Article

Identification of Precursors in InSAR Time Series Using Functional Data Analysis Post-Processing: Demonstration on Mud Volcano Eruptions

Matteo Fontana [1,†], Mara Sabina Bernardi [1,*,‡], Francesca Cigna [2,§], Deodato Tapete [2], Alessandra Menafoglio [1] and Simone Vantini [1]

[1] MOX-Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; matteo.fontana@rhul.ac.uk (M.F.); alessandra.menafoglio@polimi.it (A.M.); simone.vantini@polimi.it (S.V.)
[2] Italian Space Agency (ASI), Via del Politecnico snc, 00133 Rome, Italy; f.cigna@isac.cnr.it (F.C.); deodato.tapete@asi.it (D.T.)
* Correspondence: marasabina.bernardi@polimi.it
† Current address: Royal Holloway, University of London, Egham Hill, Egham TW20 0EX, UK.
‡ Current address: Joint Research Centre (JRC), European Commission, Via E. Fermi 2749, 21027 Ispra, Italy.
§ Current address: Institute of Atmospheric Sciences and Climate (ISAC), National Research Council (CNR), Via del Fosso del Cavaliere 100, 00133 Rome, Italy.

Abstract: One of the most promising applications of satellite data is providing users in charge of land and emergency management with information and data to support decision making for geohazard mapping, monitoring and early warning. In this work, we consider ground displacement data obtained via interferometric processing of satellite radar imagery, and we provide a novel post-processing approach based on a Functional Data Analysis paradigm capable of detecting precursors in displacement time series. The proposed approach appropriately accounts for the spatial and temporal dependencies of the data and does not require prior assumptions on the deformation trend. As an illustrative case, we apply the developed method to the identification of precursors to a mud volcano eruption in the Santa Barbara village in Sicily, southern Italy, showing the advantages of using a Functional Data Analysis framework for anticipating the warning signal. Indeed, the proposed approach is able to detect precursors of the paroxysmal event in the time series of the locations close to the eruption vent and provides a warning signal months before a scalar approach would. The method presented can potentially be applied to a wide range of geological events, thus representing a valuable and far-reaching monitoring tool.

Keywords: InSAR data; post-processing; early warning; functional data analysis; functional boxplot

1. Introduction

Since the beginning of the space age, the possibility of using orbiting "observers" able to gather information about the Earth has stimulated the interest of scientific institutions, governments and the military. Among the plethora of data that can be gathered by satellite sensor arrays, Synthetic Aperture Radar (SAR) images and their advanced interferometric processing (InSAR) stand out as some of the most valuable resources for monitoring natural geohazards [1]. This has been made possible due to the significant leap in data transmission technologies in recent years, coupled with an increase in computing power, storage capacity and sensor availability, leading to a substantial expansion in the potential applications of satellite imaging data.

The use of interferometric techniques allows the extraction of topographic and kinematic information about the measured ground surface, by comparing the phases of subsequent images, and achieving millimeter accuracy against independent geodetic monitoring data (e.g., [2–5]). As highlighted in the review article by Bernardi et al. [6], there are numerous ongoing efforts to apply conventional statistical frameworks, like time-series analysis

and geostatistics, to the output products of InSAR processing. However, the statistical analysis of these products is still in its infancy. The goal of this work is to apply advanced statistical techniques to the analysis of InSAR data to enhance the practical use of these data. In particular, we aim at tackling the problem of detecting early warning signals for geohazards.

In the last decade, various works have addressed the problem of detecting anomalous patterns in time series of ground motions obtained by InSAR processing. Berti et al. [7] classified time series of ground displacements into classes by statistically comparing them to reference displacement patterns selected by experts. Chang and Hanssen [8] proposed a similar approach based on a multiple hypothesis testing procedure to compare the time series to a comprehensive set of alternative models built from a library of canonical kinematic models. Li et al. [9] proposed an approach, based on machine learning algorithms, to classify the time series into five categories. Cigna et al. [10] defined two deviation indexes that assess, respectively, the change in the velocity and/or acceleration and the magnitude of the discontinuity, assuming a linear trend; Tapete and Casagli [11] applied one of the two deviation indexes developed by Cigna et al. [10] across the whole duration of deformation time series to detect trend variations not only in correspondence with specific time periods, but also at other epochs that may have not been known a priori. Notti et al. [12] proposed a methodology for the analysis of deformation time series including the approaches of Berti et al. [7] and Cigna et al. [10] to investigate landslides and land subsidence processes.

All the mentioned approaches are based on an a priori selection of a prespecified model for the trend. The current literature presents a gap in the application of model-free approaches based on advanced statistical methods. The current study aims at filling this gap by applying state-of-the-art statistical techniques. Indeed, the proposed approach does not rely on prior assumptions about the expected motion trend and the shape of the anomaly, and it is, therefore, able to identify anomalies presenting any kind of deviation from the regular trend, by setting as a reference the nearby curves, thus using the data themselves to define what is the "normal" behavior of the displacement evolution. Moreover, the aforementioned works use the spatial distribution of the data only for representation and interpretation of the results. In the proposed approach, the spatial dependence of the data is included in the analysis. To the best of our knowledge, there has yet to be an attempt to integrate temporal and spatial dynamics, or to utilize more sophisticated statistical methods, to extract new insights from these data.

Our aim is to develop an efficient statistically based tool capable of providing key information to feed into geohazard monitoring and early warning systems, based on a Functional Data Analysis (FDA) framework [13]. This statistical methodology is particularly suited for the problem considered, as it appropriately accounts for the intrinsic regular nature of the phenomenon under study. Indeed, geological motions follow a continuous trajectory in time (although sometimes abrupt, as in the case of paroxysmal events) with a degree regularity given by the forces involved. The proposed technique relies on the assumption that the InSAR data used for the analysis carry enough information to extract meaningful information about the phenomenon considered, with enough geographical locations to understand its spatial distribution and enough time points to extract its temporal evolution. The proposed method is here applied to an illustrative case (the detection of precursors to a mud volcano eruption in Sicily), but its general formulation makes it suitable for the detection of early warning signals of other volcano eruptions and, more generally, to the detection of early warning signals of a wide range of geological events. For example, Moro et al. [14] illustrated the possibility of detecting seismic precursor signals from InSAR data in the case of the 2009 L'Aquila earthquake.

In the eastern part of Caltanissetta, a city in Sicily Island (Italy), lies the village of Santa Barbara. This village was the site of a significant event: a mud volcano eruption, which we utilize as a test case in our study. The eruption, which occurred on 11 August 2008 was of such intensity that it caused damage to urban infrastructure up to 2 km away from

the main eruptive vent. For a more comprehensive understanding of the event and its geological characteristics, one can refer to Cigna et al. [10] and Madonia et al. [15]. The dataset exploited in this work is generated using a well-established multi-temporal InSAR processing workflow, and the initial results were briefly presented in Fontana et al. [16], where they also underwent a preliminary and exploratory analysis based on FDA. While Fontana et al. [16] already suggested the possibility of using FDA for the analysis of InSAR data, by applying a functional clustering algorithm and proposing a forecasting technique, in this work, we innovate by exploiting the FDA framework to tackle the specific problem of early warning detection and present a fully fledged algorithm able to detect precursors to the paroxysmal event in the considered application.

2. Materials

The dataset analyzed in this work, already described in Fontana et al. [16], is obtained from 32 ENVISAT Advanced SAR images along ascending track T172. The period of acquisition starts on 12 October 2002 and ends on 7 June 2008 which is the last date before the eruption of the mud volcano. The data are acquired in C-band (5.6 cm wavelength, 5.3 GHz frequency) and they are characterized by a Line-Of-Sight (LOS) with ∼23° look angle and VV co-polarization. The ground resolution is ∼20 m and the nominal site revisit is 35 days. The algorithm used for InSAR processing is the well-established Small Baseline Subset (SBAS) technique [17,18], a robust multi-temporal InSAR implementation that was originally developed in 2002 and widely exploited by the scientific community for a number of geohazard applications at the local, regional, national to continental scales (e.g., [19–22]). In the resulting dataset, 1735 coherent targets are retained and their geographical distribution covers an area of 150 km^2. The algorithm estimated, for each target, the annual LOS velocity, the time series of LOS displacements, the temporal coherence and the elevation above the reference ellipsoid. The algorithm proposed in Section 3 is applied to the displacement time series. The geographical locations of the coherent targets are represented in Figure 1, and the corresponding time series of LOS displacements are shown in Figure 2. The precision of the dataset, i.e., the standard error, is on average 0.4 mm/year for the velocity estimates and 3.6 mm/year for each displacement record, across the whole processed area. These values depend on the quality of the scatterers (e.g., their coherence) and the number of available observations (i.e., images, and small baseline interferograms) [23]. The dataset shows a generally stable scenario across the processed area, with LOS displacement velocities between ±0.5 mm/year for over 98% of the coherent targets. Some ground displacement velocity peaks of up to −12.81 and +18.56 mm/year (in the direction away from and towards the sensor, respectively) can be observed (Figure 1). These correspond to maximum cumulative values of −69 to +108 mm LOS displacement over the 2002–2008 period. The ground deformation scenario in Caltanissetta for the time period between 2002 and 2005 was described by Vallone et al. [24]. Moreover, the specific event of the mud volcano eruption, although on a different dataset than the one analyzed in this work, was considered in Cigna et al. [10], where Deviation Indices were computed to semi-automatically identify changes in InSAR time series.

The displacement data considered can be thought as a discrete sampling (at the time instants of SAR acquisitions) of a continuous phenomenon (the ground deformation). The physical constraints of the deformation dynamics suggest a degree of smoothness of the phenomenon under study. Indeed, the first and second derivatives of the trajectories represent, respectively, the velocity and the acceleration profiles of the displacement. Moreover, we assume the presence of additive error induced by the measuring process and the InSAR processing. As already explored by Fontana et al. [16], FDA is a suitable framework able to properly account for the particular features of the data considered. Indeed, FDA is the branch of statistics that considers as statistical units smooth functions depending on a continuous variable (e.g., time, space or frequency).
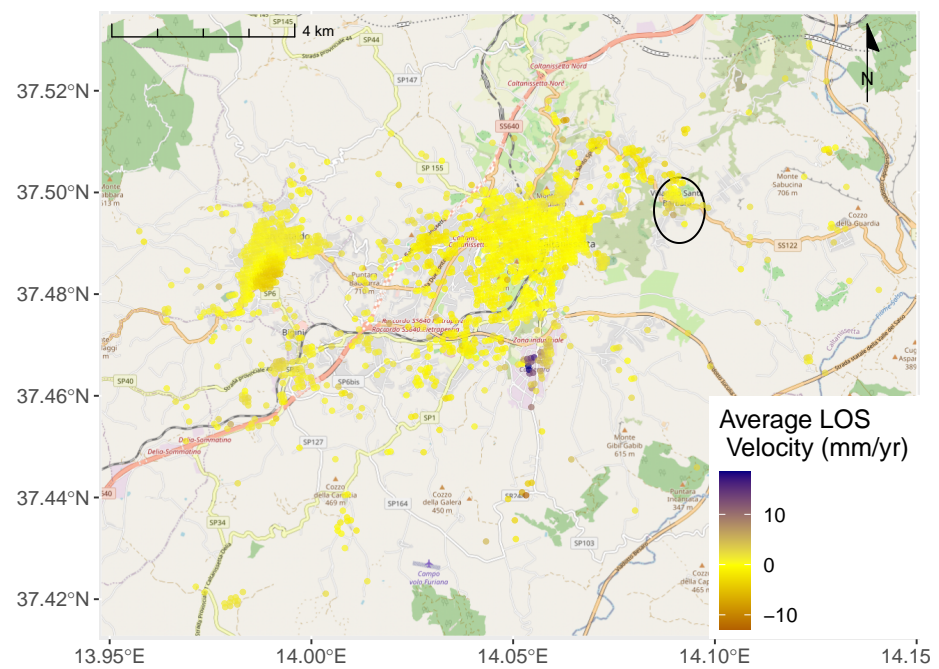
Figure 1. Map of the geographical position of the target points on an image of the area of Caltanissetta (Italy), with information about the Line-Of-Sight average displacement velocity (in mm/year) across the observation period, minimum value is −12.8, maximum value is 18.6. The area of the Santa Barbara village ($d^* < 750$ m) is outlined by a black ellipsoid east of Caltanissetta town. Copyright for the base map is owned by OpenStreetMap contributors.
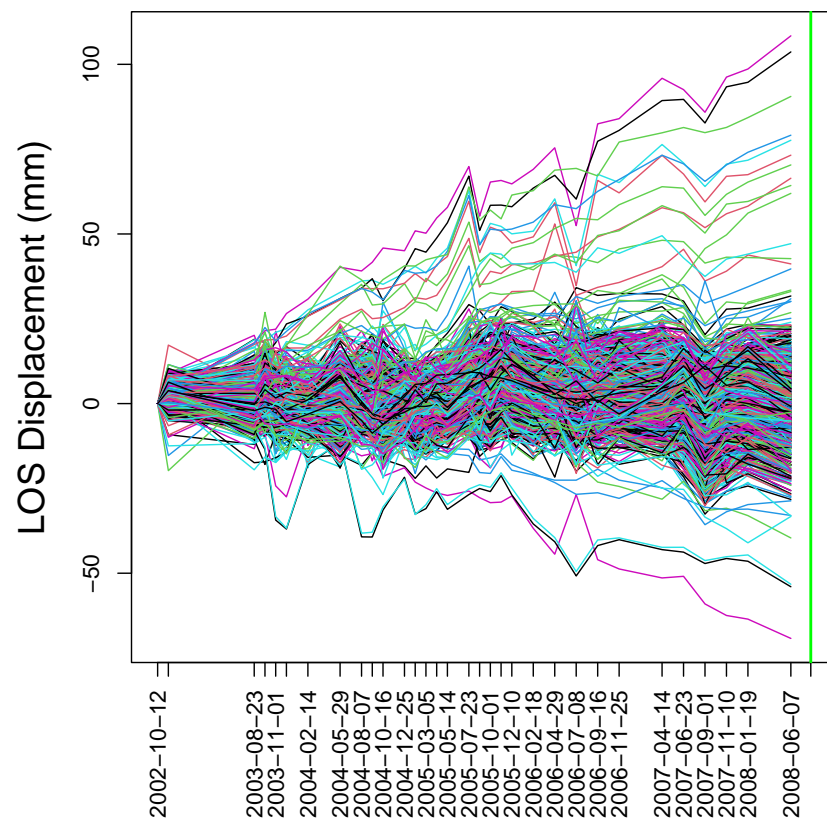
Figure 2. Plot of the time dynamics of LOS displacement for all the target points shown in Figure 1. The vertical green line marks the date of the event, 11 August 2008

3. Methods

The assumption is that, in an operational scenario of real-time monitoring of the evolving situation in an area where an event is expected to happen or has already started to occur, whenever a new satellite image is collected, this image is ingested into the processing chain, processed and converted into new information about the ongoing deformation. Therefore, apart from the time needed for the satellite image downlink to the ground station and its provision from the image provider ground segment (that for a satellite mission designed to provide imagery in emergency contexts is by definition highly shortened), the assumed operational scenario aims to approach the "real-time" performance and thus a valuable situation to address application purposes of civil protection during emergencies.

We select a calibration period of 1 year, using the six scenes from 12 October 2002 to 6 December 2003. This "burn-in" is required by the SBAS procedure to produce reliable displacement series. We then proceed to analyze each subsequent scene, namely from 6 December 2003 to 7 June 2008.

More specifically, with $y_{p_i}(t), t \in T$ being the generic deformation curve, observed on the temporal domain T at location $p_i = (\phi_i, \psi_i)\ i \in 1, \ldots, n = 1734$, where ϕ_i, ψ_i are, respectively, the latitude and the longitude of the point, and $t^* \in [6$ December 2003, 7 June 2008] being each time instant after the calibration period, the analysis performed at a given t^* considers only the data in a specific period of time T defined as [12 October 2002, t^*]. A representation of the measured points overlaid on a map of the area of interest is available in Figure 1, while their time dynamics can be seen in Figure 2.

The method identifies anomalies by comparing the time series of locations near the point to be monitored. Therefore, we restrict our attention to deformation curves measured in geographical positions that are close enough (in geographical terms) to the area in which the mud volcano is: in more mathematical terms, with p_v being the position of the mud volcano, we restrict our analysis to those geographical positions p^* for which $d_g(p_v, p^*) < d_{lim}$, where $d_g(\cdot, \cdot)$ is the geodetic distance between two geographical locations. To compute the geodetic distance, we use the inverse method by Vincenty [25] using the *WGS-84* ellipsoid.

To filter out the measurement error, and to recover the smooth structure of our subset of displacement curves, we employ a smoothing procedure based on a b-spline basis of order 6 (i.e., degree 5). This is to ensure to have cubic b-splines on acceleration curves. The smoothing parameter is then selected in a data-driven way by minimizing generalized cross-validation error, as is it commonly performed in the FDA realm [13].

After the smoothing, to focus our attention on recent variations in displacement, we restrict, for each t^*, the analysis to $y_{p^*}(t), t \in T^*$, where $T^* = [t^* - k, t^*]$, for a fixed parameter k. In other words, for each date after the calibration period in which a scene was acquired, we focus our attention on the portion of deformation curves that includes the k days that reach into the past with respect to the time point t. This parameter represents the "memory" of our method. To have data points in the period T^*, and also for the first time instants, considered, the calibration period should have a length larger or equal to k.

Moreover, since we are interested in differences in displacement that are relative, we normalize each displacement curve in the following way:

$$\tilde{y}_{p^*}(t) = y_{p^*}(t) - y_{p^*}(t^* - k), t \in T^*. \tag{1}$$

In other words, we shift each curve vertically so that its value at $t^* - k$ is zero.

We now want to determine, for each t^*, what curves $\tilde{y}_{p^*}(t)$ are outlying with respect to the other curves considered. We use the functional generalization of the classical boxplot [26]. The functional boxplot revolves around the concept of data depth: namely, a method for multi- or infinite-dimensional data that allows the establishment of a natural ordering between points that are deep in the data cloud and points that are shallow. In this specific case, the depth considered is the *Modified Band Depth* (MBD) by Lopez-Pintado and Romo [27]. Following Liu et al. [28] and Sun and Genton [26], we define the median function as

$$y_{median}(t) = y_{[1]}(t), \tag{2}$$

where $y_{[r]}$, for $r = 1, \ldots, n$, is the r-th deepest curve according to MBD, so $y_{[1]}(t)$ is the deepest curve and $y_{[n]}(t)$ the shallowest.

The 50%-central region, i.e., the functional equivalent of the univariate interquartile range, is

$$C_{0.5} = \left\{ (t, y(t)) : \min_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) \le y(t) \le \max_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) \right\}. \tag{3}$$

The upper and lower whiskers are defined as:

$$w_{upper} = \max_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) + F\left(\max_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) - \min_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) \right) \tag{4}$$

and

$$w_{lower} = \min_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) - F\left(\max_{r=1,\ldots,\lceil n/2 \rceil} y_{[r]}(t) - \min_{r=1,\ldots,\lceil n \rceil} y_{[r]}(t) \right). \tag{5}$$

where n is the number of functional observations and F is a custom parameter, usually set to 3, a standard value in the statistical literature, and commonly used as the default one in statistical packages, trading off between the type I and type II error rates of the outlier detection procedure. $F = 3$ also yields a good trade-off between data overfitting and the discrimination performance of the method, as inspected via an empirical evaluation which we omit for the sake of conciseness. All the curves that are outside of the region of space identified by w_{upper} and w_{lower} for at least one point are classified as outliers.

The essential gain provided by the use of the FDA framework and the concept of a functional boxplot is that a methodology designed like this is not only able to take into account the geographical dimension of the problem (provided by the geographical restrictions), but also the time dimension, taken into account by the use of some of the previous deformation history of a given point, instead of a single scalar value.

The algorithm was tested on the Santa Barbara (Caltanissetta) case study. In this specific case, given the characteristic dimension of the problem in geographical terms, we select $d_{lim} = 750$ m. This choice restricts the analysis to 39 points. A representation of them overlaid on a terrain map can be found in Figure 3, while their time dynamics can be seen in Figure 4. Within this area, annual LOS displacement velocities are between -2.8 and $+6.4$ mm/year, and the cumulative LOS displacements reach $+62.0$ and -33.3 mm during the 2002–2008 period.

With respect to the temporal dimension, the k parameter is set to 365 days, according to empirical observations that define the "natura" temporal scale of paroxysmal phenomena such as the ones in Caltanissetta of one year. The parameter for the smoothing procedure is set as described in the previous section, and equals $\lambda = 10^{10}$. The choice of the parameters d_{lim} and k is here based on empirical observation of the data, but it is also coherent with the dimension/scale of the mud volcano and the affected/impacted area (see geological information on the event in Cigna et al. [10], and the maps in Brighenti et al. [29]) and, from the temporal point of view, with the potential time scale of any precursors.
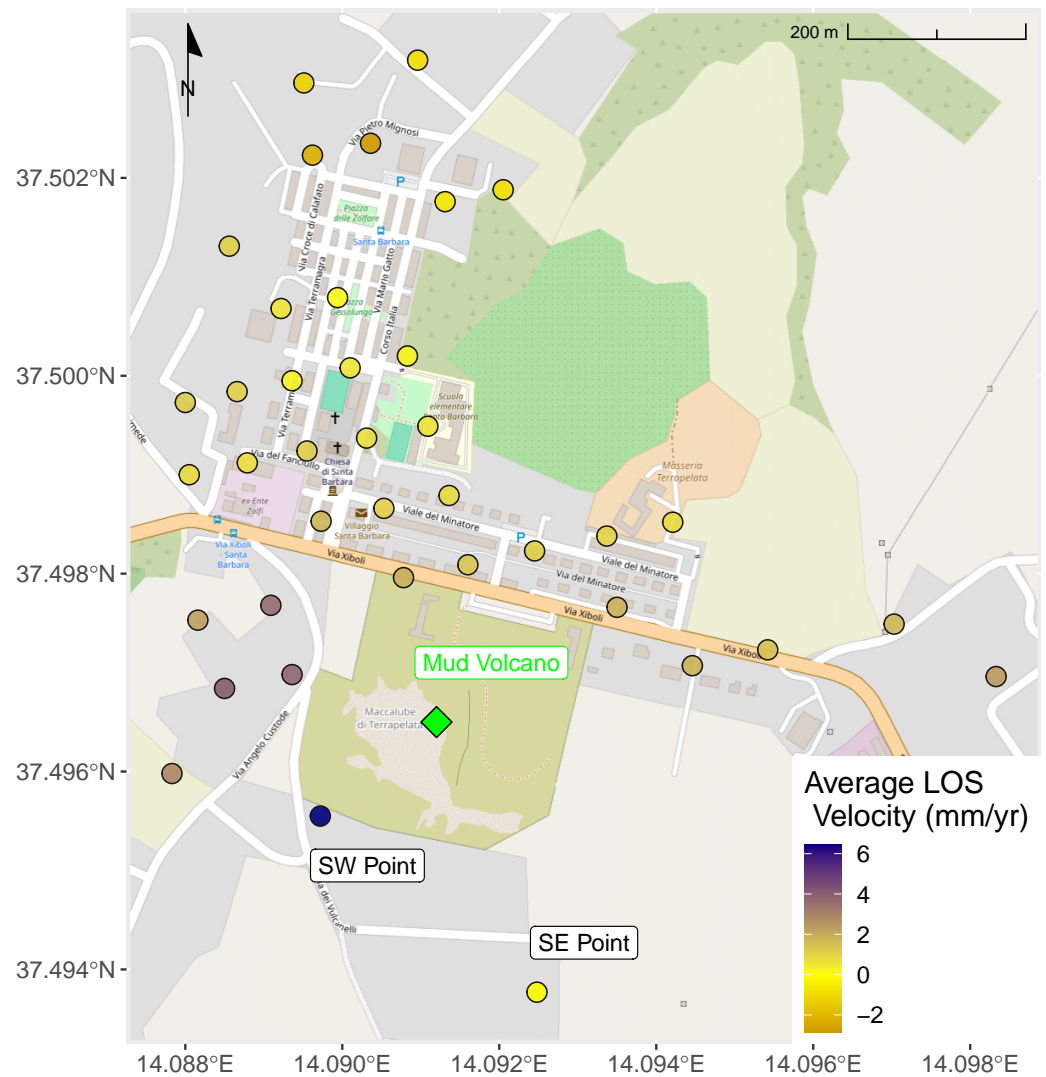
Figure 3. Detailed map of the points belonging to the Santa Barbara village ($d^* < 750$ m). Overlaid on the map, one can observe the position of the mud volcano (p_v), as well as the southwest (SW) and southeast (SE) points that will be analyzed in greater detail in the rest of the work. Color scale represents the Line-Of-Sight average displacement velocity (in mm/year) across the observation period, minimum value is -2.8, maximum value is 6.4. Copyright for the base map is owned by OpenStreetMap contributors.
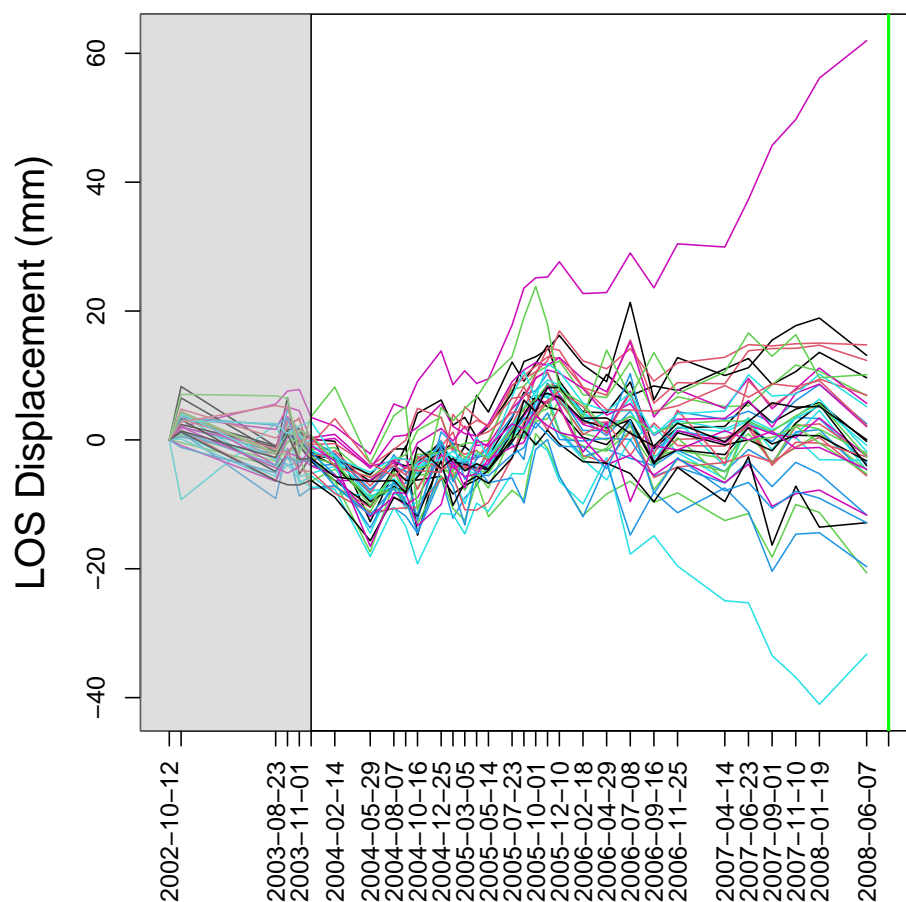
Figure 4. Plot of the time dynamics of the curves belonging to the Santa Barbara village ($d^* < 750$ m). The grey shaded area in the plot represents the calibration period, the vertical green line marks the date of the event, 11 August 2008

4. Results

Some of the sets of time series generated via the data manipulation procedure described above can be observed in the different panels of Figure 5, while the whole set of 27 figures is available in Figure S1. More specifically, Figure 5a represents data at a point in the first part of the observation window, Figure 5b at a point in the middle of the observation period and Figure 5c at a point in the last part of the observation period. In the latter two panels, the two colored curves, which are the ones located at the SW (purple) and SE (blue) points, show possibly outlying dynamics.

Having defined the different time series, the next step in the procedure is represented by computing the functional boxplots. In this case, we set $F = 3$. Of course, such a parameter can be optimized according to the outlier detection task at hand. Increasing it reduces the sensitivity of the methodology, which raises fewer alarms, while decreasing it renders the method more sensitive. A representation of the functional boxplots for the same dates explored in Figure 5 is available in Figure 6. It can be seen that in Figure 5b, the two "suspect" curves are not outside the functional whiskers shown in Figure 6b, and thus not identified as outliers, while this happens for the same curves in Figures 5c and 6c (i.e., at the end of the observation period).

The geographical representation of the outliers (Figure 7) sheds additional light on the procedure. Indeed, the two outlying curves correspond to points very close to the mud volcano (Figure 7c). Both points are characterized by anomalous values of LOS displacement but with different signs: the point southwest of the mud volcano features positive values (i.e., it is moving towards the sensor), while the point southeast of it features negative values (i.e., it is moving away from the sensor). This specific configuration, with

the two points moving in opposite directions along the ENVISAT descending LOS, suggests an inflation of the mud volcano area. The two points are located on residential structures in the urban agglomeration south of the volcano, and geological and geomorphological surveys have highlighted the development, in those and other sectors near and south of the volcano, of a series of fractures and shear lineaments which confirm the existence of high stress and strain linked to the presence of the volcano already before the August 2008 event [15,30–32].

Both sets of the 27 figures of the functional boxplots and the maps are available as Figures S2 and S3. The complete series of functional boxplots is also available on a single page as Figure S5.
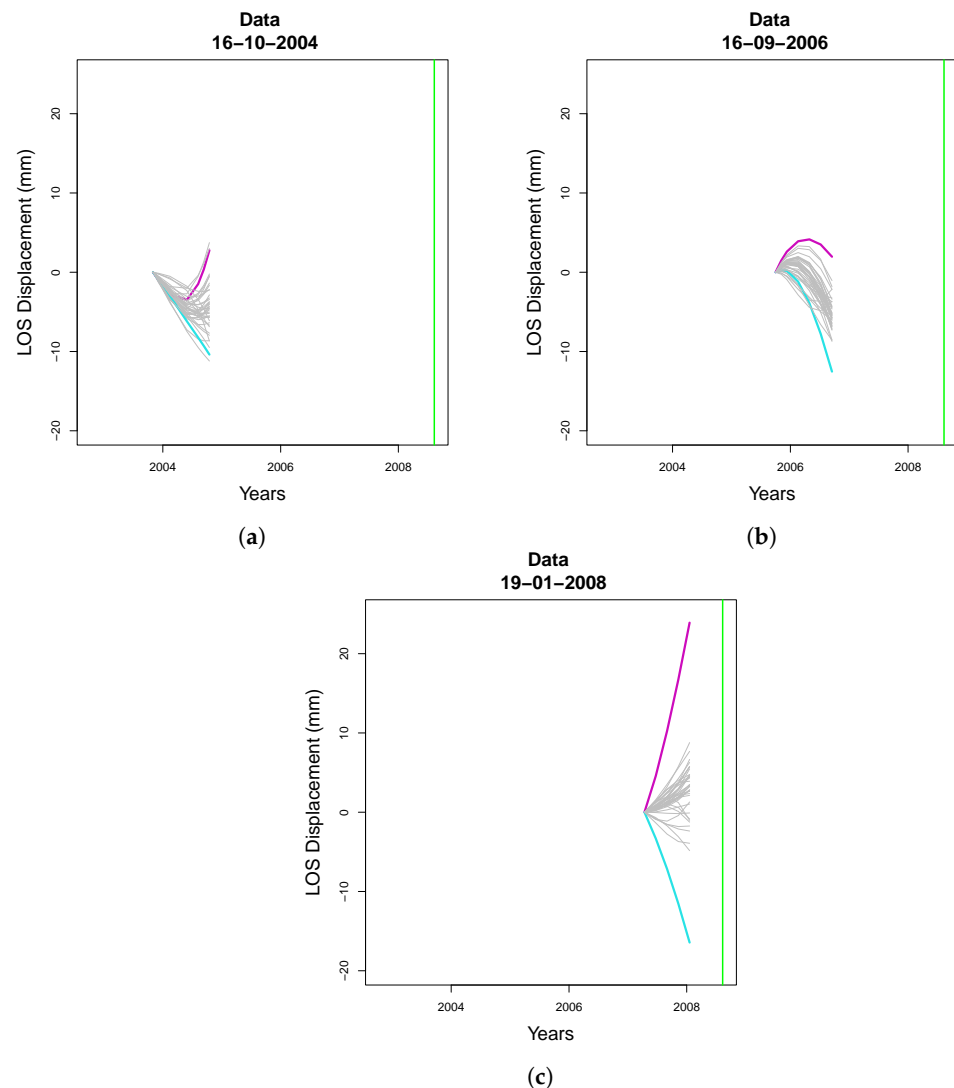

Figure 5. Representation of the set of time-series resulting from the data transformation procedure, namely (**a**) for 16 October 2004, (**b**) for 16 September 2006 and (**c**) for 19 January 2008. The green vertical line represents the date of the event.
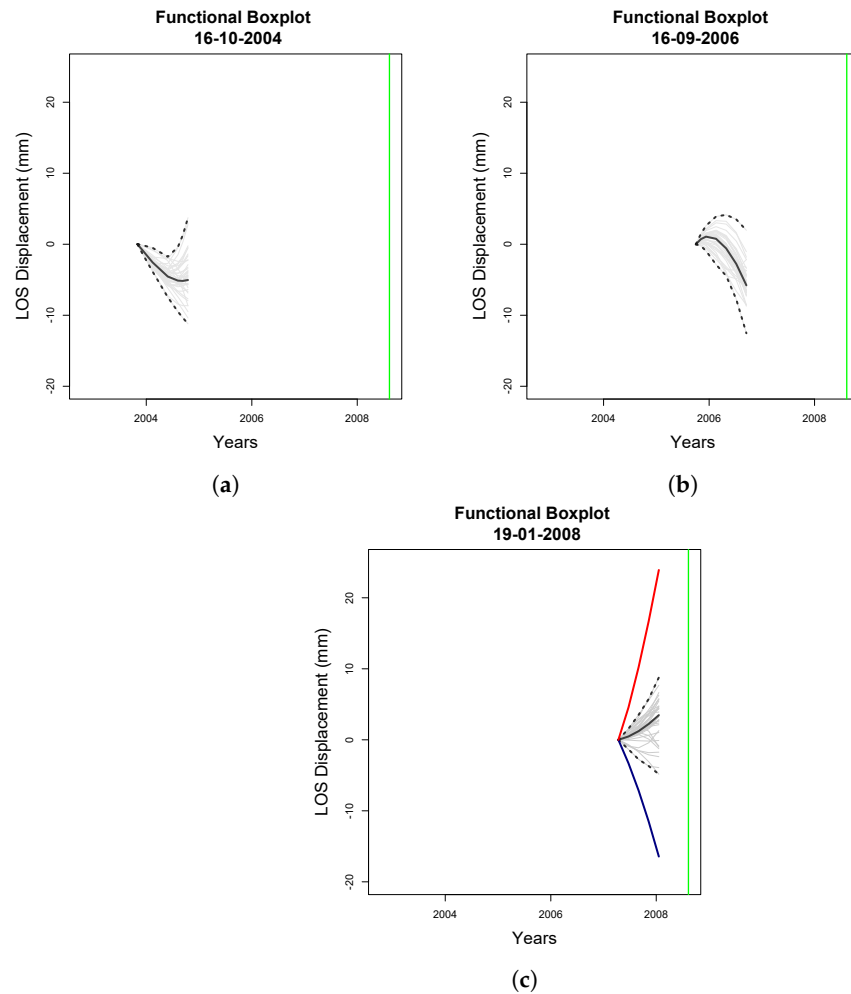
Figure 6. Representation of functional boxplots at different time instants (namely (**a**) for 16 October 2004, (**b**) for 16 September 2006 and (**c**) for 19 January 2008). The innermost black curve represents the functional median, i.e., the deepest point, while the two outermost black dashed curves are the upper and lower functional whiskers. Curves in bold colored in red or blue are, respectively, high and low outlying curves. Light grey curves are non-outlying curves. The green vertical line represents the day of the event.
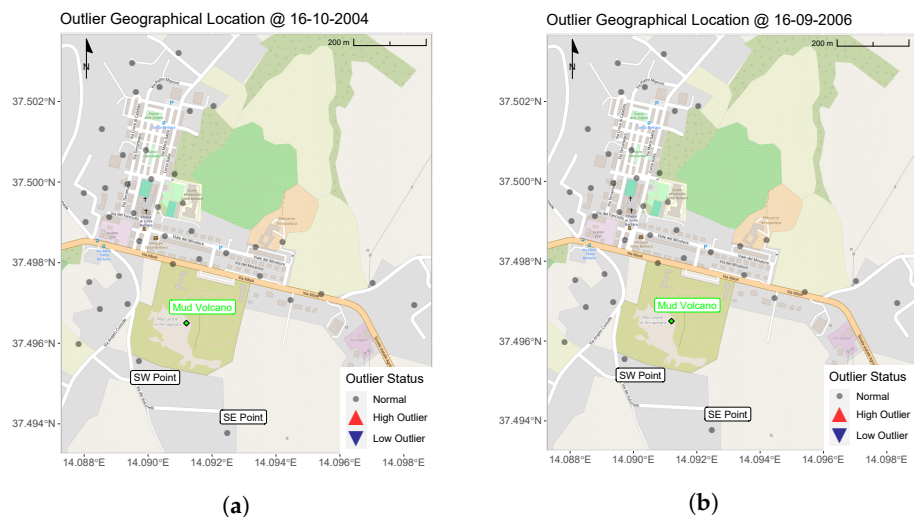
Figure 7. *Cont.*
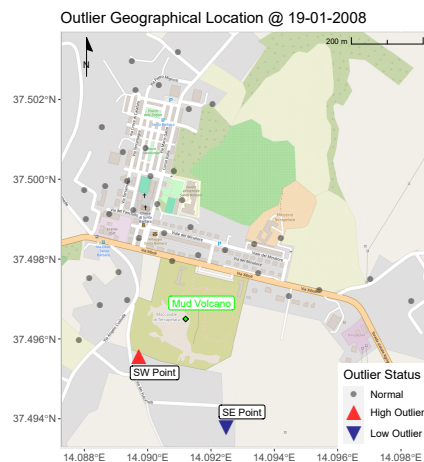
(**c**)

Figure 7. (**a**–**c**) Maps displaying geographically the outlying points for different time instants (**a**) at 16 October 2004, (**b**) at 16 September 2006, (**c**) at 19 January 2008. A triangle pointing upwards (downwards) represents a high (low) outlier. The color of the triangles is matched to the functional boxplots of Figure 6. The same reference points in Figure 3 are highlighted. Copyright for the base map is owned by OpenStreetMap contributors.

5. Discussion

By observing the time dynamics and persistence of the signals of deviation given by the procedure, we focus our attention on the period starting from 16 September 2006 to the day of the last observation, 7 June 2008. The maps and functional boxplots relative to that period can be found in Figure 8. More specifically, we observe two outlying points that switch on relatively early in time, and, for the SE outlier case, stay on for almost a year and a half before the paroxysmal event. It is immediately evident how the dimension of the "time persistence" of signals represents a fundamental aspect to be taken into consideration when analyzing deformation series using the proposed methodology.

To prove the validity of our proposal, and specifically of the use of the FDA approach to solve this kind of problem, we test our methodology against a similar one but developed without the use of functional tools. More specifically, we restrict ourselves to the same geographical area (so d_{lim} is still equal to 750 m) and we use the same memory parameter (so $k = 365$ days). We also perform the same re-centering on the first day of the time window. The essential differences between the "functional" and "scalar" proposals are thus that we do not perform any kind of smoothing on the data, and instead of using the functional boxplot, we use a standard one on the latest available observation.

The series of boxplots for the scalar approach is in Figure 9, and a summary of the comparison is available in Table 1, while the complete results alongside the boxplots are available in Figure S4. It is immediately evident, focusing on the two points close to the mud volcano, that the functional boxplot methodology is able to raise a more robust alarm than the scalar proposal, identifying two points instead of one, and in a slightly more persistent way. In fact, our suggestion to a civil protection agency in this case would have been to either increase the number of scans on the area, as something most probably was happening, or to activate an on-site monitoring of the mud volcano site.

In other words, our methodology would allow any practitioner using it to immediately detect points that had very wild and anomalous dynamics in the period of interest, and thus to provide to them very useful information to start implementing monitoring and/or mitigation strategies.

From a methodological point of view, the present work represents a step further with respect to previous studies in the field by demonstrating the effective use of FDA in early warning signal detection. In the application considered, the current work extends and complements previous studies on the same use case: as in Cigna et al. [10], precursors

of the mud volcano eruption have been identified in InSAR data, but here, no prior assumption on the shape of the anomaly was made; moreover, as in Fontana et al. [16], an FDA approach has been adopted, but here, the specific goal of early warning has been successfully addressed.

The method relies on the proper selection of the parameters d_{lim}, k and F. While the flexibility given by these parameters allows the method to adapt to a wide range of use cases with different geological conditions, the sensitiveness of the results to this choice represents a limitation of the proposed approach. The three parameters can be fixed on the basis of prior knowledge of the phenomenon under study, as performed in the current study, since they have a physical meaning and clear interpretation.
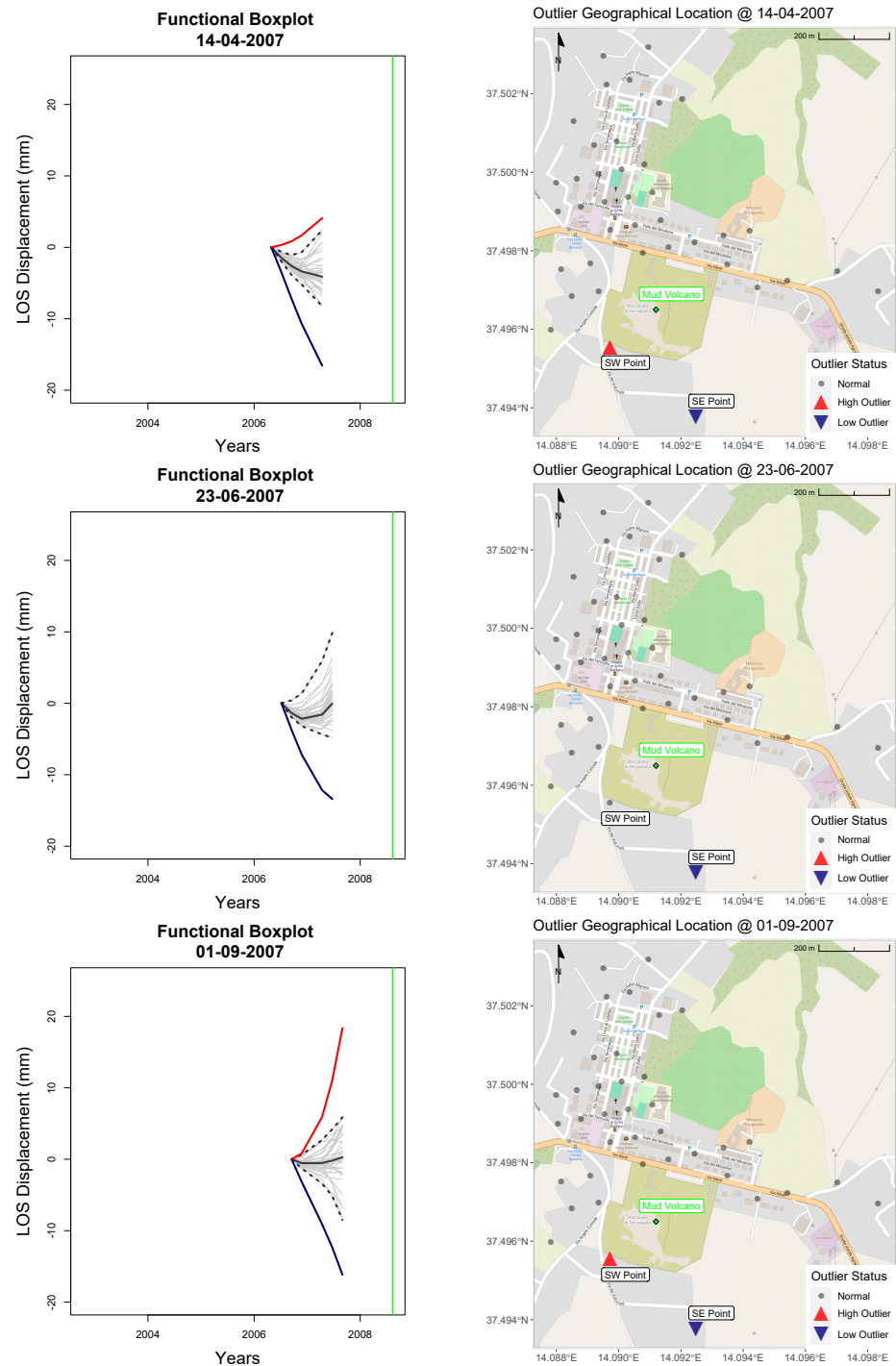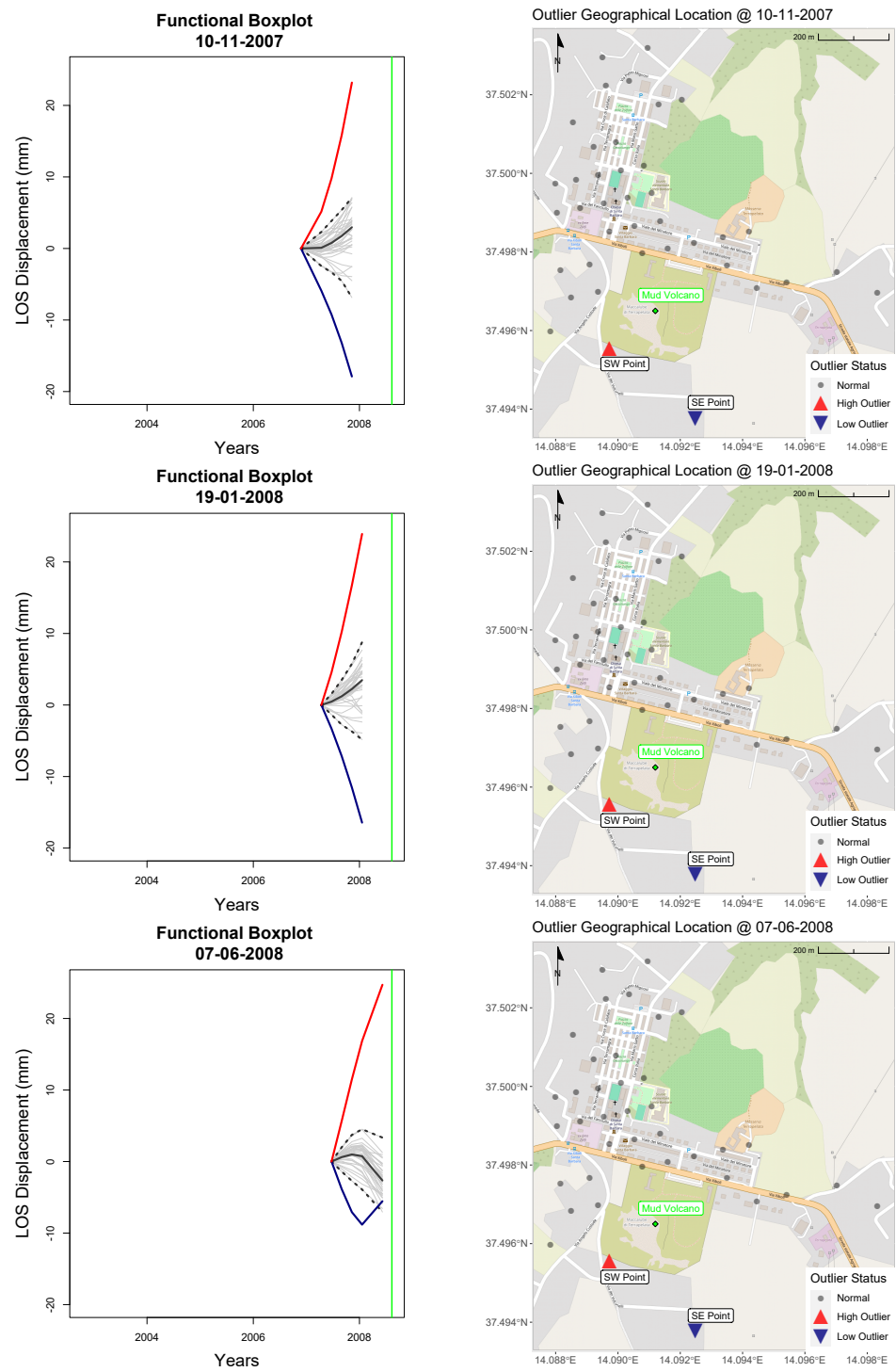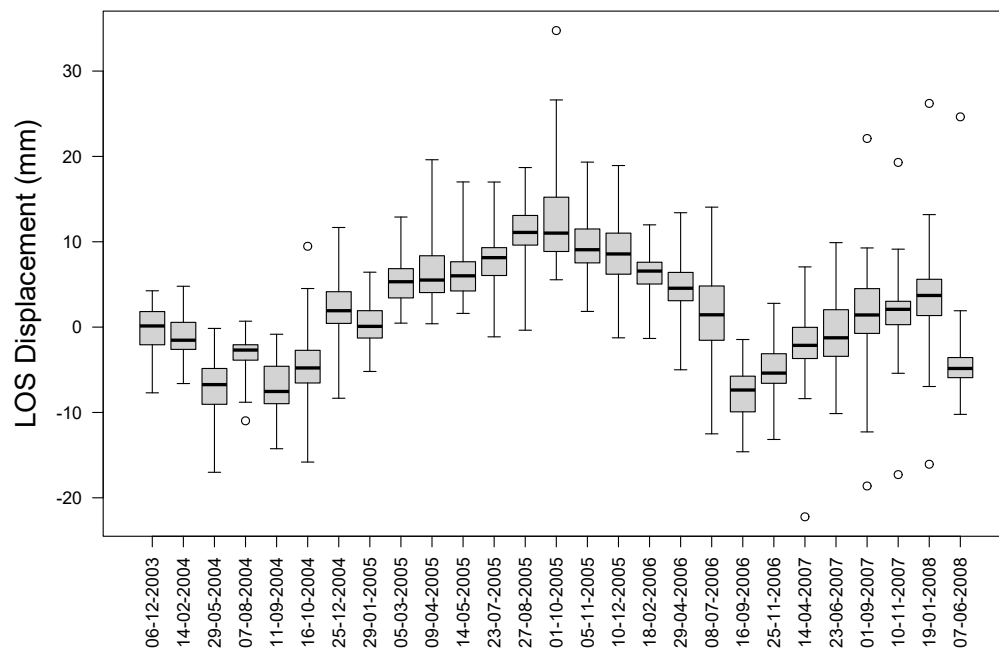

Figure 8. *Cont.*

Figure 8. Representation of the functional boxplots, alongside outlier maps, for the timeframe close to the end of the observation period. Points in red (blue) are high (low) outliers. Colour is maintained between the Functional Boxplot and the Map. Copyright for the base map is owned by OpenStreetMap contributors.

Figure 9. Series of boxplots representing the results of the scalar approach. The central box contains 50% of the observations, and is delimited by the 25% quantile (the lower bound) and the 75% quantile (the upper bound). The black line inside the box represents the scalar median (i.e., the 50% quantile). The single points are outliers, while the whiskers extend to the maximum and minimum values excluding the outliers.

Table 1. Results of the scalar approach and the functional approach for the two points closest to the mud volcano (the southwest point and the southeast point). Colored cells correspond to outliers found (red for high outliers and blue for low outliers).

Dates	Scalar Boxplot		Functional Boxplot	
	SW Point	**SE Point**	**SW Point**	**SE Point**
14 April 2007		🟦	🟥	🟦
23 June 2007				🟦
1 September 2007	🟥	🟦	🟥	🟦
10 November 2007	🟥	🟦	🟥	🟦
19 January 2008	🟥	🟦	🟥	🟦
7 June 2008	🟥		🟥	🟦

6. Conclusions

In this paper, we propose a novel approach for post-processing InSAR data in order to provide early warning signals for geological hazards. This topic is of paramount importance for the practical use of these data for geohazard monitoring to provide practitioners with useful information for implementing mitigation strategies.

The proposed methodology appropriately takes into account the spatial and temporal dimension of the problem. Indeed, to exploit the smoothness of the phenomenon, we adopt an FDA approach and we use functional boxplots to identify outlying curves. The proposed methodology does not rely on preselected patterns, unlike previous studies based on predefined deformation trends, and uses the data themselves to set the reference to detect anomalies. Therefore, the presented approach allows more flexibility and can potentially be applied to a wide variety of geological events. In the considered application, the proposed functional approach presents advantages over a scalar one, as the warnings for the paroxysmal event are provided in advance. Moreover, the analysis of the time dimension allowed by the FDA framework used in the methodology provides additional insights to practitioners. In particular, the proposed methodology allows the user to identify

precursor signs of the analyzed test case almost 5 months in advance, and with signals that are persistent. In terms of technical recommendations to a civil protection practitioner, i.e., the target user of our methodology, we believe that fundamental attention has to be paid to the geographical proximity and temporal persistence of raised warnings.

The proposed method could be tested and validated on other application cases of volcano eruptions featuring different characteristics (such as the Maccalube of Aragona, where many more events and activity have been recorded in the last years [33] compared to those of Santa Barbara) or to other paroxysmal events to assess its capabilities. Another future research direction would be the extension of the current study to find heuristic criteria that are able to guide the choice of the three parameters characterizing the proposed methodology using information on the geology of the area, or according to specific problem classes (paroxysmal vs persistent events), or for specific application tasks (volcanic eruptions, landslides, subsidence, monitoring of buildings, etc.). Another valid proposal would be to implement data-driven techniques for the selection, exploiting known techniques in uncertainty quantification such as Conformal Prediction, a novel non-parametric forecasting method based on minimal assumptions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs16071191/s. Figure S1: Series of sets of preprocessed time series; Figure S2: Series of functional boxplots; Figure S3: Series of outlier maps, the copyright of the base map is owned by OpenStreetMap Contributors; Figure S4: Series of scalar boxplots and corresponding outlier maps, the copyright of the base map is owned by OpenStreetMap Contributors; Figure S5: Series of functional boxplots on a single page.

Author Contributions: Conceptualization, M.F., M.S.B., A.M., S.V., F.C. and D.T.; methodology, M.F., M.S.B., A.M. and S.V.; software, M.F.; validation, M.F., M.S.B., A.M., S.V., F.C. and D.T.; formal analysis, M.F.; investigation, F.C. and D.T.; resources, A.M. and S.V.; data curation, M.F., F.C. and D.T.; writing—original draft preparation, M.F., M.S.B., A.M., S.V., F.C. and D.T.; writing—review and editing, M.F., M.S.B., A.M., S.V., F.C. and D.T.; visualization, M.F.; supervision, S.V. and A.M.; project administration, S.V. and A.M.; funding acquisition, S.V. and A.M. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: ENVISAT Advanced Single Look Complex Level-1 data can be freely accessed via ESA's Earth observation gateway.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rosen, P.; Hensley, S.; Joughin, I.; Li, F.; Madsen, S.; Rodriguez, E.; Goldstein, R. Synthetic aperture radar interferometry. *Proc. IEEE* **2000**, *88*, 333–382. [CrossRef]
2. Raucoules, D.; Bourgine, B.; de Michele, M.; Le Cozannet, G.; Closset, L.; Bremmer, C.; Veldkamp, H.; Tragheim, D.; Bateson, L.; Crosetto, M.; et al. Validation and intercomparison of Persistent Scatterers Interferometry: PSIC4 project results. *J. Appl. Geophys.* **2009**, *68*, 335–347. [CrossRef]
3. Manunta, M.; De Luca, C.; Zinno, I.; Casu, F.; Manzo, M.; Bonano, M.; Fusco, A.; Pepe, A.; Onorato, G.; Berardino, P.; et al. The Parallel SBAS Approach for Sentinel-1 Interferometric Wide Swath Deformation Time-Series Generation: Algorithm Description and Products Quality Assessment. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6259–6281. [CrossRef]
4. Duan, W.; Zhang, H.; Wang, C.; Tang, Y. Multi-Temporal InSAR Parallel Processing for Sentinel-1 Large-Scale Surface Deformation Mapping. *Remote Sens.* **2020**, *12*, 3749. [CrossRef]
5. Cigna, F.; Esquivel Ramírez, R.; Tapete, D. Accuracy of Sentinel-1 PSI and SBAS InSAR Displacement Velocities against GNSS and Geodetic Leveling Monitoring Data. *Remote Sens.* **2021**, *13*, 4800. [CrossRef]
6. Bernardi, M.S.; Africa, P.C.; de Falco, C.; Formaggia, L.; Menafoglio, A.; Vantini, S. On the Use of Interferometric Synthetic Aperture Radar Data for Monitoring and Forecasting Natural Hazards. *Math. Geosci.* **2021**, *53*, 1781–1812. [CrossRef]
7. Berti, M.; Corsini, A.; Franceschini, S.; Iannacone, J.P. Automated classification of Persistent Scatterers Interferometry time series. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 1945–1958. [CrossRef]
8. Chang, L.; Hanssen, R.F. A Probabilistic Approach for InSAR Time-Series Postprocessing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 421–430. [CrossRef]

9. Li, M.; Wu, H.; Yang, M.; Huang, C.; Tang, B.H. Trend Classification of InSAR Displacement Time Series Using SAE–CNN. *Remote Sens.* **2024**, *16*, 54. [CrossRef]

10. Cigna, F.; Tapete, D.; Casagli, N. Semi-automated extraction of Deviation Indexes (DI) from satellite Persistent Scatterers time series: tests on sedimentary volcanism and tectonically-induced motions. *Nonlinear Process. Geophys.* **2012**, *19*, 643–655. [CrossRef]

11. Tapete, D.; Casagli, N. Testing Computational Methods to Identify Deformation Trends in RADARSAT Persistent Scatterers Time Series for Structural Assessment of Archaeological Heritage. In Proceedings of the Computational Science and Its Applications—ICCSA, Ho Chi Minh City, Vietnam, 24–27 June 2013; Murgante, B., Misra, S., Carlini, M., Torre, C.M., Nguyen, H.Q., Taniar, D., Apduhan, B.O., Gervasi, O., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2013; pp. 693–707. [CrossRef]

12. Notti, D.; Calò, F.; Cigna, F.; Manunta, M.; Herrera, G.; Berti, M.; Meisina, C.; Tapete, D.; Zucca, F. A User-Oriented Methodology for DInSAR Time Series Analysis and Interpretation: Landslides and Subsidence Case Studies. *Pure Appl. Geophys.* **2015**, *172*, 3081–3105. [CrossRef]

13. Ramsay, J.O.; Silverman, B.W. *Functional Data Analysis*; Springer Series in Statistics; OCLC: 249216329; Springer: Berlin/Heidelberg, Germany, 2005.

14. Moro, M.; Saroli, M.; Stramondo, S.; Bignami, C.; Albano, M.; Falcucci, E.; Gori, S.; Doglioni, C.; Polcari, M.; Tallini, M.; et al. New insights into earthquake precursors from InSAR. *Sci. Rep.* **2017**, *7*, 12035. [CrossRef] [PubMed]

15. Madonia, P.; Grassa, F.; Cangemi, M.; Musumeci, C. Geomorphological and geochemical characterization of the 11 August 2008 mud volcano eruption at S. Barbara village (Sicily, Italy) and its possible relationship with seismic activity. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 1545–1557. [CrossRef]

16. Fontana, M.; Tavoni, M.; Vantini, S. Functional Data Analysis of high-frequency load curves reveals drivers of residential electricity consumption. *PLoS ONE* **2019**, *14*, e0218702. [CrossRef]

17. Berardino, P.; Fornaro, G.; Lanari, R.; Sansosti, E. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2375–2383. [CrossRef]

18. Casu, F.; Elefante, S.; Imperatore, P.; Zinno, I.; Manunta, M.; De Luca, C.; Lanari, R. SBAS-DInSAR Parallel Processing for Deformation Time-Series Computation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3285–3296. [CrossRef]

19. Yao, J.; Yao, X.; Liu, X. Landslide Detection and Mapping Based on SBAS-InSAR and PS-InSAR: A Case Study in Gongjue County, Tibet, China. *Remote Sens.* **2022**, *14*, 4728. [CrossRef]

20. Li, S.; Xu, W.; Li, Z. Review of the SBAS InSAR Time-series algorithms, applications, and challenges. *Geod. Geodyn.* **2022**, *13*, 114–126. [CrossRef]

21. Lanari, R.; Bonano, M.; Casu, F.; Luca, C.D.; Manunta, M.; Manzo, M.; Onorato, G.; Zinno, I. Automatic Generation of Sentinel-1 Continental Scale DInSAR Deformation Time Series through an Extended P-SBAS Processing Pipeline in a Cloud Computing Environment. *Remote Sens.* **2020**, *12*, 2961. [CrossRef]

22. Cigna, F.; Tapete, D. Land Subsidence and Aquifer-System Storage Loss in Central Mexico: A Quasi-Continental Investigation With Sentinel-1 InSAR. *Geophys. Res. Lett.* **2022**, *49*, e2022GL098923. [CrossRef]

23. Cigna, F.; Sowter, A. The relationship between intermittent coherence and precision of ISBAS InSAR ground motion velocities: ERS-1/2 case studies in the UK. *Remote Sens. Environ.* **2017**, *202*, 177–198. [CrossRef]

24. Vallone, P.; Giammarinaro, M.S.; Crosetto, M.; Agudo, M.; Biescas, E. Ground motion phenomena in Caltanissetta (Italy) investigated by InSAR and geological data integration. *Eng. Geol.* **2008**, *98*, 144–155. [CrossRef]

25. Vincenty, T. Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations. *Surv. Rev.* **1975**, *23*, 88–93. [CrossRef]

26. Sun, Y.; Genton, M.G. Functional Boxplots. *J. Comput. Graph. Stat.* **2011**, *20*, 316–334. [CrossRef]

27. Lopez-Pintado, S.; Romo, J. On the Concept of Depth for Functional Data. *J. Am. Stat. Assoc.* **2009**, *104*, 718–734. [CrossRef]

28. Liu, R.Y.; Parelius, J.M.; Singh, K. Multivariate analysis by data depth: descriptive statistics, graphics and inference. *Ann. Stat.* **1999**, *27*, 783–858. [CrossRef]

29. Brighenti, F.; Carnemolla, F.; Messina, D.; De Guidi, G. UAV survey method to monitor and analyze geological hazards: The case study of the mud volcano of Villaggio Santa Barbara, Caltanissetta (Sicily). *Nat. Hazards Earth Syst. Sci.* **2021**, *21*, 2881–2898. [CrossRef]

30. Bonini, M. Mud volcanoes: Indicators of stress orientation and tectonic controls. *Earth-Sci. Rev.* **2012**, *115*, 121–152. [CrossRef]

31. INGV. *Comunicato Sull'eruzione di Fango in C.da Terrapelata Santa Barbara (Cl) 11 Agosto 2008–Aggiornamento del 16 Agosto*; INGV-Sezione di Palermo: Palermo, Italy, 2008.

32. Regione Siciliana. *Emergenza "Maccalube" dell'11 Agosto 2008 nel Comune di Caltanissetta, Descrizione dell'Evento e dei Danni*; Technical Report; Dipartimento della Protezione Civile, Servizio di Caltanissetta: Caltanissetta, Italy, 2008.

33. Gattuso, A.; Italiano, F.; Capasso, G.; D'Alessandro, A.; Grassa, F.; Pisciotta, A.F.; Romano, D. The mud volcanoes at Santa Barbara and Aragona (Sicily, Italy): A contribution to risk assessment. *Nat. Hazards Earth Syst. Sci.* **2021**, *21*, 3407–3419. [CrossRef]