PRESTo, the earthquake early warning system for Southern Italy:
Concepts, capabilities and future perspectives
Claudio Satriano\textsuperscript{a,b,}\textsuperscript{*}, Luca Elias,\textsuperscript{a,b} Claudio Martino\textsuperscript{a,b}, Maria Lancieri\textsuperscript{c}, Aldo Zollo\textsuperscript{b},
Giovanni Iannaccone\textsuperscript{d}
\textsuperscript{a} RISSC-Lab, AMRA scarl, Naples, Italy
\textsuperscript{b} Department of Physics, University of Naples Federico II, Naples, Italy
\textsuperscript{c} Ecole Normale Supérieure, Paris, France
\textsuperscript{d} Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Naples, Italy

\textbf{ABSTRACT}

PRESTo (PRobabilistic and Evolutionary early warning SysTem) is a software platform for regional earthquake early warning that integrates recently developed algorithms for real-time earthquake location and magnitude estimation into a highly configurable and easily portable package. The system is under active experimentation in Southern Italy on the Irpinia Seismic Network (ISNet), which is deployed in a seismogenic area that is expected to produce a large earthquake within the next 20 years. In this paper we describe the architecture of the system and test its performances using both small earthquakes ($M_o$ $<$ 3.5) recorded at the ISNet and a large event recorded in Japan, through a simulation mode. The results show that, when a dense seismic network is deployed in the fault area, PRESTo can produce reliable estimates of earthquake location and size within 5–6 s from the event origin. Each estimate is provided as a probability density function, with an uncertainty that typically decreases with time: a stable solution is generally reached within 10 s from the origin.

Thanks to its fully probabilistic approach, PRESTo can be a powerful tool for end-users in addressing the trade-off problem of whether and when to initiate safety measures. The software makes use of widespread standards for real-time data input and output, and can be finely tuned to easily adapt it to different networks and seismogenic regions.

\textsuperscript{*} Corresponding author. Now at: Institute de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05, France. Tel.: +33 1 44 27 24 68.
E-mail address: satriano@ipgp.fr (C. Satriano).

1. Introduction

Southern Apennines is among the highest seismic risk areas in Italy. The latest strong earthquake struck the region on 23 November 1980 ($M_s$ = 6.9), and resulted in more than 3000 casualties and extensive damage. This earthquake has been associated with the rupture of a complex normal fault system, made of 3 main segments [1,2].

At present, the region is characterized by a continuous background of low magnitude seismic activity ($M_w$ $<$ 3.5, Fig. 1) probably connected to the fault system responsible for the 1980 main shock, and by occasional greater magnitude events. Nevertheless, the seismic potential of the region is considered to be high. Boschi et al. [3] indicated a probability between 20\% and 40\% for the occurrence of an earthquake of $M \geq 5.9$ in the area, in the next 20 years. Similarly, Cinti et al. [4] have provided a probability map of occurrence for $M > 5.5$ earthquakes over the next 10 years in Italy, and indicated the Campania-Lucania sector of the southern Apennines among the regions with the highest hazard.

Starting from 2005, a local seismic network, called ISNet (Irpinia Seismic Network, Fig. 1), has been deployed in the Campania-Lucania region, with the double objective of a) providing high quality data for high-resolution studies of the seismogenic faults in the area and b) testing a prototype system for earthquake early warning and post-event risk assessment, for the protection of strategically relevant infrastructures in the region.

ISNet comprises 28 6-component stations, each hosting a velocimeter (short-period or broadband) and a strong-motion accelerometer, with a 1-g dynamic range. This ensures high-quality recordings of low magnitude earthquakes ($0.1 < M < 3.5$) and should guarantee unsaturated signals for strong shakings. The telemetry has been specifically designed for real-time data transmission and analysis, and is realized through a three-layer architecture, which comprises the data loggers, four local control centers (LCC), deployed in the region and each connected to 6 or 7 stations, and the network control center (NCC) in Naples. More details on the ISNet architecture are provided in Weber et al. [5], Zollo et al. [6] and Iannaccone et al. [7].
In the framework of experimenting a regional approach to earthquake early warning (EEW), starting from 2006, we have developed real-time techniques for rapid characterization of earthquake location and size (RTLoc [9]; RTMag [8]). These methodologies are probabilistic and time-evolutionary: earthquake parameters are computed as probability density functions (PDF), and the estimates are continuously updated, as new data are available in real-time from the network, or simply as time passes.

The theoretical capabilities of each algorithm have been individually studied in the respective papers. The overall performance of the whole processing chain, in relation to large earthquakes and extended fault processes, has instead recently been investigated by Zollo et al. [10], through the computation of synthetic accelerograms for 500 scenarios, for three large earthquake mechanisms in the region. The results indicate that, while the convergence of the location and magnitude estimates is generally fast and stable, the quality of the PGV prediction is influenced by the fault extension and by directivity effects.

The early warning algorithms have been recently implemented into an integrated software platform called PRESTo (PRobabilistic and Evolutionary early warning SysTem), specifically designed for fast and efficient data acquisition, processing and dissemination of results, and built on open standards for maximum configurability and portability.

In this work we describe the overall architecture of the PRESTo software platform and discuss the most important aspects of its design. Furthermore we study the performances of the software using real events: a large earthquake (Mw 6.9) that occurred in Japan; and several small earthquakes recorded at the ISNet network in southern Italy.

2. PRESTo: a new software platform for earthquake early warning

PRESTo is the acronym of PRobabilistic and Evolutionary early warning SysTem, a new, integrated, easily deployable software at the base of the earthquake early warning system under development and testing in southern Italy [6,7,10].

In its current implementation, PRESTo follows a regional approach [11,12], i.e. it relies on the information coming from a seismic network deployed in the epicentral area to rapidly estimate source parameters (location and size) of a potentially destructive earthquake and to predict the ground motion at distant targets. It is mainly based on the RTLoc [9] and RTMag [8] algorithms for real-time earthquake location and magnitude estimation.

The novel features of the system are summarized in its name. PRESTo is “evolutionary”, i.e. the computed parameters are continuously updated and refined with time or as new data are available, and it is “probabilistic”, since the computation is based on a probabilistic earthquake location and a Bayesian approach to magnitude evaluation.

Fig. 2 shows a high level diagram of PRESTo, exemplifying the main building blocks of the system and the data flow from ground motion sensors (inputs) to sent alarms (outputs). PRESTo comprises several subsystems, integrated in a single, easy-to-manage execution, which can be compiled for Windows, Mac OS X and Linux.

The core infrastructure is based on five modules, which we will discuss in more details in the next subsections. These are as follows:

1. waveform acquisition and processing;
2. event detection;
3. real-time location;
4. real-time magnitude estimation;
5. peak ground motion prediction at targets.

The software can be easily tailored to different networks and regions, thanks to several configuration parameters and the use of widespread standards for real-time data input and output. The required configuration includes the following: the description of the seismic network (station coordinates, sensor types, IP addresses); region-specific information, like the velocity model for P- and S-waves and the regression laws for magnitude and peak ground motion prediction; control parameters for the different steps of data input, analysis and output.
PRESTo disseminates to vulnerable sites at a distance from the seismogenic areas the information about the earthquake, either through dedicated lines or through the Internet, while the network is still recording it. Since electromagnetic waves travel almost instantaneously when compared to the destructive $S$ and surface waves, which propagate at a speed of around 3.5 km/s, the first alarm can reach the target sites from a few seconds up to tens of seconds before any damage could occur, depending on the hypocentral distance of the alarm destination.

For instance, for a destructive earthquake occurring in the Irpinia region, and a target site in the city of Naples, there is a lead-time of the order of 20 s from when the alarm reaches the target, to when the destructive waves arrive there [6,10]. Such a time lapse can be sufficient to activate several automatic or personal-level safety procedures, e.g. moving elevators to the nearest floor, bringing trains to a halt, stopping gas distribution, prompting people to move to the nearest safe spot, etc. These actions can be very effective in reducing the final damages, be they direct or collateral.

The evaluation of the trade-off between the cost of activating the aforementioned procedures and the foreseeable damage prevention can only be carried out by the recipient of the alarm messages, and thus is not handled within PRESTo. The goal of the system is rather that of providing useful information on the source parameters and expected ground shaking at target sites, including uncertainties, in a very prompt and robust manner.

The problem of real-time hazard mitigation for EEW has been thoroughly studied by Iervolino et al. [13], who also discuss the importance of a trade-off analysis in the design of engineering applications of EEW.

2.1. Implementation details

We chose to implement PRESTo in C++ [14], a well-known programming language that provides optimal speed performances (a key element for an early warning system) without sacrificing the code expressiveness, thanks to its high level, object-oriented nature.

The code is easily portable to different operating systems (Windows, Linux and Mac OS X), thanks to the SDL library (Simple DirectMedia Layer, [15]), used for abstracting low-level operations, and the OpenGL libraries [16], a de facto standard for scientific and interactive visualization.

The software is organized into a main thread that implements the core processing procedures activated during an event, and some additional processing threads that handle the continuous tasks such as data acquisition and waveform analysis.

2.2. Data acquisition

PRESTo makes use of 3-component accelerometric data, provided either as real-time data streams, during normal operation, or as files to playback, during simulation mode.

The real-time ground motion data acquisition is based on SeedLink, a robust and widely used protocol for waveform data transmission [17].

In simulation mode, the input data are stored in files, one for each sensor component, using the SAC format (Seismic Analysis Code, [18]). PRESTo reads these files and converts them into simulated SeedLink data streams, where data are made available in 1-s packets, with an adjustable random delay and controllable random gaps, to simulate network latencies and failures. This operating mode is useful for rapid testing and batch processing, and has been used to run the simulations described in this paper.

An alternative way of performing a simulation test, closer to a real-time scenario, can be achieved by setting up a local or remote SeedLink server, using the SeisComP software [19] and injecting the pre-recorded waveforms in that system with the built-in SeisComP tools.

In both operational modes, each station is handled by its own processing thread, which, for each of the three channels, does the following:

- Keeps a persistent connection with the SeedLink server (in real-time mode) or simulates the incoming data streams.
- Checks the data flow and data quality of each station. Non-working stations need to be ignored by the subsequent processing steps, e.g. the fact they did not record a P-wave arrival must not be used in computing the hypocenter location.
- Performs the automatic P-wave arrival detection on the vertical component.
- Computes the mean over the last seconds of signal and removes it from the signal.
- Stores the incoming accelerations in a buffer for all the other concurrent threads to use.

2.3. Arrivals detection

We make use of a phase detector and picker algorithm optimized for real-time seismic monitoring and earthquake early
warning [20]. The basic concepts of the algorithm are similar to those of the Baer and Kradolfer picker [21] and the Allen picker [22,23]. However this new algorithm is specifically designed to operate stably on continuous, real-time, broadband signals, to avoid excessive triggering during large events, and to produce polarities and realistic time uncertainties on the picks.

The picker is controlled through five configuration parameters that affect the time windows for short and long averages and for filtering, and two thresholds for triggering and de-triggering.

2.4. Arrivals binding

This step analyzes the \( P \) arrivals at every station to determine whether they are coherent with the propagation from a common source.

So far we are using a simple criterion, based on the coincidence of a configurable number of picks (usually 3–5, depending on the network density) within a given time window, which depends on the average station spacing. Though the results are generally satisfactory, we are experimenting with more sophisticated criteria that include the temporal sequence of the arrivals and the geometry of the stations.

2.5. Earthquake location

Earthquake location is triggered at regular intervals after a new earthquake has been declared and whenever a new pick is associated with the event.

The real-time location algorithm exploits, at each time step, both the arrival times computed at the triggering stations, as well as the implicit information that can be derived from the lack of arrival detections at the other stations, which are not yet reached by the \( P \)-waves. This technique, called RTLoc [9], is based on the equal differential time (EDT) formulation [24] and on a fully probabilistic description of the hypocenter.

Given two stations and a velocity model for the subsoil, an EDT surface is an open surface in the Earth (a hyperboloid, for a constant velocity model) whose points are characterized by an equal differential travel time from the two stations. Standard EDT location algorithms (e.g. NonLinLoc. [25]) draw an EDT surface for each pair of stations and search for the hypocenter in the region crossed by the largest number of EDT surfaces (“EDT stack”). This technique is particularly resistant to outliers (false picks), since they will generate an EDT surface that is not compatible with the others and will not contribute to the EDT stack.

As illustrated in Fig. 3, the RTLoc method adds to standard EDT location the information that, at a certain time instant \( t_{\text{now}} \), some stations have not yet been reached by the \( P \) wave (not yet triggered stations). This makes it possible to draw “conditional” EDT surfaces, defined by the condition that, for each couple of triggered and not-yet-triggered stations, the latter will trigger at any time after the current clock time \( t_{\text{now}} \).

The volume bounded by these conditional EDT surfaces provides a useful constraint on hypocentral location when only a few (or just one) stations have triggered. In the very first seconds after the event declaration, the location is dominated by the conditional EDT volume, which provides, in this phase, a better constraint than the EDT surfaces. As the number of triggered stations increases, the stack of “true” EDT surfaces gets more focused and the location converges to a standard EDT location.

At each call, the RTLoc module provides:

- the probability density function for earthquake location, obtained from the stacking of EDT volumes and surfaces;
- the most likely (maximum probability) hypocenter and origin time;
- the covariance matrix, which encodes the spatial uncertainty of the earthquake location.

The algorithm makes use of 3D grids of \( P \) and \( S \) travel times from each possible location to each station. The grids are computed for a 1D or 3D velocity model, using the technique by Podvin and Lecomte [27] for the finite difference solution of the eikonal equation.

To speed up searches over the grid nodes, RTLoc employs an optimized grid walk algorithm (Oct-tree importance sampling algorithm—[25,28]) that uses a hierarchical, recursive partition of the search volume, where the spatial density of sample cells follows the regions of highest location probability.

2.6. Bayesian estimate of the magnitude

The module for real-time magnitude estimation is an implementation of the RTMag technique by Lancieri and Zollo [8].

RTMag makes use of empirical correlation laws between the \( P \) and \( S \) peak displacements (Pds), measured on the first seconds of the low-pass-filtered signal after the phase arrival, and the final earthquake magnitude \( M \). The relationship has the following form [29]:

\[
\log \text{Pd} = A + BM + C \log R / 10
\]

where \( R \) is the hypocentral distance of the station (in km). Coefficients \( A \), \( B \) and \( C \) are pre-determined from a regression analysis, and depend on the phase (\( P \) or \( S \)) and on the length of the considered time window, generally 2 or 4 s for the \( P \) phase (denoted as 2P and 4P), and 1 or 2 s for the \( S \) phase (1S and 2S).

Fig. 4 shows the regression of \( \log (\text{Pd}) \) (normalized to a reference distance of 10 km) vs. the final magnitude, which has been derived by Zollo et al. [29] using 376 records from the European Strong-Motion Database [30].

The system starts measuring the \( Pd \) after the detection of an event and its first location. For each measurement of \( Pd \), the corresponding magnitude is described as a probability density function (PDF) with a normal shape, whose average is calculated through Eq. (1) and whose standard deviation depends on the error on coefficients \( A \), \( B \) and \( C \) and on the uncertainty on distance \( R \) (provided by RTLoc). In accordance with Lancieri and Zollo [8], the PDF associated with the 2P measurements is constant above \( M = 6.5 \), to take into account the “saturation effect” observed for the corresponding regression relationship.

At each time step, the magnitude distributions for every station and time window are combined through a likelihood product. Following a Bayesian approach, the magnitude PDF retrieved at the previous time step is taken into account as “a priori” information. At the first time step the “a priori” can be optionally provided by the Gutenberg–Richter relationship.

The resulting earthquake magnitude is assigned by finding the value with the highest probability in the final distribution. The magnitude uncertainty corresponds to the magnitude range for which the probability density integral varies from 5% to 95%.

This magnitude estimation gets continuously updated to keep track of new arrivals, and new packets of signal from the stations, in addition to being updated whenever the location or uncertainty from the previous step changes.

For each triggered station:

- The theoretical \( S \)-waves arrival is computed using the most recent hypocenter, the \( S \)-waves travel-time drawn from the grids, and the automatically detected \( P \)-waves pick.
If the $P$ and $S$ arrivals are more than 2 s apart, the peak ground displacement on the 2 s following the $P$ pick ($2P$) is computed, as soon as the signal is available.

Likewise, if more than 4 s separate the $P$ and $S$ arrivals, the peak ground displacement on the 4 s following the $P$ pick ($4P$) is measured.

As soon as 2 s of acceleration after the theoretical $S$-waves arrival is available ($2S$), the peak displacement is measured for this time window too.

Each peak displacement is computed from the three-component acceleration signal, by performing a configurable band pass filter (default: 0.075–3 Hz) and a double integration of the acceleration, and finding the maximum of the vector modulus in that time window. The regression parameters of Eq. (1) and the Gutenberg–Richter coefficients are user-configurable.

2.7. Peak ground shaking at target sites

The expected peak ground acceleration and velocity (PGA and PGV), and instrumental intensity ($I$) that will be experienced at each configured target site are computed using region-specific
ground motion prediction equations, which depend on magnitude and epicentral distance.

Following the time evolution of location and magnitude estimates, the computation of PGA, PGV and $I$, and of their uncertainties, is evolutionary as well, i.e. these parameters are updated whenever the estimates of location and/or magnitude change.

The prediction laws are expressed in a generalized parameterized form to allow the needed configurability.

### 2.8. Alarm messages

In real-time mode, during the propagation of the seismic waves of an energetic event, the evolutionary estimates of location, magnitude and peak ground motion at a distance are communicated to a list of recipients.

A simplified and smaller format is available for less sophisticated receiving devices, while a more verbose, but highly structured and easily upgradable format uses QuakeML-RT [31], a standard for exchanging information on seismic events in real-time. When using the Internet for alarm transmission it is possible to opt for either UDP or TCP as transport layer, to be chosen by evaluating the trade-off between speed and reliability. Quasi-real-time alarms can also be generated in the form of text messages sent over the cellular network and e-mails.

User configurable parameters specify the sites to alert and their details and, for each site: the alarm format and transmission protocol to be used; the threshold above which to send an alarm (ground motion); the minimum variation of the previous parameter that must trigger an alarm with the updated information.

### 2.9. Graphical output

A concurrent (optional) thread of execution is in charge of displaying on screen the input data and output computations of PRESTo. This feature is useful during debugging and simulations, for presentations or in a monitoring center. It is also possible to enable the generation on disk of an animation, containing screenshots taken at a constant rate during the occurrence of an earthquake, and/or of a final screenshot. These files can be used as an additional working log of the system or be published in web reports. Fig. 5 shows an actual screenshot of PRESTo, taken after

![Fig. 4. Correlation between the logarithm of low-pass-filtered peak ground displacement $P_d$ and moment-magnitude $M_w$. The values of $P_d$ are normalized to a reference distance of 10 km. The regressions are computed measuring $P_d$ on 2-s time windows after $P$-arrival (left), and on 1-s (middle) and 2-s (right) time windows after $S$-arrival. The $P$-wave displacement is measured on the vertical component. The $S$-wave displacement is measured on root-squared sum of the horizontal components. Each panel shows the best-fit regression line (solid line) along with 1-WSE limits (dashed lines) (redrawn from [29])](image-url)

the playback of synthetic traces for the 1980 Irpinia earthquake [10]. Shown on screen are:

- The vertical acceleration at every station as seismograms. The background color indicates a problem with the data, with a dark tint given to stations characterized by a low quality signal, or that exceeded a timeout with no data received, and a lighter color highlighting stations with less severe problems such as lagging signals or data gaps.

- $P$ picks are overlaid to the seismograms as vertical red bars. During an event the picks involved in its processing are evidenced. Additionally, the 2/4 s windows used for magnitude estimation are marked in yellow ($P$ waves) or red ($S$ waves).

- A georeferenced map of the network and location grid, and a cross section of the crust. On these maps several elements are reported:
  - stations, displayed as labeled triangles, with color and overlaid icons indicating the quality of signal and the triggering state;
  - during an earthquake, the epicenter and hypocenter depth are shown as icons, labeled with the magnitude estimate;
  - around the location of the earthquake the uncertainty of the position is indicated by an ellipse, the projection of the uncertainty ellipsoid on the map plane;
  - the theoretical $P$-waves and $S$-waves propagation from the epicenter as circular wave fronts;
  - target sites, with the number of seconds remaining before the arrival of the destructive waves, and the predicted peak acceleration.

- The temporal evolution of the magnitude estimate and uncertainty, as a graph. Also marked on this graph is the computed origin time.

A set of parameters is provided to modify the graphics option or to turn graphics off for an additional performance gain.

### 2.10. Logs and reports

Detailed logs of the inner workings of PRESTo are written to files while the system runs. They help to pinpoint the source of eventual problems, by storing the timeline of every computation, which can be extracted to generate statistics on the performance of the whole system or of one of its modules. This is how the data on the performances of PRESTo during the test runs analyzed in

![Diagram](image-url)
this paper were generated. Since the main log files tend to grow unwieldy, some more focused and human readable web pages are also generated, along with simplified logs that are included, for instance, in mail messages. Another form of graphical logging is provided as a KML animation[32] written on disk after an event has ended. This file can be played back through web browser plug-ins or applications supporting KML files such as Google Earth[33]. They contain the propagation of the seismic waves from the final estimate of the hypocenter, the stations as they trigger, and the evolution of the source parameters and alarms as computed by PRESTo. See Fig. 6 for a frame of the generated KML animation rendered in Google Earth 5.

3. Testing on large earthquakes worldwide: the Mjma 7.2 (Mw 6.9) 2008 Iwate earthquake, Japan

One of the key features of our software platform for earthquake early warning is its easy adaptability to different networks and seismogenic regions.

This has proven to be particularly important for the development of the system because, though we are actively experimenting PRESTo at the ISNet, no moderate or large earthquake (M > 4) occurred in the southern Apennines region since our network became fully operative (see Fig. 1). We therefore decided to follow two distinct strategies, testing PRESTo on synthetic and real data.

The PRESTo application on synthetic data has been recently published by Zollo et al.[10]. It consists in an extensive synthetic test involving ~500 simulated events for three large earthquake mechanisms in the southern Apennines region.

The validation of PRESTo on real data has been performed using: (a) large earthquakes worldwide, recorded at networks with features similar to the ISNet (in terms of geometry, type of sensors and acquisition) and (b) small earthquakes recorded at the ISNet. The latter analysis will be illustrated in the next section. Here we present, as a case study, a simulation of the Mw 6.9, June 2008, Iwate earthquake, in Japan, from the playback of actual recorded waveforms.

The Iwate-Miyagi-Nairiku earthquake occurred on June 14, 2008, in the northern part of Japan (Fig. 7, top). The magnitude in the Japanese Meteorological Agency (JMA) scale is Mjma=7.2, while the moment magnitude has been fixed to Mw=6.9. The earthquake killed more than 20 people, while 450 were injured; about 2000 houses were damaged[34].

The K-Net and KiK-Net networks, operated by the National Research Institute for Earth Science and Disaster Prevention (NIED), provide an excellent data set for this earthquake, with 30 accelerometric stations within 60 km from the epicenter (Fig. 7, top). The mean station spacing is ~20 km, slightly larger than that of ISNet (~10 km). However the station density ensures that 3 stations are triggered within 5 s from the event origin.

We performed a playback of the recorded strong motion traces into PRESTo, in order to assess the reliability and speed of convergence for location and magnitude estimates, and to evaluate the accuracy of the ground motion prediction at selected sites.

The map view in Fig. 7 shows the available lead-time, as a function of the distance from the reference epicenter. The reference location is indicated by a star and has been calculated using accurate P and S manual pickings and the 1D velocity model employed for routine locations at the JMA[35]. The maximum theoretical lead-time at increasing distance from the reference epicenter is indicated on the map by the circles, and is defined as the difference between the S arrival at a given distance and the time at which the first magnitude estimate is available.

The yellow shaded zone is the “blind zone”, or the region where no lead-time is available, i.e. the lead-time is negative. In this area no safety actions, based on a regional EEW system, can be performed.
Depending on the specific application, the effective lead-time can be smaller, and is determined by the level of accuracy and robustness of the warning required by the recipients. For instance, in an application that requires the location and the magnitude to be stable for at least three measurements before a warning is issued, the effective lead-time would be about 5 s smaller than the maximum lead-time (see Fig. 8). The corresponding larger blind zone is indicated in orange on the map.

The bottom part of Fig. 7 shows the module of velocity recorded by the horizontal components at two sample stations. The observed peak ground velocity is compared with the prediction of PRESTo, which evolves with time. A quantitative analysis of the results of the computation is given in the ‘‘PRESTo Timeline’’ graph, shown in Fig. 8, which is a visual representation of the system log. The graph is divided into two parts, representing the input data (lower part) and the output parameters (upper part).

Inputs are of two kinds: (1) $P$-phase picks for the phase association algorithm and the real-time location module, and (2) $P$ and $S$ waveform signals from which peak displacements are measured and magnitude is estimated. In particular, the RTMag module implements regression relationships for the peak displacement ($P_d$) measured on 2 s of $P$ signal ($2P$), 4 s of $P$ signal ($4P$) and 2 s of $S$ signal ($2S$). Therefore we count, as a function of the seconds elapsed since the earthquake origin: (plot 1, from the bottom) the number of available picks and (plots 2, 3 and 4) the number of available $2P$, $4P$ and $2S$ time windows.

The evolution of output parameters is shown in the upper part of the plot. Earthquake location is represented as an error, or discrepancy, between estimated and reference depths (depth error, $DE$—plot 5) and estimated and reference epicentral locations (epicentral error, $EE$—plot 6). An uncertainty is associated with each point, as calculated by the location module. The time evolution of magnitude (plot 7) is reported with the associated uncertainty, defined as the confidence interval between 5% and 95%. The $P_d$ regression for Japan has been calculated as a function of the $M_{jma}$ magnitude [8]. Therefore the value $M_{jma}=7.2$ is given as reference.

The last three plots (8, 9 and 10) show the time evolution of the predicted peak ground acceleration (PGA) at three sample stations: AKT019 and AKT016 (also displayed in Fig. 7) and IWT025 (close to the epicenter). The predicted values ($PGA_{EST}$) are compared to the measured values ($PGA_{OBS}$) in terms of prediction error [10]:

$$PE = \log\left(\frac{PGA_{EST}}{PGA_{OBS}}\right)$$

the associated uncertainty is that on $\log(PGA_{EST})$, as provided by the ground motion prediction relationship. Here we use the attenuation law for strong ground motion in Japan derived by Kanno et al. [36].

From the analysis of the PRESTo timeline it is possible to understand how the available information from the seismic network grows with time, and how quickly the system can produce stable and reliable estimates of the earthquake parameters and the ground motion, as a function of time and amount of data.

The chosen criterion for declaring an event is to have at least 3 picks within 5 s. This condition is reached 5.03 s after the event origin, when stations IWT025, IWT026 and IWT011 trigger. The first location is available 0.3 s later, with an epicentral discrepancy, with respect to the reference location, of about 7 km, and an estimated hypocenter that is about 8 km deeper than the reference one. As explained in Section 2.5 the location is updated each second, or whenever new picks are available. One second after the first location, even though no other station has triggered,
Fig. 7. (Top) Map of the 2008 Mw 6.9 (Mjma 7.2) Iwate earthquake epicenter and of the recording stations from K-Net and KiK-Net. The star is the reference location. The circles indicate the maximum lead-time, defined as the difference between the S-arrival time and the time of the first available magnitude estimate; the yellow striped area is the corresponding blind zone. The orange area is the blind zone for a sample application that requires a more stable magnitude estimate (see text). (Bottom) Module of velocity recorded by the horizontal components at two sample stations (the peak value is marked by the red dot), compared with the time-evolving PGV estimate from the system.
the location improves, thanks to the further constraint of not-yet-triggered stations.

At the same time of the first location, the first magnitude estimate is given, since 2 s of $S$ signal is available at the nearest station (IWTH25). It is not possible to use the $P$ signal at this station, because the $P$ wave time window, before the $S$ arrival, is shorter than 2 s. Surprisingly, the estimated magnitude is exactly the reference value ($M_{jma} = 7.2$); however this is rather a combined effect of an underestimation of the 2S peak recorded at IWTH25 and the error on hypocentral location, which places the origin at a greater distance to that station, compared to the reference location. This is confirmed by the fact that after one second the location improves, while no additional Pd measurements are available yet, thus yielding a magnitude that is slightly underestimated ($M_{jma} = 7.0$).

The error on magnitude gets smaller when another 2S Pd measurement is available (7.4 s after the origin, at IWTH26 station). The first 2P measurements are made 8.1 s from the origin.
Fig. 9. Aggregate PRESTo timeline for 28 small magnitude earthquakes (Ml < 3.5) recorded at the ISNet (see Fig. 8 for the description of the plot). Each gray curve represents the time evolution of the corresponding parameter for a given earthquake. The solid curves are the average values.

Fig. 10. Stability of convergence of the real-time estimates of source parameters for the 28 small earthquakes recorded at the ISNet analyzed in this study. Each parameter (depth Z, epicenter E, magnitude M) is compared to the final value estimated by the system (Zf, Ef, Mf).
As more Pd measurements are available, the estimated magnitude settles at $M_{\text{Jma}} = 7.0$ and the uncertainty decreases.

Since the magnitude value is reasonably well estimated (the discrepancy is below 0.3 magnitude units), the PGA prediction error for stations IWTH25, AKT019 and AKT016, reported in the upper three plots, basically reflects how well the ground motion prediction relationship [36] reproduces the observed peak values.

For station IWTH25, 2.5 km from the epicenter, the prediction error is negative, i.e. the estimated PGA is lower than the observed one. This reflects the larger dispersion of peak values around the theoretical curve at short distances from the fault. This dispersion is however taken into account by the error on the attenuation law.

For stations at larger distances, the estimated PGA values are close to the observed peaks.

4. Performance tests at the Irpinia Seismic Network

We are testing PRESTo on real data at the ISNet using the small magnitude earthquakes ($M_I < 3.5$), recorded by the network. Given the network density, an MI 1.5 earthquake is recorded on average by 10 stations, while an MI 2.5 event is generally recorded by 20 stations. This makes it reasonable to test the promptness of detection on such small events (i.e. the time it takes for 3 stations to trigger) and the convergence and stability of the earthquake location. On the other hand the performance of the magnitude computation is only indicative, since for this class of magnitude the full waveform is recorded in a few seconds, and the magnitude estimation is therefore deterministic, rather than predictive.

**Fig. 11.** PRESTo timeline (top) and real-time location plot (bottom) for a Mw 3.0 earthquake inside the network. In the bottom plot, the star indicates the reference location. Non-operative stations are grayed out. $\Delta t$ is the time from the origin at which the location computation ended, while the time at which the computation started is in parenthesis.
We took for the study all the local events that have been automatically detected by the system from November 2008 to March 2010. They are 28 earthquakes with $M_l$ between 1.6 and 3.3. The system correctly detected all the 17 events with $M_l \geq 2.5$ during this period, while the detection rate is of about 10% for the smaller events with $1.5 \leq M_l < 2.5$. During the same period a false event has been declared, with an estimated $M_l$ of 2.9, and located about 20 km SW of the network.

Most of the studied earthquakes are located within the ISNet network (see Fig. 1). Six events (five to the South and one to the North) are slightly outside the network, i.e. the distance to the nearest station is less than 10 km, but the azimuthal gap is larger than 180°; two more events to the North are located at about 20 km from the nearest station.

Fig. 9 shows, as an aggregate plot, the PRESTo timelines for all the analyzed events. In each graph, the gray curves represent the time evolution of the corresponding parameter, while the solid line is the average trend. Differently from Fig. 8, in order to compare non-homogeneous magnitude values, here we indicate the magnitude error ($ME = M - M_0$), defined as the deviation from the reference value ($M_0$) obtained from the network bulletin.

Considering the averages, three picks are generally available within 4–5 s from the origin time. This is the number of arrivals required by our criterion for phase association and event detection. Therefore a first location is usually available within 4.0–5.5 s, because the time required by the location algorithm is usually below 0.5 s.

The hypocentral locations for the events inside the network are generally well constrained starting from the very first estimates. For the events outside the network the azimuth is well determined, but there is typically a larger uncertainty on the distance, as discussed in the example below.

Since we are considering small events, the magnitude is already well constrained during the first seconds. In fact 2 s of signal generally contains the final peak value; hence the magnitude estimate is deterministic. For the events outside the network, for which the location discrepancy can be as large as 15–20 km, the magnitude error stays within 0.5–0.6 magnitude units, which is reasonable for an early warning application.

Fig. 10 shows the stability of convergence for the source parameters. In this plot each value is compared to the final estimate of the system. Most of the locations are stable since the very first seconds after the event detection. The hypocentral locations for the events outside the network (see Fig. 1) need more seconds to stabilize, since the hypocenter is generally placed closer to the network when only a few triggered stations are available (see also the example below). Consequently, for these events, the magnitude (which depends on the logarithm of the hypocentral distance) stabilizes later.

Figs. 11 and 12 show the timelines and the location maps for an $M_3$ event inside the network and for an $M_2.6$ earthquake outside the network. The location map is a plot of the location probability provided by RTLoc. Each map is a snapshot at a certain time $\Delta t$ from the event origin. The reference location (obtained by the manually revised bulletin) is indicated as a star and the triggered stations are circled; non-available stations (which do not contribute to the computation) are grayed out.

The event inside the network (Fig. 11) is declared 3.99 s after the origin, when three stations have triggered (STN3, CGG3, PGN3). The first location is already quite consistent (epicentral error and depth error of about 5 km) and the magnitude is well estimated. As time progresses, the uncertainty on location and magnitude estimation decreases. The last two graphs on the upper part of the “PRESTo Timeline” are the prediction error on PGV for stations SCL3 (20 km far from the epicenter) and SNR3 (43 km far). The employed ground motion prediction equation is the “Small” regression, extracted from the ShakeMap software package [37]. As in the Japan example, since the magnitude estimate is quite consistent, the prediction error on PGV basically reflects the accuracy of the attenuation relationship.
In the second example (Fig. 12), we analyze an event outside the network. Two stations close to the epicenter (STN3, PCR3) were not operative when this earthquake occurred. The closer station, PGN3, correctly picked the \( P \) arrival, but it has not been associated with the event since the subsequent arrivals (at CGG3 and VDP3) are recorded more than 3 s later. Therefore the event is declared with the \( P \)-picks of the next three working stations (CGG3, VDP3, SRN3), 7.38 s after the origin time. As a consequence of the missed pick from station PGN3 the location is initially shifted towards CGG3. After 1.27, 8.65 s from the origin, more picks are available (including an outlier: COL3) and the location is now closer to the reference hypocenter. However the epicenter is shifted about 15 km away from the network, which implies that the magnitude is overestimated. Nevertheless, since the magnitude depends on the logarithm of the distance, the corresponding discrepancy is limited to 0.3 magnitude units. The prediction error on PGV for stations SCL3 and SNR3 is reported in the last two plots.

4.1. Data latency

A key factor to take into account for an EEW system is the data latency due to the communication system and protocol. Given the size of the data packet produced by the data-logger, we can define

\[ \text{Time from origin (s)} \]

\[ \text{N. of stations} \]

\[ \text{0} \quad 10 \quad 20 \]

\[ \text{Depth Error} \]

\[ \text{Epicentral Error} \]

\[ \text{Magnitude} \]

\[ \text{PGV Prediction Error} \]

\[ \text{PE} = \log(\text{PGV}_{\text{exp}}/\text{PGV}_{\text{opt}}) \]

\[ \text{stat: SCL3, R: 36.9 km} \]

\[ \text{stat: SNR3, R: 62.7 km} \]

\[ \text{M}=2.6 \quad 2008–11–17 (14160r) \]

\[ \Delta t \]

Fig. 12. PRESTo Timeline (top) and real-time location plot (bottom) for a Mw 2.6 earthquake outside the network. In the bottom plot the star indicates the reference location. Non-operative stations are grayed out. \( \Delta t \) is the time from the origin at which the computation ended, while the time at which the computation started is in parenthesis.
two types of latencies:

- minimum latency: the difference between the time at which a data packet is received and the timestamp of the last sample of the packet.
- maximum latency: the difference between the time at which a data packet is received and the timestamp of the first sample of the packet. It is equal to the minimum latency plus the length of the packet.

We measured the latency for more than 1,000,000 packets received from the network stations. The minimum latency is very stable (Fig. 13), with a modal value of 0.8 s. The data loggers employed in ISNet generate 1-s packets [5]; therefore the maximum latency is 1 s larger than the minimum latency.

These values are reasonable when compared with the detection times of the network.

5. Discussion and conclusions

In its current implementation, the prototype system for earthquake early warning in southern Italy uses a regional approach, built upon the Irpinia Seismic Network (ISNet), a modern, high-density and high-dynamic network, deployed around the fault system responsible for the latest strong earthquake in the region (Mw = 6.9, 1980).

Following the regional early warning paradigm, several procedures have been developed, in the last 4 years, for real-time earthquake detection, characterization of source parameters, and prediction of the expected ground shaking. These algorithms have been recently synthesized into an integrated software package, called PRESTo (PRobabilistic and Evolutionary early warning SysTem).

PRESTo has been developed keeping portability and flexibility in mind. The software is written using platform-independent libraries and can run on the major operating systems. Moreover it uses widespread seismological standards, like SeedLink, SAC and QuakeML, and is largely configurable, in order to be adapted to diverse seismic networks and seismogenic regions worldwide.

These features were crucial for the validation of the system, because no moderate or large earthquake (M > 4) has occurred in southern Italy since PRESTo has been operational. While we partly overcame this limitation by simulating several strong earthquake scenarios in the region [10], we needed to follow a double strategy in order to test the system on real earthquakes.
On the one hand, we virtually “installed” PRESTo in active seismogenic areas worldwide, where seismic networks with features similar to ISNet are available. Using waveform playback, we can inject pre-recorded traces into the system, simulating their real-time acquisition by the network, and thus check the system performances in terms of rapidity, stability and accuracy of the earthquake source estimates and of the expected ground shaking.

A test conducted using waveform data from a Mw 6.9 earthquake, recorded at the Japanese networks K-Net and KiK-Net, shows that, when a dense seismic network is deployed in the epicentral area, the system converges to a stable and accurate determination of location and magnitude within 1–2 s from event declaration. The quality of ground shaking prediction depends, as expected, on the accuracy of the employed ground motion prediction equation, and is influenced by the point source approximation made.

In the near future we will continue investigating the performances of PRESTo on large earthquakes worldwide, using the playback approach or, if possible, in real-time, by physically installing the software in the computing facilities of other networks.

A complementary type of tests has been performed using small earthquakes (M < 3.5) recorded at the ISNet. These tests are significant for what concerns the geometrical aspects: phase association and earthquake detection, and the stability and the accuracy of the hypocentral location. However they are less indicative for magnitude estimation, since the time window in which the peak displacement is measured generally includes the whole signal, thus making the estimate deterministic, rather than predictive.

The results indicate that the detection is fast, it generally requires 4–5 s from the event origin (time at which at least three picks are available), and the location is stable, from the very first seconds. However there are some “pathological” cases where the employed phase association algorithm has proven to be too simplistic. An example with an earthquake slightly outside the network has shown that, if a few stations close to the epicenter are not operational, the network geometry varies drastically (the mean station spacing increases) and the association criterion of having at least three picks within three seconds may fail. In this particular example the seismic event has been declared using the next three triggering stations, biasing the location during the first seconds.

During 17 months of operation, PRESTo correctly detected all the local earthquakes with Ml ≥ 2.5, but also declared a false event with Ml 2.9, which, again, is probably due to a phase association algorithm which is too crude. More work is therefore required on improving this aspect, using additional criteria on the distance between the early triggering stations and/or implementing a basic location as the association step, like for instance in the Earthworm “binder” [38].

The records of the retrieved magnitude and peak ground motion estimates shows, as expected, that a discrepancy on location of the order of 10–15 km has a relatively low effect on the magnitude value (± 0.3–0.4 magnitude units). Nevertheless the effect on the predicted PGV or PGA can be more important. However, the probabilistic information on hypocentral location and magnitude can be employed to fully characterize the uncertainty associated with the ground motion estimates, and carry out a real-time trade-off analysis before performing any critical safety action.

Future directions include the development of an integrated regional and on-site early warning approach, to estimate the earthquake damage potential from near- and far-source measurements of the peak displacement Pd and the period parameter τc [39] on the early P-wave signals [40]. A decision table based on thresholds for the two parameters may be used to issue an alert level depending on whether the potential damaging earthquake has occurred nearby or far away from the recording site.

Finally, integrating this kind of analysis with the real-time location provided by RTLoc and measurements of Pd and τc at sites located at increasing distances from the source would make it possible to provide an estimate of the potential damage zone within 2–3 s from the earthquake origin, increasing the lead time and reducing the blind zone.

Acknowledgments

This research was supported by the Italian DPC-S5 project. Part of the work has been carried on within the SAFER project (Seismic Early Warning for Europe), founded by the European Community via the Sixth Framework Program for Research.

Some of the earthquake traces used in this study were recorded at KiK-Net and K-Net by NIED, and downloaded through their websites (http://www.kik.bosai.go.jp and http://www.k-net.bosai.go.jp).

References