Geophysical Journal International

Geophys. J. Int. (2019) **217,** 1217–1231
Advance Access publication 2019 February 07
GJI Seismology

doi: 10.1093/gji/ggz078

Change-point analysis of V_P/V_S ratio time-series using a trans-dimensional McMC algorithm: applied to the Alto Tiberina Near Fault Observatory seismic network (Northern Apennines, Italy)

Giulio Poggiali,[1] Lauro Chiaraluce,[2] Raffaele Di Stefano[2] and Nicola Piana Agostinetti[2,3]

[1] *Geophysics section, School of Cosmic Physics, Dublin Institute for Advanced Studies, Dublin, Ireland. E-mail: giulio.p.210@gmail.com*
[2] *Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy*
[3] *Department of Geodynamics and Sedimentology, University of Vienna, Vienna, Austria*

Accepted 2019 February 6. Received 2019 January 31; in original form 2018 August 16

SUMMARY

Time-series of V_P/V_S ratio have been used to track local changes in elastic properties of rock volumes. Identifying such variations can provide information on the geophysical processes taking place inside a rock volume during the seismic cycle. A value of V_P/V_S ratio can be computed from traveltime of *P* and *S* waves generated from a single local event and it is representative of the value of the V_P/V_S ratio for the rocks traversed by the seismic ray, between the source and the receiver. It is straightforward, during a seismic sequence, to generate time-series of V_P/V_S ratio for events located close together and a single station. Such time-series should be able to monitor temporal variations of elastic parameters in the rock volume. Due to the very small nature of the expected changes in *P*- and *S*-wave velocity, the evaluation of V_P/V_S ratio time-series has been problematic in the past, and subjective choices about, for example the time-averaging scheme applied or event selection for constructing the time-series, have been proven to strongly affect the outcomes of the analysis. In this contribution, we present the application of a new methodology for a statistical evaluation of changes in V_P/V_S ratio time-series. The new methodology belongs to the wide class of 'change-point analysis' algorithms and is developed in the framework of Bayesian inference. The posterior probability distribution (PPD) of the change-point locations is obtained using a trans-dimensional Markov chain Monte Carlo (trans-D McMC) algorithm, where the existence and number of change-points is directly dictated by the data themselves. We apply the new algorithm to the seismic catalogue produced by the Alto Tiberina Near Fault Observatory seismic network (Northern Apennines, Italy). Here the high rate of background seismic release and the dense seismic network allow for a robust statistical analysis. The occurrence of change-points in V_P/V_S time-series identified with the proposed procedure is represented in space and time. The space–time distributions of change-points in the study area shows a clear peak of change-points following the occurrence of local main events, clustered along the main fault system activated. The robustness of the proposed approach makes it appropriate as an automatic, real-time tool for monitoring rock property changes related to seismic activity.

Key words: Probability distributions; Statistical methods; Time-series analysis; Seismicity and tectonics; Statistical seismology.

1 INTRODUCTION

By means of repeated velocity measurements over time, it has been possible to identify variations giving indications of the geophysical processes taking place inside a rock volume and having as a consequence elastic properties change, during the pre- (Ripepe *et al.* 2000; Lucente *et al.* 2010) co- and post-seismic phase (Li *et al.* 1998; Schaff & Beroza 2004). Among the studied quantities, the ratio between *P*- and *S*-wave velocity, so called V_P/V_S ratio, is a common and informative one giving insights into rock properties such as fluid content and porosity.

A value of V_P/V_S ratio can be computed from traveltime of *P* and *S* waves generated from a single local event and it is representative of the value of the V_P/V_S ratio for the rocks traversed by the seismic

© The Author(s) 2019. Published by Oxford University Press on behalf of The Royal Astronomical Society.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

rays, moving from the source to the receiver, in the approximation of identical ray paths. It is straightforward, during a seismic sequence or a swarm, to generate V_P/V_S ratio time-series for clusters of colocated events and a single station. Such time-series would be able to monitor the temporal variations of elastic parameters in the rock volume (Gritto & Jarpe 2014).

To search for variations in this kind of data set, two factors must be addressed. First, we have to consider how to build a time-series such that all the measurements contributing to the same time-series sample the same rock volume. This avoids misinterpretation of V_P/V_S ratio time-changes given by the migration of the seismicity, thus sampling different rocks volumes in different times. Obtaining such time-series is challenging and implies a preliminary analysis of the available data.

The second matter is related to the method we apply to identify statistically robust V_P/V_S ratio variations. The task of identifying timing of the variation in the analysed time-series and the number of such variations, falls into the category of 'change-point problems', that is to search for changes in data trends in a statistical way. Traditional algorithms for dealing with change-point problems mostly rely on moving averages (Lucente *et al.* 2010) or similar techniques. Their weakness concerns the necessity of fixing some aspects of how the algorithm works, such as defining in advance a window length or the number of change points that will be detected when such assumptions could severely affect the results.

Regarding the set up of time-series, we propose a data partition strategy based on hypocentre locations to select only seismic events with comparable ray paths and, at the same time, to keep as many data as possible for the sake of statistical robustness. For the actual change-point detection step we follow a Bayesian inversion approach. More specifically we use a Reversible Jump Markov chain Monte Carlo (RJ-McMC, Green 1995) algorithm. The RJ-McMC approach is very well suited for change-point problems, because the number of parameters is an unknown variable itself to be estimated along with the model parameters. There are many examples of successful applications of such techniques to geophysical problems (e.g. Bodin *et al.* 2012; Dettmer & Dosso 2012; Piana Agostinetti *et al.* 2015; Mandolesi *et al.* 2018) and, in particular, to the change-point analysis framework: borehole temperature data (Hopcroft *et al.* 2007, 2009), geochemical data (Gallagher *et al.* 2011), thermochronology (Gallagher 2012). We rely on the 'natural parsimony' inherent in this Bayesian approach to find the models with the least complexity necessary to explain the data (Denison *et al.* 2002; Malinverno 2002). In this specific case 'least complexity' means to place no more than the number of change-points supported by data (no overfitting). Furthermore, a scale factor for noise values is also estimated, following the 'hierarchical Bayes' methodology (Malinverno & Briggs 2004). This allows the algorithm to vary the noise level of the data set and limit the influence of fixed assumptions (underestimated or overestimated) on noise values, that otherwise would affect the number of change-points detected (Piana Agostinetti & Malinverno 2010).

Thus, the aim of the work is to find, if any, temporal and spatial variations in V_P/V_S ratio starting from raw data consisting of V_P/V_S ratio time-series automatically generated by a modular procedure able to take in continuous seismic waveforms and output earthquake locations and magnitude (Di Stefano *et al.* 2014). The time-series are not revised because the proposed methodology could contribute to innovative and heterogeneous warning systems for monitoring tectonically active regions or industrial sites potentially inducing seismic activity.

We apply our algorithm to a data set provided by the Italian National Institute of Geophysics and Volcanology (INGV) through The Altotiberina Near Fault Observatory (TABOO) research infrastructure (Chiaraluce *et al.* 2014). The study area is located in the Northern Apennines Fig. 1, where the high rate of background seismic release occurring all along a detailed normal fault system, together with the presence of a dense seismic network, fulfil the necessary conditions for a sufficient amount of data and a good spatial sampling over the area. More than 4000 time-series, spanning 4 yr, are built and analysed with RJ-McMC algorithm to highlight the occurrence of any variations in V_P/V_S ratio.

We validate retrieved change-points with an automatized method that keeps only change-points passing defined criteria and represents them in space and time, to highlight source-regions and time-intervals where V_P/V_S ratio variations are more likely to have occurred. To this end we will show how we represent the results obtained from the analysis of each single V_P/V_S ratio time-series: gathered together and represented on maps in an informative way. Finally, we discuss advantages and drawbacks of the proposed algorithm, together with new potential applications.

2 DATA

The data set used in this work comprises V_P/V_S ratios and associated error values related to 126 520 earthquakes recorded at about 79 TABOO seismic stations and occurred between April 2010 and March 2015 (Fig. 1). Such V_P/V_S ratios are calculated as $V_P/V_S = (t_S - t_P)/t_P + 1$ as in (Wadati & Oki 1933) from P and S traveltimes: t_P and t_S, respectively. Detailed information regarding the seismic network can be found in Chiaraluce *et al.* (2014).

The mean V_P/V_S ratio of the whole data set is as high as 1.9: a pretty regular value for an upper crust mainly composed by carbonate rocks pervasively traversed by fluids related to meteorological phenomena as well as released from the subducted Adriatic Plate (Chiodini *et al.* 2004; Piana Agostinetti *et al.* 2011; Piana Agostinetti *et al.* 2017). The values vary considerably in reason of hypocentral location quality derived by the number of stations recording the earthquake. As mentioned in the previous paragraph we did not edit the data set, that is we kept outliers of V_P/V_S ratio as large as 100+, in order to test the robustness of the proposed approach.

In the raw data sets each V_P/V_S value has an associated error value which is derived from an estimated P and S traveltime errors based on the original location analysis. In our work these errors on V_P/V_S values are used directly without further processing. The only exception is where unacceptable error values (i.e. error values equal to zero) are found. In such cases, we substitute zero values with the modal error value of the best-located events.

The event origin time was changed from standard date-time format to decimal days starting from January 2008. We show in Fig. 1(b) the result of this conversion reporting the occurrence of the events versus their magnitude. From Figs 1(a) and (b) we can appreciate the high rate of microseismic release characterizing such a small area (about 40 km x 30 km), with main episodes not exceeding $M_L = 3.8$. Due to the high detection capability of TABOO seismic network, Valoroso *et al.* (2017) showed that for the study area each event with $M_L > 2.9$ starts a seismic sequence activating a well defined normal fault segment.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

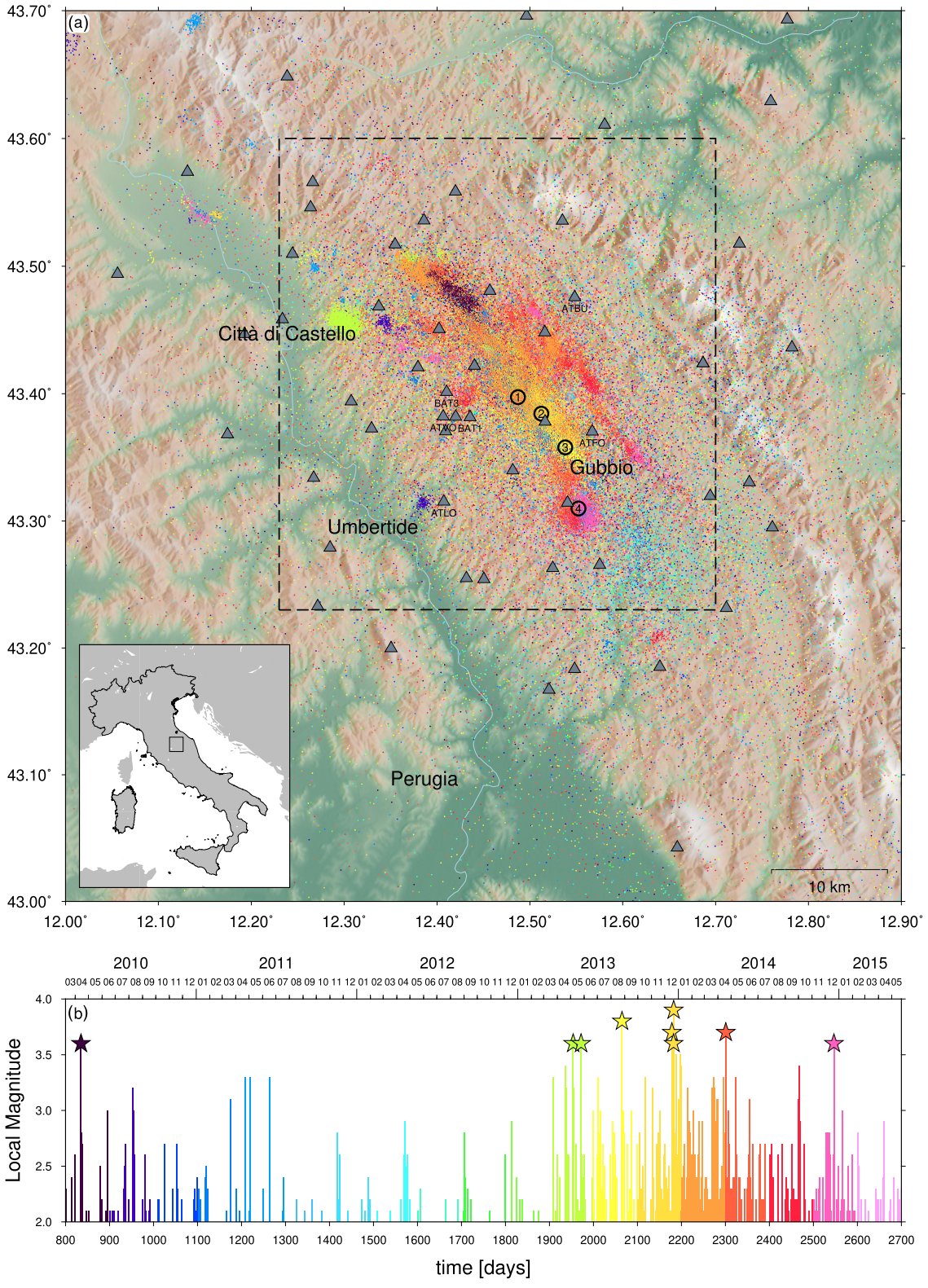

Figure 1. (a) Map of the study area. Coloured dots indicate epicentres (colours refer to origin time, see panel (b). Grey triangles show seismic stations used in this study. Numbered circles indicate the position of the cluster of seismicity cited in the text. The dashed box represents the area shown in the following figures. (b) Time-distribution of $M_L >= 2.0$ seismic events in the study area.

2.1 Data partitioning

Preparation of data involved some minor processing and, most importantly, choosing a reasonable data partition strategy to build time-series for inversion. Building V_P/V_S ratio time-series for this kind of analysis requires to arrange the data in a way such that every measurement in the same time-series is representative of the same rock volume, that is aiming to have, as much as possible, similar ray paths. If this condition is not fulfilled, changes in V_P/V_S value

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

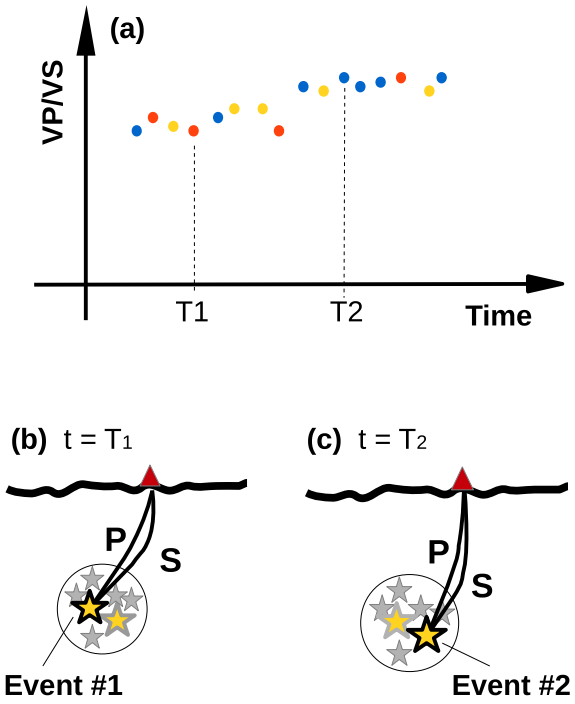

Figure 2. Simplified description of the observed data. (a) V_P/V_S ratio time-series associated with a cluster-station couple: each dot represents a measurement of V_P/V_S ratio at a given time T, different colours show different uncertainties in the measurements. $T1$ and $T2$ are the origin time of the events in panel (b). (b) For each event within a pre-defined sphere (i.e. cluster), P and S traveltimes recorded at a single seismic station (red triangle) are used to compute the measurement of V_P/V_S ratio. Such measurement is associated with the origin time of the event, here $T1$ and $T2$.

observed in a time-series could be dictated by the spatial migration of hypocentres, which leads to a change of ray paths and, thus, the sampling of different rocks (likely, with different elastic properties). Therefore we followed a selection strategy based on hypocentre positions. For each chosen station, we build a time-series selecting earthquakes originating from a volume as tiny as possible to ensure that ray paths are almost the same (see Fig. 2 for an illustrative sketch). At the same time, for robustness and reliability of the inversion process, a sufficient amount of data are needed in every time-series.

With this consideration in mind, the building of V_P/V_S ratio time-series has been made aiming to obtain the largest amount of data within the smallest source volume. The procedure for data partitioning is illustrated in Fig. 3 and detailed as follows:

(i) the whole area is gridded with grid cells size equal to 500 x 500 x 500 m;
(ii) a sphere of 1 km^3 in volume (i.e. radius of about 600 m) is centred on each node of the grid;
(iii) for each node-station pair, the number of earthquakes originating inside the sphere is counted and time-series are built from node-station pairs with a minimum number of data equal to 100.

Following these selection criteria we obtained 4200 node-station couples, of which a representation of ray coverage is visible in map in Fig. 4 (ray paths being the segments connecting each node-station pair). A time-series containing origin times (OT), V_P/V_S ratios and measurement errors is associated with each node-station couple.

Looking at the V_P/V_S values (sometimes larger than 100) and their dispersion, we are aware of the fact that the inferred location could be not optimal for all the events included in the starting catalogue (comprising more than 120 000 events), and, thus, the selection of the events belonging to a node-station couple could include some mislocated events. The choices we made for the algorithm, for example in terms of likelihood function, outliers inclusion and automatic change-point validation, are all together able to overcome this expected limitation when using automatically generated data.

3 METHOD

We organized our work-flow for data analysis in two main procedures: the first consist of applying the RJ-McMC algorithm to each time-series to obtain a change-points distribution in time for each node-station couple considered. After that, each distribution is examined to validate only change-points that are considered significant based on a list of criteria to be fulfilled. Thus, the ensemble of significant change-points is analysed considering maps of spatial and temporal density of validated change-points.

3.1 RJ-McMC implementation

3.1.1 Bayesian inference

Bayesian inference (Bayes 1763) is based on a probabilistic framework which aims to sample a posterior probability distribution (PPD) of model parameters linked to the data by means of a likelihood measure using Bayes theorem (Mosegaard & Tarantola 1995; Tarantola 2005). This relation can be written as

$$p(m|d) \propto p(m)p(d|m),$$

where $p(m|d)$ is the posterior pdf of the model m given the data d; $p(m)$ is the prior pdf on the model parameters (how the model should be based on our previous knowledge) and $p(d|m)$ is the likelihood function, which measures the probability of obtaining the data d given the model m. This measure, quantifies the fit between data and model m. Thus the posterior probability distribution describes the distribution of model parameters conditioned on data and prior information. Using a Bayesian approach to model parameter estimation is an effective way of dealing with the non-uniqueness that affects most geophysical inverse problems (Backus 1988). In this probabilistic framework, the resulting PPD is a much more complete and informative solution than any kind of search for the best fitting model (Mosegaard & Tarantola 1995).

A numerical approximation of the PPD can be obtained in an efficient way adopting a Markov chain Monte Carlo (McMC) method: a random walk is used to produce a sequence of models sampled from the PPD and such ensemble is analysed in a post-processing phase. This is a Markov process in the sense that every new model is generated conditional on the current one but accepted (or rejected) independently of any previous models. So, the basis of McMC is that starting from a point in the model parameter space, we propose a new model, *candidate model*, conditional on the current model. Different McMC implementations can be designed to control the generation of the new model and to sample efficiently a probability distribution.

The key feature of a RJ-McMC is that model complexity is inferred from the data. Thus, there is no need to impose a regularization on the solution nor to compare the results obtained with

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

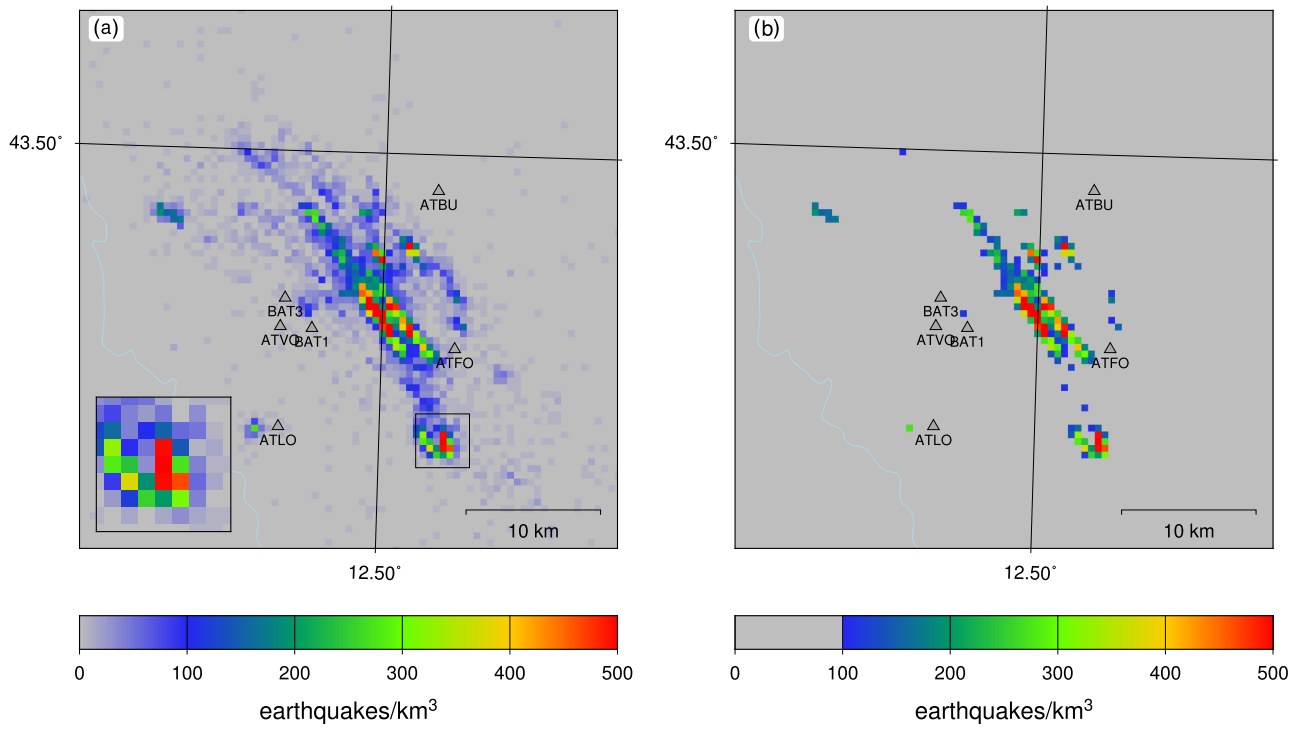

Figure 3. Example of data partitioning using cubes instead of spheres for explanatory purposes. Depth range is between 2.5 and 3.0 km. (a) number of events in a 0.5 x 0.5 x 0.5 km cube. (b) Cubes with more than 100 events that are considered in the analysis.

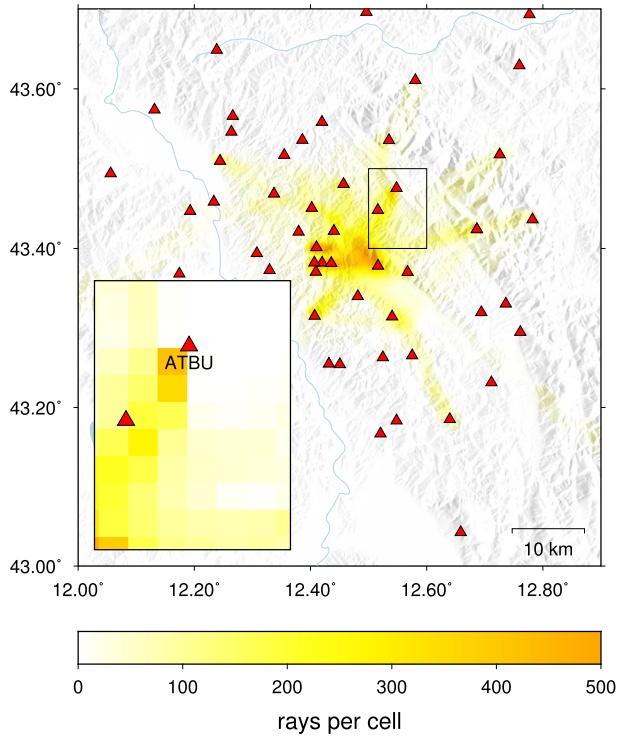

Figure 4. Total ray coverage for all the 4200 station-cluster couples selected.

different parametrizations (Sambridge *et al.* 2006). In change-point problems the ability of such techniques to treat the number of unknowns as a parameter to be estimated, is a feature of even greater importance because the number of change-points is an unknown of major interest. It could be zero if data do not support any change in time of the investigated parameters.

3.1.2 Model and priors

In our case the model is a simple step function for which we want to infer V_P/V_S values, endpoints for every partition and number of such partitions. Thus, V_P/V_S ratio time-series model is here described as a step function (left continuous) with partitions having a constant V_P/V_S ratio between change-points. The unknown model parameters are the number of change-points k, their location in time t, the V_P/V_S value for each partition and an estimated noise scale factor ω. As our step function is left continuous, each V_P/V_{Si} is the value from t_{k-1} (or t_{min}) to t_k. So an additional V_P/V_S value is needed from t_k to t_{max}. With this parametrization $\mathbf{t}$ is a vector of length k and $\mathbf{V_P/V_S}$ is a vector of length $k+1$. So a model vector can be written as: $\mathbf{m} = (k, \omega, \mathbf{t}, \mathbf{V_P/V_S})$.

The scale factor ω is introduced to follow the 'hierarchical Bayes' approach described in Malinverno & Briggs (2004). Here a common scale factor for the whole data set is estimated in a way such that $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n) = \boldsymbol{\sigma}_0 10^{\omega}$, with $\boldsymbol{\sigma}_0 = (\sigma_{0,1}, ..., \sigma_{0,n})$ the original noise estimates and n number of data points. Clearly, candidate error values $\boldsymbol{\sigma}$ also depend on the original noise level of the data set which is estimated directly from the data; in this context the original values of errors on the measurements retain only their relative importance to each other.

In Bayesian formulation prior knowledge on every parameter is expressed by means of probability distributions: typically uniform or normal distributions are used to define reasonable values that the parameters can assume. However, more complex implicit prior information can be inserted in the form of pre-selected values for some physical parameters (e.g. regional average V_P/V_S ratio) or strongly constrained parametrization (e.g. setting the same number

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

Table 1. Prior distributions of model parameters in vector **m**.

Model parameter	Prior distribution	Minimum	Maximum
Logaritmic value of scale factor for data errors ω	Uniform	–1	3
Number of change-points k	Uniform	0	100
Temporal coordinate for the change-points t_i in days	Uniform	800	2700
V_P/V_S ratio associated to each change-point	Uniform	1.5	2.5

of changepoints across different time-series). Such prior information must be carefully investigates, as clearly described in Roy & Romanowicz (2017) and Gao & Lekić (2018). To avoid any subjective choice, in our case we used uniform distributions of which minimum and maximum values for every parameter are listed in Table 1, and we invert indipendently all time-series. Uniform distributions guarantee simplicity and can be assumed from empirical obervations and field measurements. Moreover, different strategies can be easily implemented to sample uniform priors (e.g. Mosegaard & Tarantola 1995).

3.1.3 *Likelihood and acceptance probability*

For the likelihood function we adopt a sum of absolute differences measure ($l1$-norm). This choice is justified by the presence of outliers in the data set: (e.g. measurements exhibiting unreasonably high V_P/V_S ratios). One choice would have been to eliminate these values imposing a subjective threshold above which the measure is considered 'unreasonable' but as discussed earlier we decided to include all the available data exploring the possibility of creating a tool for real time automatic monitoring. In addition to this, choosing a threshold would have not been a useful choice due to the need to tune thresholds for different time-series. Furthermore an $l1$-norm is less influenced by underestimated error values that could eventually lead to unnecessary increase of the solution complexity (i.e. adding change-points not supported by data). The $l1$-norm solves these problems in an effective way and is easily implemented in the likelihood function:

$$L = \frac{1}{2\prod_{i=1}^{n} 10^{\omega}\sigma_i} \exp\left(-\sum_{i=1}^{n} \frac{\left|V_P/V_S{}_i^{obs} - V_P/V_S{}_i^{sim}\right|}{10^{\omega}\sigma_i}\right).$$

This kind of likelihood describes experimental uncertainties using a Laplacian function instead of a Gaussian function, as explained in Mosegaard & Tarantola (1995).

The crucial aspect of McMC is how we decide to accept the new model, which in this case substitute the current one, or reject it. This is the fundamental criterium that guides the sampling of the model space. Like many fixed dimension McMC algorithms, our implementation of the RJ-McMC (introduced by Green 1995) is based on Metropolis–Hastings algorithm: a model is proposed, based only on a perturbation of the current one (Markov chain), and then accepted or rejected following the acceptance probability. More specifically the probability α of a candidate model to be accepted follows the Metropolis rule:

$$\alpha = min\left[1, \frac{q(\mathbf{m}_{curr}|\mathbf{m}_{cand})}{q(\mathbf{m}_{cand}|\mathbf{m}_{curr})} \frac{p(\mathbf{m}_{cand})}{p(\mathbf{m}_{curr})} \frac{L(\mathbf{m}_{cand})}{L(\mathbf{m}_{curr})} |\mathbf{J}|\right],$$

where $p(\mathbf{m})$ is the prior distribution, $q(\mathbf{m})$ is the proposal distribution, $\mathbf{J}$ is the determinant of the Jacobian matrix of the transformation from $\mathbf{m}_{curr}$ to $\mathbf{m}_{cand}$, and $L(\mathbf{m})$ is a likelihood function which measures the fit between the observed data and the data predicted by model $\mathbf{m}$.

In this study, we adopt the approach developed in Mosegaard & Tarantola (1995), where two steps are considered for sampling solutions to geophysical inverse problems according to the PPD. In the first step, the current model is modified and the candidate model is proposed from the prior probability distributions (e.g. replacing one of the parameter in the current model with a value extracted from the prior probability distribution for such parameter, see Mosegaard & Tarantola (1995) for details on different sampling strategies). In the second step, the candidate model is accepted with a probability proportional to the ratio of the likelihood values of the candidate and current models. It can be proven that if the candidate models are generated by sampling the prior distribution, the Metropolis rule results in a random walk that samples the PPD. In this case, the expression for α reduces to

$$\alpha = min\left[1, \frac{L(\mathbf{m}_{cand})}{L(\mathbf{m}_{curr})}\right]$$

because the proposal ratio equals the prior ratio and the Jacobian term can be unity, adopting specific transformations, as showed in Piana Agostinetti & Malinverno (2010). If $L(m_{candidate}) \geq L(m_{current})$ the candidate is always accepted; otherwise the random walk moves to the candidate model with probability equal to $L(\mathbf{m}_{cand})/L(\mathbf{m}_{curr})$.

3.1.4 *Algorithm*

Our implementation of RJ-McMC falls in the so called birth–death class of algorithms, where the model dimensionality can increase or decrease only by one. More specifically the algorithm used in this work is based on Piana Agostinetti & Malinverno (2010), which in turn uses the same sampling strategy described in Mosegaard & Tarantola (1995).

After choosing a random starting model from the prior distributions of parameters, the algorithm iteratively executes the following operations to sample the PPD:

(i) pick a candidate model $\mathbf{m}_{cand}$ by applying one of the following variations (moves) on the current model $\mathbf{m}_{curr}$:

(a) perturb V_P/V_S values
(b) move a change-point in time
(c) birth move: add a new change-point
(d) death move: delete a change-point
(e) perturb noise scale value

(ii) compute the likelihood of $\mathbf{m}_{cand}$
(iii) accept or reject the candidate model according to Metropolis rule. If $\mathbf{m}_{cand}$ is accepted, then $\mathbf{m}_{curr} = \mathbf{m}_{cand}$; otherwise $\mathbf{m}_{curr}$ is left unchanged
(iv) save $\mathbf{m}_{curr}$ and restart from step (i)

Each of the different moves applied needs to be tuned to produce candidate models that lead to a reasonable acceptance ratio (ratio of accepted models to proposed models). More specifically: if perturbations are too large, we have an extensive model space exploration, but most of the proposed models would be rejected; on the contrary, if perturbations are too small, exploration is slow, but the acceptance ratio is high. In both cases sampling efficiency is compromised. Rules for tuning model perturbations applied in this work are based on acceptance ratio, considered to have 'optimal values' between 0.25 and 0.50 (Mosegaard 2006), but it can be less for trans-dimensional moves like moves (c) and (d), see Rosenthal (2011).

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

For every time-series the algorithm is iterated in 10 independent chains of 10 million iterations each. The first 5000 000 iterations of each run are discarded as a burn-in phase and every 100th model is stored. Thus, the PPD is numerically reconstructed from 500 000 models.

3.2 Automatic change-points validation

The next step of our analysis automates the check for the presence of a change-point from posterior distributions of each node-station pair. We choose an automatic way for the validation of change-points to limit the subjectivity of decisions. For a change-point to be accepted in each time window, it has to fulfil the following three criteria:

(i) Its probability in the PPD for the distribution of change-point position in time must be equal or larger than four times the prior probability (which in our case, is uniform over all the time-window). The selected change-point not only must be a peak in the posterior distribution, but it has to be a robust one, that is the position of that change-point in time is validated by many models.

(ii) Number of data before and after change-point must be larger or equal to 10 per cent of total number of data in time-series. This is to prevent a change-point being validated only because has poor constraints (few data) on one side of the time axis.

(iii) The overlapping area of the 1-D marginal PPDs for V_P/V_S ratio, calculated before (from t_{k-1} to t_k) and after (from t_k to t_{k+1}) the change-point, must be smaller or equal 10 per cent. This is to confirm only change-points that mark a significant variation of V_P/V_S value. In other words the limited overlapping of the posterior distributions of V_P/V_S ratio before and after a change-point is an indicator of a well constrained variation. This is a conservative criterium, as it could exclude short-term variation of the V_P/V_S value.

4 RESULTS

We first show some applications of the Bayesian algorithm and examples regarding change-points detection. Then we discuss some cases of the automatic validation of change-points. The occurrence of change-points is the represented in space and time with sequential maps, each of which refers to a different time window.

4.1 Inversion results

The first example (Fig. 5) shows a test where we sample the prior distributions for events originated in cluster '2' and recorded by station BAT3. Analysing samples collected from the prior distributions gives the fundamental insight that the algorithm is working properly. This is a key-point point for algorithms that follow the approach described in Mosegaard & Tarantola (1995), where the samples must be extracted from (potentially) complex priors. Locations of stations and clusters are shown in Fig. 1. More precisely cluster '2' is an ensemble of events with hypocentres inside one of the spheres described in Section 2.1; depth of sphere centre is 3.25 km in this case. Sampling from the prior distributions is obtained accepting all candidates (i.e. imposing $\alpha = 1$). Panels from (a) to (d) show sampled parameters, respectively: V_P/V_S ratio variation with time (a), change-points distribution in time (b), number of change-points (c), scale factor for data errors (d). For reference, data are shown as coloured dots in panel (a), with colours associated with data errors (σ_0).

As shown in Fig. 5, excluding the information contained in the data causes the algorithm to sample uniform distributions as expected. In particular models with no change-points or with one hundred of change-points are both samples with the same probability, without introducing any bias in the number of change-points. Moreover, change-points can be inserted at every time step with the same probability. The grey shade in Fig. 5(a) indicates that the value of the V_P/V_S ratio is also uniform within the given bounds.

In Fig. 6 we present the result of the inversion on the same data set as in Fig. 5, but sampling the PPD, that is where prior information are combined with information given by the data to obtain statistical inference on the model parameters. In panel (a), we show the PPD for V_P/V_S values in time, while in panel (b) the PPD for the time-position of the change-points is presented. We observe that the distribution in panel (b) presents three local maxima. However, only the third maximum seems to be associated with a realistic change-point. In fact, for the first one, the maximum probability is not significantly different from the prior probability value (also shown in Fig. 5b), while the V_P/V_S ratio before and after the second local maximum has similar 1-D marginal PPD. It is worth noticing that, in portions of the time-series where data are not present, the resulting PPDs tend to the prior probability distributions, as seen in panels (a) and (b) before $t = 1900$. In panel (c), we present the PPD for the number of change-points and we can appreciate an example of the 'natural parsimony' (Malinverno 2002) of trans-dimensional algorithms: the PPD of the number of change-points has a single maximum at about 8 change-points and the sampled number of change-points is never above 30, even if the maximum of change-point is as high as 100 (i.e. no over-fitting is performed). Finally, the PPD for the $\log_{10}$ value of error scaling factor is presented in panel (d), with a maximum probability for values slightly larger than 0.0 (which translates to a scaling factor slightly larger than 1.0). Such results indicates that error values in V_P/V_S ratio measurements are moderately underestimated.

4.2 Selection of significant change-points

To explain how automatic validation of change-points works, in Fig. 7 we present a collection of results from Rj-McMC inversions for different node-station pairs. Each panel shows the PPD of V_P/V_S values and change-points positions in time. All panels show some special cases to better understand how the method explained in Section 3.2 works, together with some limitations of such method.

In panel (a) we present a 'no-change-points' case, to demonstrate that change-points are actually placed only if statistically supported by data. The Rj-McMC algorithm can easily and effectively deal with a no-change-points case, unlike other methods where a predefined number of change-points are placed even if there is no need to. This is a major feature of our approach not to be underestimated because it makes the output much more reliable, knowing that if a change-point is placed is due only to the data.

Panel (b) is a zoom the same data set of Fig. 6, specifically from day 1900 to day 2300. As seen before, three change-points are seen in the results of Bayesian inversion respectively at $t = 2000$, $t = 2050$ and $t = 2200$. The first and the second are not accepted by the automatic validation: the first one because of the height of the histogram being less than four times the prior level, the second one because of the difference in V_P/V_S ratio before and after the change-point. Only the third one is classified as a 'true' change-point.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

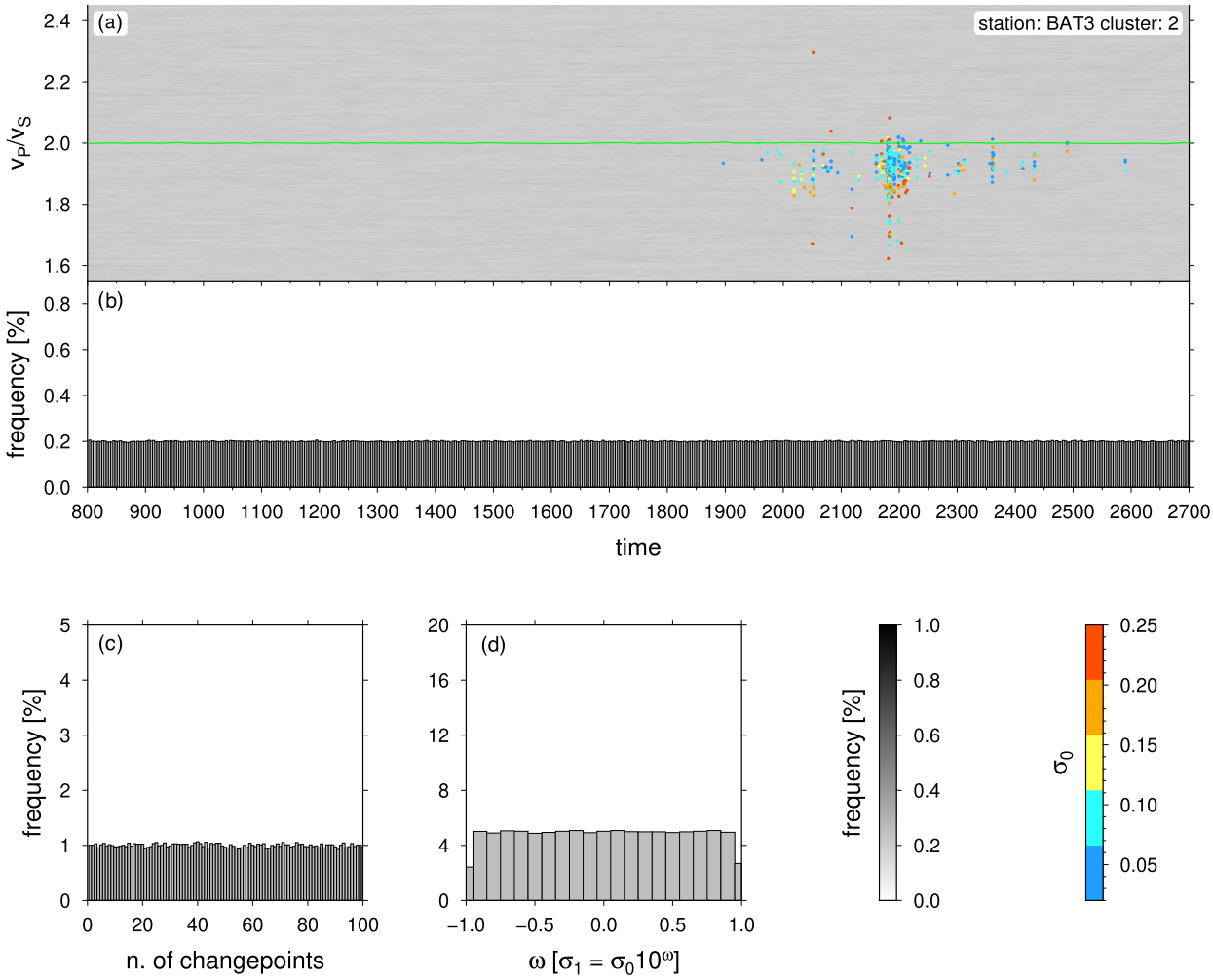

Figure 5. Results obtained sampling the prior probability distributions, for a given station-cluster couple. In the top right corner, the station and the cluster are indicated. (a) PPD for the V_P/V_S ratio. Grey shades indicate probability. A uniform grey shade is retrieved as expected for a uniform prior probability distribution. The green line indicates the mean value (i.e. the mid-point of the min/max interval for a uniform prior). Coloured dots represent data in the time-series, where colours refer to uncertainties. (b) PPD for the location of the change-points in time. As expected, a uniform distribution is retrieved. (c) PPD for the number of change-points. The algorithm uniformly sampled models with 1–100 change-points following the prior probability distribution. (d) PPD for the log value of the scale factor of the data uncertainties. A uniform distribution is also retrieved here, indicating that errors can be scaled from one-tenth to ten times.

In panel (c) we show a case where the automatic validation procedure fails. The change-point about $t = 2150$ is a false positive. It passed validation tests, but, after visual inspection we state that it is probably placed only because of a measurement far from average, with low error and few neighbouring data (especially on the right side).

Panel (d) is an example of a double V_P/V_S ratio variation (increasing at about $t = 2410$ d, and decreasing at about $t = 2450$ d) in which the increasing one is a false negative. This error in automatic selection is explained by the previous change-points (about $t = 2370$ and $t = 2380$ d) being discarded so that V_P/V_S ratio variation before and after $t = 2410$ d is not enough to pass the last criterium in Section 3.2.

4.3 Mapping change-points distribution

Each one of the 4200 node-station couples is analysed to check for change-point(s) presence with the procedure described in Section 3.2. The result is a list of times of occurrence of change-points (if any) for every node-station pair. Given the potential pitfalls illustrated in panels (c) and (d) of Fig. 7, in the following we analyse the results of the inversions as a whole, not focusing on a single node-station time-series.

First of all, we analyse the complete catalogue of change-points as a whole, considering the full investigated volume (e.g. Gritto & Jarpe 2014). Fig. 8 shows the distribution in time of all validated change-points if all station-cluster pairs are considered. Main earthquakes (magnitude ≥ 3.5) are marked with stars; distribution of 'moderate' events (magnitude ≥ 3.0) is shown as histograms. The most relevant observation is the presence of a clear maximum in the distribution of change-points at about 2200 d (early January 2014), where the distribution shows a relevant degree of positive skewness (i.e. a longer right-tail), with more than 50 per cent of all validated change-points occurring between 2150 and 2250 d (approximately November 2013 to February 2014). The distribution of both main and moderate earthquakes exhibit a maximum before 2200 d, broadly indicating that moderate and main events occur ahead of change-points.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

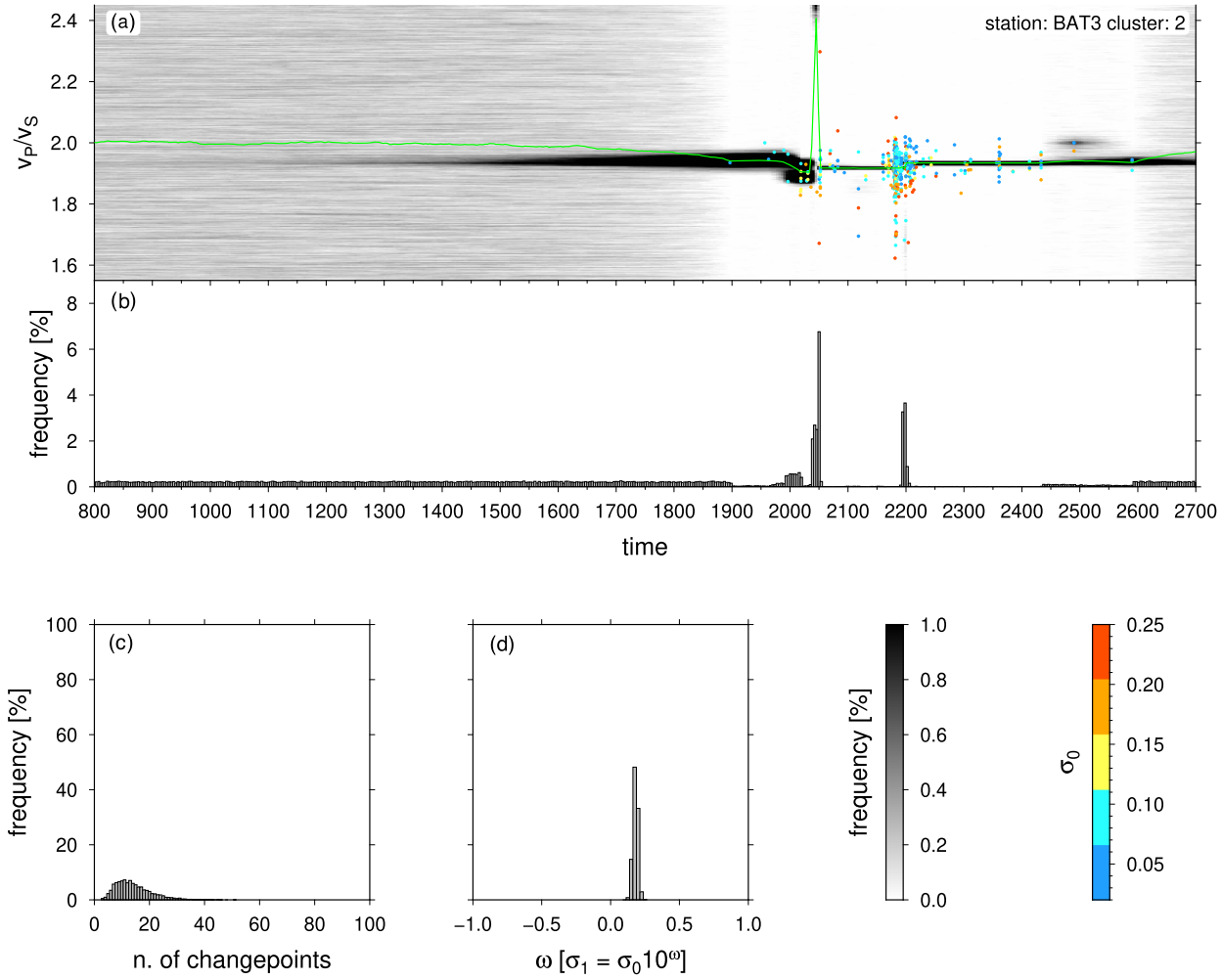

Figure 6. Results obtained sampling the posterior probability distributions, for the same station-cluster couple as in Fig. 5. Switching to posterior sampling, that is considering the data, uniform distributions are no longer obtained. Instead, distributions are shaped by the information contained in the data, so that local uniform distribution are still retrieved where data are uninformative or absent, see panel (b) between 800 and 1900 d. Details as in Fig. 5.

This feature is further investigated adding the spatial dimension to the temporal one: multiple maps of the study area have been produced at intervals of 25 d spanning 50 d, so to have snapshots of occurrence of change-points in different time windows from April 2010 and March 2015. This visual approach adds reliability to the study by overcoming the inherent limitations of analysing every time-series independently of the others. In Fig. 9 four of these time windows are shown: panels (b) (c) and (d) focus on the periods before, during and after the maximum change-points occurrence highlighted in Fig. 8; in panel (a) we show a map of a period, of considerable seismic activity (Pietralunga seismic sequence, occurred in 2010 Marzorati *et al.* 2014). All sequential maps are shown as a movie in the Supplemental On-line Materials.

The presence or absence of change-points is represented with lines of different colours: grey lines are used to connect all the analysed node-station couples (4200); black lines are used to mark those node-station couples that exhibit one (or more) change-point in that specific time window. This representation shows that not only change-points are concentrated in time (as seen also in Fig. 8), but they are also clustered in space (Fig. 9c). For an even more significant analysis we show Fig. 10, derived from Fig. 9: in this case the area is gridded and the black lines (change-point present) crossing each cell are counted and the number is normalized over the total number of rays (grey lines) crossing each cell and plotted as a contour map. The grid used as normalization is represented in Fig. 4. This normalization is used to produce a clearer picture of the areas where change-points are present by giving the same importance to highly sampled zones and poorly sampled ones: we can highlight the greater relevance of, for example a cell in which 10 out of 10 rays exhibit a change-point (100 per cent) rather than a cell where 10 out of 100 (10 per cent) exhibit a change-point. As in Fig. 8, main earthquakes are marked with stars if present in those time windows.

5 DISCUSSION

In this study, we investigated variations in V_P/V_S ratio by means of a Bayesian approach. The dense seismic network and the very high number of earthquakes recorded at the Alto Tiberina Near Fault Observatory (TABOO) seismic network, along the Northern Apennines of Italy, allowed us to analyse 4200 individual time-series related to seismic events occurred between March 2010 and May 2015.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

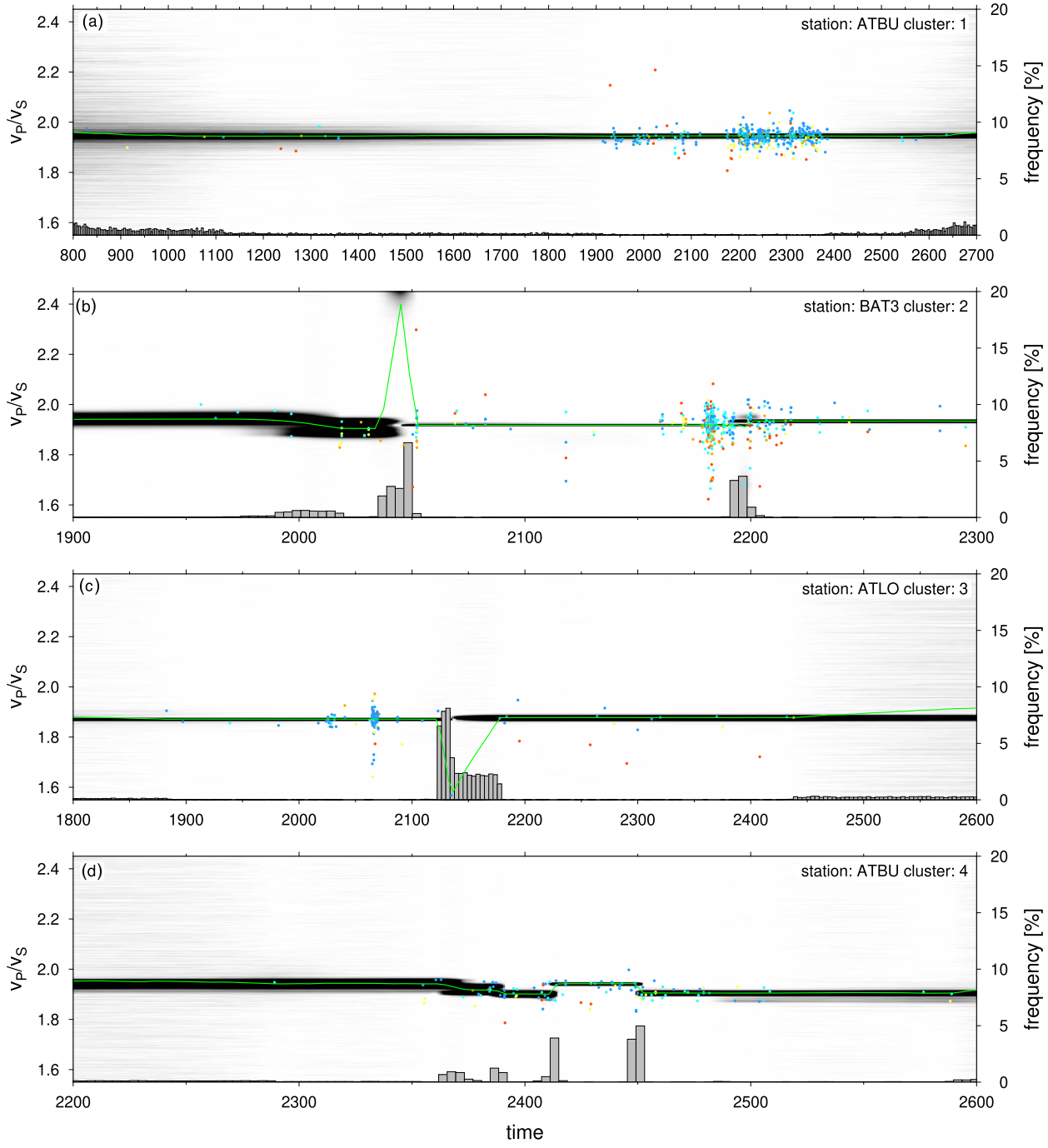

Figure 7. Selected examples of inversion results. In each panel, we summarize the results of the inversion overplotting the PPD of the V_P/V_S ratio and the PPD of the time-position of change-points, that is panels (a) and (b) in Fig. 6. (a) Example of 'no change-point' result. (b) Example of rejected and confirmed change-points. Change-points at about 2000, 2050 and 2200 d are, respectively: rejected (according to the first criterion described in Section 3.2), rejected (according to the third criterion described in Section 3.2), accepted. (c) Example of false positive change-point. (d) Example of false negative change-point at about 2420 d. See main text for more details on such cases and description of the criteria for change-point validation.

The goal is to identify temporal and spatial variations in V_P/V_S ratio starting from raw input data automatically generated by a modular procedure able to take in continuous seismic waveforms and output *P* and *S* waves arrival times, earthquakes locations and magnitude (Di Stefano *et al.* 2014). The data are considered raw because the time-series (V_P/V_S values versus time, for single event-station couple) are not filtered based on any criteria, to test the robustness of the method. In addition to this the choices we made in terms of likelihood function, outliers inclusion and automatic checks for change-point validation, limit the influence of subjective assumptions affecting change-points detection.

Our application confirms that the RJ-McMC algorithm is one of the most suitable approach to change-point problems. Among main advantages it is evident how this type of problem benefits from the

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

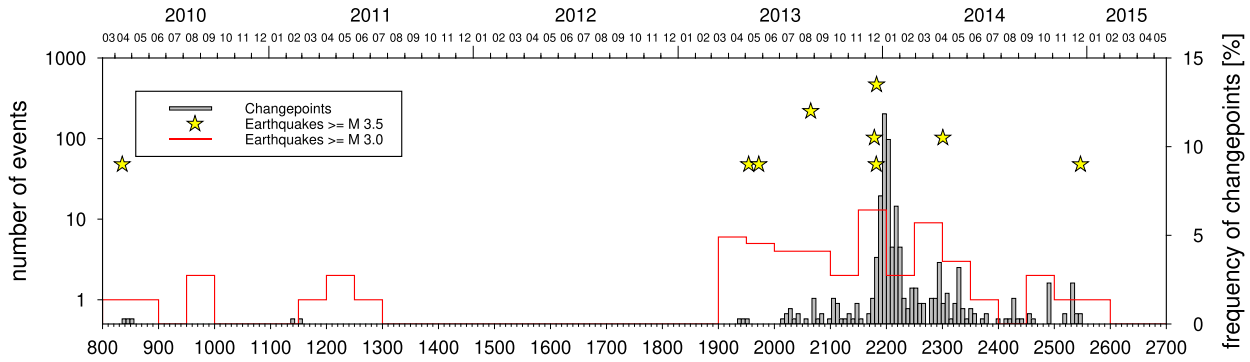

Figure 8. Change-points distribution in time (percent over total number of change-points confirmed). Each bin is 1 week in width. In red the distribution of $M_L \geq 3.0$ earthquakes (50 d bin width). Yellow stars mark main events ($M_L \geq 3.5$).

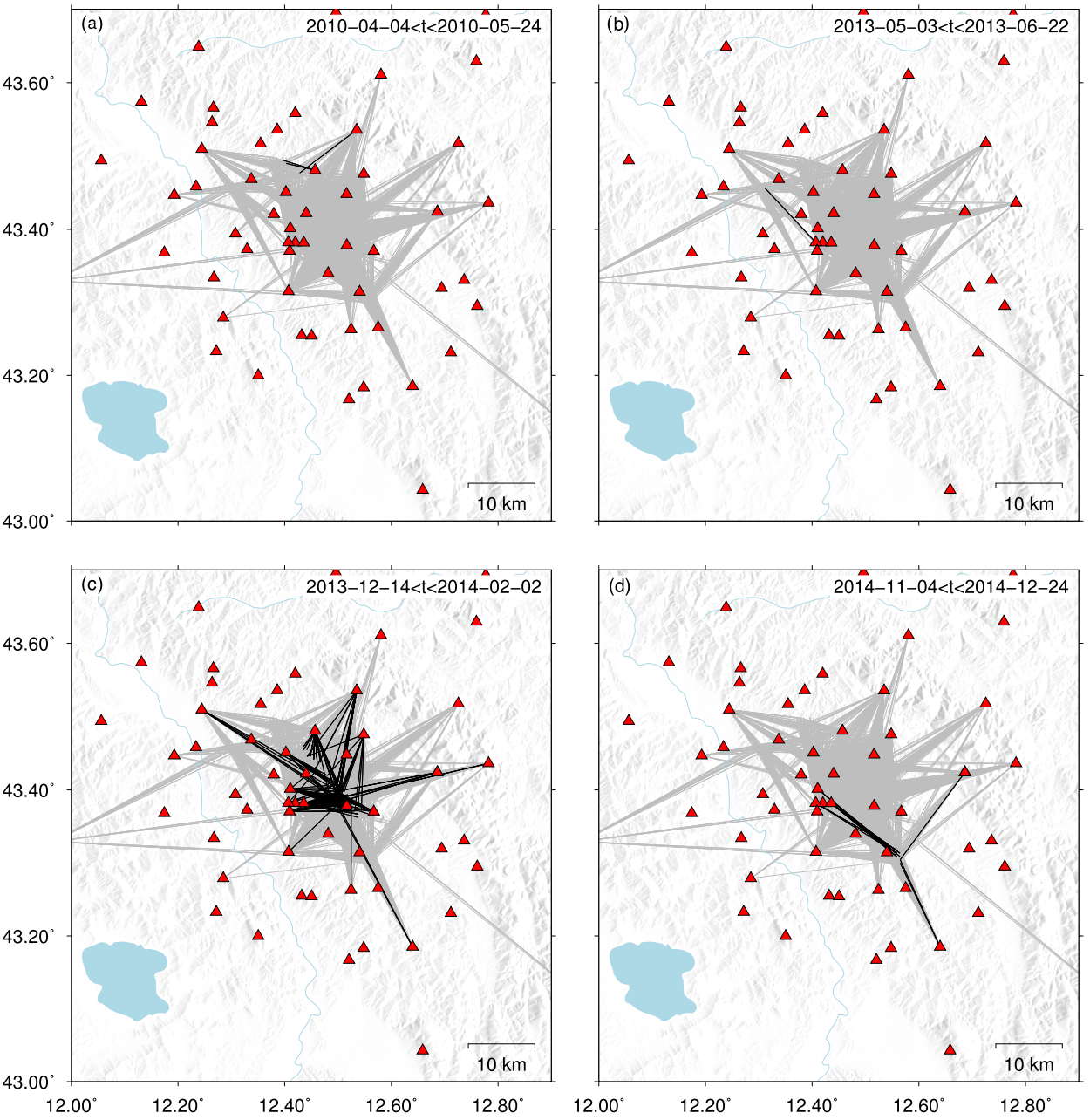

Figure 9. Node-station couples exhibiting a change-point (black lines) vs. total couples analysed (grey lines), for four different time-windows. When a change-point occurs in the time-window, the ray path between station and node is drawn with a black line. Otherwise, it is grey coloured.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

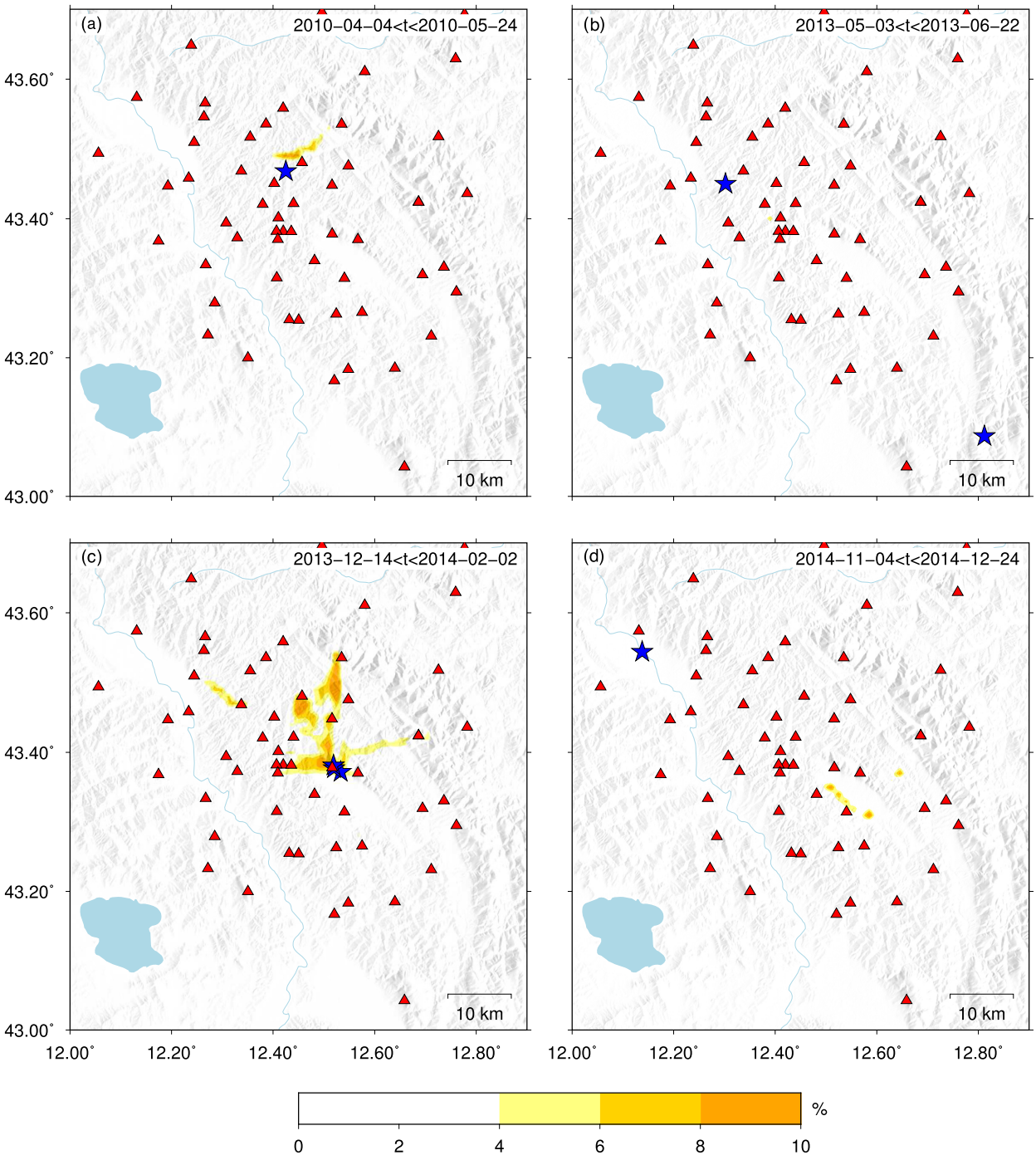

Figure 10. Count of ray paths with change-points normalized by total rays count, for four different time-windows. As in Fig. 9, but here the yellow-shade colours indicate the ratio between the number of ray paths with a change-point in the selected time-window, and the total number of ray paths.

parsimony of the algorithm (e.g. change-points are placed only if supported by the data), as showed in Fig. 6 where we have shown that the total number of change-points obtained by sampling the posterior probability distribution is very limited with respect to the maximum value (100) allowed by our setting (panel c in Fig. 6. The algorithm can also naturally deal with both no-data zones, as seen in Figs 6(a) and (b) where it tends to the prior value, and no-change-point cases, as seen in Fig. 7(a). This feature is essential in approaching very large data sets often discontinuous in time and density of observations.

In our study we find that the distribution in time of the change-points (Fig. 8), shows a robust and delayed correlation with moderate earthquakes occurrence (i.e. first a peak in frequency of moderate earthquakes and then a peak in frequency of changepoints), at the end of 2013 (around day 2000). This is the period of time when swarm like activity possibly triggered by a slow earthquake near Gubbio (Fig. 1 Gualandi *et al.* 2017). It is important to note that these seismic events occurred ahead of the change-points. In addition to this, we observe spatial directivity in the rock volumes

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

affected by the changes in the elastic properties, respect to the location of the causative events as showed in Figs 10(a), (b) and (c) (note that in Fig. 10b the area experiencing the variation is very minor). In case of fluids in the source-crustal volume, a possible interpretation for this could be related to the presence of some sort of barrier for fluids migration (Lees & Wu 2000). There are also example where the variation occurs remotely respect to the larger event or possibly without a correlation (Fig. 10d).

The velocity ratio has been used for many purposes (Nur 1972; Aggarwal *et al.* 1973; Chevrot & van der Hilst 2000; Ripepe *et al.* 2000; Lucente *et al.* 2010), such as a lithology indicator, determining degree of consolidation, identifying pore fluid, and predicting velocities. Variations in the V_P/V_S ratio are also commonly related to fluid content and their mobilization (Asada 1982), as well as to the fault-zone damage evolution (Thurber & Atre 1993; Vidale & Li 2003; Schaff & Beroza 2004). Along this sector of the Central-Northern Apennines, studies of the spatial-temporal occurrence of prolonged foreshocks and aftershock sequences have been linked to dilatancy diffusional models (Antonioli *et al.* 2005; Lucente *et al.* 2010) and fluid flow occurred in the fractured evaporites layer (Miller *et al.* 2004; Marzorati *et al.* 2014). This lithology would act like a reservoir (Trippetta *et al.* 2013) of high CO_2 pressure mobilized by a seismic event and propagating through damaged and fractured zones, leading the migration of the seismic activity usually along strike (Miller *et al.* 2004; Chiaraluce 2012). Thus, a variation in fluid content is a good candidate for the changes in the ratio we observed in TABOO area.

As mentioned earlier, it's important to emphasize that the process for defining change-points is weakly influenced by subjective decisions: Bayesian inversion controls change-points number and positioning in time and the mapping is statistically robust due to the high number of independent inversions. Nonetheless a successful application is determined by a careful choice of the modelling base function. In our case we have chosen a step function (as explained in Section 3.1.3) because we expect V_P/V_S values to maintain a certain value until a sudden change occurs and then to return to a new stable value (different from the previous one). This is strictly a data set specific choice: one can have other kinds of data sets could be considered, where oscillations or linear trends are expected (e.g. geochemical and GPS data). The algorithm is versatile and adaptable to the expected different situations, for example a sum of multiple base functions, otherwise the results will not be reliable. Two examples of this situation are presented in Fig. 7, where slow variation occurs around day 2000 in panel (b) and between days 2350 and 2400 in panel (d). With our step function modelling, the way the algorithm can deal with linear trends is by placing many change-points one close to another, which is clearly not optimal. Moreover the specific linear trend seen in Fig. 7 d is also the cause of the false negative as mentioned in Section 4.2.

There are a couple of aspects within the constructed work-flow that we may want to refine in the future, for increasing the degree of automatization of the procedure. First, the construction of the time-series has been done based on our knowledge of the TABOO data set. Both the dimension of the clusters and the minimum number of events in each cluster are strictly related to the characteristics of the seismicity pattern of the study area. A further step forward would be the introduction of a more general definition of V_P/V_S ratio time-series that could account for the ray paths (i.e. a 3-D time-series) to by-pass the definition of a cluster.

Second, the criteria for validating change-points can be further developed trying to add statistical analysis of the data-space or modifying the prior probability distribution including additional constraints. The McMC methods also give the opportunity to extract useful information with the analysis of the reconstructed full PPD, not limiting the results to a single model (best fit or maximum likelihood model). In our case the PPD is represented as 1-D marginal PPD for V_P/V_S to add a criterion in change-points validation (as explained in Section 3.2. In fact the shape of the 1-D marginal PPD for V_P/V_S is a robust indicator for the reliability of a time-variation because it contains information about the confidence on the retrieved values before and after the change-point. This fact is extremely valuable in cases, like ours, where the variations of the investigated quantity could be limited. For example, in our data set, variations in values are usually as small as 5 per cent, so an estimation of the confidence allows to rely on multiple criteria to validate change-points and not only on the mere variation in values.

The identification of transient changes occurring before, during and after the occurrence of relatively major earthquakes is definitively relevant. To the regard, the positive outcomes we gained with the proposed methodology show some potential for contributing to innovative and heterogeneous traffic lights systems monitoring tectonically active regions or industrial sites potentially inducing major episodes of seismic activity. Other kind of data sets (e.g. geochemical and geodetic data) can be integrated in the Bayesian inversion, to analyse synchronous variations on multiple observables. The availability of multidisciplinary time-series acquired by multisensor stations are the ideal data for the implementation of this new and advanced monitoring technique.

6 CONCLUSIONS

We develop a new algorithm for detecting change-points in V_P/V_S ratio time-series, based on Bayesian inference. Given the trans-dimensional nature of the algorithm, the existence and number of change-points is directly dictated by the data, and not imposed by the user. Our main findings are:

(i) the Rj-McMC algorithm has been successfully applied to V_P/V_S ratio time-series, highlighting a non random distribution of the located change-points, both in space and time;

(ii) these distributions show a correlation with the occurrence of 'moderate' earthquakes ($M_L \geq 3.0$);

(iii) the proposed procedure has potential in application to monitoring tectonically active regions, and can be used as a modular tool to be integrated with other observables.

ACKNOWLEDGEMENTS

NPA thanks the Geophysics Section of the Dublin Institute for Advanced Studies for technical and administrative support. This research has been conducted with the financial support of Science Foundation Ireland under Grant Number 14/IFB/2742. NPA research is funded by Austrian Science Foundation (FWF), project number M2218-N29. Figures are plotted using GMT (Wessel & Smith 1998).

The authors want to thank the Associate Editor, Anya Reading and the Anonymous Reviewer for the useful comments and precise corrections.

We also thank the personnel maintaining the Altotiberina Near Fault Observatory (TABOO), providing high quality raw data and high level scientific products ready to be analysed.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

REFERENCES

Aggarwal, Y.P., Sykes, L.R., Armbruster, J. & Sbar, M.L., 1973. Premonitory changes in seismic velocities and prediction of earthquakes, *Nature,* **241,** 101–104.

Antonioli, A., Piccinini, D., Chiaraluce, L. & Cocco, M., 2005. Fluid flow and seismicity pattern: evidence from the 1997 Umbria-Marche (Central Italy) seismic sequence, *Geophys. Res. Lett.,* **32**(10), doi:10.1029/2004GL022256.

Asada, T., 1982. *Earthquake Prediction Techniques: Their Application in Japan,* University of Tokyo Press.

Backus, G.E., 1988. Bayesian inference in geomagnetism, *J. geophys.,* **92**(1), 125–142.

.Bayes, T, 1763, LII. An essay toward solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S, *Philosophical transactions of the Royal Society of London,* **53,** 370–418.

Bodin, T., Salmon, M., Kennett, B. L.N. & Sambridge, M., 2012. Probabilistic surface reconstruction from multiple data sets: an example for the Australian Moho, *J. geophys. Res. Solid Earth,* **117**(B10), doi:10.1029/2012JB009547.

Chevrot, S. & van der Hilst, R.D., 2000. The poisson ratio of the australian crust: geological and geophysical implications, *Earth planet. Sci. Lett.,* **183**(1-2), 121–132.

Chiaraluce, L., 2012. Unravelling the complexity of apenninic extensional fault systems: a review of the 2009 L'Aquila earthquake (Central Apennines, Italy), *J. Struct. Geol.,* **42,** 2–18.

Chiaraluce, L. *et al.*, 2014. The Alto Tiberina Near Fault Observatory (northern Apennines, Italy), *Ann. Geophys.,* **57**(3), doi:10.4401/ag-6426.

Chiodini, G., Cardellini, C., Amato, A., Boschi, E., Caliro, S., Frondini, F. & Ventura, G., 2004. Carbon dioxide Earth degassing and seismogenesis in central and southern Italy: carbon dioxide earth degassing and seismogenesis, *Geophys. Res. Lett.,* **31**(7), doi:10.1029/2004GL019480.

Denison, D. G.T., Adams, N.M., Holmes, C.C. & Hand, D.J., 2002. Bayesian partition modelling, *Comput. Stat. Data Anal.,* **38**(4), 475–485.

Dettmer, J. & Dosso, S.E., 2012. Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains, *J. acoust. Soc. Am.,* **132**(4), 2239–2250.

Di Stefano, R., Chiaraluce, L., Valoroso, L., Waldhauser, F., Latorre, D., Piccinini, D. & Tinti, E., 2014. An automatic modular procedure to generate high-resolution earthquake catalogues: application to the Alto Tiberina Near Fault Observatory (TABOO), Italy., *AGU Fall Meeting Abstracts,* **13,** T13C–4684.

Gallagher, K., 2012. Transdimensional inverse thermal history modeling for quantitative thermochronology, *J. geophys. Res.,* **117**(B2), doi:10.1029/2011JB008825.

Gallagher, K., Bodin, T., Sambridge, M., Weiss, D., Kylander, M. & Large, D., 2011. Inference of abrupt changes in noisy geochemical records using transdimensional changepoint models, *Earth planet. Sci. Lett.,* **311**(1-2), 182–194.

Gao, C. & Lekić, V., 2018. Consequences of parametrization choices in surface wave inversion: insights from transdimensional bayesian methods, *Geophys. J. Int.,* **215**(2), 1037–1063.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika,* **82**(4), 711–732.

Gritto, R. & Jarpe, S.P., 2014. Temporal variations of vp/vs-ratio at the geysers geothermal field, usa, *Geothermics,* **52,** 112–119.

Gualandi, A., Nichele, C., Serpelloni, E., Chiaraluce, L., Anderlini, L., Latorre, D., Belardinelli, M.E. & Avouac, J.P., 2017. Aseismic deformation associated with an earthquake swarm in the northern Apennines (Italy), *Geophys. Res. Lett.,* **44,** doi:10.1002/2017GL073687.

Hopcroft, P.O., Gallagher, K. & Pain, C.C., 2007. Inference of past climate from borehole temperature data using Bayesian Reversible Jump Markov chain Monte Carlo: Bayesian inversion of borehole data for past climate, *Geophys. J. Int.,* **171**(3), 1430–1439.

Hopcroft, P.O., Gallagher, K. & Pain, C.C., 2009. A Bayesian partition modelling approach to resolve spatial variability in climate records from borehole temperature inversion, *Geophys. J. Int.,* **178**(2), 651–666.

Lees, J.M. & Wu, H., 2000. Poisson's ratio and porosity at coso geothermal area, california, *J. Volc. Geotherm. Res.,* **95**(1-4), 157–173.

Li, Y.-G., Vidale, J.E., Aki, K., Xu, F. & Burdette, T., 1998. Evidence of Shallow Fault Zone Strengthening After the 1992 M7.5 Landers, California, Earthquake, *Science,* **279**(5348), 217–219.

Lucente, F.P., De Gori, P., Margheriti, L., Piccinini, D., Di Bona, M., Chiarabba, C. & Piana Agostinetti, N., 2010. Temporal variation of seismic velocity and anisotropy before the 2009 M_w 6.3 L'Aquila earthquake, Italy, *Geology,* **38**(11), 1015–1018.

Malinverno, A., 2002. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophys. J. Int.,* **151**(3), 675–688.

Malinverno, A. & Briggs, V.A., 2004. Expanded uncertainty quantification in inverse problems: hierarchical Bayes and empirical Bayes, *Geophysics,* **69**(4), 1005–1016.

Mandolesi, E., Ogaya, X., Campanyá, J. & Piana Agostinetti, N., 2018. A reversible-jump Markov Chain Monte Carlo algorithm for 1D inversion of magnetotelluric data, *Comput. Geosci.,* **113,** 94–105.

Marzorati, S., Massa, M., Cattaneo, M., Monachesi, G. & Frapiccini, M., 2014. Very detailed seismic pattern and migration inferred from the April 2010 Pietralunga (northern Italian Apennines) micro-earthquake sequence, *Tectonophysics,* **610,** 91–109.

Miller, S.A., Collettini, C., Chiaraluce, L., Cocco, M., Barchi, M. & Kaus, B.J.P., 2004. Aftershocks driven by a high-pressure CO_2 source at depth, *Nature,* **427**(6976), 724–727.

Mosegaard, K., 2006, Monte Carlo analysis of inverse problem, *PhD thesis*, Copenhagen University.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.: Solid Earth,* **100**(B7), 12431–12447.

Nur, A., 1972. Dilatancy, pore fluids, and premonitory variations of ts/tp travel times, *Bull. seism. Soc. Am.,* **62**(5), 1217–1222.

Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.*

Piana Agostinetti, N., Bianchi, I., Amato, A. & Chiarabba, C., 2011. Fluid migration in continental subduction: the Northern Apennines case study, *Earth Pla. Sci. Lett,* **302**(3-4), doi:10.1016/j.epsl.2010.10.039.

Piana Agostinetti, N., Giacomuzzi, G. & Malinverno, A., 2015. Local 3D earthquake tomography by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.,* **201,** 1598–1617.

Piana Agostinetti, N., Giacomuzzi, G. & Chiarabba, C., 2017. Seismic swarms and diffuse fracturing within triassic evaporites fed by deep degassing along the low-angle alto tiberina normal fault (Central Apennines, Italy), *J. geophys. Res.: Solid Earth,* **122**(1), 308–331.

Ripepe, M., Piccinini, D. & Chiaraluce, L., 2000. Foreshock sequence of september 26th, 1997 Umbria-Marche earthquakes, *J. Seismol.,* **4**(4), 387–399.

Rosenthal, J.S., 2011. Optimal proposal distributions and adaptive MCMC, *Handbook of Markov Chain Monte Carlo,* Chapman and Hall/CRC .

Roy, C. & Romanowicz, B.A., 2017. On the implications of a priori constraints in transdimensional bayesian inversion for continental lithospheric layering, *J. geophys. Res.: Solid Earth,* **122**(12), 10 118–10 131.

Sambridge, M., Gallagher, K., Jackson, A. & Rickwood, P., 2006. Transdimensional inverse problems, model comparison and the evidence, *Geophys. J. Int.,* **167**(2), 528–542.

Schaff, D.P. & Beroza, G.C., 2004. Coseismic and postseismic velocity changes measured by repeating earthquakes, *J. geophys. Res.: Solid Earth,* **109**(B10), doi:10.1029/2004JB003011.

Tarantola, A., 2005. , *Inverse Problem Theory And Methods For Model Parameter Estimation,* Society for Industrial and Applied Mathematics, Philadelphia, PA.

Thurber, C.H. & Atre, S.R., 1993. Three-dimensional vp/vs variations along the loma prieta rupture zone, *Bull. seism. Soc. Am.,* **83**(3), 717–736.

Trippetta, F., Collettini, C., Barchi, M.R., Lupattelli, A. & Mirabella, F., 2013. A multidisciplinary study of a natural example of a CO_2 geological reservoir in central Italy, *Int. J. Greenhouse Gas Contl.,* **12,** 72–83.

Valoroso, L., Chiaraluce, L., Di Stefano, R. & Monachesi, G., 2017. Mixed mode slip behavior of the altotiberina low angle normal fault system

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019

(Northern Apennines, Italy) through high resolution earthquake locations and repeating events, *J. geophys. Res.: Solid Earth,* **122**(12), 10 220–10 240.

Vidale, J.E. & Li, Y.-G., 2003. Damage to the shallow landers fault from the nearby hector mine earthquake, *Nature,* **421**(6922), 524–-526.

Wadati, K. & Oki, S., 1933. On the travel time of earthquake waves. (Part II), *J. Meteorol. Soc. Jpn. Ser. II,* **11**(1), 14–28.

Wessel , P. & Smith, , W. 1998. New, improved version of Generic Mapping Tool released, *Eos, Transactions American Geophisical Union,* **79,** 579, doi:10.1029/98EO00426 .

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

time.mp4

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

Downloaded from https://academic.oup.com/gji/article-abstract/217/2/1217/5308832 by INGV user on 17 April 2019