FULL PAPER

The time lag between deformation process and seismic activity in El Hierro Island during the eruptive process (2011–2014): a functional phased approach

Sonia Pérez-Plaza[1*]　, Manuel Berrocoso[2], Belén Rosado[2], Gonçalo Prates[3] and Fernando Fernández-Palacín[1]

Abstract

On 10 October 2011, a submarine eruption occurred in El Hierro island. Thus, the eruptive process in the Canary islands was reactivated after 40 years of inactivity. The main objective of this work is to evaluate, using Functional Data Analysis, how the surface deformation phenomenon explains the seismic–volcanic activity in the island. The GNSS-GPS data are from the FRON (GRAFCAN) station, located in Frontera. These data measure, each 4 h, the distance between the FRON station and the reference station LPAL (La Palma island) from August, 2010 to December, 2013. In this study a functional correlation measure is employed to establish the relation between the deformation curve and the curve of cumulative energy released. The period of time analysed has been divided into four phases to avoid the mix of phenomena. For each phase, the correlation measure and the time lag between deformation curve and the curve of cumulative energy released have been estimated. These values show a strong relation between these curves. With respect to time lag period, the only significant lag, of about 1 month, occurred in Phase 1, which was after a long period without seismic activity. The later phases had very short, insignificant, lags. After a long period without seismic and volcanic activity in El Hierro island, the time lag between the deformation process and the beginning of the seismic activity takes approximately 1 month. In a similar situation a method to predict in real time the beginning of the seismic activity is proposed. This method, based on the changes produced in the derivative curves when there is a rapid descent in the deformation curve, could activate a warning system approximately 13 days before the beginning of seismicity.

Keywords: FDA, GPS, *P*-splines, Time series

Introduction

The Canary archipelago formation started more than 40 Ma ago (Anguita and Hernán 2000; Araña and Ortiz 1991; Carracedo et al. 1999). It is unclear if its origin is related with models based on a hotspot or mantle plume without the typical topographic and geoid heights swell (Carracedo et al. 1999). Other works as Anguita and Hernán (2000) relate its origin with magmatic intrusions from a remnant mantle thermal anomaly controlled by regional interchanging compression and extension tectonic regimes.

The subaerial morphology of El Hierro island in the Canary archipelago (27.7 °N; 18.0 °W; 278.5 km^2; and 1.501 km maximum height) results from an interpretation as triple volcanic rift (NE-rift, NW-rift and S-rift) (see Fig.1) (Carracedo et al. 1999) with well developed emplacement of dikes and fissures (Gee et al. 2001), and voluminous landslides (Masson et al. 2002; Mitchell et al. 2002).

---

*Correspondence: sonia.perez@uca.es
[1] Department of Statistics and Operations Research, University of Cadiz, Cadiz, Spain
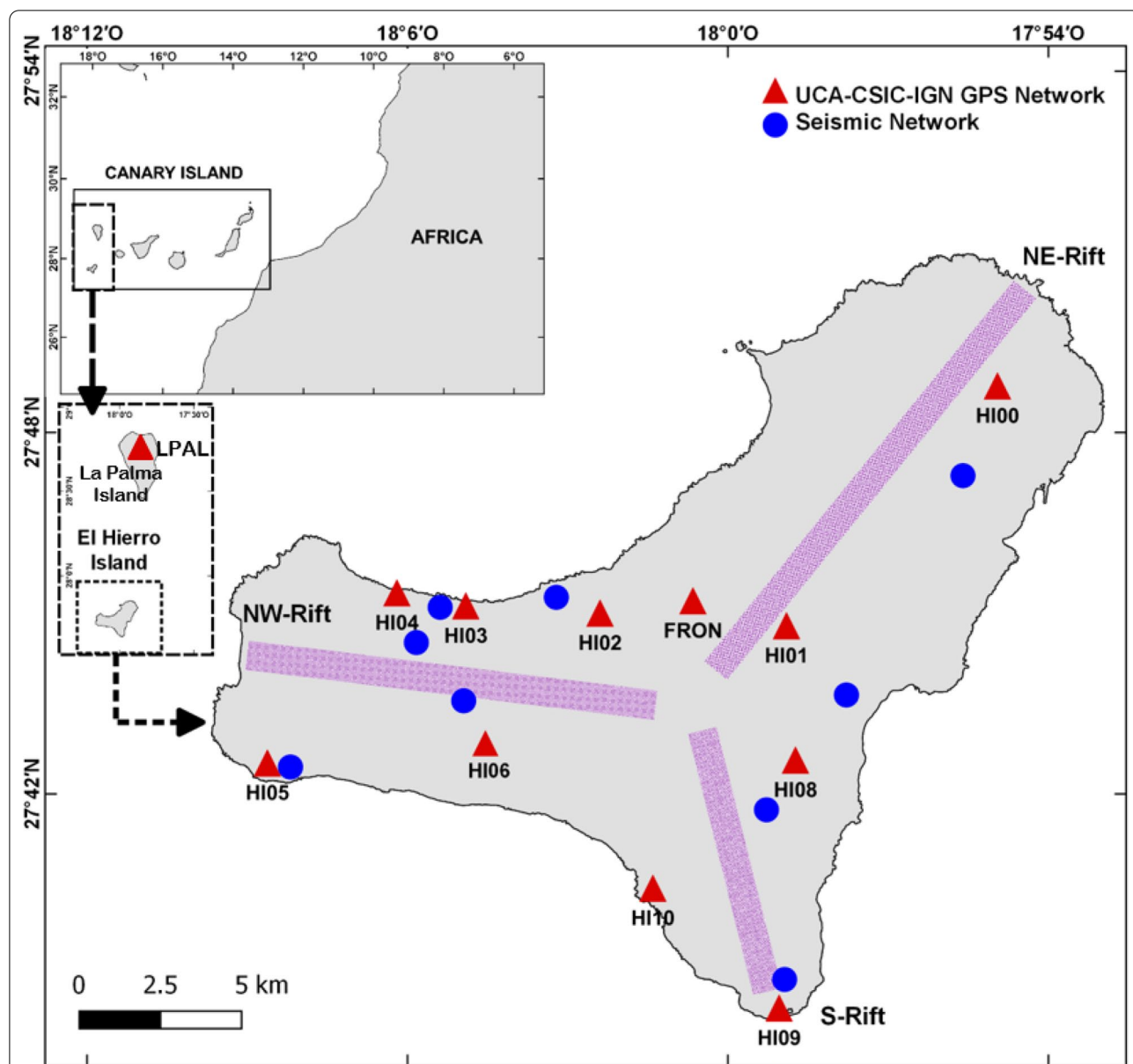Full list of author information is available at the end of the article

Pérez-Plaza *et al. Earth, Planets and Space*   (2021) 73:177

Page 2 of 17

Fig. 1 Geodynamic map of the El Hierro island. In this map the triaxial rift, the GPS stations of the geodynamic network (IGN-CSIC-UCA), and seismic stations are shown. The location of the LPAL station (EUREF) in La Palma island and FRON station (GRAFCAN) in El Hierro island can be seen in this map

The volcanic activity at El Hierro island is the result of a complex set of small insulated pockets of magma at mantle depths (Stroncik et al. 2009). According to Hernández-Pacheco (1982) and Carracedo et al. (2001), after 200 years without eruptions at El Hierro, in July 2011 the eruptive process at El Hierro was reactivated. Some warning signals were, among others, significant displacements of GNSS-GPS stations and increased seismicity. After a period without submarine eruptions, in March 2012, new episodes appeared. More seismic energy and

deformations larger than the preceding ones were caused (Prates et al. 2013).

Between 2011 and 2014 there were at least seven episodes of magmatic intrusion in El Hierro Island, but only the first one led to a submarine eruption in 2011–2012 (Lamolda et al. 2017; Prates et al. 2013). After almost 3 months of unrest, the eruption began on 10 October 2011 and it lasted almost 5 months (García et al. 2014; López et al. 2014; Prates et al. 2013). However, the magmatic activity in the island did not cease with at least six

episodes of intense seismicity and deformation occurred between June 2012 and March 2014. The study of the seismicity registered from 2011 to 2014 in El Hierro offers the opportunity to better understand the dynamics associated with the preparation, development and relaxation of a volcanic system in relation to an eruption (Telesca et al. 2016). Figure 2 shows the high number of earthquakes and their location.

Fig. 2 Daily number of earthquakes and its location in El Hierro island between 2011 and 2014
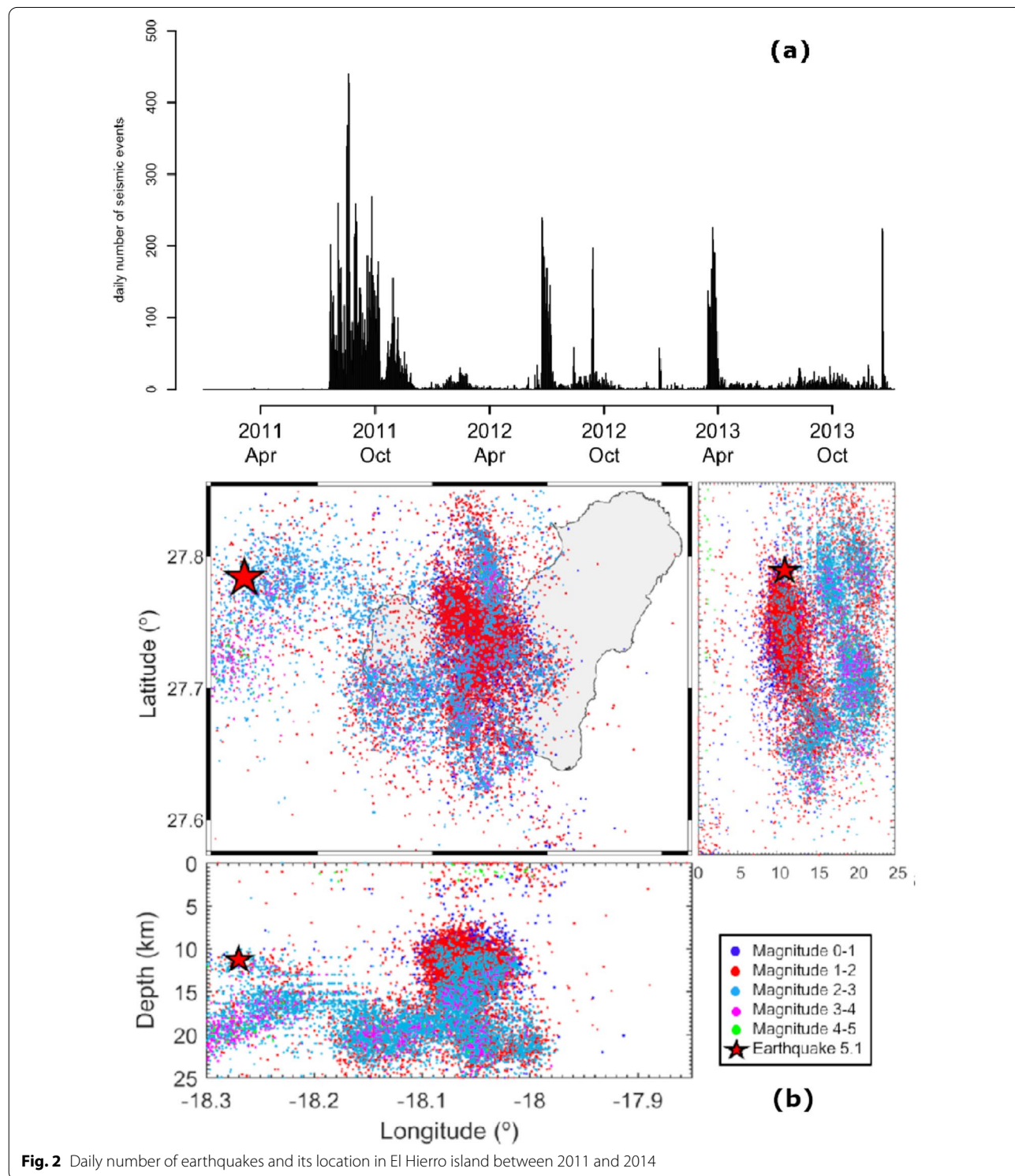
The intrusion process profoundly affects both the rising magma and surrounding rock. Variations in the magma crystallinity, composition, temperature, volatile content and vesicularity produced after the eruption caused significant changes in magma viscosity. Changes in magma rheology and dynamics should have produced by these variations (Martí et al. 2013). Moreover, in the intrusión process, magma deforms the host rock. Several models assure that the effect on the host rock after the eruption should be measurable in the surface deformation in many cases (Dzurisin 2006).

On the other hand, to evaluate the seismic activity in the island, the IGN (National Geographic Institute) seismic data catalogue (http://www.ign.es) is used. Using the IGN earthquake catalogue and the relation between the energy and the absolute magnitude of earthquake, Mg, (in the Richter scale) obtained by Choy and Boatwright (1995), the daily cumulative energy released E in joules (J) has been calculated through

$$E = 10^{1.5*Mg+4.4}$$

In the data analysis it is supposed that the released total seismic energy is constant during the day. This constant value corresponds to the amount of accumulated energy produced by all the earthquakes of that day.

In López et al. (2012) it is shown that in the eruptive process of El Hierro island the surface deformation is a very important indicator for monitoring the onset and for forecasting changes in seismic and volcanic activity during the crisis period. From GNSS-GPS observations the analysis of the changes in the deformation of the El Hierro Island regional geodynamics allow to detect the reactivation.

As shown in Prates et al. (2013) seismicity and deformation are both related to energy release, while deformation can also be related to energy accommodation.

In a volcanic system the energy source is the magma injection forcing its way from depth to the surface. This causes the volcanic edifice to inflate and volcanic unrest to occur (García et al. 2014). However, magma can retreat or reach and fill cracks in the rock changing the energy level, promoting the volcanic edifice to deflate with or without seismic energy release (Prates et al. 2013). Therefore, the volcanic edifice may inflate or deflate without energy release through seismic events or magma reaching the surface. In Prates et al. (2013) and García et al. (2014), several inflation and deflation episodes with and without associated seismic events and magma movement at depth with corresponding pressure source displacement and related ground deformation is shown. Therefore, ground deformation with respect to a reference location outside the volcanic edifice may show increase in distance on volcano inflation if the source location

projected vertically to the surface is, in general, between the locations of the reference and the ground deformation and decrease otherwise, and vice versa on volcano deflation, particularly so considering the Earth an elastic half-space (Mogi 1958).

In Berrocoso et al. (2012) and Prates et al. (2013) it is assured that after the detection of surface deformation changes, seismic activity is produced. These aspects were punctually analysed in Lamolda et al. (2017). In that paper the relationship between the sismicity and the surface deformation in 2011–2014 is analysed globally. However, the continuous nature of data promotes a processing of data from a functional point of view. The reason for proposing a Functional Data Analysis (FDA)-based treatment is that the GPS data, measured at each time interval, constitute only a discrete sequential sample of the geodetic displacements, which are continuous in nature. For this reason, when analyzing the relationship between the geodetic displacement and the seismic energy released, we will use a correlation measure adapted to the space in which we work, in this case, a functional space. Thus in this work we will calculate the cross functional correlation.

It is true that fitting the data series to a function basis creates small oscillations, especially at the beginning and the end of the data. However, working with curves has other advantages. In this work, it facilitates the calculation of the time lag between the deformation curve and the curve of cumulative energy released and this aspect had not been analyzed in other studies. Then, this information is used to establish a method that could activate a warning system. Working with curves instead of data series facilitates the analysis of the global behavior of each GPS station.

The different seismic phases that appear during the period 2011–2014 led us to carry out a more in-depth analysis by phases.

This paper is organized as follows. In "Methods" section the used methodology is described. In "Results and discussion" section, the previous procedures are applied to the time data series and lastly, in "Conclusion" section, the main results are analysed.

Methods

Data processing

In 2011 increased volcanic activity promoted a densification of the Geodetic Network placed in El Hierro island (Fig.1). However, this work is based in the FRON station which belongs to the geodetic network of permanent GPS stations of the Canary Islands GRAFCAN (Gobiermo Canario, http://www.grafcan.es/ red-de-estaciones). The reason is that FRON is stable and it is an existing station on the island before the beginning of the

eruptive process. The other stations only contain complete records after 2011. Therefore, it is not possible to analyse the relation between deformation curve and the curve of cumulative energy released before the beginning of the eruptive process with the other stations.

GNSS data were analysed using scientific software Bernese v5.0 (Dach et al. 2011). Along with the parameter estimation process, carrier phase double-difference data were used in ionospheric delay free mode. Tropospheric errors were handled using a combination of the a priori Saastamoinen model (Saastamoinen 1973) and Neill mapping functions (Niell 1996). Tropospheric parameters were estimated hourly, and ambiguities were fixed for the baseline, using the ionosphere-free observable with an a priori ionospheric model for determining the wide lane ambiguity (Mervart 1995). Ocean tide loading displacement corrections from Onsala Observatory were also introduced. Normal equations were computed for each daily solution. Finally, the episodic solution was achieved by combining the daily normal equations at the mean epoch, using the IGS final orbits (Rebischung et al. 2015), and constraining the movement of the reference benchmark, LPAL on La Palma Island, for the realization of the ITRF2008 reference frame (Altamimi et al. 2011).

The evolution of the surface deformation process was studied by using the analysis of time series of the variation of the distance between the stations FRON, located in El Hierro island, and LPAL, in La Palma island (Fig. 3a).

To obtain a cleaned series, a criterion for elimination of outliers is applied on the series of displacement data between the islands. The used criterion is appropiate for this type of data. This method considers the error variable provided by the Bernese software that measures the local error for each data. Then, for each GPS station, the values whose associated error are greater than twice the average error are removed (Pérez-Plaza et al. 2018). Then, a Kalman filter (Kalman 1960) is used to complete the non-available data.

The Kalman filter is an algorithm in two phases: prediction and correction. A new value for the system is calculated from the previous values. Then, a correction is added to this value in proportion to the prediction error, where this error is statistically minimized. Then the Kalman filter is not only a data imputation method. This method is used in this paper as a "discrete" smoothing prior to the continuous smoothing made in the FDA phase. Moreover, since it is a longitudinal model, the Kalman filter is more appropriate. This is because this model uses all the previous information of the data series and this does not occur in the case of using interpolation.
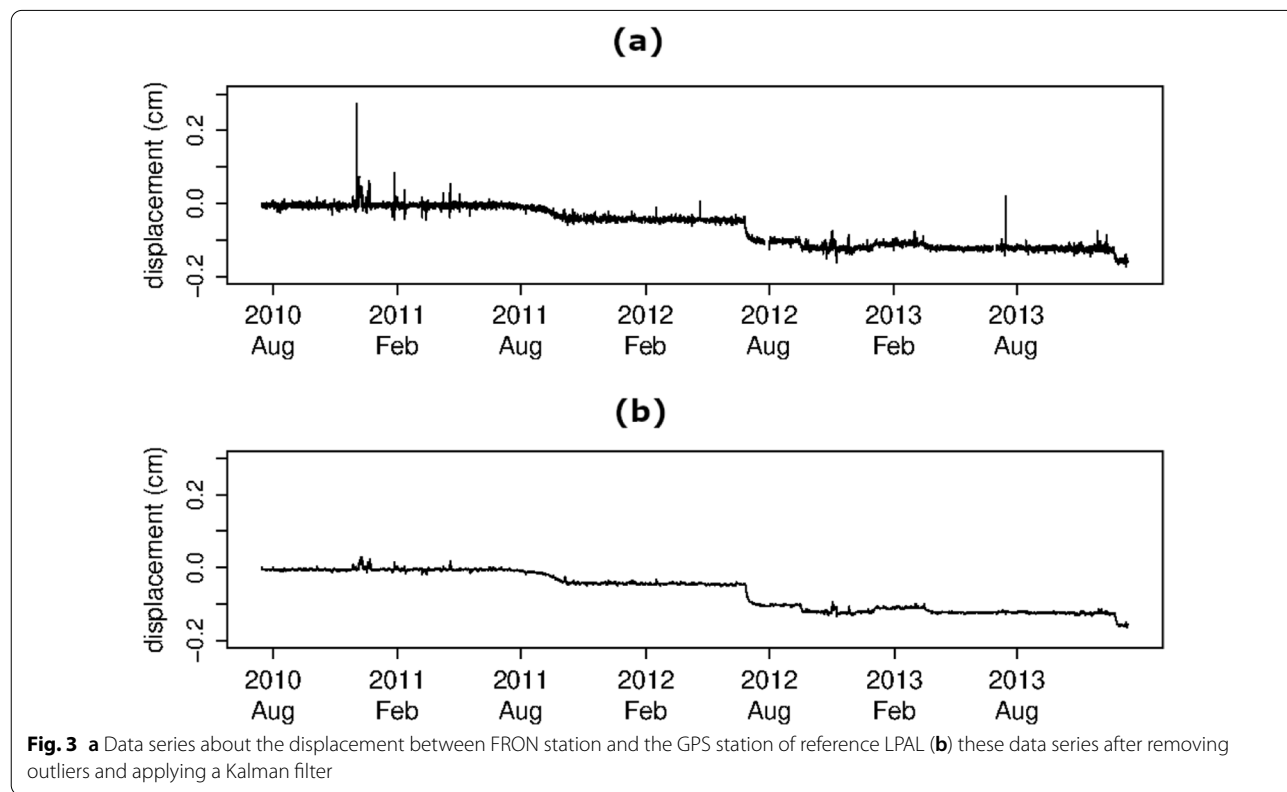

Fig. 3 **a** Data series about the displacement between FRON station and the GPS station of reference LPAL (**b**) these data series after removing outliers and applying a Kalman filter

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 6 of 17

The Kalman filter is a recursive algorithm introduced by Rudolf E. Kalman in 1960. The Kalman filter addresses the general problem of trying to estimate the state $x \in R^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation (Welch and Bishop 2002):

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{1}$$

with a measurement $z \in R^m$, that is,

$$z_k = H_k x_k + v_k \tag{2}$$

The random variables w_k and v_k represent the process and measurement noise, respectively. They are assumed to be independent (of each other), white, and with normal probability distributions.

The $n \times n$ matrix A_k in the difference Eq. (1) relates the state at time step k to the state at step $k + 1$, in the absence of either a driving function or process noise. The $n \times l$ matrix B_k relates the control input $u_k \in R^l$ to the state x. The $m \times n$ matrix H_k in the measurement Eq. (2) relates the state to the measurement z_k.

In the case of a first order model without control, since $n = 1$

$$\begin{aligned} A_k &= (1)_{1\times1}, \; x_k = (\mu_k)_{1\times1}, \\ B_k &= (0)_{1\times1}, \; \text{and} \; H_k = (1)_{1\times1}, \end{aligned} \tag{3}$$

the Eqs. (1) and (2) are transformed into the equations:

$$\begin{aligned} z_k &= \mu_k + v_k; \quad v_k \sim N(0, V) \\ \mu_{k+1} &= \mu_k + w_k; \quad w_k \sim N(0, \sigma_\mu^2) \end{aligned} \tag{4}$$

In the case of a second-order model without control, since

$$\begin{aligned} A_k &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; x_k = \begin{pmatrix} \mu_k \\ \beta_k \end{pmatrix}, \; B_k = 0, \\ w_k &= \begin{pmatrix} w_{k,1} \\ w_{k,2} \end{pmatrix}, \; \text{and} \; H_k = \begin{pmatrix} 1 & 0 \end{pmatrix}, \end{aligned}$$

the Eqs. (1) and (2) are transformed into the equations

$$\begin{aligned} z_k &= \mu_k + v_k; \quad v_k \sim N(0, V) \\ \mu_{k+1} &= \mu_k + \beta_k + w_{k,1}; \quad w_{k,1} \sim N(0, \sigma_\mu^2) \\ \beta_{k+1} &= \beta_k + w_{k,2}; \quad w_{k,2} \sim N(0, \sigma_\beta^2) \end{aligned} \tag{5}$$

Now, the use of a first-order or a second-order Kalman filter is considered. In this case, it is desirable to use a second-order Kalman filter for two reasons. First, it works better with data series, where there are sections of missing data, a common problem with this type of data. Second, it behaves better for series that present trends.

Figure 3b shows the data series after applying the cleaning process and the Kalman filter.

Phases and subphases: the change-point analysis

It is important to consider the existing seismic–volcanic phenomena when the relation between the cumulative energy released and the surface deformation is studied. Figure 4 shows four periods of transient deformation. As it is previously analysed, there is a clear relationship between deformation and seismicity. Therefore, to separate these four deformation periods and to determine the different phases, a single relevant seismo-volcanic event was considered by phase. Each phase ends the day of the highest magnitude earthquake in the period. By considering these aspects, four phases are obtained (Fig. 4).

In Fig. 5 these phases are represented on the data series about cumulative energy released. As can be observed in this figure the periods chosen by this criterion are consistent with the stabilisation periods in the cumulative energy released.

Phase 1: From the start of data series (14/07/2010) to the day (11/11/2011).
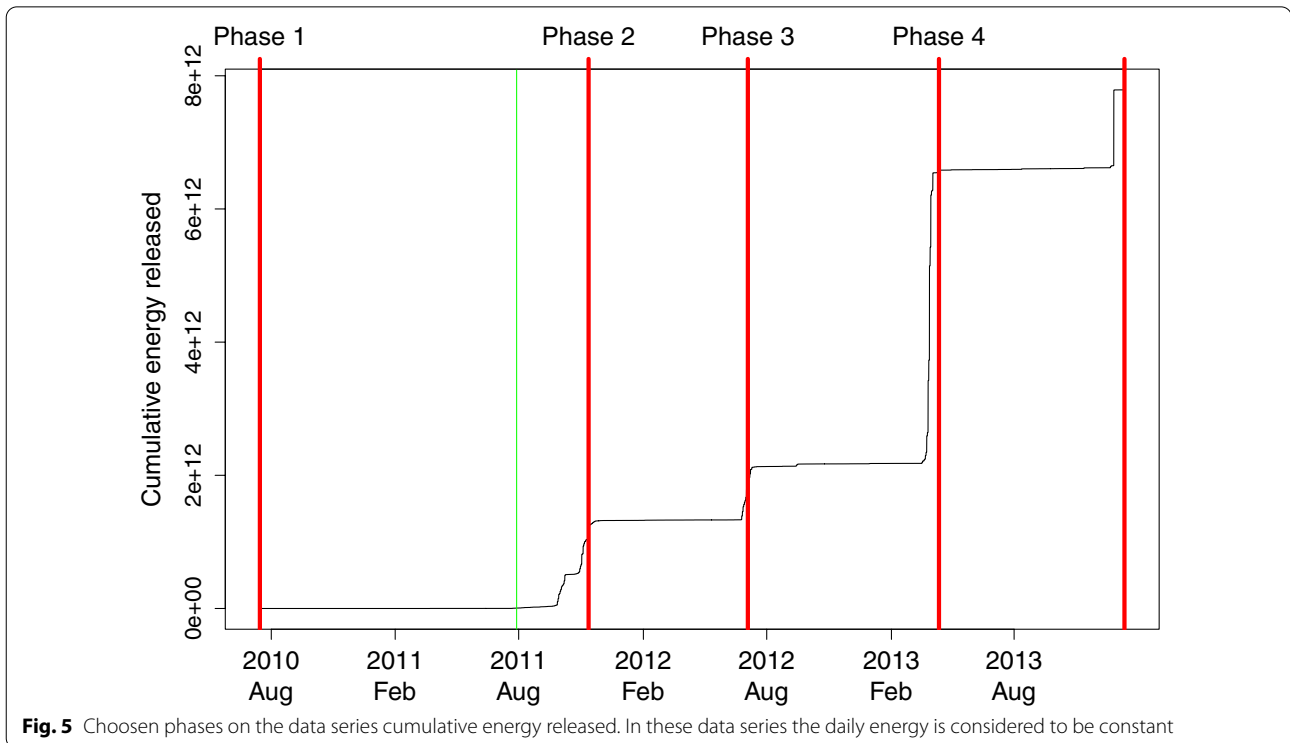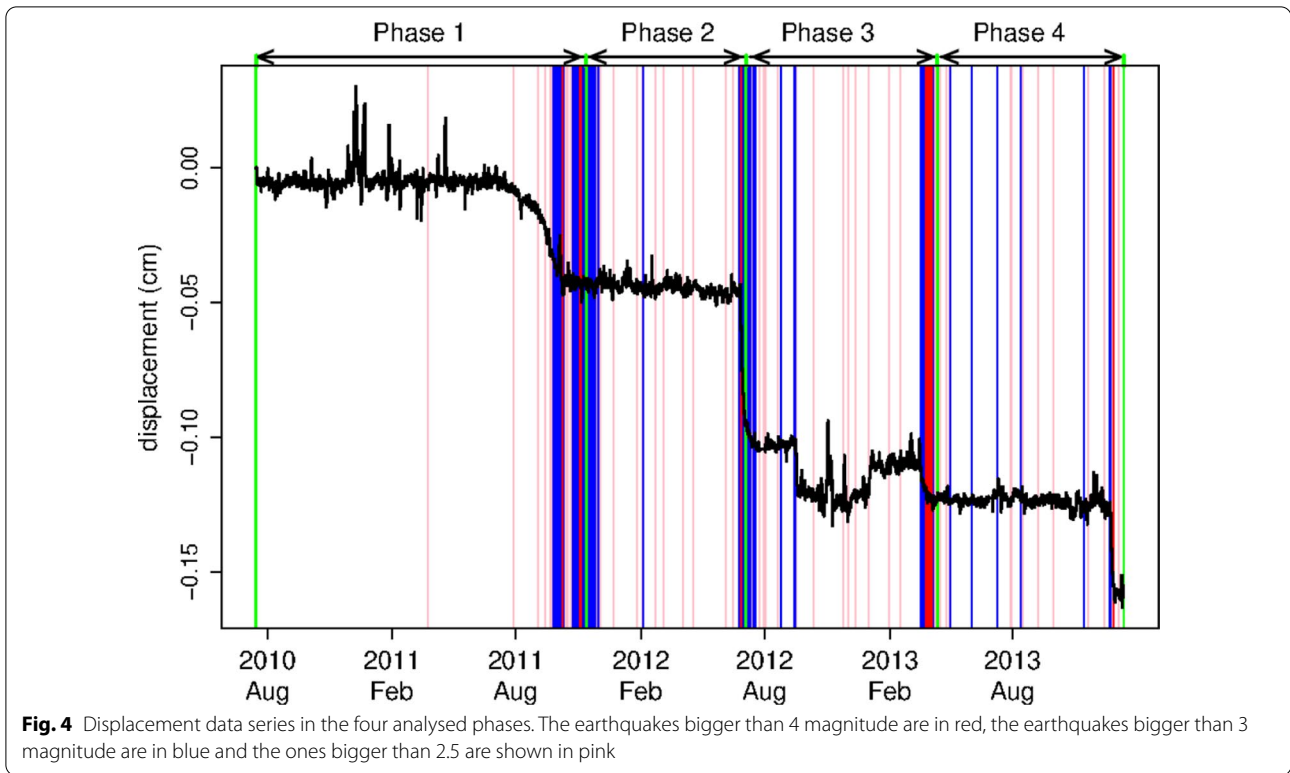
The first part of this period is characterized by the absence of earthquakes and deformation. In August 2011 the deformation reaches 24 cm/year and the cumulative energy released is increasing. On 10 October 2011, a submarine eruption occurs and on 11 November an earthquake of magnitude 4.6 occurred that marked the end of the phase.

Phase 2: From the end of the first phase (12/11/2011) to the day (03/07/2012).

Initially, the deformation is stable with some seismic events with magnitude less than 3.5. From end of June to mid-July, a reactivation process is reflected in the deformation and in the seismic events with magnitude greater than 3.5. This period ends with a earthquake of 4.2 magnitude.

Phase 3: From the end of the second phase (04/07/2012) to the day (11/04/2013).

This phase is described in García et al. (2014). As shown (see García et al. 2014 Fig. 9) the magma pressure source moves northwards becoming its vertical projection to the surface located at times between the reference location LPAL and the ground deformation location FRON, therefore, pushing the later away from the former, increasing the distance. Later in this phase, the magma pressure source moves eastwards. In all other phases the magma source projected vertically to the surface is not located, in general, between these two locations, LPAL and FRON, therefore, always decreasing the distance (García et al. 2014). This period ends with a high concentration of earthquakes greater than magnitude 4. The

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 7 of 17

Fig. 4 Displacement data series in the four analysed phases. The earthquakes bigger than 4 magnitude are in red, the earthquakes bigger than 3 magnitude are in blue and the ones bigger than 2.5 are shown in pink

Fig. 5 Choosen phases on the data series cumulative energy released. In these data series the daily energy is considered to be constant

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 8 of 17

last earthquake of magnitude 4.1 is produced on 11 April 2013.

Phase 4: From the end of the third phase (12/04/2013) to the end of data series (11/01/2014).

After a long period of stability a new reactivation of the surface deformation process is produced in December 2013. An earthquake greater than 5 magnitude is produced in the island on 27 December 2013.

The change-point analysis: subphases in the third phase

The oscillations produced in the deformation curve in the third phase could hide the relationship existing between the displacement curve and the cumulated energy released curve. There exist several methods for segmenting a data series (Aminikhanghahi and Cook 2017; Keogh et al. 2004). To isolate the oscillations produced in the deformation curve in the third phase, the change-point analysis is used. The objective is to detect whether a change has taken place for determining the points that isolate the phenomenon. This method is applied using the "changepoint" package in R software. This function implements various changepoint methods for finding single and multiple changepoints within data. In this paper the choosen method for exact changepoint detection is Segment Neighbourhood Search (SNS)(Auger and Lawrence 1989). To implement this method in R software the necessary options are the choosen "penalty method" and the "pen.value parameter", that contains the theoretical type I error. In this case, the penalty method is Asymptotic and the
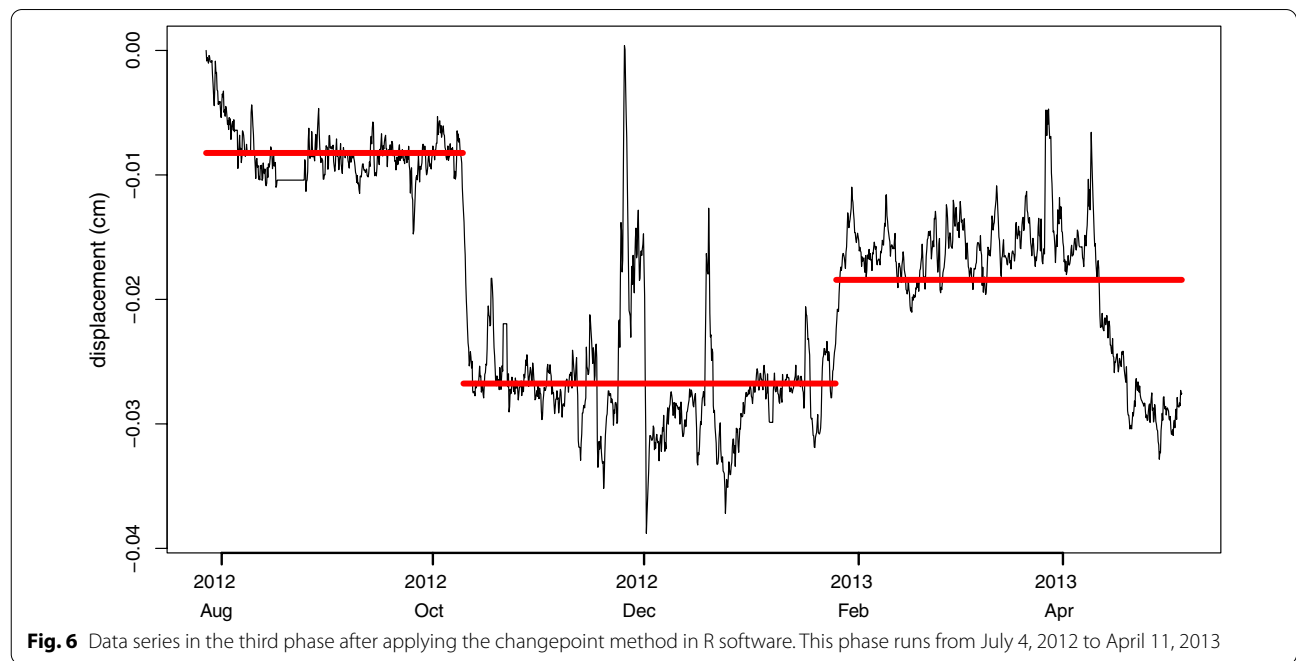
error is 0.01. The obtained segmentation using SNS method is shown in Fig. 6.

Finally, due to the continuous nature of the data, the displacement series and the series that measure the cumulative energy released have been studied from a functional point of view. Hence, our objective is studying the relation between these two curves and using this relation in Onset of Seismicity Forecast section.

Functional data analysis

The evolution of a variable over time can be seen as a functional data. Although only a sample of values of this variable are known, the functional data can be obtained by taking a linear combination of a sufficiently large number of functions from a basis of functions. Functional Data Analysis (FDA) was introduced by Ramsay (1997) and a large number of practical techniques from multivariate analysis has been extended to a functional approach. When working with functional data, reconstructing the functional form of data, from a set of known observations, is the first difficulty. For this, a basis of functions is considered. This basis is chosen in such a way that any function can be approximated from a linear combination of the K elements of the basis, $\{\phi_1, \phi_2, \ldots, \phi_K\}$. Thus:

$$x_i(t) = \sum_{m=1}^{K} c_m \phi_m(t). \tag{6}$$

Fig. 6 Data series in the third phase after applying the changepoint method in R software. This phase runs from July 4, 2012 to April 11, 2013

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 9 of 17

The choice of the basis is very important and it has been done by considering the nature of the studied curves.

Functional data can be estimated from a basis of functions by considering two elements: the number of terms in the basis K and the value of the penalty or smoothing parameter λ. The function of λ is to find a balance between the data fitting and the smoothness of the curves (Ramsay and Silverman 2005).

Hence, by following Ramsay and Silverman (2005) the curves $x_i(t)$ can be obtained minimizing the expression:

$$\sum_{j=1}^{p} [x_{ij} - x_i(t_j)]^2 + \lambda \int (D^2 x_i(t))^2 dt, \tag{7}$$

where $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is the vector of observed values at the days. $t_1, t_2, \ldots, t_p$ and D is the differential operator.

Thus, if $\lambda = 0$, there is no penalty and if λ is big, the fitted curve x_i approaches the standard linear regression to the p observed values. For this reason the data are adjusted to a polynomial of degree $p - 1$.

The reason for proposing a FDA-based treatment is that the GPS data, measured at each time interval, constitute only a discrete sequential sample of the geodetic displacements, which are continuous in nature. Using FDA, first, the original space, which is a functional space, is reconstructed. And, second, the smoothing process used to obtain the functional data filters out some of the noise that accompanies the original data. This noise is mainly due to measurement and processing errors. From the perspective of functional analysis, each sequence of GPS observations becomes a curve that can be analysed with functional techniques, developed in the last 30 years from vector multivariate techniques. In addition, the functional data can be successively derived to obtain the curves of velocity, acceleration, etc., which in turn can be treated with FDA techniques (Pérez-Plaza 2020).

To obtain the smooth curves about the displacement and about the cumulative energy released, a basis of p-splines has been used. The optimal number of terms in the basis and the optimal value for the smoothing parameter have been chosen using the generalized cross validation criterion (GCV). For mathematical details about this method see the Chapter 5 in Ramsay and Silverman (2005).

This criterion is implemented in R through the min. basis function of fda.usc package (Febrero-Bande and Oviedo de la Fuente 2012).

The splines are chunky polynomial functions on which restrictions are imposed. The junction points, called nodes, divide the range of x into regions. The splines depend on three elements: 1. Polynomial Degree 2. Number of nodes 3. Location of the nodes The general properties of a b-spline of degree q are:

- It consists of $q + 1$ polynomial pieces, each of degree q; the polynomial pieces join at q inner knots;
- At the joining points, derivatives up to order $q - 1$ are continuous;
- The b-spline is positive on a domain spanned by $q + 2$ knots; everywhere else it is zero;
- Except at the boundaries, it overlaps with $2q$ polynomial pieces of its neighbors;
- At a given x, $q + 1$ b-splines are nonzero.

In the case of using a b-splines basis with a penalty parameter, λ, is said that it is a p-splines basis. For more details about p-splines see DeBoor (1978).

From the optimal values of K and λ obtained using the generalized cross validation criterion (GCV) (implemented in the R software) and using Eqs. (6) and (7) the curves about displacement and cumulative energy released will be obtained. This procedure will be used in the entire analysis and in the analysis of each phase.

Functional descriptive statistics

Let $x_1(t), x_2(t), \ldots, x_n(t)$ be a functional random sample of size n, taken from the functional variable $\mathcal{X}$, defined in $\tau \equiv [0, T]$. Ramsay (Ramsay and Silverman 2005) defined in this case the functional descriptive statistics (functional means variances and covariances) as:

$$\bar{x}(t) = n^{-1} \sum_{i=1}^{n} x_i(t), \ \forall t \in \tau,$$

$$S_x^2(t) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i(t) - \bar{x}(t))^2; \text{ and}$$

$$Cov_x(t_r, t_s) = \frac{1}{n-1} \sum_{j=1}^{n} (x_j(t_r) - \bar{x}(t_r))(x_j(t_s) - \bar{x}(t_s)).$$

The problem is different if the objective is to calculate the functional descriptive statistics over a functional data x. In this case, if we consider the inner product:

$$< x, y > = \int_{\tau} x(s)y(s)ds.$$

the mean value function for $x \in L^2$ can be defined as $f_{\bar{x}} \in L^2$, where

$$f_{\bar{x}}(t) = \overline{x(t)} 1(t); 1(t) = 1; \forall t \in \tau$$

and

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 10 of 17

$$\overline{x(t)} = T^{-1}\langle x(t), 1(t)\rangle$$

$$= T^{-1}\int_0^T x(t)1(t)dt = T^{-1}\int_0^T x(t)dt. \quad (8)$$

Likewise the variance value function for x can be defined as $f_{S^2} \in L^2(\tau)$, where $f_{S^2}(t) = S^2_{x(t)}1(t)$

$$S^2_{x(t)} = T^{-1}\langle x(t) - \overline{x(t)}1(t),$$

$$x(t) - \overline{x(t)}1(t)\rangle = T^{-1}\int_0^T (x(t) - \overline{x(t)}1(t))^2 dt. \quad (9)$$

The covariance value function between $x, y \in L^2$ can defined as $f_{S_{XY}} \in L^2(\tau)$, where $f_{S_{XY}}(t) = S_{x(t)y(t)}1(t)$ and

$$S_{x(t)y(t)} = T^{-1}\langle x(t) - \overline{x(t)}1(t), y(t) - \overline{y(t)}1(t)\rangle$$

$$= T^{-1}\int_0^T (x(t) - \overline{x(t)}1(t))(y(t) - \overline{y(t)}1(t))dt. \quad (10)$$

Correlation measure

On the other hand, Sangalli et al. (2009) proposed, given two elements $x, y \in L^2(\tau; \mathbb{R})$, $x, y : \tau = [0, T] \longrightarrow \mathbb{R}$ and where L^2 is the space of measurable functions whose squared function has a finite integral, a measure of similarity between their associated curves given by the following formula:

$$\rho(x, y) = \frac{\int_\tau x'(s)y'(s)ds}{\sqrt{\int_\tau (x'(s))^2 ds}\sqrt{\int_\tau (y'(s))^2 ds}}. \quad (11)$$

This measure can be interpreted as the continuous version of Pearson correlation coefficient for the derivatives of the curves.

By considering the functional descriptive statistics over the functional data $x, y \in L^2$ and the classical version of the Pearson's correlation coefficient:

$$r = \frac{S_{xy}}{S_x S_y}, \quad (12)$$

and replacing (8), (9) and (10), in (12) the correlation value function $f_{r_{XY}} \in L^2(\tau)$, where $f_{r_{XY}}(t) = r_{x(t)y(t)}1(t)$ and

Hence, the measure defined by Sangalli et al. (2009), (11) is the correlation coefficient for the derivatives of the centered curves.

Now, the idea is to use this correlation measure between the curves obtained with our data. That is, our objective is to establish a relation between the cumulative energy released and the displacement of FRON station with respect to reference station in La Palma island (LPAL). The time delay between the displacement and the seismic activity is analysed too. This time delay is obtained in the following form: we displace the curve in 4-h movements and for each movement we calculate the correlation measure between the cumulative energy released and the displaced function. Then, we consider as time delay the time of traslation under which the correlation measure is maximum.

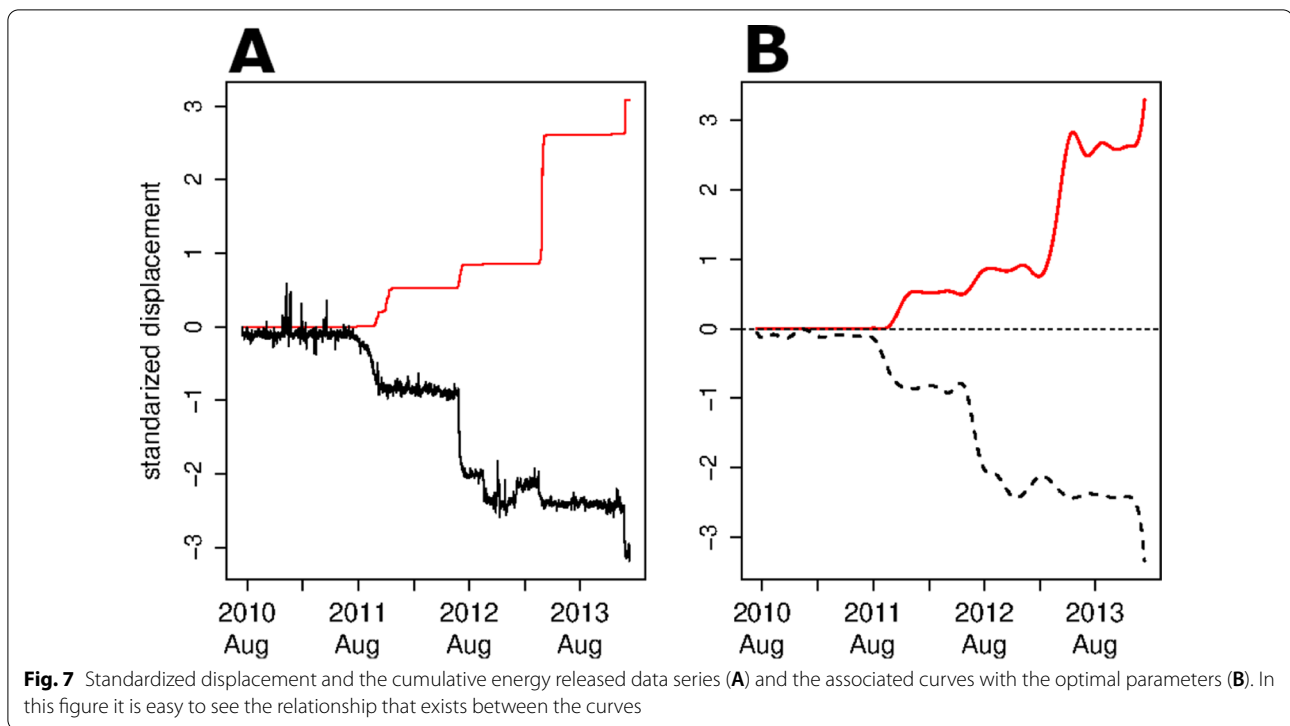A summary about the used methodology can be seen in the graphical abstract.

Results and discussion

The entire analysis period

The displacement data series and the series that measure the cumulative energy released have different value ranges and different unit. For this reason, to display together these series, these two data series were standardized, using for each curve, the mean and the standard deviation defined by (8) and (9).

The standardized displacement and the cumulative energy released data series and the associated curves with the optimal parameters are shown in Fig. 7. In this case the optimal parameters obtained using the CVG criterion in R software are $\lambda = 512$ and $K = 29$.

To evaluate the relation between these two curves, the functional correlation previously defined in (9) is used. This function was implemented in R software to obtain the lag time, where the relationship between surface deformation curve and the cumulative energy released curve was maximum. More precisely, the functional correlation was calculated between the cumulative energy released curve and the surface deformation curve, moving the last curve among 0 and 600 h in 4-h movements. Hence, the second

$$r_{x(t)y(t)} = \frac{S_{x(t)y(t)}}{S_{x(t)}S_{y(t)}}$$

$$= \frac{T^{-1}\int_0^T (x(t) - \overline{x(t)}1(t))(y(t) - \overline{y(t)}1(t))dt}{\sqrt{T^{-1}\int_0^T (x(t) - \overline{x(t)}1(t))^2 dt}\sqrt{T^{-1}\int_0^T (y(t) - \overline{y(t)}1(t))^2 dt}}$$

$$= \frac{\int_0^T (x(t) - \overline{x(t)}1(t))(y(t) - \overline{y(t)}1(t))dt}{\sqrt{\int_0^T (x(t) - \overline{x(t)}1(t))^2 dt}\sqrt{\int_0^T (y(t) - \overline{y(t)}1(t))^2 dt}}. \quad (13)$$

Fig. 7 Standardized displacement and the cumulative energy released data series (**A**) and the associated curves with the optimal parameters (**B**). In this figure it is easy to see the relationship that exists between the curves

curve was moved a maximum of 25 days. The data series of 150 correlation values are calculated.
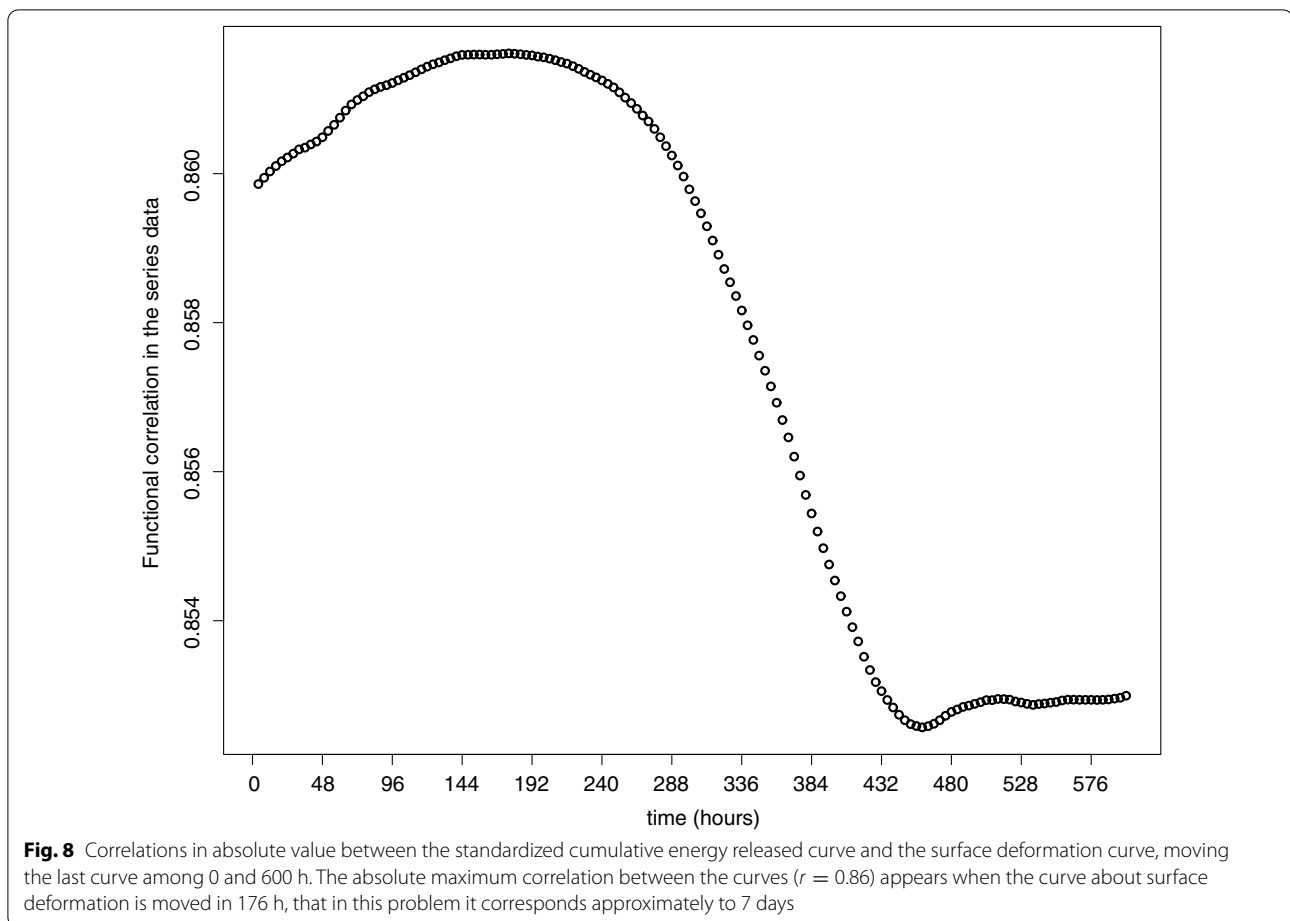
The results are shown in Fig. 8. This figure shows that the absolute maximum correlation between the curves ($r = 0.86$) appears when the curve about surface deformation is moved in 176 h, that in this problem it corresponds approximately to 7 days. This value shows a relationship between the curves, that already were studied by other authors (Domínguez Cerdeña et al. 2018; Lamolda et al. 2017). However, in the long period considered multiple sources of deformation and seismicity may exist and this fact may have influenced the results. Next, the phenomena will be examined and several time periods with an homogeneous behaviour will be considered.

The phases analysis

For each phase the standardized cumulative energy released and the standardized surface deformation curves were initialised. Figure 9 shows the displacement-data series and the data series about the cumulative energy released, initialised at the start of each phase and the smooth curves in the p-splines basis.

The functional correlations by phases and the reaction time in which the correlation measure is maximum are contained in Table 1. This table shows the high correlation values existing by phases between the displacement curve and the curve about the cumulative energy released. This trend occurs in all phases, except for the third phase, in which the pressure source location leads to increase of the distance (García et al. 2014), differently from all other phases. Moreover, in the first phase, after a long period of seismic inactivity, the response time is about 31.7 days, while than in the other phases is almost non-existing. Figure 10 shows for each phase the functional correlation between the cumulative energy released curve and the surface deformation curve, moving the last curve among 0 and 400 h in 4-h movements.

After dividing the third phase in the three parts a, b and c (Fig. 6) determined by the change point method, the correlation values obtained in the periods a and c are similar to those obtained in phases 1, 2 and 4. Then, the only one correlation value less than 0.85 is obtained in the period when the pressure source location leads to increase of the distance (García et al. 2014) (Table 2).

Pérez-Plaza *et al. Earth, Planets and Space*   (2021) 73:177

Page 12 of 17

Fig. 8 Correlations in absolute value between the standardized cumulative energy released curve and the surface deformation curve, moving the last curve among 0 and 600 h. The absolute maximum correlation between the curves ($r = 0.86$) appears when the curve about surface deformation is moved in 176 h, that in this problem it corresponds approximately to 7 days

Onset of seismicity forecast

A high relationship has been established between the sismicity and the deformation curves. Moreover, after a long period without seismic and volcanic activity, the time lag between the deformation process and the beginning of the seismic activity takes approximately 1 month.

In this section a method to predict in real time the beginning of the seismic activity will be proposed. Based on the displacement series data until the beginning of the deformation on 28 June 2011, 2-day data sequences will be added to this series data. The splines curves with the optimal parameters are considered and a level of minimal derivative is fixed. This level is calculated as the minimun of the slope of the previous 2 days. When calculating the level, the first 30 data are not considered for the oscillations that occur at the beginning of the b-splines curves.

This analysis is based on the changes produced in the derivative curves when there is a rapid descent in the deformation curve.

Figures 11 and 12 show the derivatives curves in the spline basis and the horizontal line that marks the level of minimal derivative. These figures show that after a period of 18 days approximately (16 Jul 2011) from onset of deformation (28 Jun 2011) the minimal derivative is exceeded and in this case a warning system could be activated. That is about 13 days before the beginning of seismicity (29 Jul 2011). In Fig. 5 a green line marks this day in which the cumulative energy released begins to grow significantly.
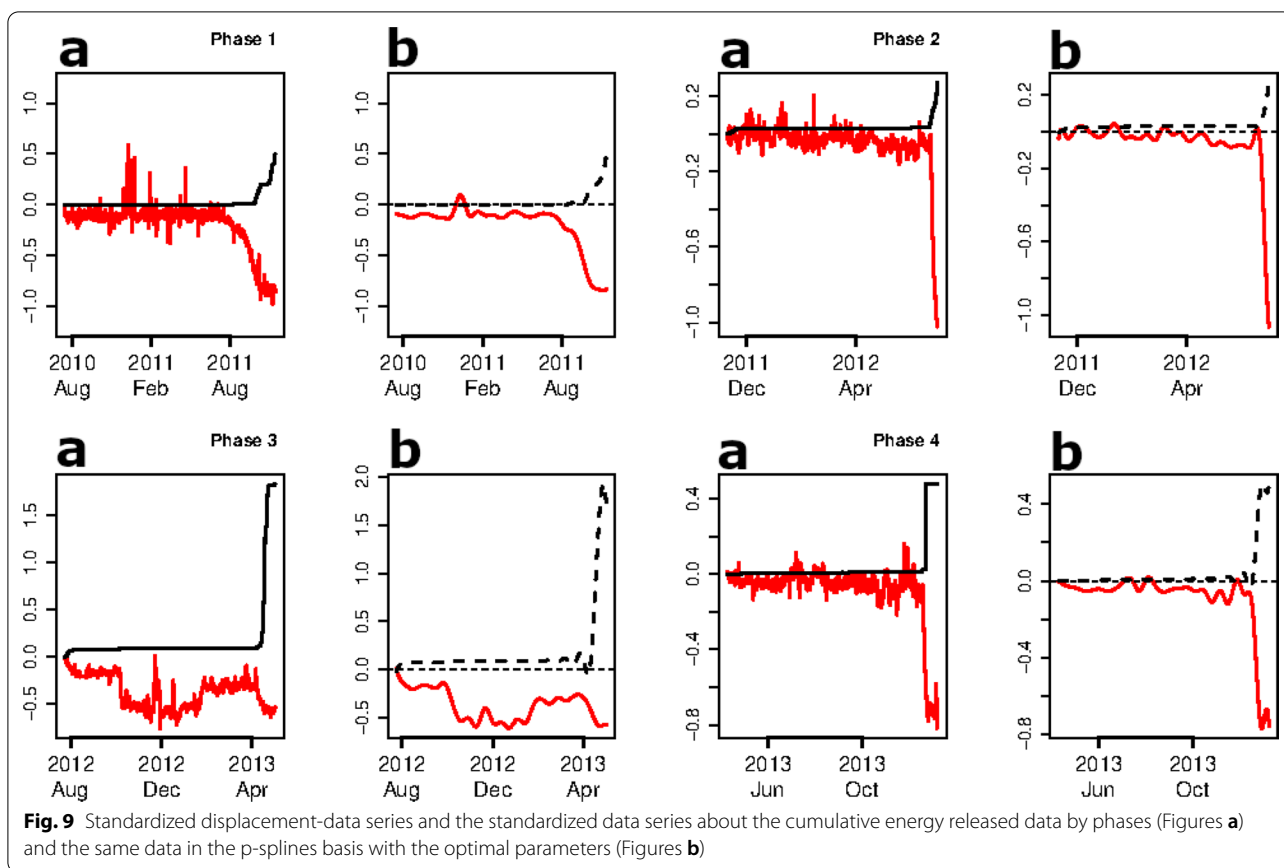
Fig. 9 Standardized displacement-data series and the standardized data series about the cumulative energy released data by phases (Figures **a**) and the same data in the p-splines basis with the optimal parameters (Figures **b**)

Table 1 Functional correlations and the time measured in days, where the correlation measure is maximum, for each phase

Phase	Correlation	Time delay
Phase 1	0.936	31.7 days (760 h)
Phase 2	0.971	0 days (0 h)
Phase 3	0.346	1.5 days (36 h)
Phase 4	0.986	2.67 days (64 h)

Conclusion

In this work, the relation between the seismicity and the deformation in El Hierro island is studied from a functional point of view. With the treatment of the data from this perspective, it is possible to analyse the evolution of the period of time between the cumulative energy released and the deformation.

Using the similarity measure proposed by Sangalli et al. (2009), a functional correlation measure is obtained. This measure is employed to establish a relation between the deformation curve (measured by the displacement between the FRON station and the reference station LPAL) and the curve of cumulative energy released.

A global analysis of these data was made to validate the above relationship. Using an analysis in several phases, the correlation measure has established a high relation in all phases ($r \geq 0.936$), except in the third one, where the pressure source location leads to increase of the distance (García et al. 2014), differently from all other phases. The different phases were considered taking into account that a single relevant seismovolcanic event was considered by phase, and each phase ends the day of the highest magnitude earthquake in the period. A detailed analysis of the third phase, that uses the change point analysis, shows that only the
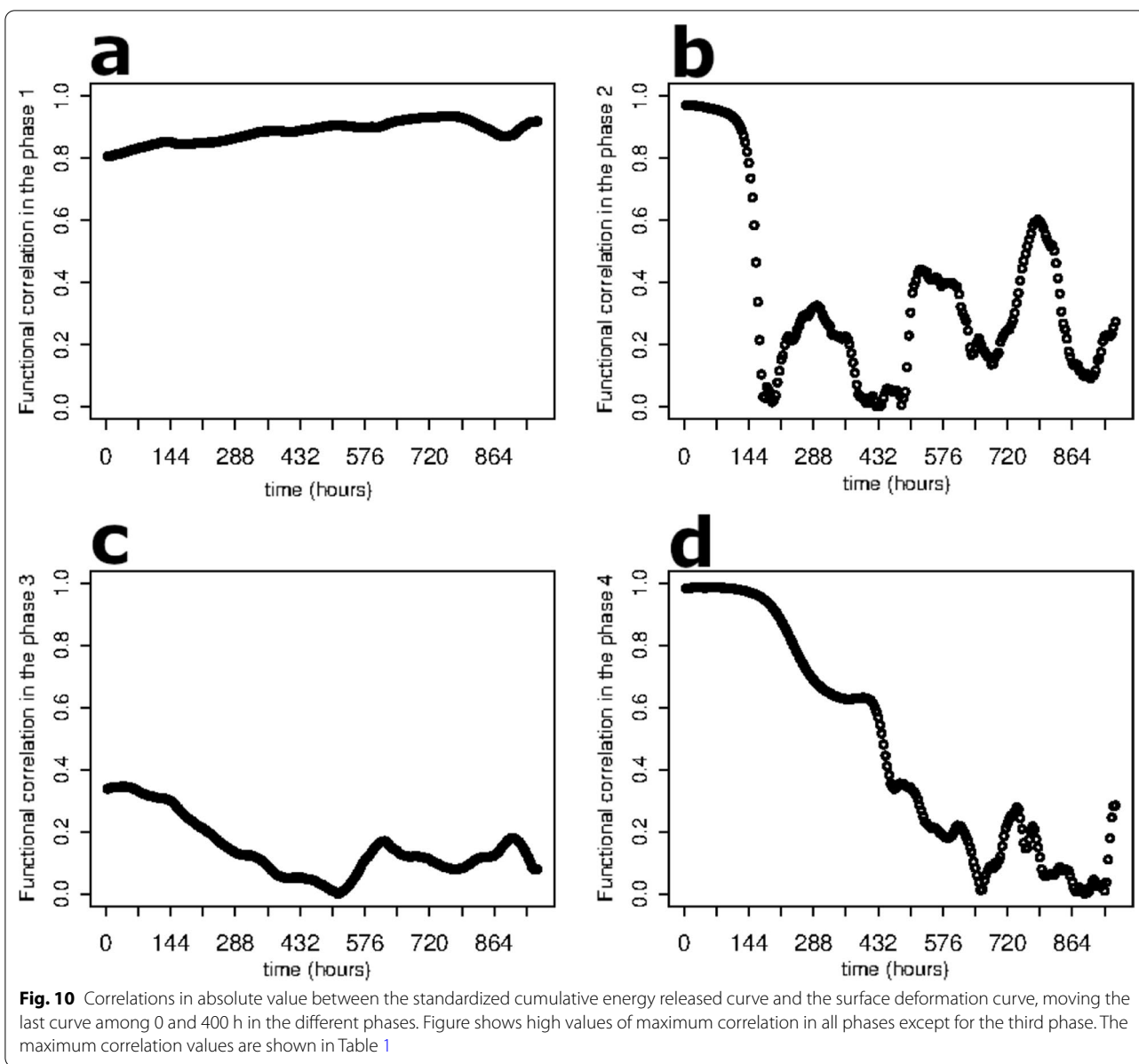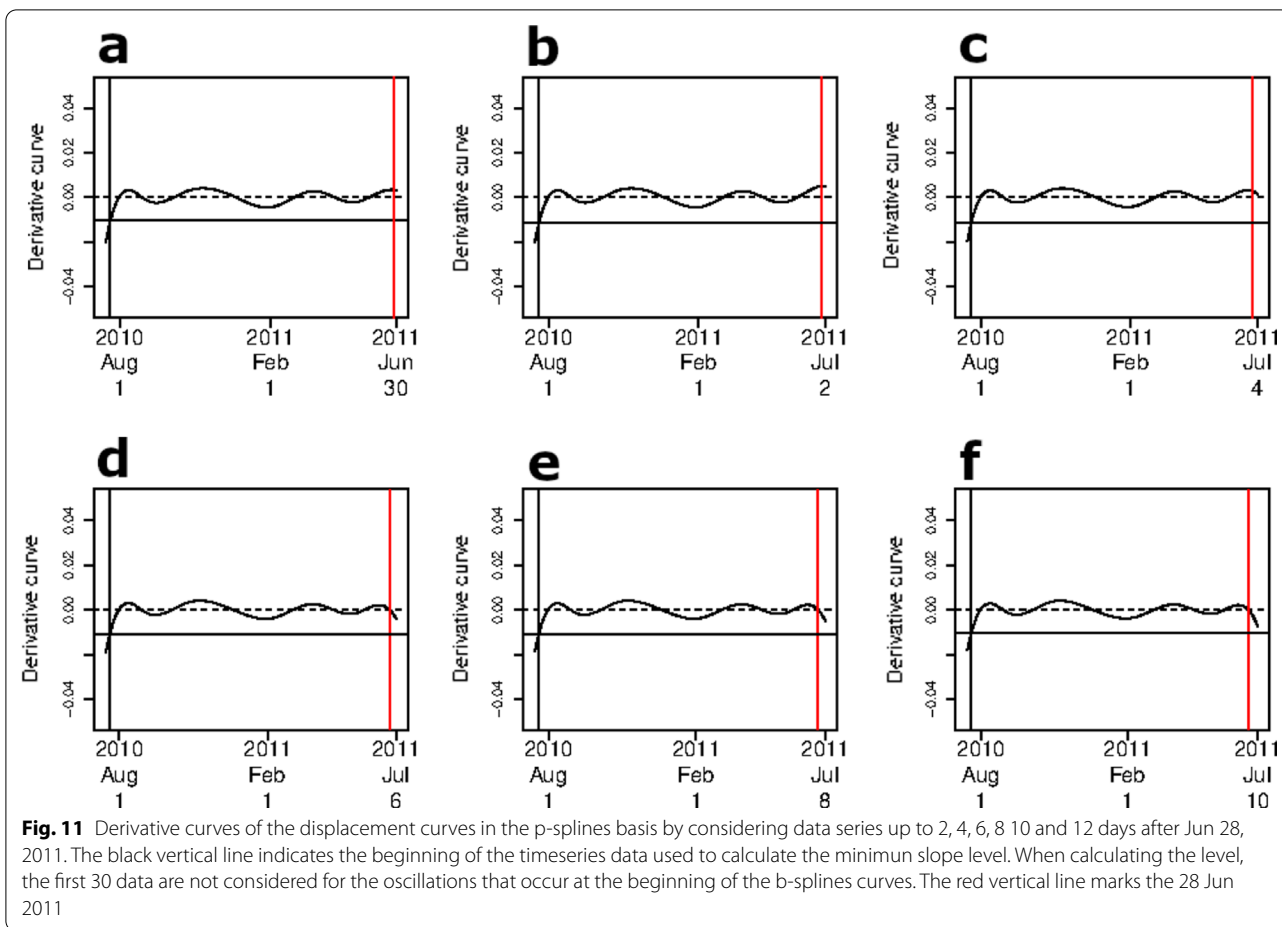
Pérez-Plaza *et al. Earth, Planets and Space*   (2021) 73:177

Page 14 of 17

Fig. 10 Correlations in absolute value between the standardized cumulative energy released curve and the surface deformation curve, moving the last curve among 0 and 400 h in the different phases. Figure shows high values of maximum correlation in all phases except for the third phase. The maximum correlation values are shown in Table 1

Table 2 Functional correlations and the time measured in days, where the correlation measure is maximum by considering three sub-phases in the third phase

Phase	Correlation	Time delay
3 (a)	0.84	0 days (0 h)
3 (b)	0.32	5.67 days (136 h)
3 (c)	0.84	0 days (0 h)

sub-phase that contains the change from decreasing distance to increasing distance and then to decreasing distance to the reference location of LPAL, due to the pressure source location moving northwards for that short period, produces a low functional correlation value. In the other sub-phases of this phase, the correlation measure values are similar to the other periods.
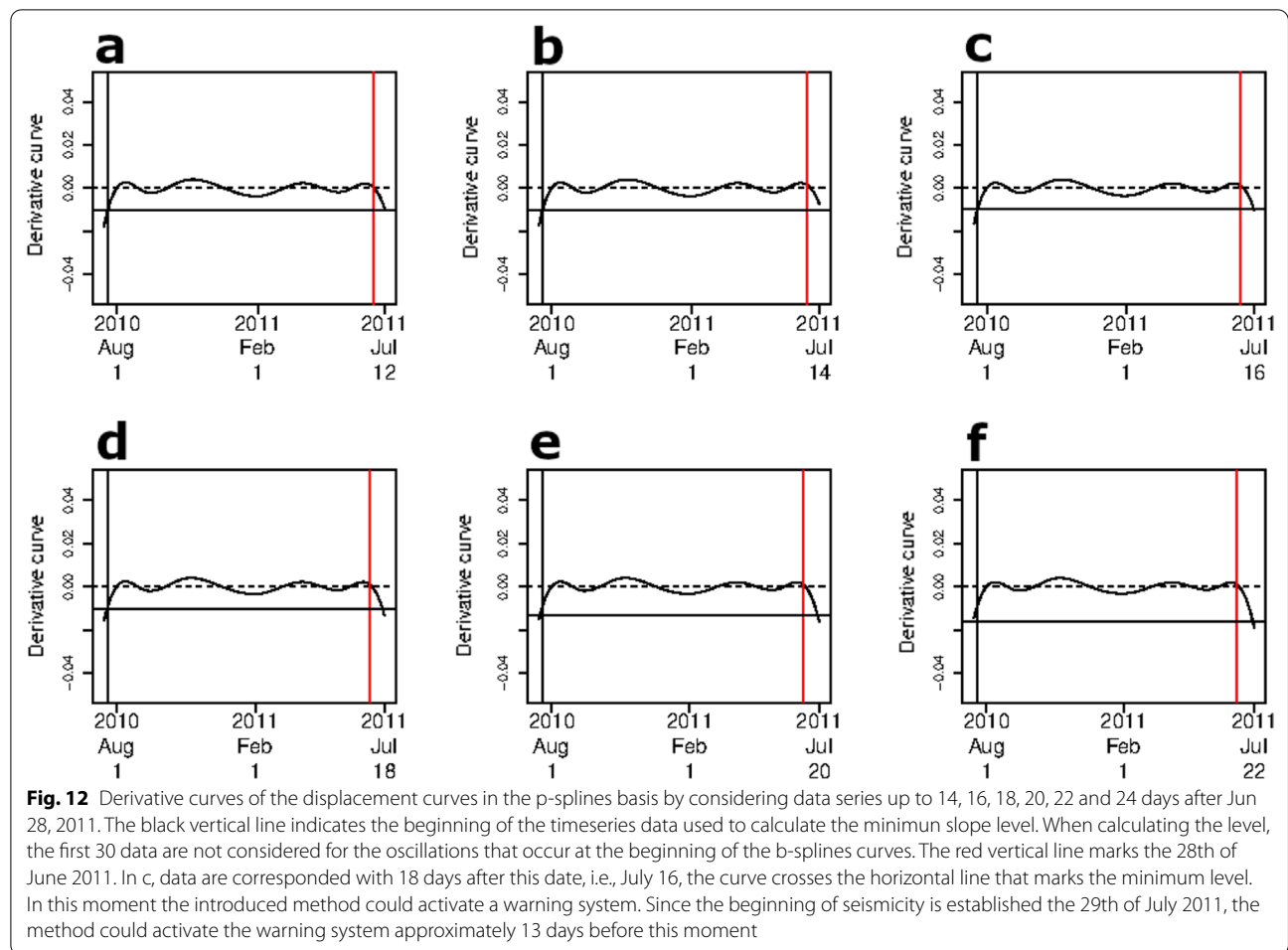
Fig. 11 Derivative curves of the displacement curves in the p-splines basis by considering data series up to 2, 4, 6, 8 10 and 12 days after Jun 28, 2011. The black vertical line indicates the beginning of the timeseries data used to calculate the minimun slope level. When calculating the level, the first 30 data are not considered for the oscillations that occur at the beginning of the b-splines curves. The red vertical line marks the 28 Jun 2011

Hence, this type of phenomenon could be detected by using the proposed correlation measure.

The synchronisation time between the deformation curve and the curve of cumulative energy released has been analysed in the different phases. It has been noted that the long period of volcanic inactivity before the first phase produces a reaction time of a month between the deformation process and the seismic data. However, in the subsequent phases, the changes produced in magma rheology and dynamics on start phases (Martí et al. 2013) produce an effect in the surface deformation curve (Dzurisin 2006). Possibly due to this effect, the deformation curve and the curve of cumulative energy released are well synchronised in their inverses movements, thus, the reaction time is almost non-existent.

After a long period without seismic and volcanic activity in El Hierro island, the time lag between the deformation process and the beginning of the seismic activity takes approximately 1 month. In a similar situation, a method to predict in real time the beginning of the seismic activity was proposed. This method, based on the changes produced in the derivative curves when there is a rapid descent in the deformation curve, could activate a warning system approximately 13 days before the beginning of seismicity.

Pérez-Plaza *et al. Earth, Planets and Space* (2021) 73:177

Page 16 of 17

Fig. 12 Derivative curves of the displacement curves in the p-splines basis by considering data series up to 14, 16, 18, 20, 22 and 24 days after Jun 28, 2011. The black vertical line indicates the beginning of the timeseries data used to calculate the minimun slope level. When calculating the level, the first 30 data are not considered for the oscillations that occur at the beginning of the b-splines curves. The red vertical line marks the 28th of June 2011. In c, data are corresponded with 18 days after this date, i.e., July 16, the curve crosses the horizontal line that marks the minimum level. In this moment the introduced method could activate a warning system. Since the beginning of seismicity is established the 29th of July 2011, the method could activate the warning system approximately 13 days before this moment

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Statistics and Operations Research, University of Cadiz, Cadiz, Spain. [2]Laboratory of Astronomy, Geodesy and Cartography. Department of Mathematics, University of Cadiz, Cadiz, Spain. [3]Higher Institute of Engineering. Department of Civil Engineering, University of the Algarve, Faro, Portugal.

References
Altamimi Z, Collilieux X, Métivier L (2011) Itrf 2008: an improved solution of the international terrestrial reference frame. J Geod 85(8):457–473. https://doi.org/10.1007/s00190-011-0444-4

Aminikhanghahi S, Cook DJ (2017) A survey of methods for time series change point detection. Knowl Inform Syst 51(2):339–367

Anguita F, Hernán F (2000) The canary islands origin: a unifying model. J Volcanol Geother Res 103:1–26

Araña V, Ortiz R (1991) The Canary Islands: Tectonics, magmatism and geodynamic framework. In Magmatismin extensional structural settings (The Phanerozoic African plate). eds A. B. Kampunzu and R. T. Lubala. Springer-Verlag, Barcelona, Spain

Auger I, Lawrence C (1989) Algorithms for the optimal identification of segment neighborhoods. Bull Math Biol 51(1):39–54

Berrocoso M, Fernández-Ros A, Prates G, Martín M, Hurtado R, Pereda J, García M, García Cañada L, Ortiz R, García A (2012) Analysis of surface deformation during the eruptive process of el hierro island (canary island, spain): Detection, evolution and forecasting. EGU General Assembly Conference Abstracts 14:43–51

Pérez-Plaza *et al. Earth, Planets and Space*   (2021) 73:177

Page 17 of 17

Carracedo J, Day S, Guillou H, Pérez Torrado F (1999) Giant quaternary landslides in the evolution of la palma and el hierro, canary islands. J Volcanol Geother Res 94:169–190

Carracedo J, Rodríguez Badiola E, Guillou H, De la Nuez J, Pérez-Torrado F (2001) Geology and volcanology of la palma and el hierro (canary islands). Estudios geológicos 57:175–273

Choy G, Boatwright J (1995) Global patterns of radiated seismic energy and apparent stress. J Geophys Res 100(B9):18205–18228

Dach R, Hugentobler U, Fridez P, Meindl M (2011) Bernese gps software ver. 5.0 http://www.bernese.unibe.ch/

DeBoor C (1978) A practical guide to splines, Applied Mathematical Sciences, vol 27. Springer-verlag, New York

Domínguez Cerdeña I, García Cañada L, Benito Sanz M, Del Fresno C, Lamolda H, Pereda de Pablo J, Sánchez Sanz C (2018) On the relation between ground surface deformation and seismicity during the 2012-2014 successive magmatic intrusions at el hierro island. Technophysics 744:422–437

Dzurisin D (2006) Volcano deformation: new geodetic monitoring techniques. Springer, Berlin

Febrero-Bande M, Oviedo de la Fuente M (2012) Statistical computing in functional data analysis: the r package fda.usc. J Stat Softw 51:1–28

García A, Fernández-Ros A, Berrocoso M, Marrero J, Prates G, De la Cruz-Reyna S, Ortiz R (2014) Magma displacements under insular volcanic fields, applications to eruption forecasting: El hierro, canary islands, 2011–2013. Geophys J Int 197(1):322–334. https://doi.org/10.1093/gji/ggt505

Gee M, Masson D, Watts A, Mitchell N (2001) Offshore continuation of volcanic rift zones. el hierro. canary islands. J Volcanol Geotherm Res 105:107–119

Hernández-Pacheco A (1982) Sobre una posible erupción en 1793 en la isla de el hierro (canarias). Estudios geológicos 38:15–26

Kalman R (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82(Series D):35–45

Keogh E, Chu S, Hart D, Pazzani M (2004) Segmenting time series: A survey and novel approach. Data mining in Time Series Databases. World Scientific, Singapore, pp 1–21

Lamolda H, Felpeto A, Bethencourt A (2017) Time lag between deformation and seismicity along monogenetic volcanic unrest periods: The case of el hierro island (canary islands). Geophys Res Lett 44:6771–6777

López C, Blanco M, Abella R, Brenes B, Cabrera Rodríguez V, Casas B, Domínguez Cerdeña I, Felpeto A, de Villalta MF, Del Fresno C et al (2012) Monitoring the volcanic unrest of el hierro (canary islands) before the onset of the 2011–2012 submarine eruption. Geophys Res Lett 39:13

López C, Martí J, Abella R, Tarraga M (2014) Applying fractal dimensions and energy-budget analysis to chatacterize fracturing processes during magma migration and eruption: 2011–2012 el hierro (canary islands) submarine eruption. Surv Geophys 35:1023–1044. https://doi.org/10.1007/s10712-014-9290-2

Martí J, Castro A, Rodríguez C, Costa F, Carrasquilla S, Pedreira R, Bolos X (2013) Correlation of magma evolution and geophysical monitoring during the 2011–2012 el hierro (canary islands) submarine eruption. J Petrol 54(7):1349–1373

Masson D, Watts A, Gee M, Urgeles R, Mitchell N, Bas LE, Canals M (2002) Slope failures on the flanks of the western canary islands. Earth-Sci Rev 57:1–35

Mervart L (1995) Ambiguity resolution techniques in geodetic and geodynamic applica-tions of the global positioning system. (phd thesis). PhD thesis, University of Bern

Mitchell N, Masson D, Watts A, Gee M, Urgeles R (2002) The morphology of the submarine flanks of volcanic ocean islands. A comparative study of the canary and hawaiian hotspot islands. J Volcanol Geotherm Res 115:83–107

Mogi K (1958) Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. Res Inst Univ Tokio 36:99–134

Niell A (1996) Global mapping functions for the atmosphere delay at radio wavelengths. J Geophys Res 101(B2):3227–3246

Pérez-Plaza S (2020) Estudio y aplicación de técnicas fda de geoposicionamiento. PhD thesis, University of Cádiz

Pérez-Plaza S, Fernández-Palacín F, Berrocoso M, Páez R, Rosado B (2018) Analysis of a gps network based on functional data analysis. Math Geosci 50:659–677

Prates G, García A, Fernández-Ros A, Marrero J, Ortiz R, Berrocoso M (2013) Enhancement of sub-daily positioning solutions for surface deformation surveillance at el hierro volcano (canary islands, spain). Bull Volcanol 75(6):1–9

Ramsay J (1997) A functional approach to modelling test data. In: der Linden WJ V, RK H (eds) Handbook of Modern Item Response Theory. Springer, New York

Ramsay J, Silverman B (2005) Functional Data Analysis. Springer, New York, NY

Rebischung P, Ray J, Benoist C, Metivier L, Altamimi Z (2015) Error Analysis of the IGS repro2 Station Position Time Series. AGU Fall Meeting Abstracts 2015:G23B-1065

Saastamoinen J (1973) Contribution of the theory of atmospheric refraction. Géodésique 107:13–34. https://doi.org/10.1007/BF02522083

Sangalli L, Secchi P, Vantini S, Veneziani A (2009) A case study in exploratory functional data analysis: geometrical features of the internal carotid artery. J Am Stat Assoc 104:37–48

Stroncik NA, Klügel A, Hansteen TH (2009) The magmatic plumbing system beneath el hierro (canary islands): constraints from phenocrysts and naturally quenched basaltic glasses in submarine rocks. Contrib Mineral Petrol 157(5):593–607

Telesca L, Lovallo M, Lopez C, Molist JM (2016) Multiparametric statistical investigation of seismicity occurred at el hierro (canary islands) from 2011 to 2014. Tectonophysics 672:121–128

Welch G, Bishop G (2002) An introduction to the kalman filter. Tech. rep., Department of Computer Science. University of North Carolina at Chapel Hill

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.