 100 Pa s. This cluster also includes most of the events generated with a viscosity of 1000 Pa s, which are relatively few in the dataset because the increase in analogue magma viscosity dramatically decreases the number of events (Spina et al., 2019). In fact, the time elapsed between consecutive explosions positively correlates with viscosity both in volcanoes environment (Dominiguez et al., 2016) and in our laboratory experiments (Spina et al., 2019). The roughness of the conduit relating to the events of Cluster 4 is mixed; indeed the

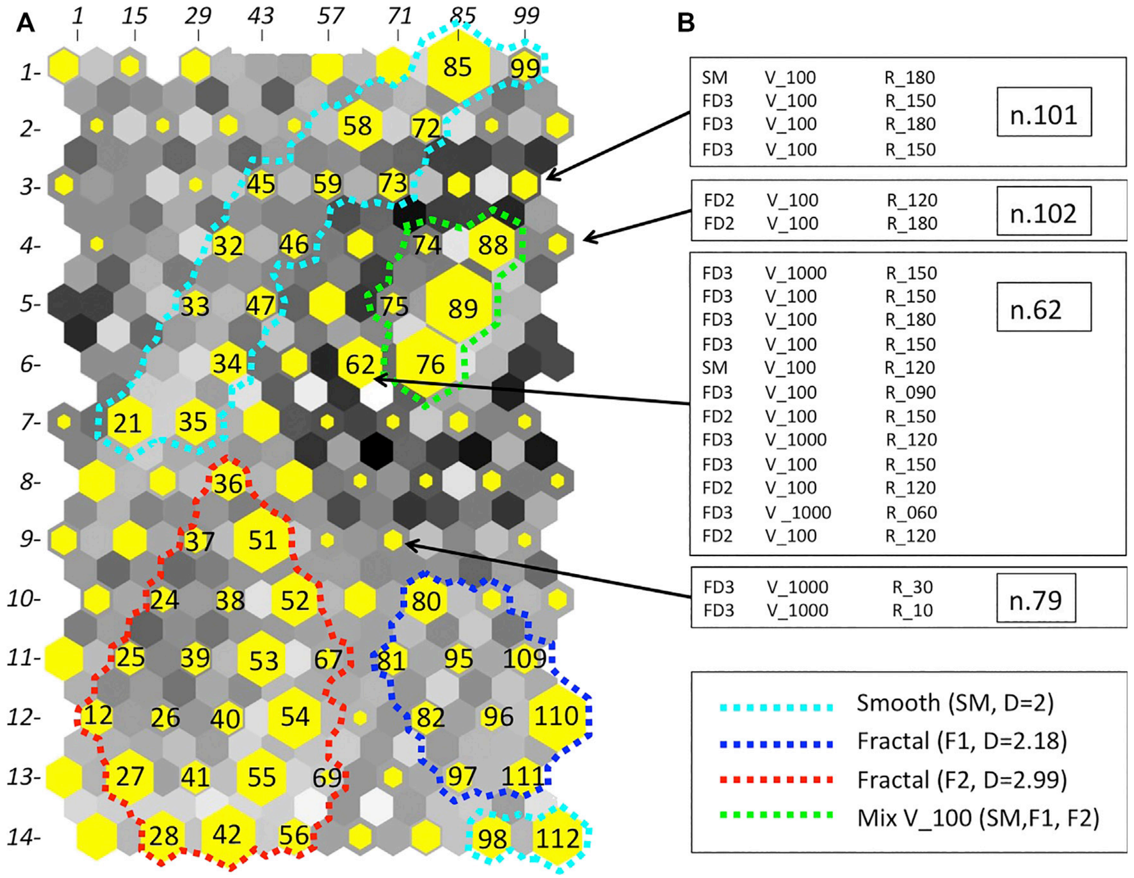


FIGURE 7 | (A) The SOM Map with 14 × 8 = 112 nodes. Not-empty nodes are shown as yellow hexagons and their size represents the data density. The numbers on the outside of the map indicate the numbering of the nodes that is also reported in some significant nodes. The gray hexagons interposed between the yellow ones indicate the Euclidean distance between the nodes according to a gray scale. Cluster 1 (cyan), Cluster 2 (blue), Cluster 3 (red) and Cluster 4 (green) are highlighted by dotted outlines. On the right **(B)** the compositions of nodes n.62, n.79, n.101 and n.102 are shown. The labels SM, F1 and F2 indicate SM (smooth), F1 (FD = 2.18) and F2 (FD = 2.99) conduit roughness, respectively. The number preceded by V indicates the viscosity in Pa s and the number preceded by R indicates the gas flow rate in 10^{-3} l/s.

cluster contains events generated with SM, F1 and F2 conduit roughness. This result suggests that in low viscosity experiments, e.g. 10 Pa s, the conduit roughness signature in the seismo-acoustic signals is clear, whereas in the experiments conducted with higher viscosity (e.g. 100 or 1000 Pa s) the conduit roughness imprint is less distinguishable in the signal features and the effect of viscosity prevails. However, in some nodes there are events produced with high viscosity (100 or 1000 Pa s) that were characterized by a prevailing type of conduit roughness. This is the case of the nodes 62, 79, 101 and 102 shown in **Figure 7B**. For the analysis of the other nodes, see the **Supplementary Material**, which contains the results of all the nodes of the SOM map.

Analyzing the distribution of the flux rate (5, 10, 30, 60, 120, 150, 180 × 10^{-3} l/s) for all the nodes of the four main clusters we find that it is less characterizing than the distribution of the analogue magma viscosity and the conduit roughness. However, we can note that in Cluster 4 the low flux rate classes (5, 10, 30 × 10^{-3} l/s) are missing. Furthermore, the few experimental events characterized by viscosity of 100 Pa s in Cluster 2 (most of the events that fall into this cluster were generated in experiments with analogue magma viscosity of 10 Pa s) systematically exhibit flux rate equal or below 60 × 10^{-3} l/s.

Figure 8 shows a summary of the distribution of experimental events of the four main clusters in the analogue magma viscosity (**Figure 8A,D,G,J**), conduit roughness (**Figure 8B,E,H,K**) and flux rate (**Figure 8C,F,I,L**) classes. An overall picture of the distribution of the 435 events contained in the four main clusters with respect to the different classes of conduit roughness, viscosity and gas flow rate is shown in **Figure 9**. We can see that seismo-acoustic signals produced in experiments conducted with low analogue magma viscosity (10 Pa s) are separated into Clusters 1 (cyan), 2 (blue) and 3 (red), which are characterized by different degrees of conduit roughness (SM, F2 and F3, respectively), whereas the events generated in high viscosity condition (100 Pa s and a few examples with 1000 Pa s), belonging to Cluster 4 (green), are distributed in all three conduit roughness classes (SM, F2 and F3). Furthermore, most of the events that fall into Cluster 4 were generated with flow rate > = 60 × 10^{-3} l/s.

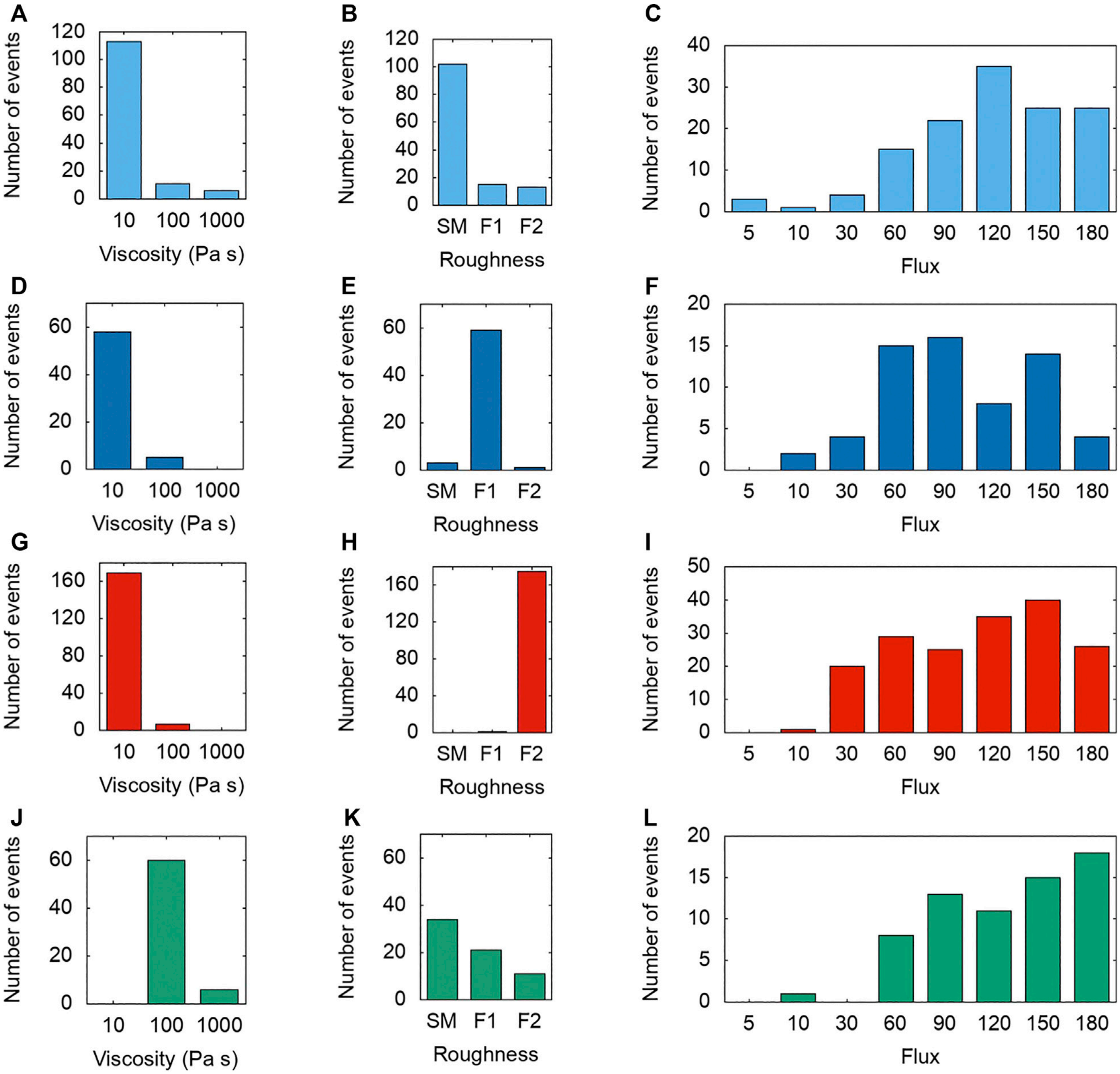


FIGURE 8 | Histograms of the distribution of Cluster 1 (cyan), Cluster 2 (blue), Cluster 3 (red) and Cluster 4 (green) events in the analogue magma viscosity, analogue conduit roughness and flux rate classes. Plots **(A)**, **(D)**, **(G)**, and **(J)** show the distribution of events belonging to the four clusters in the classes of viscosity, expressed in Pa s **(B)**, **(E)**, **(H)**, and **(K)** represent the conduit roughness classes (SM: smooth; F1: FD = 2.18; F2: FD = 2.99). The flux rate (expressed in 10^{-3} l/s) histograms are shown in **(C)**, **(F)**, **(I)**, and **(L)** plots.

DISCUSSION

In the present work, we applied SOM to perform the analysis of experimental events and investigate the links between seismo-acoustic signals and their source processes, which in the case of analogue experiments are known and controlled. In this way, we benefit of joint information deriving from an efficient seismo-acoustic signal clustering method and the experiments conducted under controlled physical conditions. This approach can help us to increase our understanding of the degassing processes from basaltic open conduit systems and of their elastic markers. For this reason, we developed a parameterization strategy of seismo-acoustic experimental signals inspired by the one we have successfully used to cluster natural events (Ham et al., 1999; Del Pezzo et al., 2003; Esposito et al., 2007; Esposito et al., 2013a; Giudicepietro et al., 2017; Esposito et al., 2018; Nuha et al., 2019), but with appropriate modifications to better extract the features of the experimental data. Furthermore, we chose to jointly use the acoustic and seismic information that together cover the radiation emitted by the entire process of rising in the conduit and bursting to the free surface of the gas bubbles.

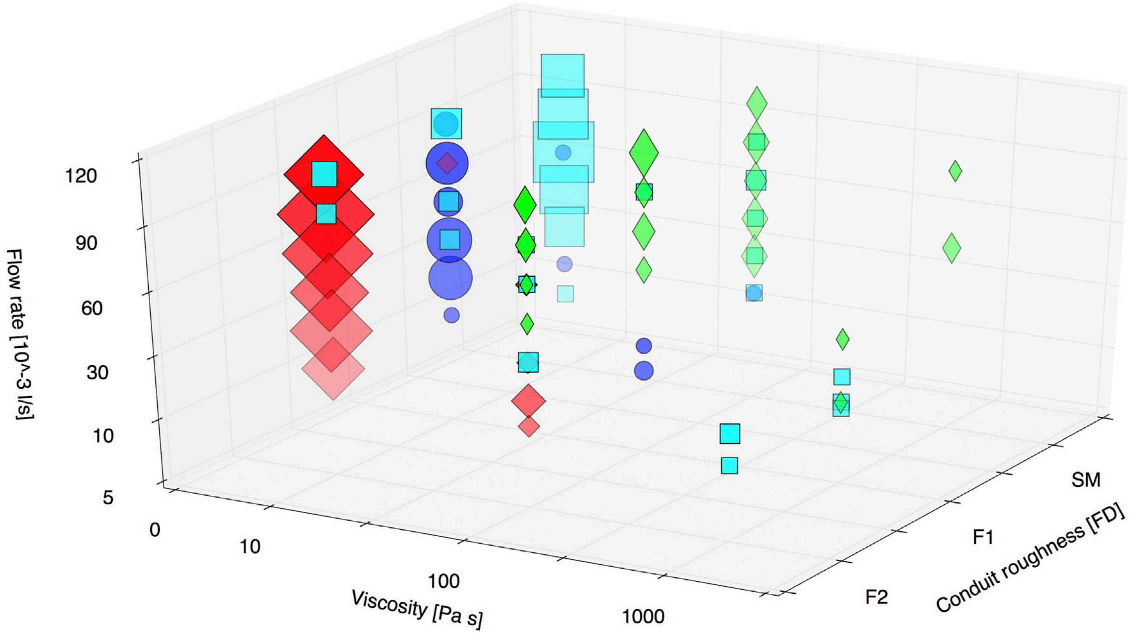


FIGURE 9 | Overall picture of the distribution of events belonging to the four main clusters with respect to the viscosity of the analogue magma, roughness of the analogue conduit and the flow rate. Cluster 1 is indicated in cyan (squares), Cluster 2 in blue (circles), Cluster 3 in red (diamond) and Cluster 4 in green (thin diamond). The size of the squares is proportional to the number of events that share the same viscosity, roughness and flux values and belong to the same cluster.

For the extraction of features, we used the LPC, which provides a compressed representation of the signal in the frequency domain of the seismic and acoustic recordings of each event. We have also adopted a parameterization in the time domain using a waveform function and rearranging the features in descending order, in order to avoid the problems related to the picking of the onset of the seismo-acoustic transients. This waveform parameterization technique, used for the first time to extract features for a seismic event classification problem applied to Mt. Vesuvius data, improved by 5% the performance obtained with this method (Scarpetta et al., 2005).

Experimental results of the degassing pattern have shown that at constant viscosity the increase in gas flux rate generates progressively longer and less-spaced slug bubbles which, at a given threshold of flux rate, merge leading to the transition from slug flow to churn-annular flow (e.g. Fabre and Linè, 1992; Paglianti et al., 1996). All the same, at medium-low gas flux rate an increase in liquid viscosity favors the coalescence of bubbles and the transition from slug flow to churn-annular flow (e.g. Pioli et al., 2012; Spina et al., 2019).

Based on the analysis of our SOM map, we identified four main clusters (**Figure 7A**) characterized by events generated under different experimental conditions, highlighting differences in the degassing patterns. In particular, most of the events of Cluster 1 are generated in experiments with low viscosity magma and smooth conduit and are associated with elongated slugs (**Figure 10A**). Notably, the few cases exhibiting a viscosity of 100 Pa s are characterized by relatively low gas flow rate ($<60 \times 10^{-3}$ l/s). Most of the events included in Clusters 2 and 3 were also generated in low viscosity conditions (10 Pa s), but they are clearly separated according to the roughness of the conduit which for Cluster 2 is of type F1 (FD = 2.18), whereas for Cluster 3 it is of type F2 (FD = 2.99). On the opposite, events of Cluster 4 are characterized by higher analogue magma viscosity (100 Pa s and few cases with 1000 Pa s) and gas flux generally above the threshold of 60×10^{-3} l/s. The degassing pattern in such cases is mostly represented by a chain of over-pressurized slugs that tend to coalesce with an irregular profile, suggesting the onset of churn-annular flow (**Figure 10B**). Increasing the viscosity of the liquid phase, the threshold of gas flux rate marking the transition between degassing regimes decreases (e.g. Pioli et al., 2012). Hence, the clusters associating 10 Pa s events (Cluster 1, 2, 3) and the one (Cluster 4) that groups high viscosity (100–1000 Pa s) and moderate to high gas flux events might actually reflect different degassing regimes, i.e. slug or churn-annular flow. Moreover, the features of the seismic and acoustic signals can be influenced by the flow regime that is generated in the experiment, which can be determined by different combination of viscosity of the analogue magma and gas flow rate. This can explain the characteristics of the events generated with different analogue magma viscosity and grouped in the same cluster.

Notably, gas flow rate and viscosity are the main controlling factors in determining the transition between different degassing patterns (e.g. Pioli et al., 2012). Clustering of seismo-acoustic events, here performed, demonstrated that different degassing regimes are clearly linked to different features of the radiated elastic energy, with fundamental implications for monitoring purposes.

The clustering results suggest also that the viscosity of the liquid, where gas flows through, plays an important role in determining the features of the seismic and acoustic signals generated by such a gas

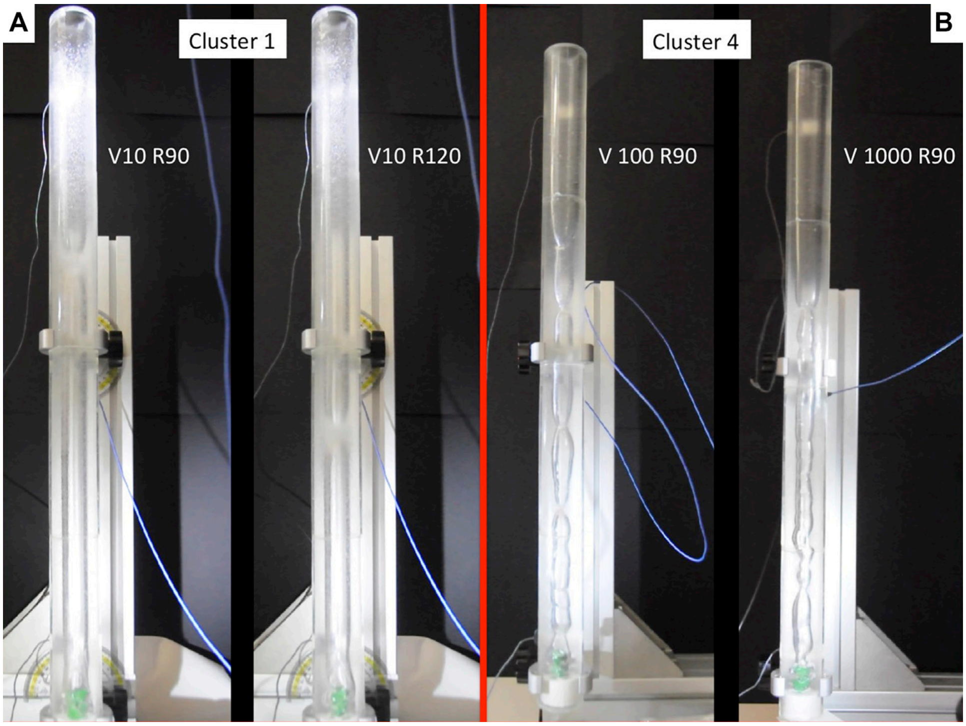


FIGURE 10 | Examples of images of the experiments that generated events belonging to Cluster 1 **(A)** and Cluster 4 **(B)**. V indicates the analogue magma viscosity in Pa s and R indicates the gas flow rate in 10^{-3} l/s. The tubes in the picture are 80 cm high and are smooth.

uprising. Lyons et al. (2013) also found that fluid viscosity plays a major role in controlling the character of seismic and acoustic source generation. The same conclusion was reached by Clarke et al. (2019), who recorded in laboratory acoustic emissions, resulting from fracturing and fluid depressurization through the fractured rock. They concluded how the viscosity of volcanic fluids may affect the spectral content of the volcano seismicity. Actually, magma viscosity is a factor that significantly affects the eruptive style of volcanoes (e.g. Cassidy et al., 2018). Also field studies have shown that the waveform and spectral characteristics of the seismic signals acquired on volcanoes are influenced by the properties of the fluids filling the plumbing system (e.g. Kumagai and Chouet, 2000; Kumagai et al., 2002; Ichihara et al., 2013). This relationship between magma properties and explosive mechanisms is also highlighted by studies conducted at Stromboli volcano (Esposito et al., 2008; Giudicepietro et al., 2019; Giudicepietro et al., 2020; Witsil and Johnson, 2020), which is considered a natural laboratory for this type of activity, where the characteristics of the seismic signals produced by the explosions are influenced by the physical properties of the magma, e.g. gas content. However, our experiments show that the SOM map also identifies clusters (e.g. Cluster 1 Cluster 2 and Cluster 3) characterized by the prevalence of events generated with a specific analogue conduit roughness (e. g. smooth, fractal FD = 2.18 and fractal FD = 2.99). As mentioned above, the roughness of the conduit seemed to be a most prominent feature for clustering in low viscosity runs (10 Pa s). Basing on image analysis only (evaluation of slug velocity and rate), Spina et al. (2019) has also hypothesized a more dominant effect of conduit roughness at 10 Pa s compared to higher viscosity runs (10 and 1000 Pa s). In fact, the effect of conduit roughness is predominant in the inertia driven regime (i.e. for significant high values of superficial gas to liquid velocity ratio) and for higher values of two phase gas flux and gas content (Shannak, 2008; Bhagwat and Ghajar, 2016). This is promising to better understand the contribution of conduit roughness in natural systems.

The method tested in this article can be applied to seismic and/or acoustic events recorded in other laboratory experiments in order to analyze the characteristics of the produced data and their similarities. The necessary requirement to apply the method to other laboratory experiments is that the data are similar to those here examined and the dataset is composed of a sufficient number of events. Moreover, this method can easily be applied to analyze naturally recorded seismic and acoustic events, appropriately scaling the duration of the signal windows and the other parameters. Of course, experimental data are simple signals compared to real world observations. For example they lack of seismo-acoustic coupling, whereas acoustic wave propagation is known to couple into the seismic wavefield (e.g., Jolly et al., 2017; Matoza et al., 2019). These characteristics must be taken into account when addressing the feature extraction strategy of real signals. It is worth noting that each experiment was performed under steady conditions, in terms of conduit roughness, analogue magma viscosity and air flow rate, while on volcanoes such parameters can be space- and time-dependent. For instance, the material filling the conduit is assumed to be stratified, with a high-viscosity layer of melt in the uppermost part, whose features can gradually evolve in time (e.g. Lautze and Houghton, 2005; Capponi et al., 2016). In spite of these differences between laboratory experiments and real volcanoes, the findings of this work suggest how seismo-acoustic signals can contain plenty of information about the conditions of the uppermost part of the plumbing system and its variations over time.

CONCLUSION

In this work we present an application of the unsupervised SOM network on a dataset of experimental events obtained under different physical conditions. For the first time SOM network are used on an analogue experimental data set in order to better understand the relationship between degassing behavior and seismo-acoustic signals. The seismo-acoustic signals have been obtained by varying three parameters: the roughness of the epoxy conduits (SM, F1 and F2), the viscosity of the analogue magma (10, 100 and 1000 Pa s) and the gas flow rate (5, 10, 30, 60, 90, 120, 150, 180 × 10^{-3} l/s). An automatic trigger technique has been applied to identify the events. In order to group the data, we have first performed a processing phase in which the events have been represented through a 112-feature vectors encoding both spectral, by using the LPC method, and time (WP) information from seismo-acoustic event pairs generated by experimental runs, conducted under different physical conditions, and included in the dataset (616 pairs). The obtained feature vectors became the input of the selected neural network, the SOM, to perform clustering. The SOM is a projection technique able to identify and visualize at the same time on a bi-dimensional plane the hidden structure of the data. The results are very promising as the SOM was capable to separate events generated with low viscosity analogue magma from those generated with high viscosity magma and to discriminate among those generated with different roughness of the conduit. The promising results obtained from the data generated by the analogue experiments demonstrate that the method could be a valuable tool in volcano monitoring to provide interpretative tool on clustering that can be obtained on real data. Moreover, our study confirms that seismic and acoustic signals recorded on volcanoes can provide insights into the characteristics of the fluids in the plumbing system and their temporal evolution. In particular, viscosity, which most affects the clustering results, is one of the main parameters controlling the style of degassing and eruptions.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

FG conceptualization, article writing, data analysis procedures, interpretation of results, article revising; AE article writing, data analysis procedures, interpretation of results, article revising; LS article writing, dataset preparation, interpretation of results, article revising; AC article writing, dataset preparation, interpretation of results, article revising; DM article writing, dataset preparation, interpretation of results, article revising; LL data analysis procedures, interpretation of results, article revising; GM conceptualization, data analysis procedures, interpretation of results, article revising.

FUNDING

This work benefited from funds of the EU (DG ECHO) Project EVE n. 826292 and was partially supported by the project Progetto Strategico Dipartimentale INGV 2019 “Forecasting eruptive activity at Stromboli volcano: timing, eruptive style, size, intensity and duration” (FIRST) and by the project INGVFISR-2017 “Sale Operative Integrate e Reti di Monitoraggio del Futuro: l’INGV 2.0”. This work is also supported by a Marie Sklodowska-Curie Innovative Training Network Fellowship of the European Commissions Horizon 2020 Programme under Contract Number 765710 INSIGHTS. This work was also funded by the project AEOLUS (Fondo Ricerca di Base MORGABASE2015 funded by Department of Physics and Geology, University of Perugia). This work was partially supported by the project Progetto Strategico Dipartimentale INGV 2019 “FIRST - ForecastIng eRuptive activity at Stromboli volcano: Timing, eruptive style, size, intensity, and duration”.

ACKNOWLEDGMENTS

We thank the Reviewers and the Editors, who have helped us improve our work. We wish to thank the Project ROUGHER (Progetti di Ricerca Libera) funded by INGV.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2020.581742/full#supplementary-material.

REFERENCES

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.* 68 (5), 1521–1532.

Ambrosino, F. C., Sabbarese, C., Roca, V., Giudicepietro, F., and Chiodini, G. (2020). Analysis of 7-years radon time series at Campi Flegrei area (Naples, Italy) using artificial neural network method. *Appl. Rad. Isotop.* 163, 109239. doi:10.1016/j.apradiso.2020.109239

Arciniega-Ceballos, A., Alatorre-Ibargüengoitia, M., Scheu, B., and Dingwell, D. B. (2015). Analysis of source characteristics of experimental gas burst and fragmentation explosions generated by rapid decompression of volcanic rocks. *J. Geophys. Res. Solid Earth*, 120 (7), 5104–5116. doi:10.1002/2014jb011810

Arciniega-Ceballos, A., Alatorre-Ibargüengoitia, M., Scheu, B., Dingwell, D. B., and Delgado-Granados, H. (2014). Seismological analysis of conduit dynamics in fragmentation experiments. *J. Geophys. Res. Solid Earth*, 119 (3), 2215–2229. doi:10.1002/2013jb010646

Bhagwat, S. M., and Ghajar, A. J. (2016). *Effect of pipe surface roughness on frictional pressure drop in gas-liquid two phase flows, fluid mechanics and thermodynamics 2016*, in “12th International Conference on Heat Transfer”

Burzynski, A. M., Anderson, S. W., Morrison, K., Patrick, M. R., Orr, T., and Thelen, W. (2018). “*Lava lake thermal pattern classification using self-*

organizing maps and relationships to eruption processes at Kīlauea Volcano, Hawai 'i". in *Field Volcanology: A Tribute to the Distinguished Career of Don Swanson*, 538, 307.

Capponi, A., James, M. R., and Lane, S. J. (2016). Gas slug ascent in a stratified magma: implications of flow organisation and instability for Strombolian eruption dynamics. *Earth Planet Sci. Lett.* 435, 159–170. doi:10.1016/j.epsl.2015.12.028

Carniel, R., Jolly, A. D., and Barbui, L. (2013). Analysis of phreatic events at Ruapehu volcano, New Zealand using a new SOM approach. *J. Volcanol. Geoth. Res.* 254, 69–79. doi:10.1016/j.jvolgeores.2012.12.026

Carniel, R. (1996). Neural networks and dynamical system techniques for volcanic tremor analysis. *Ann. Geophys.* 39 (2), 496.

Cassidy, M., Manga, M., Cashman, K., and Bachmann, O. (2018). Controls on explosive-effusive volcanic eruption styles. *Nat. Commun.* 9, 2839. doi:10.1038/s41467-018-05293-3

Chouet, B. A., and Matoza, R. S. (2013). A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *J. Volcanol. Geoth. Res.* 252, 108–175. doi:10.1016/j.jvolgeores.2012.11.013

Clarke, J., Adam, L., Sarout, J., van Wijk, K., Kennedy, B., and Dautriat, J. (2019). The relation between viscosity and acoustic emissions as a laboratory analogue for volcano seismicity. *Geology* 47 (6), 499–503. doi:10.1130/G45446.1

Del Pezzo, E., Esposito, A., Giudicepietro, F., Marinaro, M., Martini, M., and Scarpetta, S. (2003). Discrimination of earthquakes and underwater explosions using neural networks. *Bull. Seismol. Soc. Am.* 93 (1), 215–223.

Divoux, T., Vidal, V., Melo, F., and Géminard, J. C. (2008). Acoustic emission associated with the bursting of a gas bubble at the free surface of a non-Newtonian fluid. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* 77 (5), 056310. doi:10.1103/PhysRevE.77.056310

Dominguez, L., Pioli, L., Bonadonna, C., Connor, C. B., Andronico, D., Harris, A. J. L., et al. (2016). Quantifying unsteadiness and dynamics of pulsatory volcanic activity. *Earth Planet Sci. Lett.* 444, 160–168. doi:10.1016/j.epsl.2016.03.048

Ersoy, O., Aydar, E., Gourgaud, A., Artuner, H., and Bayhan, H. (2007). Clustering of volcanic ash arising from different fragmentation mechanisms using Kohonen self-organizing maps. *Comput. Geosci.* 33 (6), 821–828. doi:10.1016/j.cageo.2006.10.008

Esposito, A. M., Giudicepietro, F., Scarpetta, S., D'Auria, L., Marinaro, M., and Martini, M. (2006a). Automatic discrimination among landslide, explosion-quake, and microtremor seismic signals at Stromboli volcano using neural networks. *Bull. Seismol. Soc. Am.* 96 (4A), 1230–1240.

Esposito, A. M., Scarpetta, S., Giudicepietro, F., Masiello, S., Pugliese, L., and Esposito, A. (2006b). Nonlinear exploratory data analysis applied to seismic signals. *Lect. Notes Comput. Sci.* 3931/2006, 70–77. doi:10.1007/11731177_11

Esposito, A., Esposito, A. M., Giudicepietro, F., Marinaro, M., and Scarpetta, S. (2007). "*Models for identifying structures in the data: a performance comparison*", in *International conference on knowledge-based and intelligent information and engineering systems*. Berlin, Heidelberg: Springer, 275–283.

Esposito, A. M., Giudicepietro, F., D'Auria, L., Scarpetta, S., Martini, M. G., Coltelli, M., Marinaro, M., et al. (2008). Unsupervised neural analysis of very-long-period events at Stromboli volcano using the self-organizing maps. *Bull. Seismol. Soc. Am.* 98 (5), 2449–2459. doi:10.1785/0120070110

Esposito, A. M., D'Auria, L., Giudicepietro, F., Caputo, T., and Martini, M. (2013a). Neural analysis of seismic data: applications to the monitoring of Mt. Vesuvius. *Ann. Geophys.* 56 (4), 0446. doi:10.1371/journal.pone.0203210

Esposito, A. M., D'Auria, L., Giudicepietro, F., and Martini, M. (2013b). *Waveform variation of the explosion-quakes as a function of the eruptive activity at Stromboli volcano, Neural nets and surroundings*. Berlin, Heidelberg: Springer, 111–119.

Esposito, A. M., D'Auria, L., Angelillo, A., Giudicepietro, F., and Martini, M. (2014). "*Predictive analysis of the seismicity level at Campi Flegrei volcano using a data-driven approach*", in *Recent advances of neural network models and applications*. Cham: Springer, 133–145.

Esposito, A. M., Giudicepietro, F., Scarpetta, S., and Khilnani, S. (2018). "*A neural approach for hybrid events discrimination at Stromboli volcano*", in *Multidisciplinary approaches to neural computing, Series "Smart Innovation and Springer International Publishing AG, Systems and Technologies (SIST)*, (New York: Springer), 69. doi:10.1007/978-3-319-56904-8_2 http://www.springer.com/us/book/9783319569031.

Esposito, A. M., Alaia, G., Giudicepietro, F., Pappalardo, L., and D'Antonio, M. (2020a). "Unsupervised geochemical analysis of the eruptive products of Ischia, Vesuvius and Campi Flegrei", in *Progresses in Artificial Intelligence and Neural Systems*. (Singapore: Springer), 175–184.

Esposito, A. M., De Bernardo, A., Ferrara, S., Giudicepietro, F., and Pappalardo, L. (2020b). SOM-based analysis of volcanic rocks: an application to Somma-Vesuvius and Campi Flegrei volcanoes (Italy), in *Neural approaches to dynamics of signal exchanges*. (Singapore: Springer), 55–60.

Fabre, J., and Liné, A. (1992). Modeling of two-phase slug flow. *Annu. Rev. Fluid Mech.* 24 (1), 21–46. doi:10.1146/annurev.fl.24.010192.000321

Giudicepietro, F., Calvari, S., Alparone, S., Bianco, F., Bonaccorso, A., Bruno, V., et al. (2019). Integration of ground-based remote-Sensing and in Situ multidisciplinary monitoring data to analyze the eruptive activity of Stromboli volcano in 2017–2018. *Rem. Sens.* 11 (15), 1813. doi:10.3390/rs11151813

Giudicepietro, F., Esposito, A. M., and Ricciolino, P. (2017). Fast discrimination of local earthquakes using a neural approach. *Seismol Res. Lett.* 88 (4), 1089–1096. doi:10.1785/0220160222

Giudicepietro, F., López, C., Macedonio, G., Alparone, S., Bianco, F., Calvari, S., et al. (2020). Geophysical precursors of the July-August 2019 paroxysmal eruptive phase and their implications for Stromboli volcano (Italy) monitoring. *Sci. Rep.* 10 (1), 1–16. doi:10.1038/s41598-020-67220-1

Ham, F. M., Leeney, T. A., Canady, H. M., and Wheeler, J. C. (1999). "*An infrasonic event neural network classifier IJCNN'99*", in International Joint Conference on Neural Networks Proceedings (Cat. No. 99CH36339), 6, 3768–3773.

Ichihara, M., Lyons, J. J., and Yokoo, A. (2013). Switching from seismic to seismo-acoustic harmonic tremor at a transition of eruptive activity during the Shinmoe-dake 2011 eruption. *Earth Planets Space* 65, 14 doi:10.5047/eps.2013.05.003

James, M. R., Lane, S. J., and Chouet, B. A. (2006). Gas slug ascent through changes in conduit diameter: laboratory insights into a volcano-seismic source process in low-viscosity magmas. *J. Geophys. Res. Solid Earth* 111 (B5) doi:10.1029/2005JB003718

James, M. R., Lane, S. J., Chouet, B., and Gilbert, J. S. (2004). Pressure changes associated with the ascent and bursting of gas slugs in liquid-filled vertical and inclined conduits. *J. Volcanol. Geoth. Res.* 129 (1–3), 61–82.

Johnson, J., Aster, R., Jones, K. R., Kyle, P., and McIntosh, B. (2008). Acoustic source characterization of impulsive Strombolian eruptions from the Mount Erebus lava lake. *J. Volcanol. Geoth. Res.* 177 (3), 673–686. doi:10.1016/j.jvolgeores.2008.06.028

Jolly, A. D., Matoza, R. S., Fee, D., Kennedy, B. M., Iezzi, A. M., Fitzgerald, R. H., et al. (2017). Capturing the acoustic radiation pattern of strombolian eruptions using infrasound sensors aboard a tethered aerostat, Yasur volcano, Vanuatu. *Geophys. Res. Lett.* 44 (19), 9672–9680. doi:10.1002/2017GL074971

Kobayashi, T., Namiki, A., and Sumita, I. (2010). Excitation of airwaves caused by bubble bursting in a cylindrical conduit: experiments and a model. *J. Geophys. Res. Solid Earth* 115, B10. doi:10.1029/2009JB006828

Köhler, A., Ohrnberger, M., and Scherbaum, F. (2009). Unsupervised feature selection and general pattern discovery using Self-Organizing Maps for gaining insights into the nature of seismic wavefields. *Comput. Geosci.* 35 (9), 1757–1767. doi:10.1016/j.cageo.2009.02.004

Kohonen, T., Hynninen, J., Kangas, J., and Laaksonen, J. (1996). SOM_PAK: the self-organizing map program package, Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, *Finland.* Available at: www.cis.hut.fi/research/som_lvq_pak.shtml

Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, M., Caudron, C., et al. (2015). Obspy: a bridge for seismology into the scientific python ecosystem. *Comput. Sci. Discov.* 8, 014003. doi:10.1088/1749-4699/8/1/014003

Kumagai, H., and Chouet, B. A. (2000). Acoustic properties of a crack containing magmatic or hydrothermal fluids. *J. Geophys. Res.* 105 (B11), 25493–25525. doi:10.1029/2000JB900273

Kumagai, H., Chouet, B. A., and Nakano, M. (2002). Temporal evolution of a hydrothermal system in Kusatsu-Shirane Volcano, Japan, inferred from the complex frequencies of long-period events. *J. Geophys. Res.* 107 (B10), 2236. doi:10.1029/2001JB000653

Langer, H., Falsaperla, S., Masotti, M., Campanini, R., Spampinato, S., and Messina, A. (2009). Synopsis of supervised and unsupervised pattern classification techniques applied to volcanic tremor data at Mt Etna, Italy. *Geophys. J. Int.* 178 (2), 1132–1144. doi:10.1111/j.1365-246X.2009.04179.x

Lautze, N. C., and Houghton, B. F. (2005). Physical mingling of magma and complex eruption dynamics in the shallow conduit at Stromboli volcano, Italy. *Geology* 33 (5), 425–428. doi:10.1130/G21325.1

Lyons, J. J., Ichihara, M., Kurokawa, A., and Lees, J. M. (2013). Switching between seismic and seismo-acoustic harmonic tremor simulated in the laboratory: insights into the role of open degassing channels and magma viscosity. *J. Geophys. Res. Solid Earth* 118, 277–289. doi:10.1002/jgrb.50067

Macedonio, G., Dobran, F., and Neri, A. (1994). Erosion processes in volcanic conduits and application to the AD 79 eruption of Vesuvius. *Earth Planet Sci. Lett.* 121 (1–2), 137–152. doi:10.1016/0012-821x(94)90037-x

Makhoul, J. (1975). Linear prediction: a tutorial review. *Proc. IEEE* 63, 561–580.

Marple, L. (1980). A new autoregressive spectrum analysis algorithm, *IEEE Trans. Acoust. Speech Signal Process.* 28, 4.

Masiello, S., Esposito, A. M., Scarpetta, S., Giudicepietro, F., Esposito, A., and Marinaro, M. (2006). Application of self organized maps and curvilinear component analysis to the discrimination of the Vesuvius seismic signals. *In WSOM*.

Matoza, R. S., Arciniega-Ceballos, A., Sanderson, R. W., Mendo-Pérez, G., Rosado-Fuentes, A., and Chouet, B. A. (2019). High-broadband Seismoacoustic signature of Vulcanian explosions at popocatépetl volcano, Mexico. *Geophys. Res. Lett.* 46 (1), 148–157.

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., et al. (2015). "Librosa: audio and music signal analysis in python", in Proceedings of the 14th Python in Science Conference, 18–25.

Messina, A., and Langer, H. (2011). Pattern recognition of volcanic tremor data on Mt. Etna (Italy) with KKAnalysis—a software program for unsupervised classification. *Comput. Geosci.* 37 (7), 953–961. doi:10.1016/j.cageo.2011.03.015

Nuha, H. H., Balghonaim, A., Liu, B., Mohandes, M., and Fekri, F. (2019). *Seismic data compression using deep neural network predictors*, in "*SEG technical program expanded abstracts 2019*". (New York: Society of Exploration Geophysicists), 258–262.

Paglianti, A., Giona, M., and Soldati, A. (1996). Characterization of subregimes in two-phase slug flow. *Int. J. Multiphas. Flow* 22 (4), 783–796. doi:10.1016/0301-9322(96)00019-5

Pioli, L., Bonadonna, C., Azzopardi, B. J., Phillips, J. C., and Ripepe, A. M. (2012). Experimental constraints on the outgassing dynamics of basaltic magmas. *J. Geophys. Res. Solid Earth* 117 (B3), 19.

Ren, C. X., Peltier, A., Ferrazzini, V., Rouet Leduc, B., Johnson, P. A., and Brenguier, F. (2020). Machine learning reveals the seismic signature of eruptive behavior at Piton de la Fournaise volcano. *Geophys. Res. Lett.* 47, e2019GL085523doi:10.1029/2019GL085523

Rowe, C. A., Aster, R. C., Kyle, P. R., Dibble, R. R., and Schlue, J. W. (2000). Seismic and acoustic observations at mount Erebus volcano, ross Island, Antarctica, 1994–1998. *J. Volcanol. Geoth. Res.* 101 (1–2), 105–128. doi:10.1016/S0377-0273(00)00170-0

Scarpetta, S., Giudicepietro, F., Ezin, E. C., Petrosino, S., Del Pezzo, E., Martini, M., et al. (2005). Automatic classification of seismic signals at Mt. Vesuvius volcano, Italy, using neural networks. *Bull. Seismol. Soc. Am.* 95 (1), 185–196.

Shannak, B. A. (2008). Frictional pressure drop of gas liquid two-phase flow in pipes. *Nucl. Eng. Des.* 238 (12), 3277–3284. doi:10.1016/j.nucengdes.2008.08.015

Simons, B. C., Jolly, A. D., Eccles, J. D., and Cronin, S. J. (2020). Spatiotemporal relationships between two Closely-spaced strombolian-style Vents, Yasur, Vanuatu. *Geophys. Res. Lett.* 47 (5), e2019GL085687. doi:10.1029/2019GL085687

Spina, L., Cannata, A., Morgavi, D., and Perugini, D. (2019). Degassing behaviour at basaltic volcanoes: new insights from experimental investigations of different conduit geometry and magma viscosity. *Earth Sci. Rev.* 10, 28. doi:10.1016/j.earscirev.2019.03.010

Spina, L., Morgavi, D., Cannata, A., Campeggi, C., and Perugini, D. (2018). An experimental device for characterising degassing processes and related elastic fingerprints: analogue volcano seismo-acoustic observations. *Rev. Sci. Instrum.* 89, 055102. doi:10.1063/1.5020004

Spina, L., Taddeucci, J., Cannata, A., Gresta, S., Lodato, L., Privitera, E., and Palladino, D. M. (2015). Explosive volcanic activity at Mt. Yasur: a characterization of the acoustic events (9–12th July 2011). *J. Volcanol. Geoth. Res.* 302, 24–32. doi:10.1007/s00445-017-1108-x

Trnkoczy, A. (2012). "*Understanding and parameter setting of STA/LTA trigger algorithm*", in *IASPEI New manual of seismological observatory practice 2 (NMSOP-2). Deutsches GeoForschungs Zentrum GFZ, Potsdam, Potsdam*, 1–20. doi:10.2312/GFZ.NMSOP-2_IS_8.1

Unglert, K., Radić, V., and Jellinek, A. M. (2016). Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. *J. Volcanol. Geoth. Res.* 320, 58–74. doi:10.1016/j.jvolgeores.2016.04.014

Vidal, V., Ichihara, M., Ripepe, M., and Kurita, K. (2009). Acoustic waveform of continuous bubbling in a non-Newtonian fluid. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* 80 (6), 066314. doi:10.1103/PhysRevE.80.066314

Vidal, V., Géminard, J. C., Divoux, T., and Melo, F. (2006). Acoustic signal associated with the bursting of a soap film which initially closes an overpressurized cavity. *Eur. Phys. J. B Condens. Matter Complex Syst.* 54 (3), 321–339.

Watson, L. M. (2020). Using unsupervised machine learning to identify changes in eruptive behavior at Mount Etna, Italy. *J. Volcanol. Geoth. Res.* 46, 107042.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, s., et al. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bull. Seism. Soc. Am.* 88, 95–106.

Witsil, A., and Johnson, J. B. (2020). Analyzing continuous infrasound from Stromboli volcano, Italy using unsupervised machine learning. *Comput. Geosci.* 140, 494. doi:10.1016/j.cageo.2020.104494

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.

frontiers
in Earth Science

METHODS
published: 03 March 2021
doi: 10.3389/feart.2020.579923

Application of Subspace-Based Detection Algorithm to Infrasound Signals in Volcanic Areas

Mariangela Sciotto * *and Placido Montalto*

Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo, Catania, Italy

Infrasonic signals investigation plays a fundamental role for both volcano monitoring purpose and the study of the explosion dynamics. Proper and reliable detection of weak signals is a critical issue in active volcano monitoring. In particular, in volcanic acoustics, it has direct consequences in pinpointing the real number of generated events (amplitude transients), especially when they exhibit low amplitude, are close in time to each other, and/or multiple sources exist. To accomplish this task, several algorithms have been proposed in literature; in particular, to overcome limitations of classical approaches such as short-time average/long-time average and cross-correlation detector, in this paper a subspace-based detection technique has been implemented. Results obtained by applying subspace detector on real infrasound data highlight that this method allows sensitive detection of lower energy events. This method is based on a projection of a sliding window of signal buffer onto a signal subspace that spans a collection of reference signals, representing similar waveforms from a particular infrasound source. A critical point is related to subspace design. Here, an empirical procedure has been applied to build the signal subspace from a set of reference waveforms (templates). In addition, in order to determine detectors parameters, such as subspace dimension and detection threshold, even in presence of overlapped noise such as infrasonic tremor, a statistical analysis of noise has been carried out. Finally, the subspace detector reliability and performance, have been assessed by performing a comparison among subspace approach, cross-correlation detector and short-time average/long-time average detector. The obtained confusion matrix and extrapolated performance indices have demonstrated the potentiality, the advantages and drawbacks of the subspace method in tracking volcanic activity producing infrasound events. This method revealed to be a good compromise in detecting low-energy and very close in time events recorded during Strombolian activity.

Keywords: infrasound signal, subspace detector, trigger algorithm, Infrasound volcano monitoring, strombolian activity, Infrasound events, Etna volcano, Infrasonic tremor

Edited by:
Reik Donner,
Hochschule Magdeburg-Stendal,
Germany

Reviewed by:
Stephen Arrowsmith,
Southern Methodist University,
United States
Oliver D. Lamb,
University of North Carolina at
Chapel Hill, United States

Correspondence:
Mariangela Sciotto
mariangela.sciotto@ingv.it

Specialty section:
This article was submitted to
Volcanology,
a section of the journal
Frontiers in Earth Science

Received: *03 July 2020*
Accepted: *26 November 2020*
Published: *03 March 2021*

Citation:
Sciotto M and Montalto P (2021)
Application of Subspace-Based
Detection Algorithm to Infrasound
Signals in Volcanic Areas.
Front. Earth Sci. 8:579923.
doi: 10.3389/feart.2020.579923

INTRODUCTION

Amplitude transient detection plays a fundamental role in volcano monitoring, allowing counting amplitude transients, identifying amplitude and occurrence rate variations. Besides, it is an essential step to localize seismic sources and their possible migration, which could be related to changes in volcano state and dynamics.

Classical methods for signal detection in seismology and volcano-seismology are grouped into two main categories: energy and correlation detectors. The former, called incoherent energy detectors, include algorithms searching for signals which are not or poorly known, such as STA/LTA (short-time average/long-time average) algorithm (Allen, 1978; Trnkoczy, 2012). These techniques, routinely used in volcano seismology, need no data pre-processing, beside the filtering applied to identify the desired signals, such as volcano-tectonic (VT) earthquakes or long period (LP) events. This approach suffers from high rate of false alarms or even of missed detections, due mainly to the background noise strongly affecting the reliability of this technique. This is especially true on active volcanoes with continuous volcanic tremor, which could dramatically reduce the signal (intended as the amplitude transients) to noise ratio. The latter group, correlation detectors, consists of algorithms based on the cross-correlation between a known waveform and the continuously recorded signal. These algorithms are very sensitive, give low false alarm rate but have the disadvantage of being able to detect only signals which are very similar to the template waveform, which in turn needs to be well known (Withers et al., 1999; Gibbons et al., 2007). In volcano acoustics, similar techniques (e.g., Cannata et al., 2013a; Cannata et al.,2013b; Thompson, 2015; Hotovec-Ellis and Jeffrines, 2016; Matoza et al., 2019a; Senobari et al., 2019), or methods making use of advanced signal processing techniques (Bueno et al., 2019), are implemented to identify and extract amplitude transients from the real-time streaming of signals, that characterize explosive or degassing activity. In particular, energy detectors, such as STA/LTA, are efficient algorithms when multiple infrasound sources are active (as at multi-vent volcanoes) and exhibit space-time variations, while correlation detectors are a powerful tool when we want to identify amplitude transients produced by a single and/or stable infrasound source in order to study its physical properties (Montalto et al., 2010; Sciotto et al., 2013; Cannata et al., 2013a; Hotovec-Ellis and Jeffrines, 2016; Yokoo et al., 2019).

Subspace-based detectors overcome the aforementioned limitation, in that they operate a comparison between the continuous signal and a set of reference waveforms hereafter called templates (Harris, 2006). One of the strong points of this method is the assumption on noise statistical features: it is supposed to be uncorrelated zero-mean gaussian noise. Signals acquired on active volcanoes generally are affected by band overlapped noise (e.g., volcanic tremor and volcanic infrasound tremor; e.g., Cannata et al., 2013b; Cannavò et al., 2019). In the light of it, without loss of generality, well known sources of noise, like infrasound tremor, can be preventively filtered. While in correlation detectors the waveform is a single template or a stacked waveform (Gibbons and Ringdal, 2006), in the subspace approach the designed set of templates is built by means of the Singular Value Decomposition (SVD) of a matrix whose columns are a variable number of templates. Subspace methods have been carried out mainly in seismology, where they have been applied for earthquakes tracking, especially in case of aftershock sequences (Harris and Dodge, 2011; McMahon et al., 2017), as well as to identify low-frequency earthquakes in non-volcanic tremor (Maceira et al., 2010).

Volcano acoustic plays a fundamental role for both monitoring purpose and the study of the explosion dynamics and revealed to be a reliable tool to characterize eruptive activity and shed a light into the shallow plumbing structure system at Etna (Cannata et al., 2013a; Sciotto et al., 2013, Spina et al., 2015; Cannavò et al., 2019; Sciotto et al., 2019). Proper detection of signal of interest is a crucial, and at the same time critical, issue in volcano seismology, in that it allows extracting and collecting amplitude transient waveforms (events), which are therefore analyzed to provide information about spectral content and source location. In particular, in volcano monitoring, events detection has direct consequences in pinpointing the real number of generated events and identifying amplitude and occurrence rate variations. This information can be of support to follow the explosive activity and to improve the assessment of volcanic activity. This is particularly true on Etna, where multiple open-conduit vents exist, whose activity often consists of persistent Strombolian explosions, producing low amplitude and very close in time infrasound events. In order to accomplish the detection task, several algorithms have been proposed in literature; in particular, to overcome limitations of classical approaches such as short-time average/long-time average and cross-correlation detector, in this paper a subspace-based detection technique has been implemented.

In this paper, we attempt to clear the way to the application of subspace detection method in volcano-acoustics, previously investigated in Sciotto et al., (2011) in a preliminary study, comparing its performance with correlation and STA/LTA detectors. In particular, we test this technique on signals recorded by the infrasound permanent network deployed at Mt. Etna, which represents an ideal dataset to lead tests on this matter. Indeed, infrasound activity at Mt. Etna is almost continuous, and is also characterized by both discrete amplitude transients and continuous tremor, produced by several summit craters and eruptive fractures often opening on the flanks of the summit cones. Moreover, the infrasound signals are generated by different source mechanisms related to explosive activity, such as Strombolian activity and lava fountaining, as well as to degassing phenomena (Cannata et al., 2013a; Sciotto et al., 2013; Spina et al., 2015; Sciotto et al., 2019). Therefore, in a multi-vent and open-conduit volcano such as Mt. Etna, where volcanic activity is almost persistent and prone to eruptive fracture opening, infrasound signal can consist of a superposition of signals from different time-variant and stationary infrasound sources. If on the one hand each infrasound source is repetitive, on the other hand it can undergo modifications in time. In these cases, correlation detector may fail in detection of the variation in infrasound waveforms caused by these factors. Subspace detector is supposed to accomplish these two tasks: high sensibility and high flexibility.

DATA AND METHOD

For the purpose of subspace-based detection implementation, theory of detection problem is first introduced. Successively subspace approach is explained and an empirical procedure used to build and design signal subspace described. Other two sub-sections are dedicated to discuss the statistical analysis of

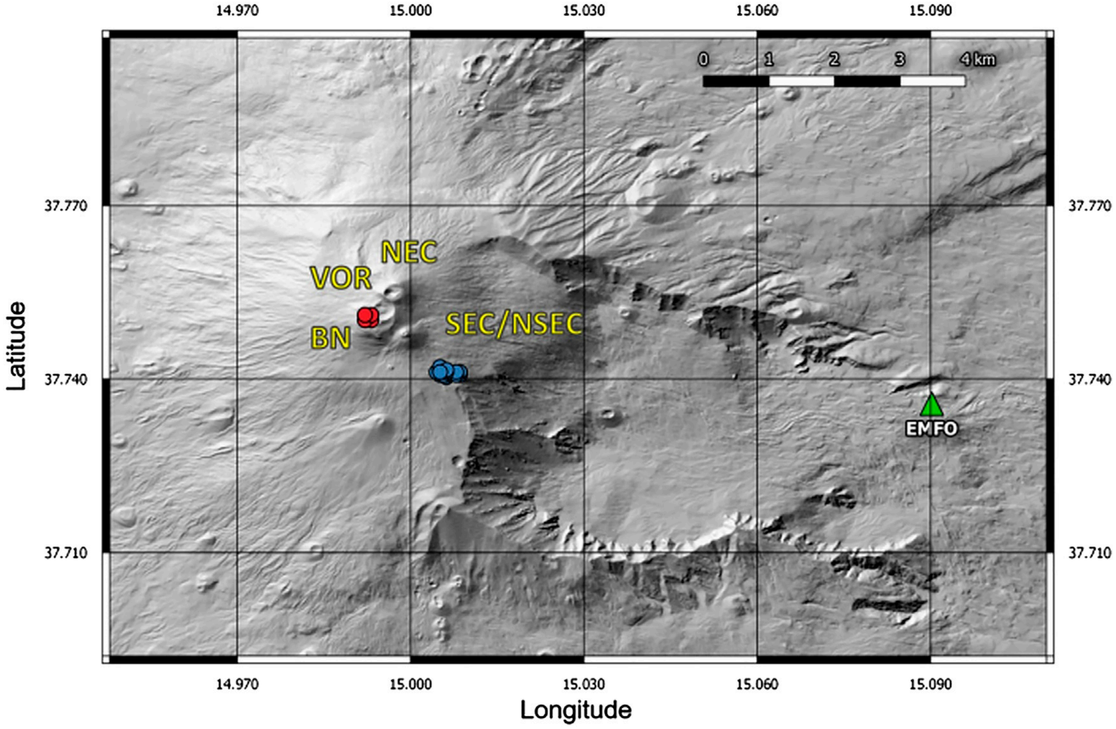


FIGURE 1 | Digital elevation model of Etna (Tarquini et al., 2007) with the location of the infrasonic station used in this work (green triangle EMFO), summit crater acronyms (VOR, Voragine; BN, Bocca Nuova; NEC, North-East Crater; SEC, South-East Crater; NSEC, New South-East Crater), infrasonic tremor source locations (red circles) and infrasonic event source locations (blue circles).

noise related to parameter estimation, both for subspace, correlation and STA/LTA detectors.

Subspace Detection Problem

Detectors usually implement a binary hypothesis test on the presence or absence of the signal of interest in a data observation window (Van Trees, 1968). In particular, the test is aimed to choose between the null hypothesis H_0, where noise only is present, and the alternative hypothesis H_1, where both the signal of interest and noise are present.

$$x[n] = \eta \quad \text{under hypothesis } \mathrm{H}_0, \tag{1}$$

$$x[n] = s + \eta \quad \text{under hypothesis } \mathrm{H}_1. \tag{2}$$

where $x[n]$ is the n-long window of continuous data, s is the signal of interest and η is the background noise assumed zero-mean Gaussian and temporally and spatially uncorrelated. In general, considering a multi-channel data acquisition, if N_t is the number of samples of observation window and N_c is the total number of data channel streams, the total number of samples N of a multiplexed data stream vector $x[n]$ is:

$$N = N_t. \tag{3}$$

In our framework, infrasound sensor acquires only one channel, so in **Eq. 3** $N = N_t$.

Signal s in **Eq. 2** is assumed deterministic and dependent on a vector of an unknown parameters a and expressed by unknown linear combination of a basis waveform:

$$s = Ua, \tag{4}$$

where U is a $N \times d$ matrix of d unknown signals that represent the subspace bases. The subspace dimension d takes value from 1 to length of vector $x[n]$: $d \in [1\ N]$. Without loss of generality, U can be made orthonormal:

$$U^T U = I, \tag{5}$$

where I is $d \times d$ matrix.

Under these assumptions, the probability densities function (pdf) for the recorded data under the null hypothesis H_0 (no events present) is:

$$p(x[n]|H_0) = \left[\frac{1}{2\pi\sigma^2}\right]^{N/2} \exp\left(-\frac{1}{2\sigma^2} x^t[n] x[n]\right), \tag{6}$$

while, under the null hypothesis of H_1 (events present), pdf can be expressed as:

$$p(x[n]|H_1) = \left[\frac{1}{2\pi\sigma^2}\right]^{N/2} \exp\left(-\frac{1}{2\sigma^2}(x[n] - Ua)^t (x[n] - Ua)\right). \tag{7}$$

As formulated by Harris (2006), the detection rule is a likelihood ratio test comparing the probability that the observed data are due to signal and noise to the probability that they are due to noise alone:

$$\Lambda(x[n]) = \frac{p(x[n]|H_1)}{p(x[n]|H_0)} < \gamma, \tag{8}$$

using **Eqs. 6**, **7**, the likelihood ratio test expressed in **Eq. 8** can be rewritten as a Generalized Likelihood Ratio Test (GLRT; Van Trees, 1968):

$$\Lambda(x[n]) = \frac{\underset{\{a.\sigma\}}{\max} p(x[n]|H_1)}{\underset{\{\sigma\}}{\max} p(x[n]|H_0)} < \alpha. \tag{9}$$

Using natural logarithm, **Eq. 9** can be rewritten as:

$$\begin{aligned} l(x[n]) = l(\Lambda(x[n])) &= -\frac{N}{2}\left(\frac{x^t[n]x[n] - x_p^t[n]x_p[n]}{x^t[n]x[n]}\right) \\ &= -\frac{N}{2}(1 - c[n]), \end{aligned} \tag{10}$$

where $l(x[n]) = l(\Lambda(x[n]))$ when the pdfs are in the exponential family, $x_p[n]$ is the least-squares estimate of the signal $x[n]$ in the detection window:

$$x_p[n] = UU^T x[n], \tag{11}$$

and $c[n]$, known as the subspace detection statistics, represents the ratio of the energy projected into the signal subspace U to the energy in the original data, and is given by:

$$c[n] = \frac{x_p^T[n]x_p[n]}{x^T[n]x[n]} \in [0\,1]. \tag{12}$$

The generalized likelihood ratio test (**Eq. 9**) detects an event of interest if the generalized log likelihood ratio (**Eq. 10**) exceeds a certain threshold α:

$$l(x[n]) = -\frac{N}{2}(1 - c[n]) > \alpha. \tag{13}$$

Considering the subspace detection statistics $c[n]$, an event is detected if:

$$c[n] > \gamma, \tag{14}$$

where γ is the threshold for the subspace and needs to be defined. In order to apply subspace detector based on **Eq. 14**, the first step is the construction of the signal subspace U, starting from the template matrix.

This matrix has peculiar characteristics, which are described in the *Template Matrix*, and consists of templates, representing previously observed events of interest, and is a fundamental tool for building the subspace. The number and type of templates needed to build the matrix depends on detector design. Signal subspace is the core of the algorithm, since it is the vector subspace used to represent the reference templates to be found into the signal. In order to extract orthonormal bases, Singular Value Decomposition (SVD) is applied to the template matrix, and then the dimension of the subspace is chosen. The dimension determines the amount of energy that the subspace is able to capture. Once the SVD is applied, d singular values are used to build the subspace (**Eqs. 4**, **11**). A few approaches have been implemented in literature to set this parameter, aiming to gain a compromise between detecting weak and less represented events (characterized by waveforms quite different from the reference templates) and having low false alarm or loss of significant events. In this paper, following Harris (2006), we selected the dimension parameter by means of an empirical approach making use of two different graphs as explained in *Subspace Design*.

Regarding the definition of the threshold γ, Harris (2006) studied the distribution of $c[n]$ statistics and derived the threshold using the Neyman-Pearson criterion (Van Trees, 1968). Under this criterion, the subspace dimension d is firstly determined by maximizing the probability of detection P_D for a fixed false alarm rate P_F using the following equations:

$$1 - F_{d,N-d}\left(\frac{\gamma}{1-\gamma}\frac{N-d}{d}\right) = P_F, \tag{15}$$

$$1 - F_{d,N-d}\left(\frac{\gamma}{1-\gamma}\frac{N-d}{d}, \overline{f}_c.N.SNR, \left(1 - \overline{f}_c\right).N.SNR\right) = P_D. \tag{16}$$

where P_F is evaluated from the cumulative central F distribution with d and N-d degrees of freedom under the null hypothesis H_0 and P_D is expressed in terms of the cumulative doubly non-central F distribution (Mudholkar et al., 1976) with the same degrees of freedom, $\overline{f}_c.N.SNR$ is the non-centrality parameter for the numerator, $[(1 - \overline{f}_c).N.SNR]$ is the non-centrality parameter for the denominator, fc is the average fraction of energy for all design set events, N is the embedding space dimension, and (N-d) is the dimension of the orthogonal complement of the signal subspace; finally SNR is the signal-to-noize ratio in the detection window.

Dataset

In order to design a dataset for the analysis, we chose a 1-h-long time interval (13:30–14:30 of May 30, 2019) of infrasound continuous signal recorded at EMFO station. This station belongs to the Infrasound Permanent Network run by Istituto Nazionale di Geofisica e Vulcanologia (INGV), is equipped with a GRAS 40AN microphone with a flat response at a sensitivity of 50 mV/Pa in the frequency range of 0.3–20,000 Hz and sampling rate of 50 Hz, and is located about 8 km far from Etna summit craters and about seven from the eruptive fracture (**Figure 1**). This station, together to ESLN (which is deployed at about the same distance from the summit area), was the only station able to record the explosive activity, is one of the less noisy stations among the permanent network, and, if compared with summit stations (located at higher altitude) is less affected by wind noise that can hide weak amplitude transients.

Signal buffer is characterized by infrasound events generated by an intense Strombolian activity that occurred at an eruptive fracture opened southeast of New Southeast Crater on the firsts

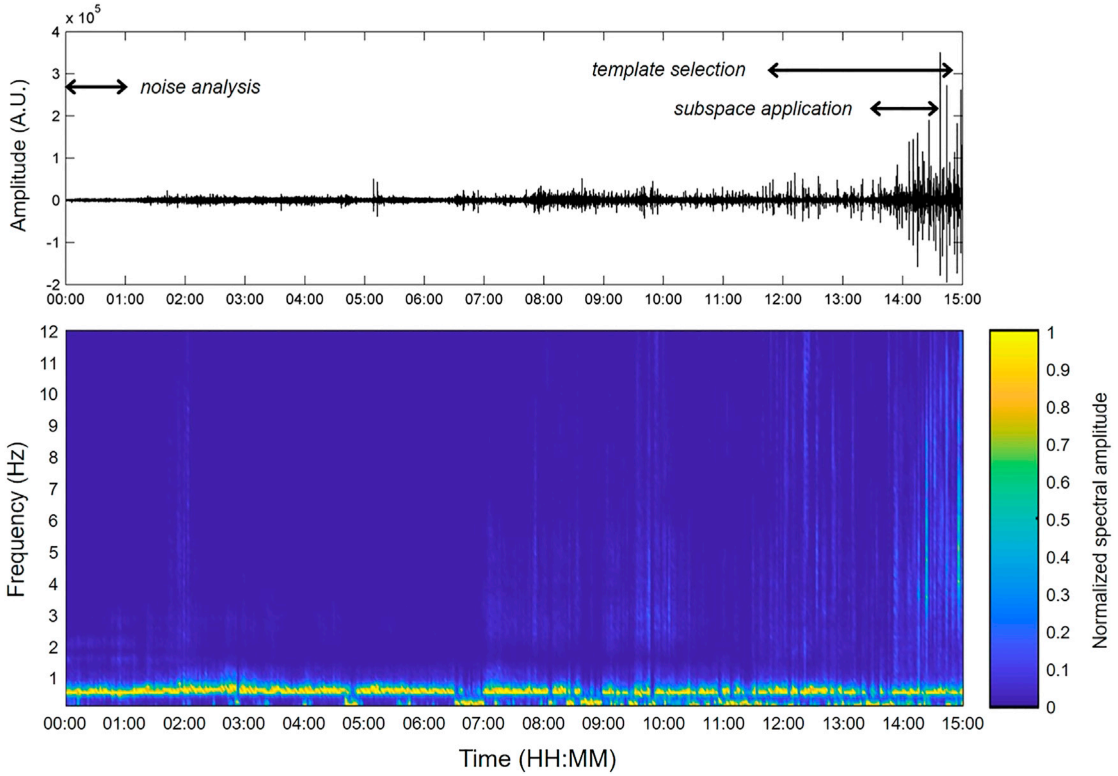


FIGURE 2 | Infrasound signal recorded at EMFO station on May 30, 2019 including the three buffers took into account in this work for statistical parameter estimation (00:00 -01:00), templates selection (12:00 -15:00) and subspace detector application (13:30 -14:30) (top) and normalized spectrogram, where spectra are averaged in 2-min-long windows (bottom).

hours of May 30, 2019 (here after NSEC, INGV-OE Internal Report, 2019). Here lava flows, ash emission, Strombolian and spattering activity took place. Explosive activity produced infrasound amplitude transients characterized by most of energy in the band 2.5–10 Hz, which are identifiable in the spectrogram from about 07:00 UTC, and with amplitude varying in a wide range (**Figure 2**). From about 14:00 UTC infrasonic activity at this fracture became more energetic, explosion generated infrasound events were more energetic and very close in time (**Figure 2**). Higher amplitude infrasonic events were detected and located by the real-time automatic system in force at INGV-OE (INGV-OE Internal Report, 2019) in correspondence of the eruptive fracture, as shown by blue circles in **Figure 1**.

In addition to infrasound events from the eruptive fracture, an overload continuous low frequency infrasonic tremor (~0.6 Hz, **Figure 2**), whose source was located at Bocca Nuova crater (BN; red circles in **Figure 1**), was recorded. These characteristics make the dataset particularly useful to be used as test for an automatic detection algorithm. In particular, the chosen signal is suitable for verifying the subspace capability to detect the maximum number of infrasound events, especially of low amplitude ones, and to compare its performance with other detection algorithms. Furthermore, it allows us to verify this triggering technique in presence of noise, which is represented by the overlying low frequency infrasonic tremor.

Subspace Algorithm Implementation

The subspace algorithm for event detection needs several key steps to be accomplished in order to be efficiently implemented, which are examined in following subsections and are summarized as follows:

- events of interest selection and pre-processing of template matrix (*Template Matrix*);
- statistical analysis of noise aiming to choose the threshold value (*Subspace Design*);
- subspace design (SVD and setting up of required parameters for subspace building) (*Threshold Setting*).

Three buffers of infrasound signal were selected for subspace method application (**Figure 2**): 1) a 1 h-long time interval of signal consisting of background noise, and with no infrasound events, recorded during the same day of the dataset of analysis, to carry out statistical parameter estimation (00:00–01:00 of May 30, 2019; all times are in GMT); 2) a 3 h-long time interval of signal characterized by infrasound events of interest, for waveform templates selection (12:00–15:00 of May 30, 2019); and 3) 1 h-long time interval to test the

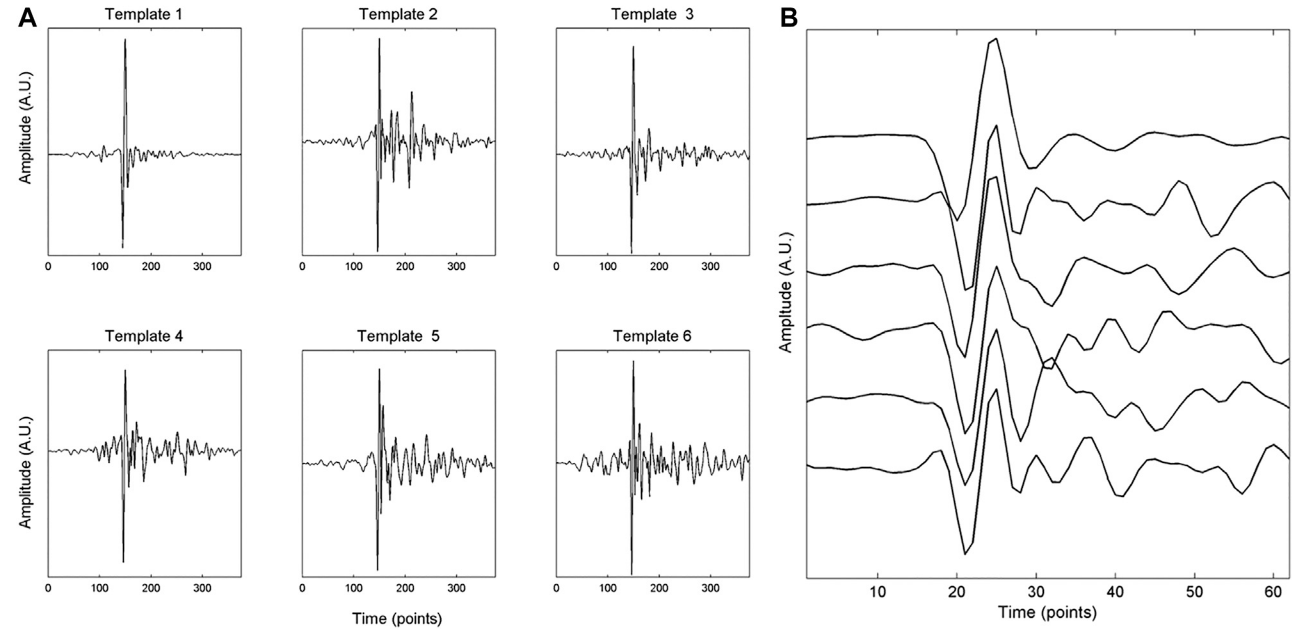


FIGURE 3 | Event waveforms composing the template matrix before **(A)** and after **(B)** pre-processing.

subspace detector and searching for events of interest (13:30–14:30 of May 30, 2019).

We performed a first test by detrending and filtering each signal window between 1 and 10 Hz, to get rid of noise such as wind and low frequency tremor generated by a second infrasound source (**Figures 1**, **2**), and set to zero mean and unit variance. A second test was performed by filtering signal between 0.5 and 10 Hz, in order to include the low frequency infrasonic tremor, and verify its influence on the detection.

Template Matrix

Building the template matrix is the preparatory step for subspace design. The template matrix is thought to consist of events of interest we are searching into the continuous signal. These can be manually selected, or, for a more robust procedure, waveforms can be automatically detected by a trigger algorithm (Maceira et al., 2010; Song et al., 2014). We made use of this last approach, and first triggered the events by means of STA/LTA energy detector. Secondly, we applied waveform cross-correlation, choosing an appropriate threshold, and selected the first event of each family, related to the infrasound source of interest (**Figure 3A**) in which the events were grouped. Once extracted, waveforms were aligned (**Figure 3B**); the algorithm has been designed to allow the operator to choose the alignment method. Waveforms can be aligned based on maximum or minimum amplitude value, or by means of manual picking. Successively, they were placed as columns in the template matrix.

Subspace Design

Signal subspace (U in **Eq. 4**) is the vector subspace used to represent the reference templates we want to find into continuous signal. The SVD provides the singular values allowing to build the subspace of the signal, that is a low-dimension representation of signal. Meaning of the dimension of subspace relies in the amount of energy that it is able to capture, and hence in the degree of waveform variation the algorithm is capable to detect. A few approaches have been implemented in literature (e.g., Harris, 2006; Song et al., 2014) to set this parameter, aiming to gain a compromise between detecting weak and less represented events (that is events having waveform quite different from reference template) and having low false alarm rate and possible loss of significant events. In this paper, following Harris (2006), we selected the dimension parameter by means of an empirical approach, making use of two different graphs. First, the fractional energy captured for each event is calculated:

$$f_c^i = a_d^{iT} a_d^i, \tag{17}$$

where f_c^i is the fraction of energy captured by the *ith* template and a_d^i is the *ith* unknown parameter of the coefficient matrix (**Eq. 4**, see Song et al., 2014 for further details). **Figure 4A** shows f_c of each template in function of the dimension of representation. The second plot is built by calculating the difference between the average captured energy in function of the dimension (**Figure 4C**):

$$\Delta\overline{f_c} = \overline{f_c}(d+1) - \overline{f_c}(d), \tag{18}$$

where $\overline{f_c}$ is the average fraction of energy captured by each of the d templates. Values of f_c^i and $\Delta\overline{f_c}$ determined for our dataset of analysis are plotted in **Figure 4**. In particular, in **Figure 4A** the dimension d, instead of maximizing P_D value (**Eq. 16**), is graphically determined as the lower value so that all curves (or the average curve) lay above the assumed percentage of energy we want to capture (e.g., 80 or 90%). As an alternative, in **Figure 4C**

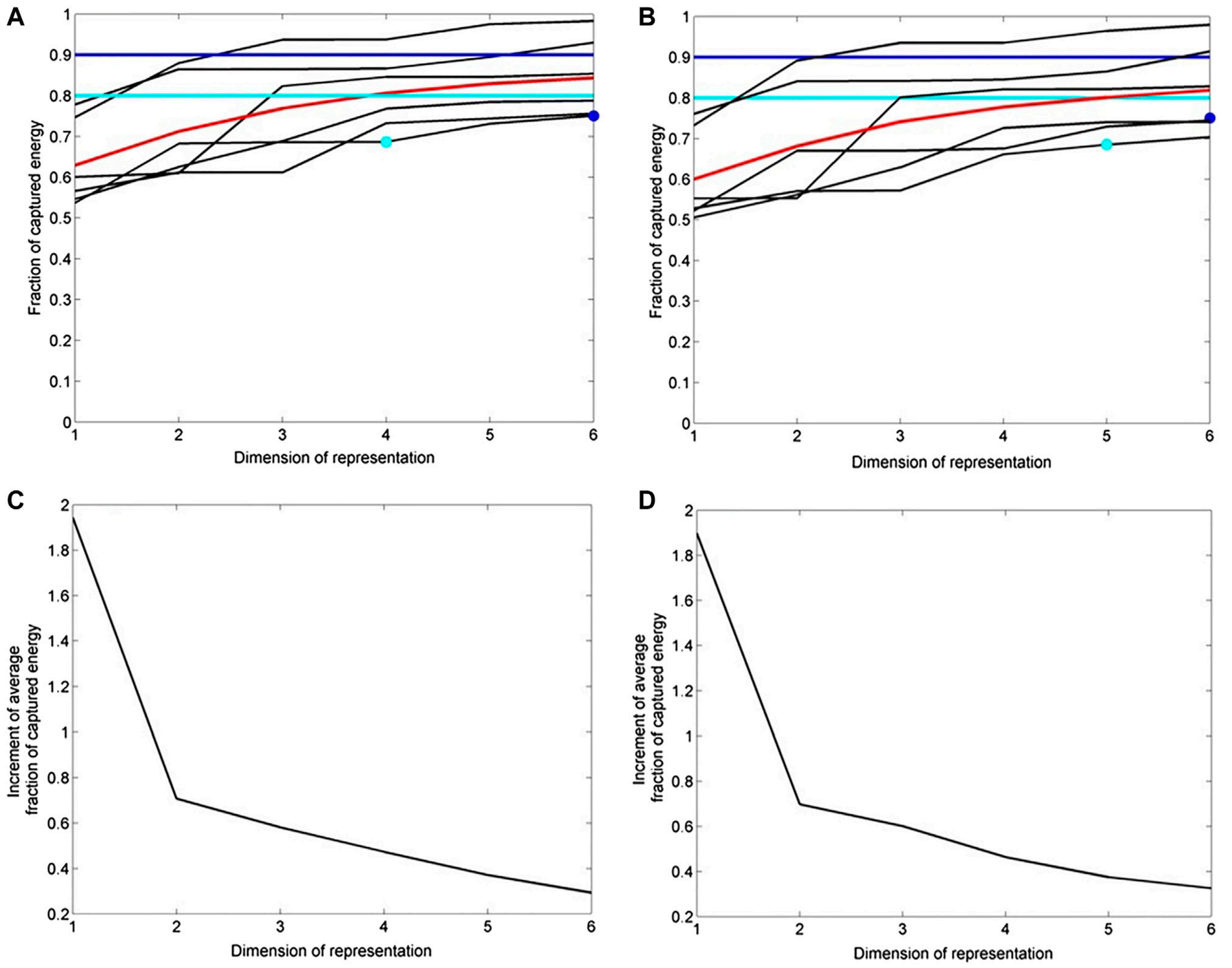


FIGURE 4 | Fraction of captured energy by each template of the designed matrix in function of the subspace dimension for signal filtered in the band 1–10 Hz **(A)** and in the band 0.5–10 Hz **(B)**. Black curve represents captured energy by each template, red curve the average captured energy, while, blue and light blue lines represent the 90 and 80 percentage of captured energy, respectively. Increment of average fractional captured energy (Δf_c), in function of subspace dimension for signal filtered in the band 1–10 Hz **(C)** and in the band 0.5–10 Hz **(D)**.

an adequate dimension is the value beyond which the amount of energy increase is negligible. Once the dimension is selected, the subspace is built.

The sufficient statistic for the subspace can be now calculated by implementing **Eqs. 11**, **12**, by means of windows of signal subspace sliding against the continuous signal, and compared against the threshold (**Eqs. 13**, **14**) to declare if an event is present.

In subspace detector algorithm, signal in a detection window is projected into a subspace spanned by the d columns of the subspace representation. The statistic is therefore the ratio of the squared norm of the projected vector to the squared norm of the original data vector (**Eq. 12**). It ranges between 0 and 1 and is a measure of the linear dependence between the signal and the orthonormal bases constituting the signal subspace. Every time the sufficient statistic exceeds the given threshold (γ) a detection is declared.

Threshold Setting

The choice of the threshold γ is always a compromise between an aggressive value, with the highest number of detections, even of less energetic events and leading to a high number of false detections, and a conservative value, when we want a minimum number of false detections at the cost of less true detections. Usually, threshold choice is based on the operator background experience and signal characteristics, that makes its value pretty subjective. In this paper, we tried to derive an empirical threshold value based on the data statistics and on the nature of the problem, e.g., waveform of events to be identified into incoming signals. In order to objectively compare the performance of subspace detector, we determined threshold of the three different applied triggering algorithms (subspace detector, correlation detector and *STA/LTA*) by means of the same approach.

Following Harris (2006) and Song et al. (2014), we implemented the Neyman-Pearson decision criterion (Van Trees, 1968). In the criterion, the threshold γ is derived from the false alarm rate with **Eq. 15**. In order to obtain the detection threshold, a few parameters need to be determined: 1) the false alarm probability (**Eq. 15**), 2) subspace dimension and 3) N. Regarding this latter, we should discuss about noise. Indeed, noise in the detection windows is assumed to be statistically uncorrelated. As Wiechecki-Vergara et al. (2001) point out, the effective dimension of the embedding space can be significantly lower than N if the data are filtered prior to detection. Noise could be correlated and could reduce the effective dimension of the embedding space even if data are not filtered. As Song et al. (2014) suggested, we applied the correction for the influence of the correlated noise, and estimated the effective embedding space dimension of detection windows $\widehat{N}$ used in subspace/correlation, and in STA/LTA detectors. According to Wiechecki-Vergara et al. (2001), the effective dimension of the embedding space $\widehat{N}$ is related to the variance of the sample correlation coefficient $\widehat{c_{ij}}$ between noise data η^j and event signal s^i.

$$\widehat{c_{ij}} = \frac{s^{iT}\eta^j}{\sqrt{\left(s^{iT}s^i\right)\left(\eta^{jT}\eta^j\right)}}. \tag{19}$$

In particular, once calculated the cross-correlation values using the specific window length N, the variance is obtained and the effective embedding space of the respective detection window (subspace, correlation, STA/LTA) are calculated by means of:

$$\widehat{N} = 1 + \sigma^{-2} \leq N, \tag{20}$$

in the light of it, **Eq. 15** can be rewritten as:

$$1 - F_{d,\widehat{N}-d}\left(\frac{\gamma}{1-\gamma}\,\frac{\widehat{N}-d}{d}\right) = P_F. \tag{21}$$

Hence, simple correlation detector can be written as:

$$\widehat{c} = \frac{s_m^T x}{\sqrt{\left(s_m^T s_m\right)\left(x^T x\right)}}, \tag{22}$$

where s_m is the master event data, x is data to be detected. A comparison between **Eq. 12** and **Eq. 22** shows that the correlation coefficient $\widehat{c}$ is equivalent to the square root of the subspace detection statistics $c[n]$ with a signal subspace dimension of $d = 1$. Here cross-correlation threshold is γ_c and both detectors have a false alarm rate:

$$1 - F_{1,\widehat{N}-1}\left(\frac{\gamma_c}{1-\gamma_c}\,\frac{\widehat{N}-1}{1}\right) = P_F. \tag{23}$$

After obtaining $\widehat{N}$ for subspace, next step is related to estimation of false alarm probability P_F by means of **Eq. 23**.

In order to accomplish this task, threshold γ_c is obtained by cross-correlating each template with noise data by means of **Eq. 19** (see Song et al., 2014 for further details). Cross-correlation value distribution is then plotted and the detection threshold is set aiming to obtain the minimum number of false detection (**Figure 5**). In this paper, we empirically estimated the correlator threshold as the value corresponding to a chosen percentile of the distribution.

Once the cross-correlation threshold γ_c and the effective embedding space of detection window ($\widehat{N}$) are known, false alarm probability can be estimated by inverting **Eq. 23**.

At this stage of the processing, we own all parameters needed to derive γ from **Eq. 21**, that is: 1) the false alarm probability P_F (**Eq. 23**), 2) the effective dimension of the embedding space $\widehat{N}$, and 3) the subspace dimension (whose method of derivation is exposed in following sub-section).

With the aim of evaluating the advantages/effectiveness of the subspace-based detector, we make a comparison with the performance of STA/LTA and the simple correlator trigger algorithms. As regards the former, the detection statistic is calculated:

$$r[n] = \frac{x_{STA}^T[n]x_{STA}[n]/N_{STA}}{x_{LTA}^T[n]x_{LTA}[n]\,/N_{LTA}}, \tag{24}$$

where x are data to be scanned, and the detection problem is formalized as:

$$r[n] > \gamma_r. \tag{25}$$

The STA/LTA threshold was determined by means of γ_r the same approach implemented for subspace, once computed $\widehat{N}$STA, $\widehat{N}$LTA (obtained by **Eqs. 19**, **20**) and P_F (obtained by **Eq. 23**), solving the following equation:

$$P_F = 1 - F_{\widehat{NSTA},\widehat{NLTA}}\left(\gamma_r\right). \tag{26}$$

Regarding the STA/LTA detector, based on event spectral content, we empirically set *STA* and *LTA* window length equal to 3 and 25 times the dominant period (Withers et al., 1998) respectively, corresponding with the lower frequency characterizing infrasound events.

Correlation detector was instead implemented as a subspace detector in which $d = 1$, by means of **Eq. 22**. Detection window length chosen for subspace and correlator scan was set to 4 times the dominant period to include the entire waveform (**Figure 3B**).

RESULTS

Subspace-based algorithm, as well as correlation detector and STA/LTA, were applied to the dataset of 30/May/2019 (13:30–14:30), which consists of the infrasound signal recorded by EMFO station, and is characterized by infrasound events located at the eruptive fracture and infrasonic tremor located at BN (**Figures 1**, **2**). Concerning the subspace method, we used, as event templates, waveforms extracted from signal recorded at the same station (**Figure 3A**), by means of the approach described in *Template Matrix*, setting a cross-correlation threshold equal to 0.6. Once

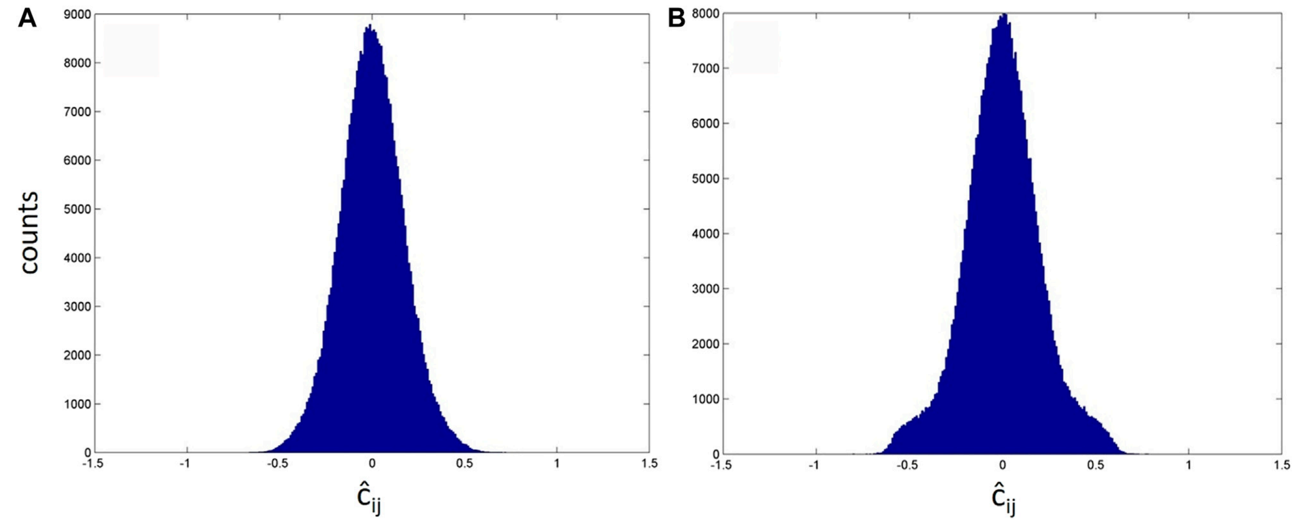


FIGURE 5 | Histograms of cross-correlation value distributions between event templates and noise data where infrasonic tremor is filtered out **(A)**, and with presence of infrasound tremor **(B)**.

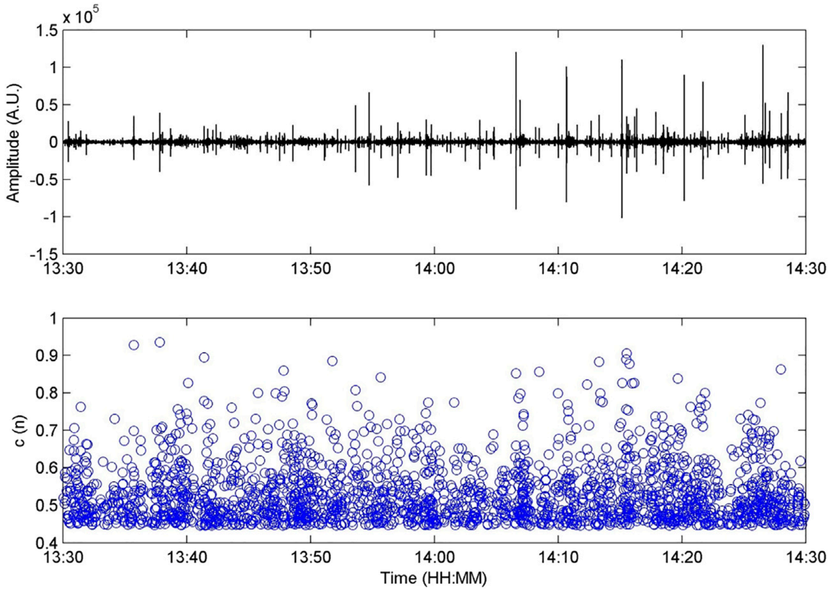


FIGURE 6 | Signal recorded at EMFO station and filtered in the band 1–10 Hz (top) and subspace detection statistic values (c(n)) above the estimated threshold (ɣ).

template matrix has been designed, waveforms were cut in 62 points-long windows, filtered, normalized and aligned by their positive peaks (**Figure 3B**). A first test was carried out by filtering signal in the frequency band 1–10 Hz. This frequency band was chosen with the aim of filtering out the correlated low frequency tremor.

We performed two computations, by using 99.9 and 99.99, as percentile for statistic threshold estimation (**Eqs. 22**, **23**, **Figure 6**), and built the subspace by using a dimension of representation equal to 4 (**Figures 4A,C**). The estimated thresholds and other setting parameters are reported in **Table 1**.

One subspace detection statistic $c[n]$ value is calculated in each detection window, with a sliding step of three points, we obtain $N_{tot}/3\ c[n]$, where N_{tot} is the signal buffer length. In **Figure 6**, the recorded signal and the sufficient statistics of subspace detector ($c[n]$) above the estimated threshold are shown.

With the aim of avoiding more values of detection statistics for each triggered event, we extrapolated one detection in a 1 s long

TABLE 1 | Setting parameters estimated or fixed in the detection.

Parameter	Estimated (E) or fixed (F) value		
Filtering band	1–10 Hz	1–10 Hz	0.5–10 Hz
Percentile	99.9 (F)	99.99 (F)	99.9 (F)
Probability of false alarm (PF)	7.4 e–04 (E)	6.02 e–05 (E)	8.9 e–04 (E)
Subspace dimension (d)	4 (E)	4 (E)	5 (E)
Threshold for subspace (ϒ)	0.45 (E)	0.53 (E)	0.63 (E)
Threshold for STA/LTA (ϒr)	2.33 (E)	2.74 (E)	2.18 (E)
Threshold for correlator (ϒc)	0.28 (E)	0.37 (E)	0.38 (E)
Detection window length for subspace/correlation	62 pt (F)	62 pt (F)	62 pt (F)
Detection window length for STA	43 pt (F)	43 pt (F)	43 pt (F)
Detection window length for LTA	375 pt (F)	375 pt (F)	375 pt (F)
Number of detections with subspace	881 (E)	483 (E)	201 (E)
Number of detections with STA/LTA	101 (E)	47 (E)	93 (E)
Number of detections with correlator	581 (E)	278 (E)	164 (E)

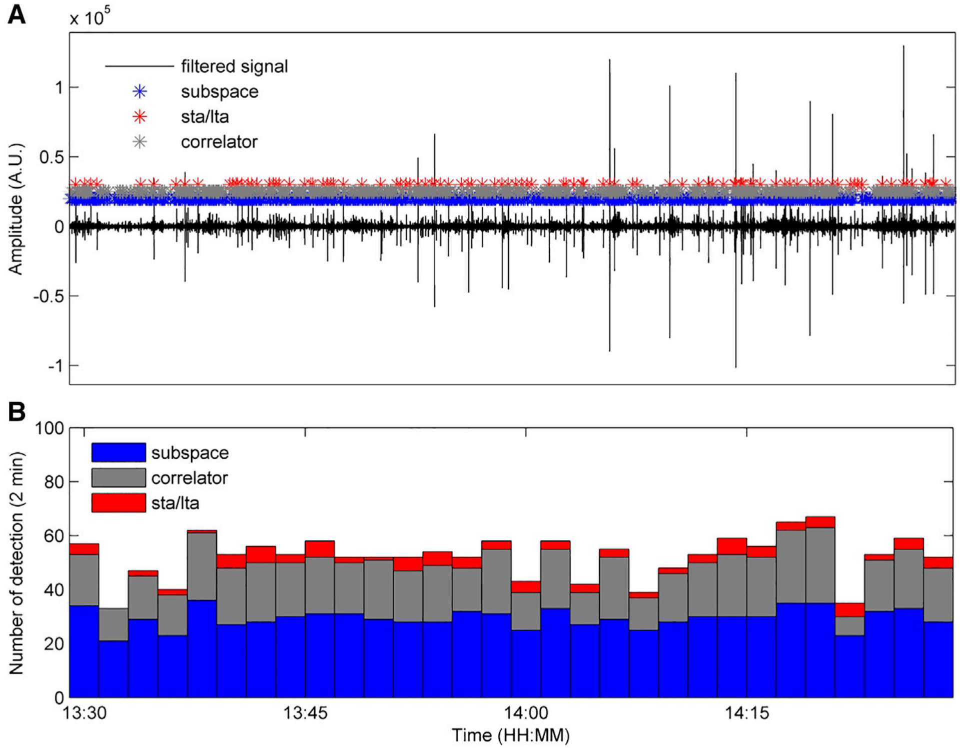


FIGURE 7 | Signal recorded at EMFO station and filtered in the band 1–10 Hz with the overlapped positions of detections obtained in 1-s long window by subspace (blue asterisks), correlation (green asterisks) and STA/LTA (red asterisks) detection algorithms for 99.9 percentile and using four SVD **(A)**. Event occurrence rate resulting from STA/LTA, correlation and subspace (bottom) detection algorithms counted in 2-min-long window **(B)**.

window. For the purposes of the comparison among subspace, correlator and STA/LTA performance, the detections computed by the aforementioned algorithms are overlapped to the continuous signal in **Figure 7A**. Histograms, obtained by counting detections in 2 min long windows for all algorithms, are shown in **Figure 7B**.

Plots of the occurrence rates show that subspace succeeds in detecting a higher number of amplitude transients, especially if compared with the STA/LTA triggered events (**Figure 7B**).

Results reveal the capability of the subspace-based algorithm to detect infrasound events of lower amplitude, while STA/LTA algorithm is able only to detect high amplitude transients. **Figures 8A,B**, which reports a zoom of continuous signal and the detected event positions, shows that subspace-based algorithm detected even more transients than correlator.

Successively, with the aim of highlighting the influence and the importance of parameter setting in this kind of approach, we run the detectors by setting the percentile equal to 99.99

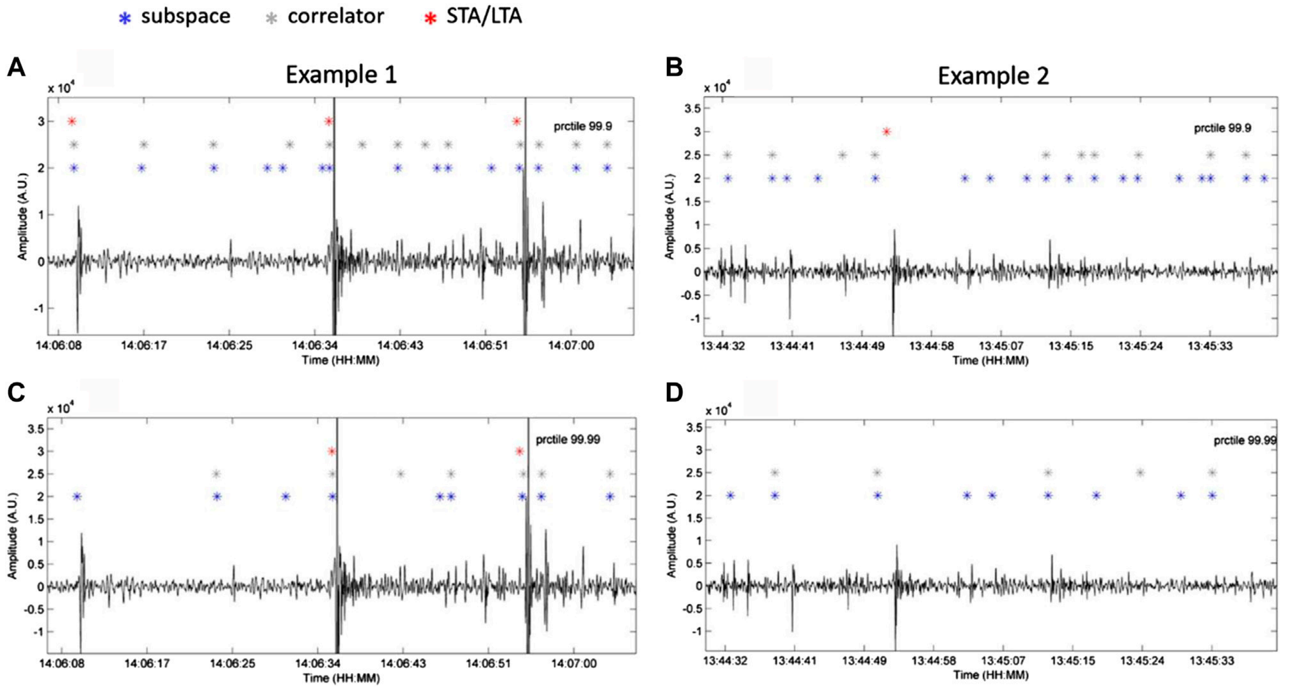


FIGURE 8 | Zoom in of **Figure 7A** (example 1, **A** and example 2 in **B**) by using 99.9 as percentile and respective comparison with results of **Figure 9A** using 99.99 as percentile **(C,D)**.

instead of 99.9, used to obtain the respective detection threshold. The estimated parameters are reported in **Table 1**. By using a 99.99 percentile, detection threshold of the three detectors raises and, as a consequence, we observe a decrease of detection number as also demonstrated by the occurrence rates (**Figure 9** and **Table 1**). Furthermore, subspace and correlator exhibit a similar trend of the event occurrence rate.

In the first case (percentile 99.9), the three detectors trigger more events with respect to the second one (percentile 99.99) (**Figures 8**, **9**). In particular, the subspace succeeds in the detection of very low amplitude transients. Nevertheless, a lower threshold can imply the identification of a higher number of false detections. By an inspection of signal buffer reported in **Figures 8**, almost all the detections are real and not false positives.

Reliable estimation of the detection capability when correlated noise, due to low frequency infrasonic tremor, is overlapped, was tested using the three described methods. In the light of it, signal was previously filtered in the band 0.5–10 Hz. In such a way, signal to be scanned is characterized by both infrasound events, exhibiting a frequency content in the band 2.5–10 Hz, and the infrasonic tremor, whose spectral peak is at ~0.6 Hz (**Figure 2**).

The resulting cross-correlation value distribution is shown in **Figure 5B**. A value of 99.9 as percentile was chosen, and the subspace was built by means of five SVD (**Figures 4B,D**). By introducing a correlated continuous tremor on the signal, the subspace and correlator thresholds result higher, due to the increase of variance, as expected (**Table 1**).

In this case, comparison with STA/LTA detector is inconsistent due to a not well defined statistics (**Eqs. 19**, **20**, **26**) used in the threshold γ_r computation. Despite this, simple correlation detector and subspace detector, from a practical point of view, exhibit robustness in the detection of waveforms transient as reported in **Figure 10**.

Comparison between **Figures 7B**, **10B**, whose results were obtained by using similar false alarm probability (**Table 1**), reveals that, if the low frequency infrasonic tremor is overlapped to the signal and not filtered out, the number of detections by means of both subspace and correlation based methods is lower. Furthermore, even in this case, subspace method succeeds in event detection with respect to the correlator detector.

DISCUSSION AND CONCLUSION

In the present work, we applied a subspace-based trigger algorithm for the automatic detection of infrasound amplitude transients in volcanic area. A 1 h-long buffer of continuous infrasound signal characterized by amplitude transients related to Strombolian activity, taking place at an eruptive fracture, was analyzed (**Figure 1**). In order to test the feasibility and performance of this technique, we made a comparison among

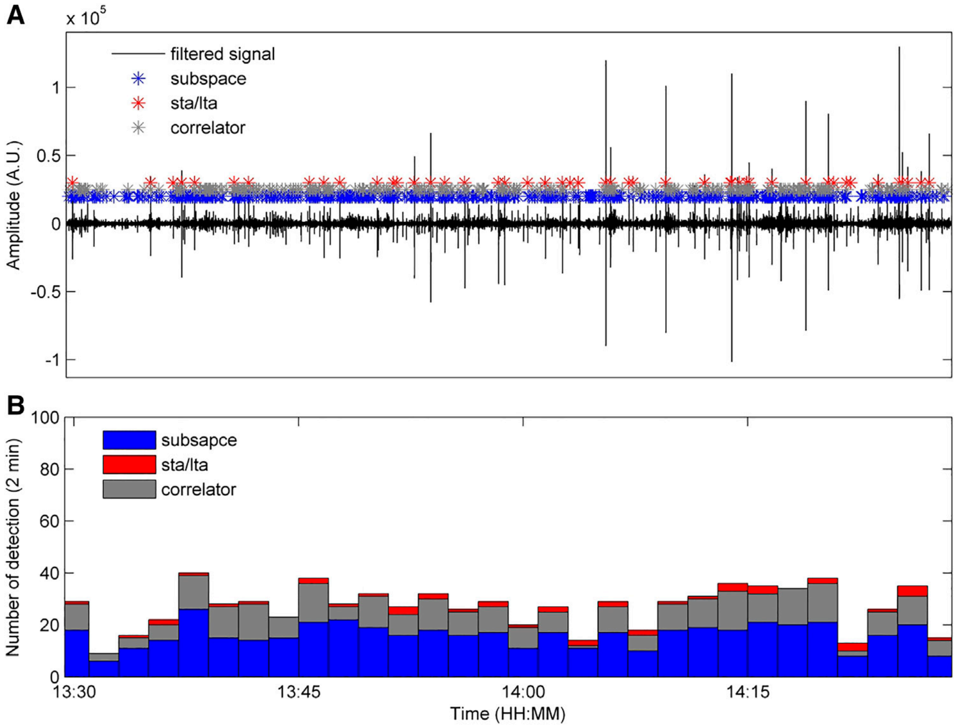


FIGURE 9 | Signal recorded at EMFO station and filtered in the band 1 -10 Hz with the overlapped positions of detections obtained in 1-s long window by subspace (blue asterisks), correlation (green asterisks) and STA/LTA (red asterisks) detection algorithms for 99.99 percentile and using four SVD **(A)**. Event occurrence rate resulting from subspace, correlation and STA/LTA detection algorithm counted in 2-min-long window **(B)**.

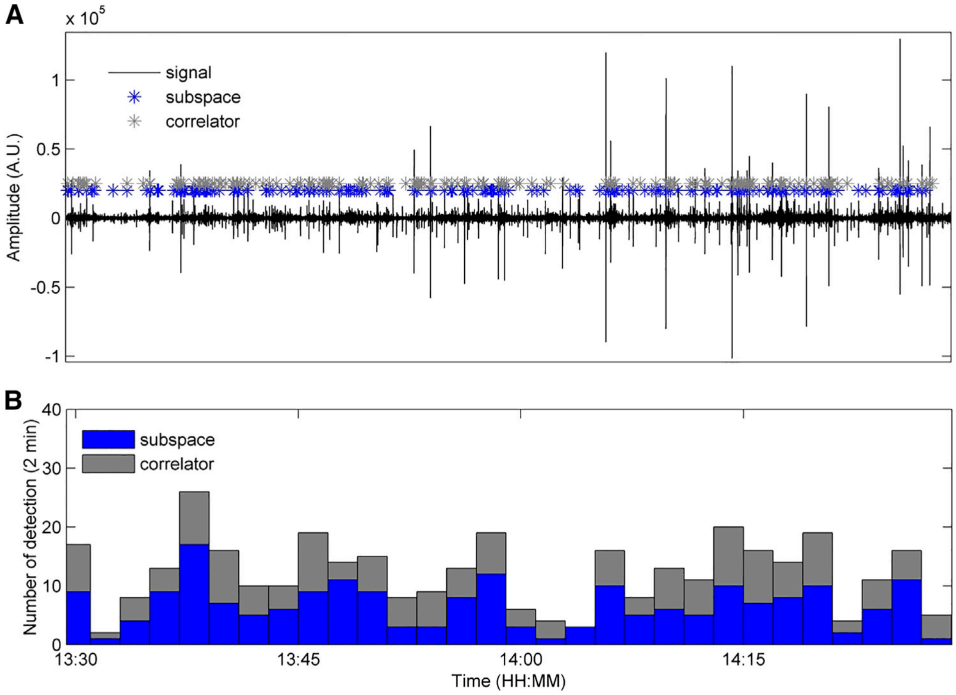
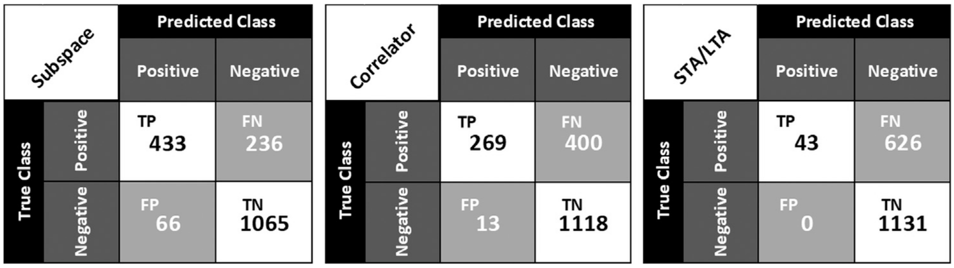


FIGURE 10 | Signal recorded at EMFO station and filtered in the band 0.5–10 Hz with the overlapped positions of detections obtained in 1-s long window by subspace (blue asterisks) and correlation (green asterisks) detection algorithms for 99.9 percentile and using five SVD **(A)**. Event occurrence rate resulting from subspace and correlation detection algorithm counted in 2-min-long window **(B)**.



Subspace		Predicted Class: Positive	Predicted Class: Negative
True Class	Positive	TP 433	FN 236
True Class	Negative	FP 66	TN 1065

Correlator		Predicted Class: Positive	Predicted Class: Negative
True Class	Positive	TP 269	FN 400
True Class	Negative	FP 13	TN 1118

STA/LTA		Predicted Class: Positive	Predicted Class: Negative
True Class	Positive	TP 43	FN 626
True Class	Negative	FP 0	TN 1131

FIGURE 11 | Confusion matrices of subspace, correlator and STA/LTA triggering algorithms. True positive (TP), false positive (FP), true negative (TN) and false negative (FN) values were obtained by an accurate analysis consisting of a visual inspection of the 1 h-long buffer of continuous infrasound signal and its comparison with the detections.

results obtained by implementing STA/LTA, correlation and subspace detectors. Several computations were carried out by using different setting parameters (**Table 1**).

Results highlight that subspace detector succeeds in detection of explosive activity related infrasound events. Indeed, **Figures 7B** and **9B**, show that subspace detector turns out to detect a higher number of amplitude transients with respect to both correlation and STA/LTA trigger algorithms. This is particularly true for a lower false alarm probability (that is choosing 99.99 as percentile of cross-correlation distribution values in **Eq. 22**, **Figure 5A**). Indeed, the ratio between the total number of detections of subspace with respect to correlation and STA/LTA is higher than for a higher false alarm probability (that is choosing 99.9) (**Table 1**).

It is worth noting that subspace detector is able to detect even infrasound events of low amplitude (**Figure 8**), which, especially in presence of noise, are difficult to be triggered by means of an energy detector. Several runs were performed by using variable setting parameter values. These tests demonstrate that the choice of setting parameter plays a fundamental role in the outcomes of elaborations. In particular, an in-depth statistical analysis of noise needs to be carried out, in that distribution value of cross-correlation between template waveform and noise determines the threshold of the detector.

Succeeding in the detection of all amplitude transients allows to monitor the time variation of occurrence rate, and thus to follow the evolution of explosive activity. Indeed, energy-based trigger algorithm (as STA/LTA) often fails in detection of low amplitude transients, events with low signal to noise ratio and events too close in time to each other.

With the aim of performing a quantitative estimate of the subspace effectiveness, especially in terms of false alarms and missed detections, and in general, to validate the results, a visual inspection of the 1 h-long signal buffer (**Figure 6**) and its comparison with the results obtained by means of the three algorithms were carried out. In particular, concerning the performance assessment goal, we chose to analyze the results of the second test presented into the work (**Figures 8C,D**, **9**), that is the one with percentile equal to 99.99 (which is the most conservative test), and inspected the true and false positive, and true and false negative events. We collected a dataset of 669 true events belonging to the same family (same source). Successively, by comparing waveforms with the detections (**Figure 9A**), a confusion matrix was calculated, for each of the three algorithms (**Figure 11**). The performance indices were then extrapolated from each confusion matrix (**Table 2**).

The validation results show that the subspace detector is generally characterized by the best indices among those derived from the confusion matrix (**Figure 11**; **Table 2**). In particular, it performs the best error rate (number of all incorrect predictions divided by the total number of dataset), accuracy (correct predictions divided by the total number of the dataset), and sensitivity, also called true positive rate, which is the number of correct positive predictions divided by the total number of positive, compared to the other two methods. In particular, STA/LTA gives the poorest results, as expected (**Table 2**). Nevertheless, it has to be highlighted that the

TABLE 2 | Performance indices, with relative equations, calculated from confusion matrices (see **Figure 11**) for subspace, correlator and STA/LTA triggering algorithms.

	Performance indices						
	Error rate (ERR %)	**Accuracy (ACC %)**	**Precision (PR %)**	**Sensitivity (SN %)**	**Specificity (SP %)**	**False positive rate (FPR %)**	**F-score (FS %)**
EQUATION	$\frac{FP+FN}{TP+TN+FP+FN}$	$1-ERR$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$	$\frac{TN}{TN+FP}$	$\frac{FP}{TN+FP}=1-SP$	$\frac{2\times SN\times PR}{SN+PR}$
Subspace	16.78	83.22	86.77	64.72	94.16	5.83	74.14
Correlator	22.94	77.06	95.39	40.20	98.85	1.15	56.57
STA/LTA	34.78	65.22	100.00	6.43	100.00	0.00	12.08

correlator and STA/LTA show the best precision, which quantifies the number of correct positive predictions divided by the total number of the positive prediction, specificity (representing the true negative rate) and false positive rate, due to the lower false positive values and higher true negative values (**Figure 11**; **Table 2**). Finally, F-Score, providing a single score balancing both precision and sensitivity, shows the highest value in case of the subspace detector, due to the high rate of the actual triggered waveforms (**Figure 11**; **Table 2**). In the light of this, the subspace detector method can be considered a good compromise in recognizing low-amplitude waveforms, particularly useful in tracking volcanic activity producing low-energy and very close in time events.

As regards the better performance of subspace over correlation detector, in terms of error rate, accuracy, and sensitivity, this can be ascribed to the events waveform variability. Indeed, subspace method allows making a comparison between the continuous signal and a set of waveforms of interest (templates) and the linear combination among them, instead of a single one or multiple in few cases (Hotovec-Ellis and Jeffrines, 2016; Senobari et al., 2019), as the correlation detector does. This makes the subspace detector particularly attracting and suited to detect infrasound events undergoing slight modifications in waveforms due for example to geometrical characteristic variations (e.g., vent/crater enlargement) (e.g., Cannata et al., 2011; Fee et al., 2017; Matoza et al., 2019b).

Automatic detection of seismo-volcanic events (seismic and infrasound) in monitoring framework and/or analyses of huge dataset is successfully currently carried out by means of energy detectors, correlation-based algorithms or methods making use of advanced signal processing techniques (e.g., Cannata et al., 2013a, Cannata et al., 2013b; Thompson, 2015; Hotovec-Ellis and Jeffrines 2016; Bueno et al., 2019; Matoza et al., 2019b; Senobari et al., 2019). Among more recent developed trigger algorithms, for example, VINEDA, designed for infrasound event detection (Bueno et al., 2019), parses the original signal into a characteristic function, whose amplitude is proportional to the sharpness of the original explosion onset. This latter method is very useful to increase signal to noise ratio, much better than the STA/LTA approach, and is able to detect even low-amplitude transients. Another recent method, REDPy (Hotovec-Ellis and Jeffrines, 2016), designed for earthquake detection, consist in performing a STA/LTA, storing the triggered events, and then cross-correlating waveforms to find similar events. REDPy is an effective tool in order to discover events belonging to the same family and offers the advantages and the disadvantages of both STA/LTA and cross-correlation. In the framework of our goal, that is to identify and extracting waveforms similar to each other (hence, belonging to the same source), subspace detector exhibits the advantages of being able to detect low-amplitude transients associated to a specific source, as well as transients showing slight waveform modifications. The latter feature is related to the use of SVD and the subspace method, which consists of the scanning of the signal buffer with all the linear combinations among basis vectors derived from the template matrix (previously designed). An additional feature of the subspace method lies in its reliability in the detection of events that are close in time to each other.

Furthermore, in the proposed subspace-based detection method, algorithm parameters are automatically tuned by implementing a statistical analysis of the background noise.

In general, as reported in Song et al., 2014, the limitation of the subspace detector is the complexity and relatively large computation cost in building the signal subspace. Nevertheless, the aim of our work focused on the statistical analysis of noise in order to optimize the quality of the trigger, in terms of low-amplitude event number.

In future developments, the subspace detector method has the potentiality to be adapted to multi-station framework, considering multiple instances or by multiplexing the data (e.g., Song et al., 2014; McMahon et al., 2017). In the perspective of a real-time implementation, a detector should scan a 2-min-long buffer. Actually, using simple Matlab scripts, it takes 0.4 s on average to process 2 min of infrasound data, seeming suitable for real-time processing. Future work will be dedicated to optimization and code compilation to build a module working in a real-time framework.

Findings of this work reveal the potentiality of subspace-based method in infrasound event detection. Advantages of this technique are particularly interesting in infrasound recorded in open-conduit volcanoes such Mt. Etna, where activity at summit craters often consists of persistent Strombolian explosions, producing infrasound events very close in time, that can take place at several vents, thus giving rise to multiple and time varying infrasound sources.

This first application of a subspace-based detection algorithm to infrasound signal proves that this is an efficient technique for identification and triggering of events in volcanic area.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MS is the main investigator of this research. She assembled and implemented the method of subspace-based detector and performed data elaborations. PM contributed to the design of the detection problem methodology and to the statistical analysis and contributed to the manuscript writing.

FUNDING

This work partially funded by the FISR project “SALE OPERATIVE INTEGRATE E RETI DI MONITORAGGIO DEL FUTURO: L’INGV 2.0 (S.O.I.R.).”

ACKNOWLEDGMENTS

We are indebted to the technicians of the INGV, Osservatorio Etneo and Italian Civil Protection Department (DPC) for funding and enabling the acquisition of infrasonic data. Also, we thank the reviewers and editors for their useful suggestions.

REFERENCES

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces. *Bull. Seismol. Soc. Am.* 68 (5), 1521–1532.

Bueno, A., Diaz-Moreno, A., Álvarez, I., De la Torre, A., Lamb, O. D., Zuccarello, L., et al. (2019). VINEDA—volcanic INfrasound explosions detector algorithm. *Front. Earth Sci.* 7, 335. doi:10.3389/feart.2019.00335

Cannata, A., Di GraziaAliotta, G. M., Cassisi, C., Montalto, P., and Patanè, D. (2013a). Monitoring seismo-volcanic and infrasonic signals at volcanoes: Mt. Etna case study. *Pure Appl. Geophys.* 170 (11), 1751–1771. doi:10.1007/s00024-012-0634-x

Cannata, A., Montalto, P., and Patanè, D. (2013b). Joint analysis of infrasound and seismic signals by cross wavelet transform: detection of Mt. Etna explosive activity. *Nat. Hazards Earth Syst. Sci.* 13, 1669–1677. doi:10.5194/nhess-13-1669-2013

Cannata, A., Sciotto, M., Spampinato, L., and Spina, L. (2011). Insights into explosive activity at eruptive fissure closely-spaced vents by infrasound signals: example of Mt. Etna 2008 eruption. *J. Volcanol. Geoth. Res.* 208, 1–11.

Cannavò, F., Sciotto, M., Cannata, A., and Di Grazia, G. (2019). An integrated geophysical approach to track magma intrusion: the 2018 christmas eve eruption at mount Etna. *Geophys. Res. Lett.* 46, 8009. doi:10.1029/2019GL083120

Fee, D., Izbekov, P., Kim, K., Yokoo, A., Lopez, T., Prata, F., et al. (2017). Eruption mass estimation using infrasound waveform inversion and ash and gas measurements: evaluation at Sakurajima Volcano, Japan. *Earth Planet Sci. Lett.* 480. doi:10.1016/j.epsl.2017.09.043

Gibbons, S. J., and Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophys. J. Int.* 165, 149–166. doi:10.1111/j.1365-246X.2006.02865.x

Gibbons, S. J., Sørensen, M. B., Harris, D. B., and Ringdal, F. (2007). The detection and location of low magnitude earthquakes in northern Norway using multi-channel waveform correlation at regional distances. *Phys. Earth Planet. In.* 160, 285–309. doi:10.1016/j.pepi.2006.11.008

Harris, D. B. (2006). *Subspace detectors: theory*. Livermore, CA: Lawrence Livermore National Laboratory Internal Report UCRL-TR-222758

Harris, D. B., and Dodge, D. A. (2011). An autonomous system for grouping events in a developing aftershock sequence. *Bull. Seismol. Soc. Am.* 101, 763–774. doi:10.1785/0120100103

Hotovec-Ellis, A. J., and Jeffries, C. (2016). *Near real-time detection, clustering, and analysis of repeating earthquakes: application to mount st. Helens and redoubt volcanoes*. Reno, Nevada: Seismological Society of America Annual Meeting

INGV-OE Internal Report (2019). 23/2019. BollettinoEtna20190604. Available at: www.ct.ingv.it.

Maceira, M., Rowe, C. A., Beroza, G., and Anderson, D. (2010). Identification of low-frequency earthquakes in non-volcanic tremor using the subspace detector method. *Geophys. Res. Lett.* 37, L06303. doi:10.1029/2009GL041876

Matoza, R. S., Arciniega-Ceballos, A., Sanderson, R. W., Mendo-Perez, G., Rosaldo-Fuentes, A., and Chouet, B. (2019a). High-Broadband seismoacoustic signature of vulcanian explosion at Popocatépetl volcano, Mexico. *Geophys. Res. Lett.* 46 (1), 148–157. doi:10.1029/2018GL080802

Matoza, R. S., Fee, D., Green, D., and Mialle, P. (2019b). *Volcano infrasound and the international monitoring system: challenges in middle atmosphere dynamics and societal benefits Book: infrasound monitoring for atmospheric studies*. Switzerland, Europe: Springer Nature. doi:10.1007/978-3-319-75140-5_33

McMahon, N. D., Aster, R. C., Yeck, W. L., McNamara, D. E., and Benz, H. M. (2017). Spatiotemporal evolution of the 2011 Prague, Oklahoma, aftershock sequence revealed using subspace detection and relocation. *Geophys. Res. Lett.* 44, 7149–7158. doi:10.1002/2017GL072944

Montalto, P., Cannata, A., Privitera, E., Gresta, S., Nunnari, G., and Patanè, D. (2010). Towards an automatic monitoring system of infrasonic events at Mt. Etna: strategies for source location and modeling. *Pure Appl. Geophys.* 167, 1215–1231. doi:10.1007/s00024-010-0051-y

Mudholkar, G. S., Chaubey, Y. P., and Ching-Chuong, L. (1976). Approximations for the doubly noncentral-F distribution. *Commun. Stat. Theor. Methods* 5, 49–63. doi:10.1080/03610927608827331

Sciotto, M., Cannata, A., Gresta, S., Privitera, E., and Spina, L. (2013). Seismic and infrasound signals at Mt. Etna: modelling of North-east Crater conduit and its relation with the feeding system of the 2008-2009 eruption. *J. Volcanol. Geoth. Res.* 254, 53–68. doi:10.1016/j.jvolgeores.2012.12.024

Sciotto, M., Cannata, A., Prestifilippo, M., Scollo, S., Fee, D., and Privitera, E. (2019). Unravelling the links between seismo-acoustic signals and eruptive parameters: Etna lava fountain case study. *Sci. Rep.* 9, 16417. doi:10.1038/s41598-019-52576-w

Sciotto, M., Rowe, C. A., Cannata, A., Arrowsmith, S., Privitera, E., and Gresta, S. (2011). Investigation of volcanic seismo-acoustic signals: applying subspace detection to lava fountain activity at Etna volcano. AGU–American Geophysical Union, Fall Meeting. San Francisco, CA.

Senobari, N. S., Funning, G. J., Keogh, E., Zhu, Y., Yeh, C. M., Zimmerman, Z., et al. (2019). Super-efficient cross-correlation (SEC-C): a fast matched filtering code suitable for desktop computers. *Seismol Res. Lett.* 90, 322–334. doi:10.1785/0220180122

Song, F., Warpinski, N. R., Nafi Toksőz, M., and Sadi Kuleli, H. (2014). Full-waveform based microseismic event detection and signal enhancement: an application of the subspace approach. *Geophys. Prospect.* 62, 1406–1431. doi:10.1111/1365-2478.12126

Spina, L., Cannata, A., Privitera, E., Vergniolle, S., Ferlito, C., Gresta, S., et al. (2015). Insights into Mt. Etna's shallow plumbing system from the analysis of infrasound signals. *Pure Appl. Geophys.* 172, 473–490. doi:10.1007/s00024-014-0884-x

Tarquini, S., Isola, I., Favalli, M., and Battistini, A. (2007). *TINITALY, a digital elevation model of Italy with a 10 m-cell size*. Ercolano, Italy: Istituto Nazionale di Geofisica e Vulcanologia (INGV).

Thompson, G. (2015). "Seismic monitoring of volcanoes," in *Encyclopedia of earthquake engineering*. Editors M. Beer, I. A. Kougioumtzoglou, E. Patelli, and S. K. Au (Berlin, Heidelberg: Springer). doi:10.1007/978-3-642-35344-4

Trnkoczy, A. (2012). *IASPEI New manual of seismological observatory practice 2 (NMSOP-2). Deutsches GeoForschungsZentrum GFZ, Understanding and parameter setting of STA/LTA trigger algorithm*. Berlin, Germany: Helmholtz-Zentrum. doi:10.2312/GFZ.NMSOP-2_IS_8.1

Van Trees, H. L. (1968). *Detection, estimation and modulation theory*. Hoboken, NJ: John Wiley & Sons.

Wiechecki-Vergara, S., Gray, H. L., and Woodward, W. A. (2001). *Tech. Rep. DTRA-TR-00–22, Statistical development in support of CTBT monitoring*. Dallas, TX: Southern Methodist University.

Withers, M., Aster, R., and Young, C. (1999). An automated local and regional seismic event detection and location system using waveform correlation. *Bull. Seismol. Soc. Am.* 89, 657–669.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., et al. (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bull. Seismol. Soc. Am.* 88 (1), 95–106.

Yokoo, A., Ishii, K., Ohkura, T., and Kim, K. (2019). Monochromatic infrasound waves observed during the 2014–2015 eruption of Aso volcano, Japan. *Earth Planets Space* 71, 12 doi:10.1186/s40623-019-0993-y

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.

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: frontiersin.org/about/contact

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