

---

# Deliverable

---

## D3.1 New Perspectives in OEF Models through the analysis of candidate precursors

Report information	
<b>Work package</b>	WP3: Advance: Advancing operational earthquake forecasting and earthquake predictability
<b>Sub-package</b>	WP3.1
<b>Authors</b>	C. Voisin, CNRS, UGA
<b>Reviewers</b>	W. Marzocchi (UniNa), M. Herrmann (UniNa)
<b>Approval</b>	Management Board
<b>Status</b>	Final
<b>Delivery deadline</b>	28.02.2022
<b>Submission date</b>	28.02.2022
<b>Intranet path</b>	[DOCUMENTS/DELIVERABLES/File Name]



---

## Table of contents

<b>1.</b>	<b>Scope of this report</b>	<b>3</b>
<b>2.</b>	<b>A rapid overview of past and recent candidate precursors</b>	<b>4</b>
2.1	List of candidate precursors to impending earthquakes	4
2.2	New perspectives offered by enhanced observational capabilities	5
<b>3.</b>	<b>Hydrogeochemical monitoring</b>	<b>6</b>
3.1	The IRON network	6
3.2	Recent results for the Amatrice-Visso-Norcia sequence	8
3.3	Other geochemical observations	11
<b>4.</b>	<b>Continuous seismic monitoring in Italy</b>	<b>14</b>
4.1	Observations based on seismic activity	14
4.2	Observations of crustal velocity changes based on ambient seismic noise interferometry	14
	15	
4.3	Observations of crustal attenuation changes based on ambient seismic noise coherence	16
<b>5.</b>	<b>Perspectives in OEF enhancement with new candidate precursors</b>	<b>25</b>
5.1	Towards a comprehensive physical model of the earthquake preparation	25
5.2	A possible path towards the improvement of operational earthquake forecasting	27
5.3	On the use of machine learning to improve the forecasting	27
<b>6.</b>	<b>Conclusions</b>	<b>28</b>

---

## Summary

Advancing Operational Earthquake Forecasting (OEF) as a key element of dynamic risk assessment is achieved through a range of coordinated activities. Using community accepted retrospective and fully prospective testing as tools for performance evaluation (-->WP7), WP3 will have a measurable impact on advancing the state-of-the-art and state of practical OEF implementation in Europe and worldwide. Focus areas are:

- Improving process understanding: By conducting targeted experiments, we will contribute to advancing high-quality earthquake predictability research. This includes a multi-parameter search for precursory signals and operationalising ambient noise time-series analysis.
- Transfer knowledge from other disciplines to OEF, such as rock-deformation labs, underground labs, induced seismic sequences, and adopt novel statistical methods from ecology that combines geological, tectonic and seismic data for developing innovative spatio-temporal triggering models, with full quantification of uncertainty in a Bayesian framework (T7.1).
- Develop next generation of physics-based earthquake forecasting models and techniques (T7.2); this includes models mostly based on continuum mechanics and on statistical physics (e.g., network theory) which may benefit from the availability of high-quality seismic data (T2.4).
- Develop next generation of stochastic and hybrid earthquake forecasting models (T7.2); improve description of space- time variability in the frequency-magnitude distribution and earthquake clustering properties, exploiting advances in observational capabilities.
- Develop workflows to formally integrate expert-based OEF assessments into dynamic risk assessment.

In this deliverable we describe the work made in Task 3.1 where we explore the existence of possible seismic and non-seismic precursors which may contribute to increase the current earthquake forecasting capabilities. The largest and most innovative part of the deliverable (section 4.3) is related to a draft of one paper that will be submitted soon and it describes a novel set of anomalies that may precede large earthquakes.

---

## 1. Scope of this report

This document is part of the deliverables of the H2020 European project RISE – Real Time Earthquake-risk reduction for a Resilient Europe. It is intended to serve as a Public Report on the new perspectives in Operational Earthquake Forecasting through the analysis of candidate precursors.

The purpose of this report is to describe earthquake precursor candidates in detail, and to present their relevance to improving our societies' ability to predict the occurrence of strong earthquakes. Before going any further on this subject, it seems particularly useful to clarify what is understood in this report by a precursor candidate to an earthquake. By precursor we mean any geophysical and/or geochemical anomaly that could be undoubtedly linked to the preparation of a major, potentially destructive earthquake. The search for precursory phenomena to earthquakes dates back to the beginning of seismology. Despite many a posteriori observations and description of

precursors of different origins (geochemistry, geodesy, seismology, electromagnetic, ...), the possibility to recognize these punctual anomalies before the impending earthquakes has remained out of scope. The last decades have seen a worldwide effort to extend the observational capabilities, through the deployment of dense recording networks and the intensification of the measurements. These new observational capabilities open the possibility to continuously monitor the Earth's crust.

Although the scope of the RISE project is Europe-wide, the main part of this report focuses on Italy, for two reasons. The first is that this country is regularly hit by violent earthquakes that are either unique (L'Aquila in 2009, Emilia in 2012) or organized in sequences (Colfiorito in 1997, Amatrice, Visso and Norcia in 2016). The second reason, obviously related to these numerous earthquakes, is that Italy has equipped itself with very dense monitoring networks, including a seismological network and a network of Radon measurement stations. These conditions make this country, and more particularly the central Apennines region, a natural laboratory where the search for precursors is easier and allows the detection of continuous phenomena, which could help forecasting the imminence of a large seismic event. Background seismicity is sustained, allowing the implementation, integration and possibility to test the relevance of old and new precursor candidates within a time-varying risk analysis approach both in time and space: the objective of WP3 of the RISE project.

---

## **2. A rapid overview of past and recent candidate precursors**

### **2.1 List of candidate precursors to impending earthquakes**

The development of the plate tectonic theory joined with the developments of crack theory (among others) created a general framework for the understanding of earthquake physics and their potential precursors (Rikitake, 1968). The precursors to earthquakes are classified in space-based (remote sensing) and ground-based. In this report, we will restrict ourselves to only some of the ground based precursor candidates. A recent study has listed the claimed ground based precursors as follows (Conti et al., 2021):

- 1) Seismicity and seismicity changes are the most extensively studied phenomenology before, during and after earthquakes (Mignan 2008; Hong et al., 2018; De Santis et al., 2019b). Extreme low frequency acoustic emissions have been claimed (see for example Ihmlé and Jordan, 1994) to constitute earthquake precursors before the main rupture at a higher frequency.
- 2) Lithospheric mechanical deformations, such as those detectable with creep- and stain-meters (Niu et al., 2008; Langbein, 2015).
- 3) Variation of the groundwater level and composition, reported some weeks up to few hours before earthquakes (Hayakawa et al., 1997; Koizumi et al., 1999).
- 4) Gas exhalations, mainly (but not only) of radon or radioactive ions induced by gas-water release from earthquake preparation zone into the atmosphere (Khilyuk et al., 2000; Pulinets et al., 2003).

- 5) Fluctuations of temperature observed in temporal correlation with some earthquakes and possibly reconciled with variation of groundwater circulation and uplift or more recently with vapour condensation on surface (Tramutoli et al., 2005).
- 6) Propagation of acoustic gravity waves (AGW) (Molchanov and Hayakawa, 2008), A physical mechanism of seismo-ionospheric coupling including both AGW and radon exhalation has been recently suggested (Rapoport et al., 2020).
- 7) Fluctuation of electric and magnetic field components in a large range of frequencies [from ULF (Uyeda, et al., 2009b; Han et al. (2014))] to VHF (Sorokin et al., 2020). Many observations have been reported on ground and in space of (direct, induced and secondary) electromagnetic emissions localized on the earthquake area or measured along the related field line or spread around it.
- 8) Ground based observations of ionospheric parameters [such as Total Electron Content (TEC) (Liu et al., 2004; Liu, 2009), VLF reflection height (Hayakawa et al., 1996, 1997; Rodger et al., 1999), whistler dispersion (e.g. Hayakawa et al., 1993), critical frequency foF2 (Hobara and Parrot, 2005), etc.].

We invite the reader interested into all this candidate precursors to turn towards the existing literature on this topic (e.g. Conti et al., 2021). In the following, we will address only hydrogeochemical changes (points 3 and 4 above) prior to the Italian earthquakes and, as described below, the new candidate precursors coming from the analysis of the continuous seismic noise.

## **2.2 New perspectives offered by enhanced observational capabilities**

Ambient noise seismology is a recent field that started in the early 2000s with the observation by the physicists that the correlation of a randomly fluctuating field measured at two points would lead to the retrieval of the Green's function between them (Weaver and Lobkis, 2001). In seismology, the great effort made to improve the quality of the seismological station networks as well as the data infrastructures and their distribution makes it possible to use these techniques to exploit what was previously called seismic noise. The first application was to perform a seismic noise tomography of the western US area (Shapiro et al. 2004). But very rapidly the idea to use the seismic noise for a continuous monitoring of the crust rose in the community. Originally introduced by the GFZ (German Research Centre for Geosciences in Potsdam) for the monitoring of the Merapi volcano (Sens-Schönfelder & Wegler, 2006), the technique of measuring velocity changes rapidly spread over all tectonic environments, including landslides (Mainsant et al., 2012; Voisin et al., 2016) and faults (Brenguier et al., 2008; Wegler et al., 2009). A large body of literature exists now, and it would be fastidious to explore here all the subtleties of the various techniques aimed at enhancing the quality and the stability of the measurement with respect to the variations of the seismic sources of noise both in time and space. For the sake of clarity, the reader has to keep in mind that these techniques are very precise but that they are subjected to environmental parameters (Larose et al., 2015) such as temperature changes of the near surface, or groundwater changes (Garambois et al., 2019; Voisin et al., 2016).

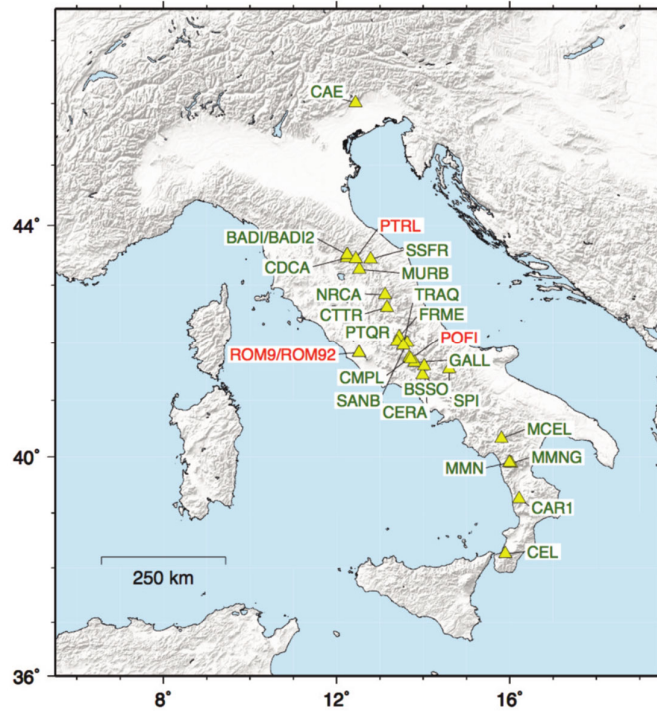
---

### 3. Hydrogeochemical monitoring

Terrestrial gases in groundwater and soil air have been extensively studied over the years in seismically active areas, especially in ex-USSR, China, Japan, and United States, in search of premonitory changes that might be useful for earthquake prediction. Concentrations of radon, helium, hydrogen, mercury, carbon dioxide, and several other volatiles have generally been found to be anomalously high along active faults, suggesting that the faults may be paths of least resistance for the terrestrial gases generated or stored in the earth to escape to the atmosphere. Temporal gas concentration changes with durations of a few hours to several months have been observed before many large earthquakes at a relatively small number of favorably situated stations at epicentral distances of up to several hundreds of kilometers. These “sensitive” stations are generally located along active faults, especially at intersections or bends of faults, or some other structurally weak zones, possibly because of tectonic strain concentration at such places. However, gas concentrations, especially those measured at or near the ground surface, may also be significantly affected by meteorological, hydrological, and other nontectonic changes in the environment (King, 1986).

#### 3.1 The IRON network

The Italian Radon mOnitoring Network (IRON) is a nationwide permanent network for near-real-time measurements of soil radon emissions in Italy. In 2018, IRON consists of 26 stations mainly concentrated in the Central-Southern Apennines, but marginally covering the whole Italian peninsula. At present, most IRON stations have recorded radon concentration timeseries for more than 7-8 years. The network exhibits a large heterogeneity in terms of sites, installations types and collected time series. The amount of data, together with the systematic methods of measurements, allowed us to evaluate some significant aspects related both to the measurement methodology and to the complex dynamics of soil radon emanations.



**Figure 1 The Italian radon monitoring map (in 2018)**

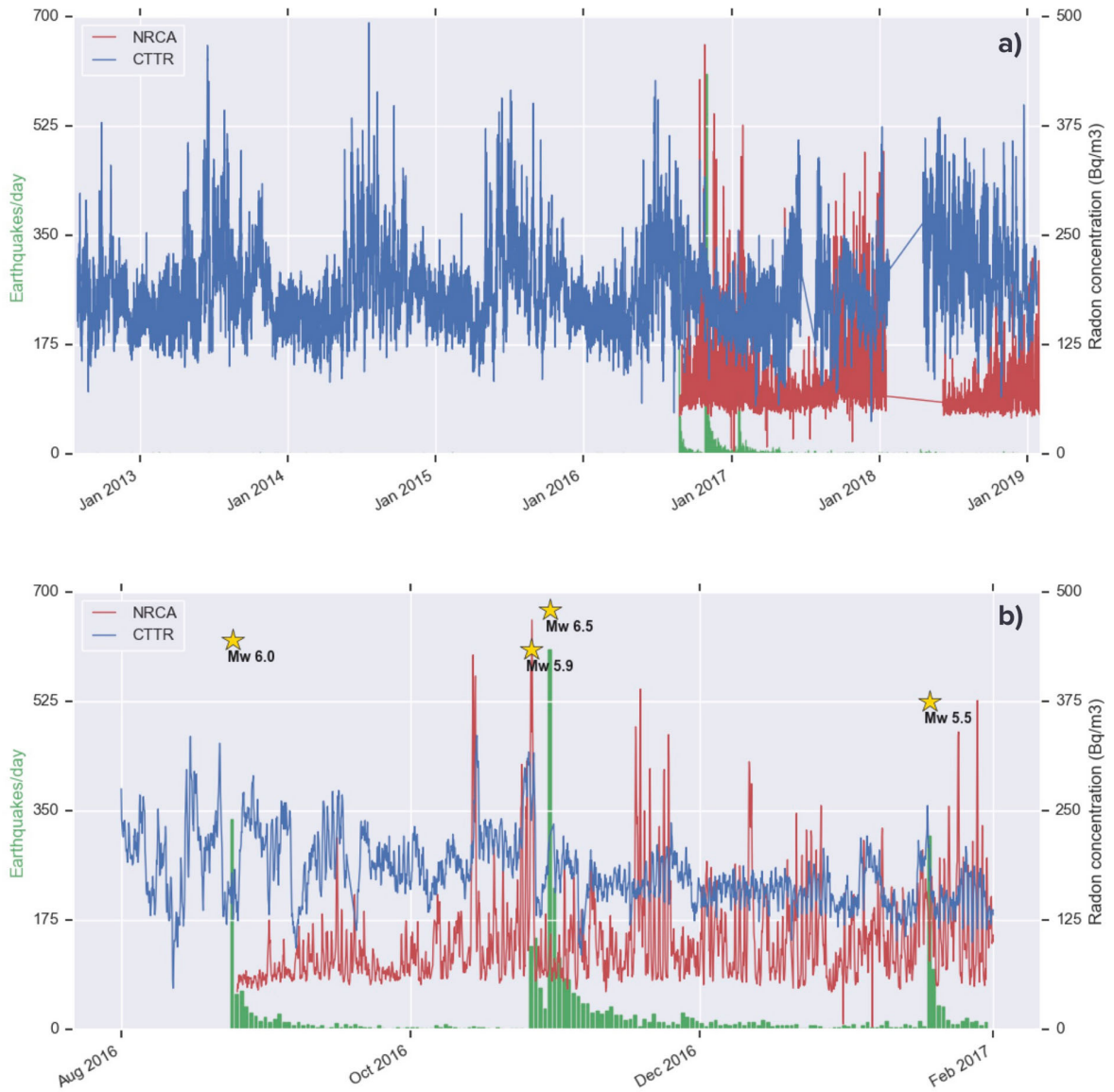
The aim of IRON is designing and developing a network of permanent stations for the continuous monitoring of radon concentration in seismically active areas, in order to explore the possibility of a physical link between seismogenic processes and temporal variability in radon emissions. The different observational setups have a direct impact on the features of the recorded signal, and observed fluctuations in radon concentration may be ascribed to geophysical processes taking place at depth in the crust. An essential requirement for a radon monitoring network to be an effective research tool in seismotectonics is the ability to provide long-term records of data. The main limitation when trying to associate radon measurements with Earth's internal processes is the difficulty of ruling out the possibility that observed variations are due to environmental effects, such as meteorological parameters, and/or random noise because seismotectonically and meteorologically induced radon anomalies may look strikingly similar. Meteorological parameters (temperature, pressure and precipitation variations) generally play an important role in modulating soil radon emanations. The relative importance among the main relevant variables (temperature, precipitation, pressure) in modulating the radon emissions cannot be uniquely determined and it is likely to be site dependent, since different analyses led to different results. An empirical correction procedure to take into account (i.e. remove, or at least reduce) the effect of meteorological parameters on the radon measured concentration has been developed (Piersanti et al., 2016). This approach goes far beyond the simpler and widely attempted "one to one" association between a single radon anomaly detected by a single station and a specific earthquake (which is intrinsically less robust). Additionally, this strategy offers the possibility to study the behaviour of radon emanation evolution under controlled conditions, since standardized shelter and borehole installation types grant steady and predictable features in the recorded time series. Several IRON stations are co-located with INGV-INSN stations that contain seismometers, accelerometers, and often GPS

instrumentation. This ideally guarantees a multi-parametric, multi-site, and multi-level long term measurement approach, as will be discussed later in the report.

### **3.2 Recent results for the Amatrice-Visso-Norcia sequence**

The radioactive nature of radon makes it a powerful tracer for fluid movements in the crust, and a potentially effective marker to study processes connected with earthquakes preparatory phase. To explore the feasibility of using soil radon variations as earthquakes precursor, the radon concentration data recorded by two stations located close to the epicentre of the strongest mainshock (Mw 6.5 on October 30, 2016) of the seismic sequence which affected central Italy from August 2016 are analysed (Soldati et al., 2020). Two stations (CTTR and NRCA, Figure 1) operate in the framework of the permanent Italian Radon monitoring Network almost continuously since 2012 and 2016, respectively, the latter being installed just after the first mainshock of the sequence (Mw 6.0 on August 24, 2016). An increase of radon emanation is clearly visible about 2 weeks before the Mw 6.5 event on both the time series, more pronounced on NRCA, nearer to the epicentre, suggesting the possibility of a direct association with the earthquake occurrence (Figure 2).





**Figure 2 (a)** Time series of soil radon concentration (Bq m<sup>-3</sup>) from NRCA (red) and CTTR (blue) stations. While CTTR registered almost continuously since August 2012, NRCA was installed the day right after the occurrence of the Amatrice Mw 6.0 mainshock. **(b)** Zoom over the period corresponding to the major seismic activity of the central Italy sequence of 2016–2017. Green bars indicate the number of earthquakes per day occurred within a radius of 40 km from the two stations, with yellow stars in correspondence of the four Mw ≥ 5.5 mainshocks. CTTR shows a clear annual variability linked to environmental processes. The zoom in (b) shows also for CTTR and NRCA a large number of peaks in [Rn], which makes it difficult to link any one of them to the occurrence of an earthquake. After (Soldati et al., 2020)

Figure 2 shows the difficulty to link any one of the peaks in [Rn] to the occurrence of a large earthquake: Even though some of the peaks may be recognized a posteriori as precursors, the time lag between the peak and the mainshock is difficult to assess. Furthermore, in an attempt to efficiently forecast the occurrence of a future mainshock, the peaks in [Rn] have to be correctly identified and their origin clarified. An independently developed detection algorithm (Piersanti et al., 2016) aimed at highlighting the connections between radon emission variations and major earthquakes occurrence succeeds in forecasting the Mw 6.5 mainshock on NRCA time series (Figure 3).

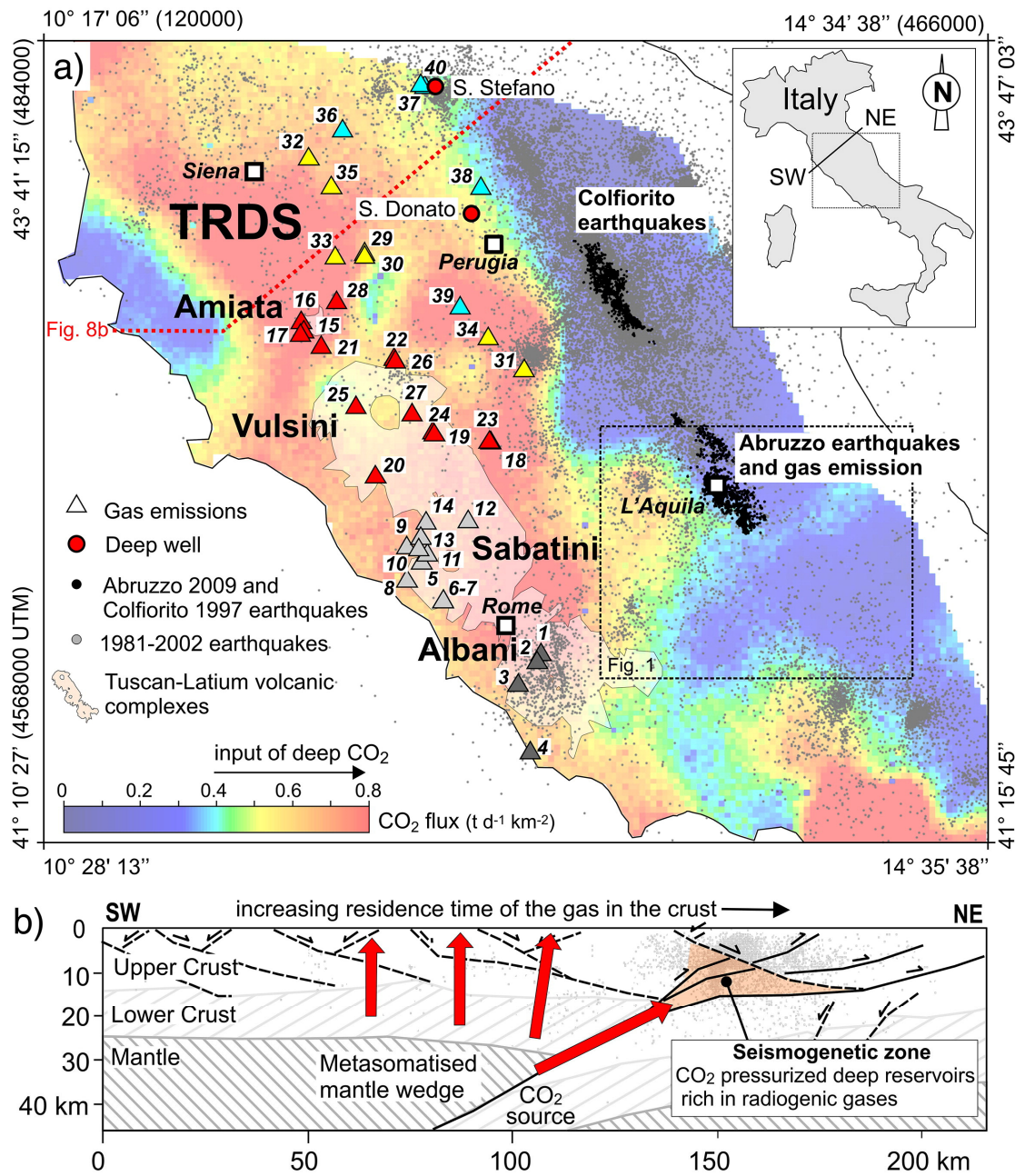


**Figure 3** Alarms issued by the detection algorithm applied to the time series of radon concentration of NRCA **(a)** and CTTR **(b)**, from the beginning of the central Italy seismic sequence. Alarms are considered true (red triangles) if followed within 40 days by a seismic event (yellow stars) with magnitude  $M_w \geq 4.0$ , otherwise they are regarded as false (blue triangles). After Soldati et al., 2020.

The resulting time advance of the alarm is consistent with that obtained using a Bayesian approach to compute the a posteriori probability of multiple change points on the radon time series of NRCA. Moreover, it is in agreement with the delay time that maximizes the correlation between radon and seismic anomalies. Applying the detection algorithm to CTTR time series returns alarms for both the Mw 6.0 event, with epicentre closer to this station, and the stronger Mw 6.5 event, but with a higher number of false detections. Finally, we found that a preliminary correction of the bias introduced by variations of meteorological parameters does not affect our main finding of an increase in radon concentration before the major mainshocks. The study of Soldati et al., 2020 confirms that, although much work is still needed especially about the false detections, a monitoring approach based on a permanent dense network is crucial for making radon time series analysis an effective complement to traditional seismological tools.

### **3.3 Other geochemical observations**

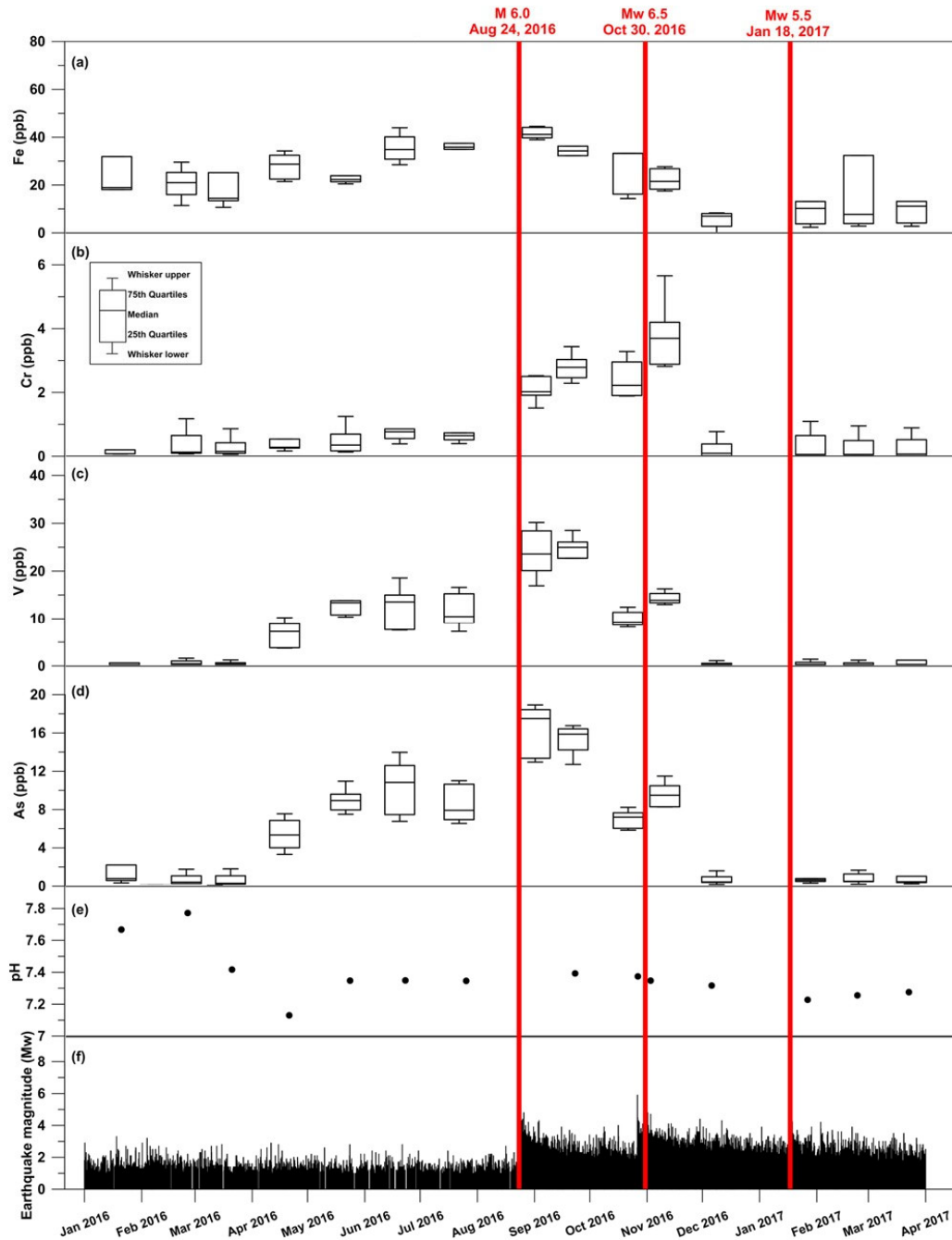
In addition to the Radon monitoring, there are a number of local initiatives to monitor the water concentration in different elements, including CO<sub>2</sub> (Chiodini et al., 2004, 2011), and to link them with the desorption of Arsenic and Vanadium (Barberio et al., 2017; Boschetti et al., 2019). Italy is a collage of two tectonic provinces: the Tyrrhenian plate to the West and the Adriatic plate to the East. These two blocks differ strongly in terms of their properties of CO<sub>2</sub> transport from depth to the surface (Figure 4, after Chiodini et al. 2011).



**Figure 4 a)** The location of the gas emissions is compared with the extension of the Tuscan Roman Degassing Structure (TRDS) which was previously defined on the basis of the deeply derived CO<sub>2</sub> dissolved in the main aquifers of the region (Chiodini et al., 2004). The CO<sub>2</sub> fluxes fed by deep sources are highlighted by green-red colours. The gas emissions are reported with different colours on the basis of their location. The location of gas emissions and the TRDS is compared with the seismicity occurred in the area from 1981 to 2002 (Castello et al., 2006), with the location of the Abruzzo earthquakes (earthquakes from December 2008 to December 2009,  $M > 2.5$ ; ISIDE, 2009) and with the 1997 seismic sequence of Colfiorito (best relocated earthquakes from 26 September–3 November 1997,  $M > 2.5$ ; Chiaraluce et al., 2003); **b)** Conceptual model of the degassing process and of its relation with the structural setting and the earthquake locations (grey dots). The cross section is based on CROP03 deep seismic reflection profile (modified from Chiodini et al., 2004).

The more mature Tyrrhenian crust allows for a free degassing of deep CO<sub>2</sub> up to the surface, hence a very high probability to capture and measure this CO<sub>2</sub> in thermal waters and springs. On the contrary, the much less mature Adriatic crust holds traps for the deep CO<sub>2</sub> where pressure could easily build up. Interestingly, the seismicity concentrates into a lineament along the boundary between the two provinces.

Different mechanisms at work to destabilize the faults in the area have been proposed, and they all rely onto the continuous monitoring of the quality of thermal waters and springs. Concerning the recent Amatrice-Visso-Norcia sequence, Figure 5 presents the results of the concentration changes in different elements measured at the Sulmona Plain Test Site (SPTS), south of the sequence.



**Figure 5** Time series (January 1st, 2016–March 31st, 2017) of element concentration, pH (Sulmona test site springs), and earthquake magnitude. Concentration of elements is shown as box-and-whisker plots representing interquartile ranges (25th–75th quartiles). Each plot includes geochemical data from the Sulmona test site springs (Supplementary Tables S2 and S3). Seismic data are from the INGV database (available online at <http://cnt.rm.ingv.it/>). (a) Iron, Fe, concentration. (b) Chromium, Cr, concentration. (c) Vanadium, V, concentration. (d) Arsenic, As, concentration. (e) Potential of hydrogen, pH. (f) Magnitude (Mw) of earthquakes belonging to the 2016–2017 central Apennines sequence. After Barberio et al. (2017).

There is a positive anomaly in the concentration of Arsenic [As] and Vanadium [V] that starts nearly 5 months before the occurrence of Amatrice earthquake. What is really important here, and that

makes a significant difference with the Radon monitoring for instance, is that this anomaly appears several months before the mainshock and exists until the occurrence of the rupture. There is a duration to this anomaly, and the continuity of the measurements and the sustained high concentrations all increase the confidence in the possibility of an earthquake preparation process at work. Later in this report, we will confront these continuous geochemical observations with the most recent seismological results.

---

## **4. Continuous seismic monitoring in Italy**

### **4.1 Observations based on seismic activity**

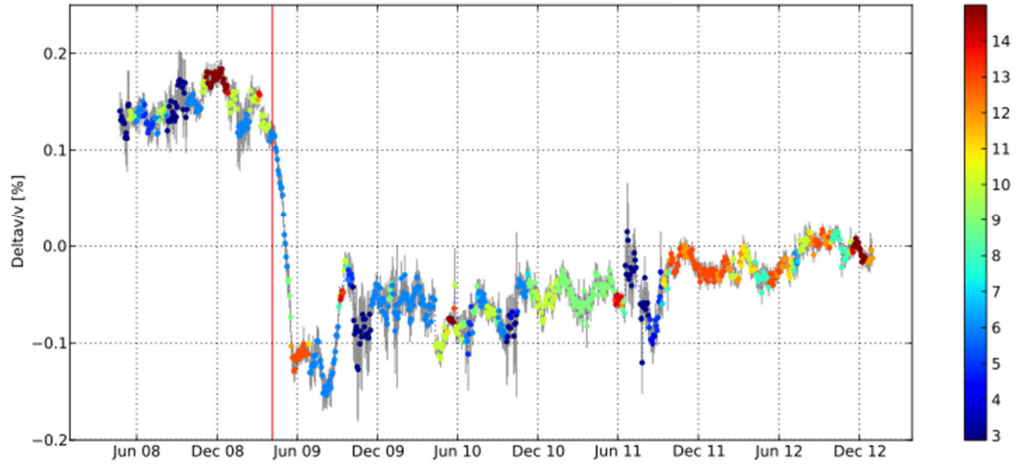
A classical framework for earthquakes to occur is that they are the result of long-term strain accumulation on active faults and complex transient triggering mechanisms. Laboratory experiments show accelerating deformation patterns before failure conditions are met. Imaging similar preparatory phases in nature remains difficult because it requires dense monitoring in advance (Chiarabba et al., 2020). The 2016 Amatrice-Visso-Norcia sequence provided a unique testing ground to image the preparatory phase of a large event. The crustal volume of the Norcia incipient fault was densely illuminated by seismic rays from more than 13,000 earthquakes that occurred within the 3 months before the main shock. Seismic tomography realized in distinct time windows revealed the precursory changes of elastic wave speed, signalling (1) the final locked state of the fault, and (2) the rapid fault-stiffness alterations near the hypocentre just a few weeks before the event. The results show the systematic loading of the main asperity of the Mw 6.5 earthquake, revealed by the systematic increase in P-wave velocity during an apparent 2-months long preparatory phase. An abrupt velocity reduction due to accelerated creep close to the hypocentre that occurred a few weeks or days before the rupture. These results provide the first evidence of precursory velocity changes before fault failure in nature that confirm laboratory observations. The geophysical imaging of such changes in a useful time frame before the earthquake provides new perspectives for understanding earthquake nucleation mechanisms. A systematic documentation of those changes, and their statistical significance over long time frames, will be essential for physics-based earthquake forecasts (Chiarabba et al., 2018, 2020). Of course, these significant results depend entirely on the density and location of seismicity, which is required for tomography to resolve the processes well. The Amatrice aftershock sequence offered a one-of-its-kind opportunity to image the crustal volume around Norcia. This effort and the tools developed for this purpose might be of valuable help to discriminate the case of a single mainshock/aftershock sequence (L'Aquila, Irpinia, Emilia) and the case of a multi-mainshocks sequence (such as the Colfiorito or the Amatrice-Visso-Norcia).

### **4.2 Observations of crustal velocity changes based on ambient seismic noise interferometry**

Ambient seismic noise interferometry has been used to investigate the 2009 L'Aquila earthquake case (Soldati et al., 2015; Zaccarelli et al., 2011). In 2009, the small number of seismic stations available in the region and the method of measuring velocity changes in the crust from coda waves



resulted in a low spatial resolution on the one hand, and a poor temporal resolution of two months on the other hand. Figure 1 shows the measured velocity changes over the period 2008-2012.



**Figure 6** Relative velocity variations obtained from cross-correlation of noise data recorded by the six stations within 40 km of L'Aquila, in the 5-yr period from 2008 January to 2012 December. Colours represent the number of station pairs considered in the computation, with a minimum of three pairs to get an estimate of velocity variation in any time window. Vertical bars indicate the uncertainties of the measurements. The vertical red line highlights the time of occurrence of the L'Aquila main shock (Soldati et al., 2015).

The two months of time averaging is imposed by the choice to make measurements on the coda waves, that need more time to be reconstructed especially in a very perturbed environment with multiple sources (the aftershocks) changing in space and amplitude. This very severe constraint prevents a fine analysis of the evolution of the crust before the earthquake. The velocity drop prior to the earthquake cannot be related to the imminence of rupture. Moreover, the average distance of 40 km between stations dilutes the possible information in a very large and globally unchanged volume, whereas we are looking for the trace of phenomena of low amplitude and which most probably take place in a very limited volume. Therefore, the velocity variations measured before the earthquake are probably due to environmental disturbances (change in water table level for example). This hypothesis is supported by a more recent study that uses the auto-correlation of the ambient seismic noise recorded at the L'Aquila station (AQU), which shows the very high sensitivity of the  $dv/v$  measurement on the coda to these environmental disturbances (Poli et al., 2020). Using 17 years of data, it is demonstrated for the preseismic period that the 5-years hydrological cycle largely dominate the velocity changes that can be measured in the middle crust. This study highlights the possibility of using seasonal and multiyear perturbations to probe the physical properties of seismogenic fault volumes in the region or L'Aquila. However, the time resolution of 90 days to obtain this result implies that the velocity changes measurement ( $dv/v$ ) remains unable to track the fast changes related to the preparation of the earthquake.

### **4.3 Observations of crustal attenuation changes based on ambient seismic noise coherence**

The attenuation of seismic waves depends on several factors, including the presence of fluids in the crust. Measuring the attenuation of seismic waves and their variations before earthquakes has become possible thanks to a method built on the ambient seismic noise. This method allows to measure attenuation through correlated waveform changes induced by fluid movements. Although the results are not yet peer-reviewed or published, we present here the application of this method to recent earthquakes that occurred in Italy: L'Aquila, Amatrice and the Visso-Norcia pair. Little is known on the preparation of large earthquakes. A consensus states that earthquake nucleation consists schematically of a slipping zone of restricted dimensions that develops on the fault plane, sometimes indirectly underlined by a localized swarm of seismicity which is recognized in hindsight as foreshock sequence. The role of the surrounding crustal volume in the activation of the fault is largely underestimated: Using a monitoring method based on seismic waveforms reconstructed from noise correlations (CCW), we show that the crustal volume hosting the normal faults of the Central Apennines underwent a non-seismic anomaly for several months before the L'Aquila, Amatrice, and Visso-Norcia ruptures, over a distance larger than 100 km. This anomaly consists of a series of increasingly pronounced and frequent disturbances that progressively focus on the future rupture zone. These disturbances persist after the rupture, although much less powerful and frequent. These results considerably broaden the understanding of the phenomenon that precedes and follows a major earthquake and places it at the apex of a process that begins long before nucleation and continues long after rupture.

#### **4.3.1 Attenuation and changes of seismic noise correlated waveforms (CCW)**

Contrarily to the other studies that exploit the changes of arrival time of either ballistic (P, S, or surface) or coda waves reconstructed by correlations of successive time windows, the CCW method exploits the subtle changes of the shape of waveforms, regardless of their arrival time (18 57560, 2018). Some studies have investigated the changes of seismic attenuation before and after some large earthquakes (Malagnini et al., 2019; Rohrbach et al., 2013) and these changes were related to mechanical damages in the volume surrounding the faults. The crack density within a fault's damage zone is thought to vary as seismic rupture is approached, as well as during the postseismic period. However, external stress loads, such as seasonal or tidal loadings, may also change the crack density in rocks and all such processes can leave detectable signatures on seismic attenuation that may hide the changes related to the preparation of a large earthquake. As an example, the San Andreas Fault at Parkfield is affected by seasonal loading cycles, which precluded the detection of changes before the last 2006 earthquake. Based on the analysis of thousands of small seismic events, Malagnini et al. (2019) observed random fluctuations over the period from the San Simeon earthquake up to the 2004 M6 Parkfield main shock.

Seismic velocity variations associated with the seismogenic process have been studied world-wide for almost half a century since the dilatancy hypothesis proposed in early 1970s (Rohrbach et al., 2013). Seismic ambient noise is an ideal source for monitoring temporal variations of



both the attenuation and of the seismic velocity in a seismogenic zone. The seismic velocity variations associated to the 2008 Wenchuan earthquake that occurred on the Longmenshan fault were retrieved from the analysis of 10 years of seismic events (Wang et al., 2011), and compared with synthetic continuous recordings. To extract the seismic velocity changes from the ambient noise empirical Green's function, the observables are the surface wave phase velocity dispersion relationship. For amplitude changes, the observables are the horizontal to vertical spectral ratio (H/V). The processing of these synthetic ambient noise records showed a large velocity drop and a recovery phase directly before the mainshock. Using real continuous seismic recordings, Ohmi et al. studied the temporal variation of the source region of the 2007 Noto Hanto (Japan) earthquake (Ohmi et al., 2008). They computed the 1-day autocorrelation functions (ACF) of band-pass filtered seismic noise portion recorded by short-period seismometer at four seismic stations. The comparison of the daily ACF showed gradual changes in the arrival time of some waves in the ACF of one station in the preceding 2 weeks of the mainshock, which were interpreted as the change of seismic velocity structure in the volume considered. These results stimulated the interest in investigating the ACF of ambient seismic noise of the stations located in the vicinity of the faults prone to rupture. In the following, we introduce a different approach to seismic noise monitoring that focuses on the changes of waveforms (CCW) observed in the central part of the ACF. We propose this measurement to be a proxy for the seismic waves attenuation variations related to fluid changes.

#### 4.3.2 The CCW computations

##### 4.3.2.1 Data Preprocessing

The vertical component (Z) of the seismic data is downloaded on a daily basis. The instrumental response is removed and the signal is decimated down to 5 Hz. The signal is detrended and the mean is removed. In order to remove as much as possible, the contribution of local and teleseismic earthquakes, we apply a classical Short Term Average versus Long Term Average (STA/LTA) filter in the frequency bands of 1-2 Hz for the local earthquakes and 0.05-0.2 Hz for the teleseismic earthquakes. The LTA is computed on the entire signal; the STA is computed over 60s for teleseismic earthquakes and 3s for local ones. The detection threshold of the STA/LTA ratio is set to 1.5 for the teleseismic earthquakes and 3 for the local ones. All time windows where the STA/LTA ratio is higher than the threshold are removed from the analysis. Finally, to further reduce the transient sources, which impact the seismic signal, and to converge towards a stationary dataset, we apply a spectral whitening of the data that equalizes the spectrum over the entire frequency band of interest.

##### 1.1.1.1 Autocorrelation Computations

The daily preprocessed signal is chopped into 288 time windows of 5 minutes each. Each individual window is apodised using a Tukey filter. Then the ACF of each window is computed and normalized by the energy of the signal.

##### 1.1.1.2 Coherence of Correlated Waveforms (CCW)

We define a reference by averaging all the 5 min ACFs of the first month during the period of analysis. In a second step, each 5 min ACF is compared to this reference. As we are interested in the central part of the ACFs, we cut the 5 min ACFs and the reference at  $\pm 5s$ . Finally, the comparison is made by computing the covariance between each 5 min ACFs and the reference, similarly to what was proposed to compare earthquake waveforms (Roux & Ben-Zion, 2013).

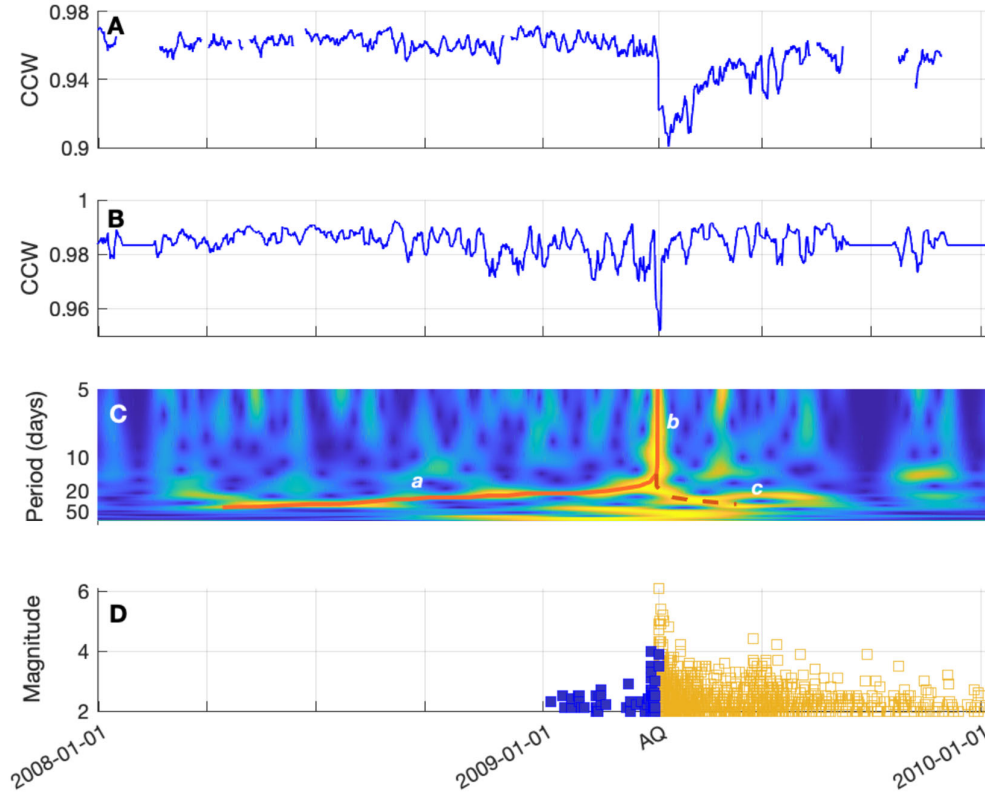
#### 1.1.1.3 Post-processing and graphical representations

A visual inspection of the CCW outcome is performed to eliminate possible outliers that may appear because of the preprocessing. The remaining CCW are interpolated to obtain a complete timeseries. Finally, the results are smoothed with a Savitsky-Golay filter to 5 days. The continuous wavelet transform is computed with a Morlet wavelet of varying period.

#### 1.1.2 Application to the L'Aquila earthquake

The L'Aquila earthquake (6 April 2009) occurred as the mainshock of a long sequence of small magnitude foreshocks (Sugan et al., 2014). The rupture occurred on the Paganica fault (Chiarabba et al., 2009; Chiaraluce et al., 2009) and generated severe damages and losses to the town of L'Aquila and to the surrounding villages. The closest seismic station to the M6.2 L'Aquila earthquake is the L'Aquila station (code: AQU) operated by the Mediterranean seismic network (Med-Net) since 1989. In the following, we use the continuous records of the ambient seismic noise to monitor the state of the crust beneath the L'Aquila seismic station. As the goal is to detect a precursory signal in the vicinity of the fault, we will use hereafter the ACF for the seismic records of AQU station. Former studies have investigated the changes of velocity and of decorrelation using the coda of (auto)correlation functions (Hillers et al., 2015). Contrarily to these studies, we focus here on the subtle changes of waveforms in the central (i.e., ballistic) part of the ACF. The data are processed and autocorrelated to monitor the local state of the crust beneath the L'Aquila station. Figure 2A presents the evolution of CCW over the period 2008-2010, computed in a frequency range from 0.03 to 1 Hz, which corresponds to a depth range from 1 to 30 km. The L'Aquila earthquake (AQ) is marked by a coseismic CCW drop of 4%, followed by a phase of progressive recovery superimposed with strong oscillations of the measurement. Although some data are missing, we note that the recovery is not complete by the beginning of 2010. The main earthquake itself is preceded by several months of a progressive and quasi linear decrease of the CCW. However, the change in CCW is of the same order of magnitude as those recorded earlier in 2008, making it difficult to turn the decrease of CCW in a neat signature of the impending earthquake. The depth of nucleation of the L'Aquila earthquake is in the range of 9-12 km, as indicated by the relocation of the foreshock seismicity (Valoroso et al., 2013). Assuming that the ambient seismic wavefield is dominated by surface Rayleigh waves, and according to the velocity model of the L'Aquila area (Herrmann et al., n.d.) and the corresponding sensitivity kernels of the Rayleigh waves, the 0.1-0.3 Hz frequency range appears most sensitive to this 9-12 km depth range. The vertical component of the seismic records of the L'Aquila seismic station are filtered in the 0.1-0.3 Hz band to focus only on the possible changes of the Earth's crust at the nucleation depth of

the L'Aquila earthquake. Figure 2B presents the evolution of CCW for the 0.1-0.3 Hz frequency range. The L'Aquila earthquake is marked by a coseismic CCW drop of 3%, similar to what is measured in the broadband evolution. The coseismic drop is followed by an extremely rapid post-seismic recovery associated with profound oscillations of the measurement. This is one of the main differences between the two curves presented in Figures 2A and 2B. Another strong difference concerns the period preceding the mainshock, which is characterized by oscillations of the CCW that apparently grow in amplitude as we go closer to the fault failure. A trained eye may distinguish that the oscillations have a decreasing period as we approach the time of the mainshock. As this growth of oscillations of CCW may constitute a precursory pattern, dating the beginning of these growing oscillations and characterizing them turns into a crucial issue. For that, we compute the continuous wavelet transform of the CCW. This time versus period representation of the CCW provides a robust estimate of the dominant period present in the CCW measurement at any given time (Figure 2C). The strong oscillations that precede the April 6 mainshock are characterized by a dominant period that continuously decreases from 40 days to 5 days (limited by the choice of the 5 days smoothing window) over the 300 days before the mainshock. This very particular feature is interpreted as the signature of the preparation process of the L'Aquila earthquake. The strong oscillations of the CCW that follow the mainshock are characterized by a reverse pattern, with an increasing period and a decreasing amplitude. The L'Aquila mainshock appears as the culminating point of a suite of oscillations that reveal changes of the state of the crust in the nucleation zone. Figure 2D presents the timing and magnitude of the sequence of foreshocks and aftershocks of the L'Aquila earthquake. The L'Aquila CCW precursory pattern has a duration that exceeds by far the duration of its foreshock sequence, which now appears as a secondary phase occurring at the end of the preparation process, when the CCW oscillations grow in strength and the decrease of the dominant period marks an acceleration.



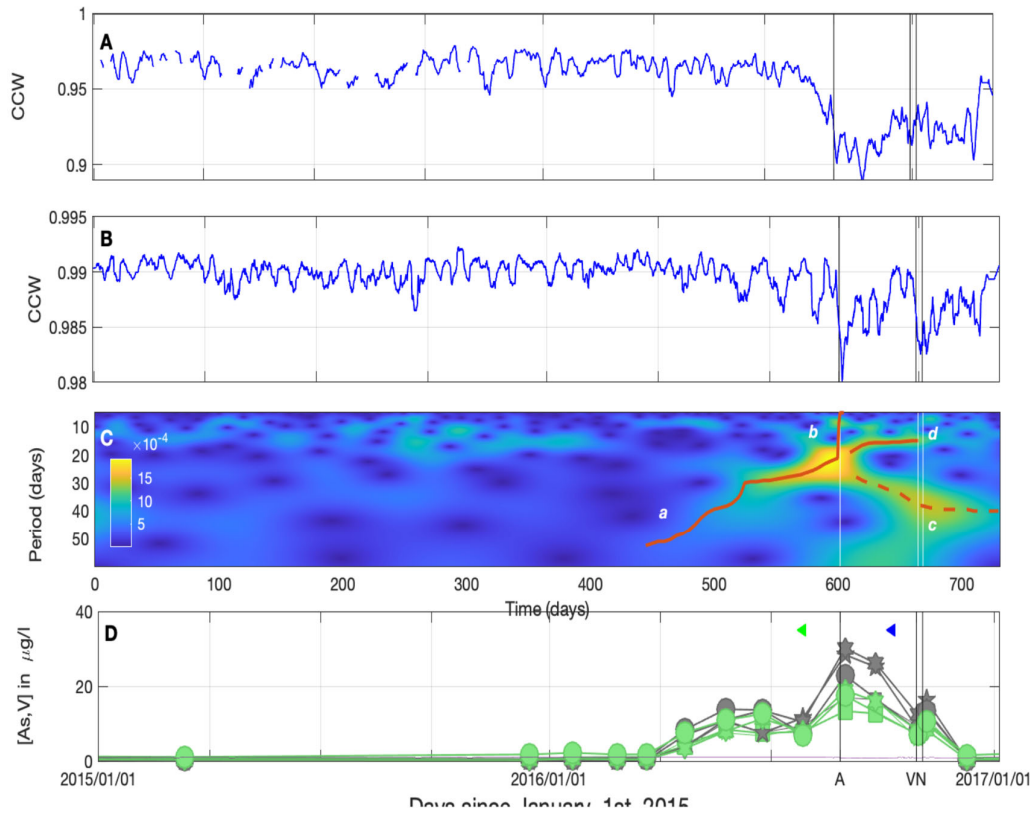
**Figure 7** Evolution of CCW measured at L'Aquila seismic station in 2008-2009. **(A)** CCW measured at station AQU over the years 2008 and 2009, computed in the 0.03 – 1 Hz frequency range and smoothed at 5 days. The reference is computed over the first 30 days of available data of 2008. The mainshock (AQ) is marked by the large decrease of the CCW measure and is followed by remarkable post-seismic recovery. Before the mainshock, the CCW marks a period of decrease that starts at the beginning of 2009 and keeps going until the mainshock occurrence. **(B)** CCW measured at station AQU over the years 2008 and 2009, computed in the 0.1 – 0.3 Hz frequency range and smoothed at 5 days. The reference is computed over the first 30 days of available data of 2008. The mainshock (AQ) is marked by the large decrease of the CCW measure and is followed by extremely rapid post-seismic recovery superimposed with large oscillations. The months preceding the mainshock is characterized by oscillations of growing amplitude and decreasing period. **(C)** Normalized continuous wavelet transform of the CCW of Fig. 2B. The yellow patches indicate the dominant period of the signal at any given time. The red continuous line (branch *a*) indicates the evolution of the dominant period from 40 days down to 5 days over the 350 last days before the L'Aquila earthquake. The red discontinuous line (branch *c*) underlines the reverse pattern that develops after the mainshock (branch *b*). **(D)** Sequence of foreshocks (blue squares) and aftershocks (orange squares) of the L'Aquila mainshock (data from the ISIDE database).

### 1.1.3 Application to the Amatrice-Visso-Norcia earthquake sequence

The M6.2 Amatrice earthquake (24 August 2016) occurred 50 km north of L'Aquila on a complex fault system involving many different segments including the Monte Vettore fault (Perouse et al., 2018). The rupture generated severe damages and losses to the town of Amatrice and to the surrounding villages. The closest seismic station is the Norcia station (code: NRCA, Figure 1) operated by the Italian Istituto Nazionale di Geofisica e Vulcanologia (INGV) since 2010. The station is located 16 km north of the Amatrice epicenter, close to the northern termination of the

rupture. The continuous records of the ambient seismic noise are used to monitor the state of the crust beneath the NRCA seismic station. Figure 3A presents the evolution of CCW over the period 2015-2016, computed in the 0.03 to 1 Hz frequency range, which corresponds to a depth range from 1 to 30 km assuming the same velocity model as for L'Aquila (Figure S1). The Amatrice earthquake is marked by a large coseismic CCW drop of 4%, followed by short recovery rapidly interrupted by the following of M5.9 Visso and M6.5 Norcia mainshocks. The preseismic period shows smooth oscillations of the CCW measure above the 0.95 threshold with no particular pattern. However, the last 5 weeks preceding the Amatrice earthquake are marked by a CCW decrease below the 0.95 value. The duration of this decrease exceeds the 5 days smoothing of the curve, which makes this CCW decrease significant and a potential precursor candidate. The depth of nucleation of the Amatrice earthquake is in the range of 4 to 8 km; assuming that Rayleigh waves dominates the seismic wavefield and the same velocity model than for the L'Aquila area, we present in Figure 3B the evolution of CCW computed in the 0.5-0.7 Hz frequency band, corresponding to a depth range of 3-4 km slightly above the nucleation. At first, the CCW measurements oscillates in narrow range of values, from 0.995 to 0.98. This means that the changes revealed by the technique are very subtle. Then, the Amatrice earthquake is marked by a sudden decrease of the measurement of 1%, followed by a rapid recovery superimposed with strong oscillations of the CCW. The recovery is interrupted by interferences generated by the subsequent Visso and Norcia mainshocks, themselves followed by their own post-seismic recovery. The year 2015 presents smooth oscillations, similar to what was observed in the broadband frequency range (Figure 3A). The 5 months period preceding the Amatrice earthquake presents a different behavior though. The CCW measurement exhibits remarkable oscillations that grow in amplitude until the Amatrice rupture. Figure 3C presents the continuous wavelet transform of the CCW measurement presented in Figure 3B. It shows the progressive decrease of the dominant period of the oscillations preceding the Amatrice earthquake, a similar signature to the one that preceded the L'Aquila earthquake. The branch *a* in Figure 3C underlines this evolution and allows to traced back this precursory phase until nearly 150 days before the earthquake. We can clearly see two phases in this signal, the first with a low amplitude from 450 to 520 days, and the second, marked by increasingly stronger amplitudes from day 520 to the day of the earthquake (day 602). This remarkable precursor pattern revealed by the continuous monitoring of the ambient seismic noise is corroborated by other observables from different disciplines, mainly geochemistry. Two stations (CTTR and NRCA) operate in the framework of the permanent Italian Radon monitoring Network (Cannelli et al., 2018) and recorded almost continuously since 2012 and 2016, respectively, the latter being installed just after the Amatrice mainshock of the sequence (Mw 6.0 on August 24, 2016) the radon concentration close to the epicentre of the Norcia mainshock (Mw 6.5 on October 30, 2016). The blue triangle in Figure 3D marks the significant increase of radon emanation and the associated detection, about 2 weeks before the Mw 6.5 Norcia on both the time series (Soldati et al., 2020). The green triangle in Figure 3D indicates the detection of an anomaly at station CCTR, about 30 days before the Amatrice earthquake. The inherent difficulty with radon measurements is their high sensibility to environmental effects such as precipitations, temperature, air pressure, and local soil conditions, making difficult to assign a detected anomaly to the preparation of a large earthquake. In addition, environmental effects impact the radon measurements too strongly, making it difficult to show a continuous evolution of the state of the crust until the

rupture. Consequently, the detection of a radon anomaly is considered a posteriori as real if it is followed by an earthquake of magnitude larger than 4, occurring in a window of 40 km and in the following 40 days; otherwise it is classified as a false detection, related to the environmental effects. Other geochemical markers are much less sensitive to these environmental effects, and may help to track the evolution of the fault system prone to rupture. Figure 3D presents the evolution of the concentration of Arsenic (As) and Vanadium (V) in a selection of water springs located in the Sulmona Plain Test Site, south of the Amatrice rupture (Barberio et al., 2017; Boschetti et al., 2019). These springs are the principal outlet of a large hydrological system that drain the aquifers of the Central Apennines ranges. The concentrations are measured every month. All the geochemical measurements show an increase in the concentration of As and V that begins about 5 months before the Amatrice earthquake, together with changes in the pH of waters (not shown). This significant increase is maintained until the earthquake, before gradually decreasing and returning to normal in early 2017. The near perfect match between the onset of this concentration increase and the onset of the precursor observed using continuous seismological data (Figure 3A, B and C) supports the idea that the state of the crust is profoundly altered in the few months prior to the Amatrice rupture. The geochemical measurements following the Amatrice earthquake show a steady decrease until the Visso and Norcia earthquakes, when a secondary and short-lived peak is observed. In contrast, the evolution of CCW (Figure 3B) over the same period between the main earthquakes of the sequence shows a rapid recovery marked by deep oscillations until the Visso and Norcia earthquakes. The analysis of the dominant frequency content by the wavelet approach (Figure 3C) shows the coexistence of two trends in the signal. The first one (branch *c*) dips towards increasingly longer periods. The second one (branch *d*) evolves on the contrary towards shorter and shorter periods and abruptly stops at the dates of the Visso and Norcia earthquakes. The behavior of branch *d* is very similar to that of branch *a*, and we therefore interpret this branch *d* as the precursor signal to the Visso and Norcia earthquakes. This interpretation is bolstered up by a positive  $V_p/V_s$  anomalies that indicate high fluid pore pressure in the volume surrounding the hypocentral area (Chiarabba et al., 2018). Moreover, using the seismic rays from more than 13,000 earthquakes that occurred within the 3 months before the main shock nucleation to perform time-lapse seismic tomographies revealed the precursory changes of elastic wave speed, signaling the final locked state of the fault, and the rapid fault-stiffness alterations near the hypocenter just a few weeks before the event (Chiarabba et al., 2020).



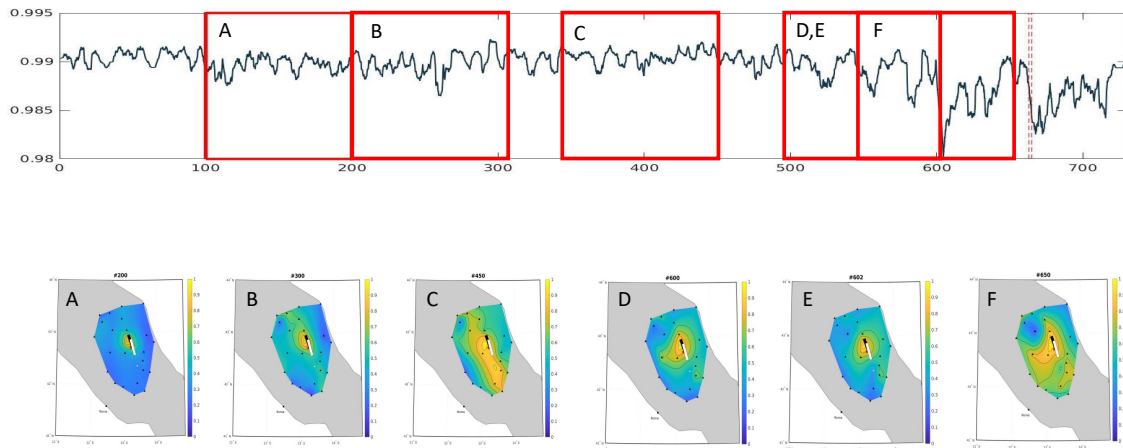
**Figure 8** Changes of CCW measured at Norcia seismic station in 2015-2016. **(A)** CCW measured at station NRCA over the years 2015 and 2016, computed in the 0.03 – 1 Hz frequency range and smoothed at 5 days. The reference is computed over the first 30 days of available data of 2015. The Amatrice mainshock (A) is marked by the large decrease of the CCW measure, and is followed by the post-seismic recovery, rapidly interrupted by the interferences due to Visso (V) and Norcia (N) earthquakes. Before the Amatrice mainshock, the CCW marks a period of decrease that starts 5 weeks before and keeps going until the mainshock occurrence. **(B)** CCW measured at station NRCA over the years 2015 and 2016, computed in the 0.5 – 0.7 Hz frequency range and smoothed at 5 days. The reference is computed over the first 30 days of available data of 2015. The Amatrice mainshock (A) is marked by a decrease of the CCW measure, and is followed by the post-seismic recovery superimposed with large oscillations. The period before the Amatrice mainshock is characterized by oscillations of growing amplitude and decreasing period. **(C)** Continuous Wavelet Transform of the CCW of Fig. 3B. The color scale indicates the distribution of energy of CCW as a function of the period. The red continuous line (branch *a*) indicates the evolution of the dominant period from 50 days down to 20 days over the 150 last days before the Amatrice earthquake. The red discontinuous line (branch *c*) underlines the reverse pattern that develops after the Amatrice mainshock (branch *b*). Simultaneously with branch *c*, another branch *d* emerges with decreasing periods as for branch *a*. It seems to begin very early after the Amatrice earthquake and develops until the occurrence of the Visso and Norcia earthquakes, where it stops abruptly. **(D)** Concentration in Arsenic (green symbols) and Vanadium (grey symbols) measured in different water springs located in the Sulmona basin, south of the rupture. All curves show a strong and sustained increase of these concentrations during the 5 months preceding the amatrice mainshock. These concentrations peak at the time of the Amatrice mainshock and slowly decrease over the months that follows, at the exception of a minor secondary peak at the time of Visso and Norcia earthquakes. Green and blue triangles indicate the radon anomalies detected by the CTTR and NRCA stations.

The previous results suggest the detection of a precursor signal to the Central Apennine earthquakes over periods of several months. The seismic stations of L'Aquila and Norcia were selected

to illustrate these results because of their proximity to the faults that ruptured during the earthquakes. If the station of L'Aquila was located just above the rupture in a very favorable position to observe the precursor, the station Norcia is located more than 15 km North of the hypocentral zone of Amatrice, and even if it is a little closer to the ruptures of Visso and Norcia earthquakes, the question of the maximum distance at which the precursor is detectable is raised.

Following the earthquake of L'Aquila, a significant effort has been made to develop the permanent seismological network in Italy. Thus, in 2016, the area around Amatrice benefited from the deployment of 30 stations spread within 100 km, allowing a spatial investigation of the precursor. In the following, we show the evolution of the CCW for each of these stations and compare them to the evolution of the CCW measured at the NRCA station, in the 0.5-0.7 Hz frequency range, where the precursor is well observed. Figure 4 presents a selection of 100-day wide snapshots in the time and space evolution of the precursory pattern prior to the Amatrice earthquake. Each image represents a 100-day window during which the CCW of each station is correlated to the CCW of the NRCA station. The result are then interpolated to create the maps. If during this time window, either station records the same signal as NRCA then the value for that station will be 1. On the contrary, if either station records a different signal than NRCA then the value for that station will be low. Figure 4A corresponds to a time window from day 100 to day 200, long before the beginning of the precursor. The map shows that station NRCA is only correlated to itself, which means that at this time, there is no specific spatial pattern that could correspond to a precursor. Figure 4B represents the time window from day 200 to day 300. Again station NRCA is correlated to itself. However, the stations oriented in a Northwest-Southeast trend seem to correlate more than the other stations located away from this orientation. Figure 4C, obtained for the time window from day 350 to day 450 (i.e. 250 to 150 days before the Amatrice earthquake), confirms the previously observed trend. Here, all stations aligned with NRCA along the northwest-southeast axis correlate very well with the NRCA station. This is in stark contrast to the stations located outside this azimuth, which correlate poorly or not at all with NRCA. Figure 4C shows a very peculiar pattern oriented along the tectonic lineament and the principal faults that ruptured in the sequence of 2016. By the end of the time window of figure 4C (150 days before the earthquake), the precursor (branch a in Figure 3C) starts to be detected with the CWT. Figure 4D obtained for the time window from day 500 to day 600 and Figure 4E, obtained for the time window from day 502 to day 602 (day of the Amatrice earthquake) illustrate the progressive localization of the highly correlated pattern. By the time of the earthquake, only the closest seismic stations to the impending rupture (NRCA, LNSS, FDMO) share the information on the precursor. Finally, Figure 4F is obtained for a time window that goes from day 550 to day 650, respectively before the Amatrice earthquake and before the Visso-Norcia earthquakes. The map shows a very different pattern where the stations to the west and to the south of the Amatrice rupture are now correlated with NRCA. During this period, Figure 3C shows the coexistence of two evolutions: the post-seismic phase of Amatrice and the precursor phase to Visso-Norcia earthquakes. In space however, it is harder to distinguish these two evolutions: the images appear dominated by the Amatrice's post-seismic evolution that is located on the hanging wall side of the ruptured fault, to the west.



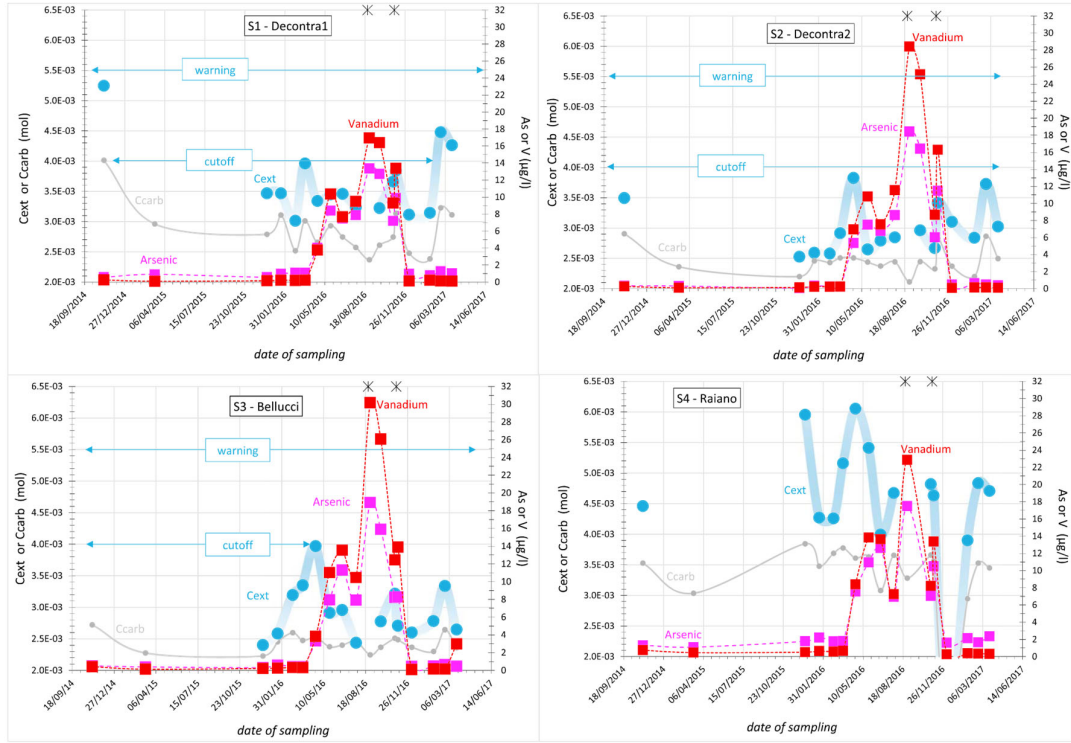


**Figure 9** Space and time evolution of the Amatrice precursory pattern. **(Top)** CCW measured at station NRCA over the years 2015 and 2016, computed in the 0.5 – 0.7 Hz frequency range and smoothed at 5 days. The reference is computed over the first 30 days of available data of 2015. The red frames labeled A to F indicate several 100 days windows used to compare the seismic stations with respect to NRCA. **(Bottom)** Each map represents the similarity of the individual CCW measurements of the seismic stations to the CCW measured at station NRCA. A high value of correlation indicates a station that share a common information with NRCA. In frame A, the station NRCA is only correlated to itself, meaning that the stations are independent from each other. In frame B, the seismic stations to the North of the fault zone have higher value of correlation than the rest of the stations. In frame C, all stations aligned with fault zone strike angle are correlated to NRCA. This time window corresponds to the maximal extent of the precursory pattern. Frames D and E show a progressive focalization of the pattern towards the future rupture area. Frame F shows the post-seismic pattern of the Amatrice earthquake.

## 5. Perspectives in OEF enhancement with new candidate precursors

### 5.1 Towards a comprehensive physical model of the earthquake preparation

The concomitance of geochemical and CCW anomalies that precede the Amatrice earthquake suggests a common phenomenon at the origin of these anomalies. The increased concentrations of As and V are accompanied by changes in the pH of the sampled waters that are interpreted as the intrusion of deep CO<sub>2</sub> into the system (Boschetti et al., 2019). Based on the geochemical data, this deep CO<sub>2</sub> intrusion into the fault system appears to be continuous. However, the last measurements available one month before the Amatrice earthquake indicate a significant decrease in As and V concentrations (Figure 5), although their level remains very high. This seems to suggest that the process may not be as continuous as one might think, and that it develops in successive intrusions. This view is reinforced by the shape of the precursor revealed by the CCW (Figure 8; measured with a far better temporal resolution of 5 days), which shows a succession of increasingly strong and frequent disturbances.



**Figure 10** Time series for Cext-Ccarb (big blue-small gray dots joined by solid lines) and As-V (pink-red squares joined by dashed lines) of the S1-S4 springs. The dimension of the data points is greater than analytical accuracy (5%). Cutoff and warning threshold refer to Cext values of  $4.0\text{E-}03$  and  $5.5\text{E-}03$  m, respectively (Chiodini et al., 2004; Martini, 2016). These values are purely indicative and only shown for comparison. Cext gaps are due to a lack of pH data. Asterisks (\*) denote  $M_w \geq 5.0$  primary shock events: 24 August 2016 ( $M_w = 6.0$ ) and 30 October 2016 ( $M_w = 6.5$ ). After (Boschetti et al., 2019).

Deep  $\text{CO}_2$  shifts geochemical equilibria and these changes are measured as far as 60 km from the epicentral zone. The CCW measurement thus seems sensitive to either geochemical changes or to the evolution of  $\text{CO}_2$  in the system, the latter being a more reasonable hypothesis. Indeed, several studies have demonstrated the impact of fluid substitutions on seismic attributes, namely wave velocity and especially seismic attenuation (O Amoroso et al., 2017; Ortensia Amoroso et al., 2014; Dupuy, Garambois, & Virieux, 2016; Dupuy, Garambois, Asnaashari, et al., 2016). At depths larger than a few kilometers, the  $\text{CO}_2$  will mostly affect the seismic attenuation rather than the P or S wave velocity (O Amoroso et al., 2017). This suggests that the CCW measurement is a proxy for the changes of seismic attenuation related to fluid substitutions at depth.

A strong argument in favor of the sensitivity of CCW on deep fluid movements lies in its elongated shape in the direction of the faults at the onset of the precursor (Figure 4, Frame C). This highly anisotropic shape might be related to the much higher permeability properties along the tectonic lineament than perpendicular to it. In addition, the progressive focusing of the precursor towards the future epicentral zone is compatible with a process of fluid diffusion away from the source (the hypocentral zone) towards the periphery. Indeed, the diffusion distance depends on the period of the disturbances as well as on their amplitude. Since the period of the perturbations decreases more rapidly than their amplitude increases, this makes it possible to observe the focusing of the diffusion pattern towards the future Amatrice earthquake.

The link between CCW changes and deep CO<sub>2</sub> is reinforced by the observations made before and after the L'Aquila earthquake of 2009. Many of the foreshocks that preceded the mainshock are located in a region of high  $V_p/V_s$  ratio that suggests the presence of fluids in the hypocentral area (Lucente et al., 2010). The intrusion of fluids culminated with the occurrence of a magnitude 4 foreshock a week before the event, which could have destabilized the fault lead to the rupture. These indirect observations of fluids are backed up by direct measurements of the amount of CO<sub>2</sub> in the thermal waters of the area (Bonfanti et al., 2012). The measurements began a few days before the M4 foreshock that occurred a week before the L'Aquila mainshock and lasted during the 70 days following the earthquake. They show that the CO<sub>2</sub> amount obeys a complex dynamic made of peaks and troughs, similar to the dynamic of the CCW changes. This similarity of behaviors fosters the relation between the CO<sub>2</sub> and the CCW measurement.

## **5.2 A possible path towards the improvement of operational earthquake forecasting**

A previous milestone of the project, MS23 – Scheme of OEF model to include anomalies, described the backbone to incorporate the identification of possible anomalies with the current earthquake forecasting models. In a nutshell, the Bayesian framework will allow us to get a posterior earthquake forecasting distribution, starting from a prior given by the current stochastic models (e.g., ETAS) and using a likelihood that contains all information about the anomalies. To achieve this, we need to quantify some observables that are not available at a certain point in time, such as i) the frequency of earthquakes that have been anticipated by the anomaly in a time window  $\tau$ , and ii) the fraction of time in which the anomaly has not anticipated a large earthquake in a time window  $\tau$ , i.e.,  $\tau N / T$  (assuming that the anomalies are not overlapping), where  $N$  is the number of anomalies, and  $T$  is the total length of the time period investigated. These two quantities can be empirically obtained by applying the above techniques for searching anomalies in the whole Italian earthquake catalog and/or to catalogs of other regions. Once we get these two quantities, we can deliver the first OEF model that incorporates anomalies.

## **5.3 On the use of machine learning to improve the forecasting**

One of the possible shortcomings of the CCW technique is the sensitivity of the measurement to the slight variations of the spectrum of the seismic noise induced by their variations in time and space. A close inspection of the spectrograms recorded at station AQU and NRCA in the years and months before the earthquakes of L'Aquila and Amatrice respectively didn't show significant changes of the features of the seismic noise. However, applications of recent developments in unsupervised machine learning have tried to identify weak sources that are potentially hidden in the noise and therefore hard to detect with classical analyses (Seydoux et al., 2020). The seismic data recorded at L'Aquila and at the neighboring stations from 2008 to 2009 are processed together through an array covariance matrix from which it is possible to extract frequency-dependent features that evolve with time before and after the earthquake (Shi et al., 2020). Despite the classification of the seismic data in clusters, their direct interpretation in terms of fault physics is

far from being straightforward. An analysis of the L'Aquila data alone and direct comparison with the fluctuations of the CCW measurement preceding and following the L'Aquila earthquake could help to grasp further details of the physics of faulting. Note that once the clusters are established, there is a possibility to treat them as classical earthquakes into any forecasting algorithm such as ETAS.

---

## 6. Conclusions

The existence of the precursory pattern that develops in time and space before the large M6 earthquakes of the central Apennines might be of some help improving the time-varying forecasting of impending events. The spreading of the pattern and the progressive focusing towards the epicentral area progressively define the faults that are prone to rupture within a given radius that is governed by the relative locations of the seismic stations with respect to the faults. In the case of the Amatrice earthquake, the size of the pattern at the verge of the rupture is the yellow area in Figure 9D and corresponds to a 20x20 km<sup>2</sup> area. The dynamics of the precursor is apparently nonlinear, with a slow start and a final acceleration, associated with foreshocks and/or geochemical anomalies. The duration of this phase of acceleration is 5 months for the L'Aquila and Amatrice earthquakes. The duration of the precursor preceding M6.5 Norcia earthquake is only two months, i.e., starting after the Amatrice earthquake. This suggests that there is no clear correlation between the duration of the precursor and the magnitude of the impending earthquake.

The coherence of correlated waveforms allows dabbling into the physics of the nucleation process by revealing successive transient pulses which peak above the hypocentral area. These pulses occur at an accelerating pace and define a sweep-like precursory pattern observed before the L'Aquila, Amatrice, and the Visso-Norcia earthquakes, suggesting the similarity of the preparation stage of the large earthquakes of the Central Apennines. The case of the L'Aquila earthquake, for which the station is located just above the fault that ruptured, shows a preparation phase that begins much earlier than the foreshock sequence. This suggests that the nucleation of strong earthquakes, marked in this case by the change in seismicity in the hypocentral region, represents only the final phase of preparation for the major earthquake.

Reciprocally, these major seismic events are followed by transient oscillations that define a reverse sweep like pattern. Hence, each major M6+ seismic events occurred as the culminating point of a process that started a few months before and kept going after their occurrence, superimposed to the post-seismic recovery measured by geodesy. The continuous wavelet analysis allows discriminating between different and simultaneous trends such as the post-seismic trend of Amatrice and the precursor to Norcia. This ability to distinguish coexisting evolutions in the state of the crust could be of valuable help to assess the risk of having a sequence of large mainshocks rather than the more classical single mainshock-aftershocks event.

The present CCW observations are the cornerstone to a new path that, together with other observables, will hopefully lead to reliable and deterministic earthquake prediction in the Central Apennines area, and possibly in all tectonic environments where fluids are particularly pervading.

**Liability Claim**

The European Commission is not responsible for any that may be made of the information contained in this document. Also, responsibility for the information and views expressed in this document lies entirely with the author(s).