A Bayesian approach to the real-time estimation of magnitude from the early *P* and *S* wave displacement peaks

M. Lancieri[1] and A. Zollo[2]

Received 18 September 2007; revised 2 August 2008; accepted 7 August 2008; published 3 December 2008.

[1] It has been shown that the initial portion of *P* and *S* wave signals can provide information about the final earthquake magnitude in a wide magnitude range. This observation opens the perspective for the real-time determination of source parameters. In this paper we describe a probabilistic evolutionary approach for the real-time magnitude estimation which can have a potential use in earthquake early warning. The technique is based on empirical prediction laws correlating the low-frequency peak ground displacement measured in a few seconds after the *P* and/or *S* phase arrival and the final event magnitude. The evidence for such a correlation has been found through the analysis of 256 shallow crustal events in the magnitude range M_{jma} 4–7.1 located over the entire Japanese archipelago. The peak displacement measured in a 2-s window from the first *P* phase arrival correlates with magnitude in the range $M = [4–6.5]$. While a possible saturation effect above $M \simeq 6.5$ is observed, it is less evident in an enlarged window of 4 s. The scaling of *S* peaks with magnitude is instead also observed at smaller time lapses (i.e., 1 s) after the first *S* arrival. The different scaling of *P* and *S* peaks with magnitude when measured in a 2-s window is explained in terms of different imaged rupture surface by the early portion of the body wave signals. We developed a technique to estimate the probability density function (PDF) of magnitude, at each time step after the event origin. The predicted magnitude value corresponds to the maximum of PDF, while its uncertainty is given by the 95% confidence bound. The method has been applied to the 2007 (M_{jma} = 6.9) Noto Hanto and 1995 (M_{jma} = 7.3) Kobe earthquakes. The results of this study can be summarized as follows: (1) The probabilistic algorithm founded on the predictive model of peak displacement versus final magnitude is able to provide a fast and robust estimation of the final magnitude. (2) The information available after a few seconds from the first detection of the *P* phase at the network can be used to predict the peak ground motion at a given regional target with uncertainties which are comparable to those derived from the attenuation law. (3) The near-source *S* phase data can be used jointly with *P* data for regional early warning purposes, thus increasing the accuracy and reliability of magnitude estimation.

Citation: Lancieri, M., and A. Zollo (2008), A Bayesian approach to the real-time estimation of magnitude from the early *P* and *S* wave displacement peaks, *J. Geophys. Res.*, *113*, B12302, doi:10.1029/2007JB005386.

1. Introduction

[2] The final aim of an earthquake early warning system (EEWS) is to evaluate in real-time the probability of losses and to activate automatic and/or manual procedure to mitigate earthquake effects. An EEWS is a modern real-time information system able to provide rapid notification of an occurring earthquake through the fast telemetry and processing of data from dense instrument arrays deployed in the source region of the event (regional) or surrounding the target infrastructure (site-specific) [*Kanamori*, 2005; *Nakamura*, 1988].

[3] A site-specific EEWS consists of a single sensor or an array deployed in the proximity of the target site, and whose measurements of amplitude and predominant period on the initial *P* wave motion [*Wu and Kanamori*, 2005a, 2005b] are used to predict the ensuing peak ground motion (mainly related to the arrival of late *S* and surface waves) at the same site. A regional EEWS is based on a dense sensor network covering a part of or the entire seismic source area. The relevant source parameters (event location and magnitude) are estimated from the early portion of the recorded signals and are used to predict, with a quantified confidence, a ground motion intensity measure at a distant location where a target site of interest is located. The success of a regional EEWS critically depends on the capability to provide in a

[1]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Napoli, Naples, Italy.
[2]Dipartimento di Scienze Fisiche, Università di Napoli Federico II, Naples, Italy.

0148-0227/08/2007JB005386$09.00

fast, secure and reliable way the size and location of an earthquake while it is still occurring [*Kanamori*, 2005].

[4] Standard techniques used to determine the source parameters of an earthquake are generally not suited for early warning applications, since they typically use the complete event waveforms from all the stations of a local seismic network, with a consequent large time delay for data transmission and analysis. For this reason, a different strategy is required, where the computation starts when a few seconds of data from a small number of recording stations is available, and the source parameters estimates are progressively updated with time.

[5] Assuming that a dense seismic network is deployed around the earthquake source zone, an evolutionary estimation of source parameters requires that earthquake location and magnitude, and their associated uncertainty, are determined as a function of time, i.e., of the number of recording stations and of the length of ground motion signal recorded at each station.

[6] Recently, robust and effective algorithms for the real time and evolutionary event location have been developed [*Horiuchi et al.*, 2005; *Satriano et al.*, 2008; *Cua and Heaton*, 2007]. In particular, *Cua and Heaton* [2007], generalized the concept of not yet arrived data, described by *Horiuchi et al.* [2005] in order to start the location with one single triggering station.

[7] *Satriano et al.* [2008] further extended and generalized the *Horiuchi et al.* [2005] approach by developing a fully probabilistic technique where the hypocenter is described as a PDF instead of as a point. Their technique makes use of the equal differential time formulation [*Font et al.*, 2004] to incorporate both triggered arrivals and not yet triggered stations and applies a nonlinearized, global search for each update of the location process.

[8] On the other hand the feasibility of determining in real time the final earthquake magnitude using the information contained in the initial portion of the recorded *P* and *S* signals is matter of investigation and debate [*Olson and Allen*, 2005; *Rydelek and Horiuchi*, 2006; *Rydelek et al.*, 2007; *Hellemans*, 2006; *Zollo et al.*, 2006, 2007].

[9] In the ElarmS methodology [*Allen*, 2007] the magnitude is rapidly estimated using the frequency content in a window of 4 s of signal after the first *P* wave arrival. Founded on a method described first by *Nakamura* [1988], the predominant period, τ_p of the vertical component waveform, is based upon a recursive calculation of squared ground velocities and displacement

$$\frac{1}{\tau_p} = \frac{1}{2\pi}\sqrt{\frac{\int_0^{t_0} \dot{u}^2(t)dt}{\int_0^{t_0} u^2(t)dt}} \tag{1}$$

where u is the vertical component of displacement; and t_0 is the time window duration (i.e., 4 s). The maximum τ_p value within 4 s is found to scale with event magnitude [*Allen and Kanamori*, 2003; *Olson and Allen*, 2005; *Wu and Kanamori*, 2005a].

[10] An alternative to the predominant period is the analysis of low-frequency peak displacement measured in a few seconds after the *P* phase arrival, PD. *Wu et al.* [2006] investigated the attenuation relationship between PD, the hypocentral distance and the magnitude (estimated as local magnitude) on southern California [*Wu et al.*, 2006; *Wu and Zhao*, 2006] and Taiwan events.

[11] On the basis of the analysis of the Mediterranean, near-source strong motion database, *Zollo et al.* [2006] showed that the peak displacement amplitude of initial *P* and *S* scales with magnitude in the moment magnitude range $4 \leq M_\omega \leq 7$. This relationship may be of potential use for EEWS application since the earthquake size can be, therefore, estimated using only a couple of seconds of signal from the *P* or *S* wave onset, i.e., while the rupture itself is still propagating and before the final rupture dimension is achieved. The potential use of *S* waves for early warning purposes is one of main implications of the results of *Zollo et al.* [2006, 2007], despite that their use for early warning purposes is usually excluded [e.g., *Rydelek et al.*, 2007].

[12] In this work we develop a Bayesian real-time magnitude algorithm based on empirical prediction laws correlating the low-frequency peak displacement measured on the initial portion of *P* and *S* wave of strong motion signals with the final earthquake magnitude.

[13] This correlation has been pointed out by *Zollo et al.* [2006] for the Euro-Mediterranean earthquake records and then confirmed on Japanese earthquakes recorded by the K-Net and KiK-Net strong motion network [*Zollo et al.*, 2007].

[14] At any time *t*, after the first event detection the conditional probability density function (PDF) of magnitude given the observed data vector is expressed, through the Bayes' theorem, as the product between the conditional PDF of data, given the magnitude, and a priori PDF based on the prior information available before the time *t*. In the proposed methodology the a priori distribution is derived by the Gutenberg-Richter occurrence law, similarly to *Iervolino et al.* [2006] and *Cua and Heaton* [2007].

[15] In order to investigate the feasibility of early warning in Japan using the proposed approach, we have applied it to two large events that occurred recently in Japan (Kobe 1995 and Noto Hanto 2007 [*ABS Consulting*, 2007]) and whose records have not been used to determine the prediction law.

2. Peak Ground Displacement and Earthquake Magnitude

2.1. Empirical Correlation Law for Japan

[16] With the aim to determine an empirical correlation law between early peak displacement and magnitude for shallow crustal Japanese earthquakes, we analyzed the strong motion records from 256 Japanese events, with a maximum depth of 50 km and magnitude ranging from 4 to 7.1 distributed over the entire Japanese archipelago (Figure 1).

[17] The data are preliminary grouped into magnitude bins of width 0.3, which approximates the standard error on magnitude determination [*Katsumata*, 1996]. The magnitude bin grouping has the additional advantage to average out possible effects on peak amplitudes due to the fault mechanism and rupture directivity.

[18] For earthquake waveforms extracted from K-Net and KiK-Net catalogues, the information concerning the event locations and magnitudes (M_{jma}) are provided by the Japan Meteorological Agency (JMA). For earthquakes shallower

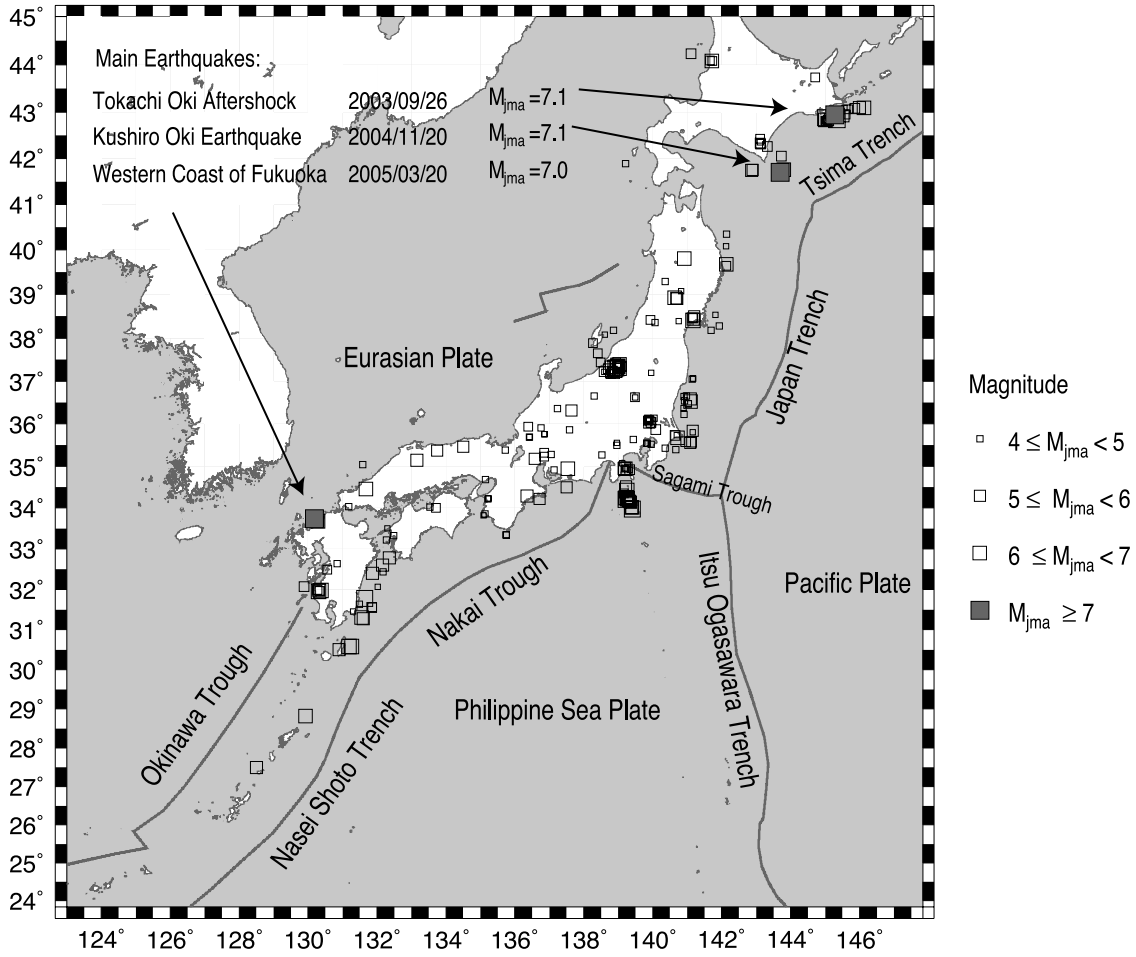

Figure 1. Geographic distribution of the selected event on the Japanese archipelago. The grey lines refer to plate boundaries. The Pacific plate is subducting from the east beneath the North American and Eurasian plates in eastern Japan; the Philippine Sea plate is descending from the south beneath the Eurasian plate in southwest Japan. Large intraplate earthquakes occur along the plate boundaries off the Pacific coast of the Japan islands. Intraplate earthquakes within the continental plate take place in the upper crust beneath the Japan islands and along the coast of the Sea of Japan. Although the crustal intraplate earthquakes do not occur as frequently as the intraplate earthquakes, they generally inflict greater damage because they are shallow and near densely populated areas [*Zhao et al.*, 2000].

than 60 km M_{jma} is determined by *Tsuboi*'s [1954] relationship.

[19] By a comparison between the M_{jma} and moment magnitude M_ω retrieved from the CMT solution [*Dziewoski et al.*, 1981], *Katsumata* [1996] found negligible differences between the two units for shallow earthquakes in the magnitude range 5 to 7. Hereinafter we simply indicate magnitude as M.

[20] A total of 2640 records, with hypocentral distances HD < 60 km have been extracted from the K-Net seismic network. We integrated this waveform database for events with $M > 6.1$ using 68 additional records from the KiK-Net borehole strong motion network.

[21] The distance distribution of analyzed records for Japan earthquakes [*Zollo et al.*, 2007, Figure 1] is rather different from that for the Euro-Mediterranean (MED) earthquakes in the work by *Zollo et al.* [2006, Figure 1]. Whereas most of MED records occur at an HD < 25 km, most Japanese records are at HD > 40 km. Despite a much larger number of records for Japan, we would, therefore, expect a stronger distance attenuation effect on peak amplitudes than for Euro-Mediterranean data.

2.2. Data Processing

[22] The procedure of data analysis involved sensitivity correction, P and S phase identification, double integration and filtering in a frequency band [0.075–3] Hz. P and S peaks are read on the modulus of displacement, expressed in meters, defined as

$$D = \sqrt{\mathrm{NS}^2 + \mathrm{EW}^2 + \mathrm{UD}^2} \tag{2}$$

where NS, EW, and UD are the north-south, east–west, and vertical components, respectively.

[23] A preliminary P arrival identification has been performed through an automatic first-P picker [*Allen*, 1978],

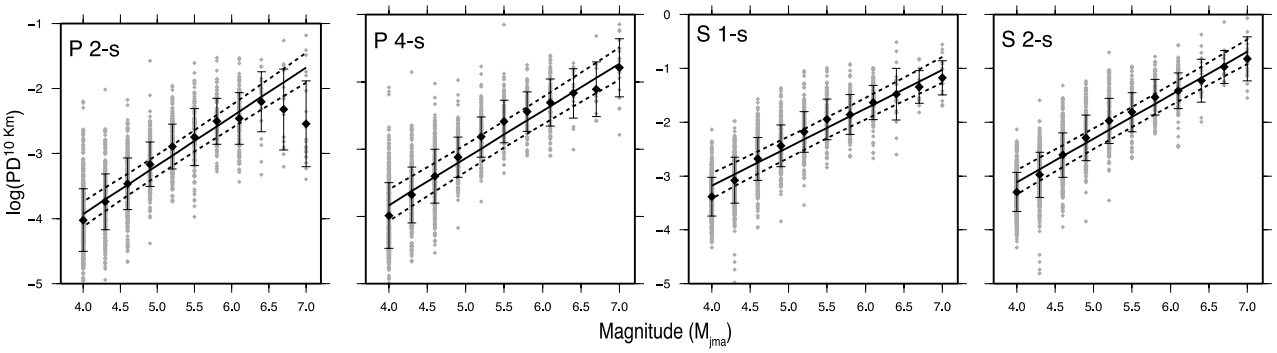

Figure 2. Correlation between low-pass-filtered peak ground motion value and magnitude. The logarithm of peak ground displacement normalized at a reference distance of 10 km as a function of M_{jma} is shown. Black diamonds represent the peak average on each magnitude bin with the associated standard deviation, while the grey dots are the peak values read on each record. The black solid lines represent the best fit line evaluated through a linear regression. The dashed lines show the prediction bounds, i.e., the 95% confidence interval for a new observation which will be centered on $\widehat{Y} = \log(PD^{10\ \mathrm{km}})_k$ whose length is given by $V[\widehat{Y}] = s\{(1/n) + [(M_k - \bar{M})^2/(M_k - \bar{M})^2]\}^2$, where s is the root of the mean square error, n is the number of observation, and $\bar{M}$ is the mean value of magnitude.

then each pick has been manually reviewed. The first S arrivals from the horizontal components of all the selected strong motion records have been identified and picked manually.

[24] Starting from the P wave and S wave picked arrivals, we measure the maximum amplitude on increasing time windows, with duration of 2–4 s for the P phase and 1–2 s for the S phase.

[25] In order to correct the P and S wave peak amplitude for the distance attenuation effect, we assume an attenuation relationship of the form [*Campbell*, 1985; *Wu and Zhao*, 2006; *Zollo et al.*, 2006]:

$$\log(\mathrm{PD}) = A + BM + C\log(R) \tag{3}$$

where PD is the observed peak displacement amplitude, M is the magnitude, and R is hypocentral distance. The parameters A, B, and C have been estimated by a standard best fit regression analysis and are reported along with their uncertainties in Table 1. Using the attenuation relationship (3) retrieved for shallow Japanese earthquakes, the P and S peaks have been normalized to a reference distance of 10 km [e.g., *Zollo et al.*, 2007]. Figure 2 shows the relation between the logarithm of distance normalized, peak displacement for P and S waves in the different time windows and the final magnitude of the analyzed events.

[26] The hypothesis of a Gaussian distribution of peak readings in each magnitude bin has been verified using a chi-square test with confidence level of 95%. We can, therefore, consider that the mean and the standard deviation of peak values are adequate estimators of the statistical distribution.

[27] Looking in detail at P data distribution for the 2-s window, one can note that peak amplitudes are log linearly correlated with magnitude up to about M = 6.5, while at larger magnitudes the peaks scatter around a nearly flat trend indicating a saturation effect. We point out the rather poor P data sampling for magnitude ranging between M = 6.5 and M = 7.1 (only 8 events with $M > 6.5$).

[28] Using a larger time window (4 s) the P peak displacement appears to correlate with magnitude all over the whole explored magnitude range. Interestingly, the S peaks show an excellent correlation with magnitude in the full considered range for 1- and 2-s windows after the first S arrival, while no saturation effect is visible at $M > 6.5$ as for the P data in the 2-s time window.

3. Real-Time Bayesian Estimation of Earthquake Magnitude

[29] On the basis of the observed correlation relationships, we implemented a probabilistic, evolutionary algorithm for real-time magnitude evaluation (RTMag).

3.1. A Bayesian, Evolutionary Approach

[30] At any time step t from the first event detection, one can define the conditional probability density function (PDF) of magnitude, given the observed data vector $\underline{d}_t$ (i.e., the distance-normalized peak displacement read in the early portion of P and S phase):

$$f(m|\underline{d}_t) = \frac{f(\underline{d}_t|m)f_t(m)}{\int_{M_{\min}}^{M_{\max}} f(\underline{d}_t|m)f_t(m)dm} \tag{4}$$

where $f(m|\underline{d}_t)$ is the conditional PDF of M given data $\underline{d}_t$, $f(\underline{d}_t|m)$ is the conditional PDF of P and/or S data d_t given magnitude M, and $f_t(m)$ is the "prior" probability density function of magnitude M, which expresses the probability of earthquake magnitude based on the available prior information before the current time t.

Table 1. Coefficients of Estimated Curves Along With Retrieved Standard Errors

Phase	Seconds	A_c'	$\Delta A_c'$	B_c'	$\Delta B_c'$	SE_c	C_c	ΔC_c
P	2	−6.93	0.48	0.75	0.09	0.32	−1.13	0.06
P	4	−6.46	0.48	0.70	0.09	0.40	−1.05	0.10
S	1	−6.03	0.54	0.71	0.10	0.38	−1.40	0.05
S	2	−6.34	0.48	0.81	0.09	0.37	−1.33	0.05

[31] The time step is defined as the minimum window length of transmitted data packets, here assumed as 1 s.

[32] Assuming that at a given time t, the observed data $\underline{d}_t$, from an event of given magnitude, have a lognormal distribution and are statistically independent, the PDF $f(\underline{d}_t|m)$ is expressed by the likelihood

$$f(\underline{d}_t|m) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_{\log(d_{t,k})} d_{t,k}} \times \exp\left[-\frac{\left(\log(d_{t,k}) - \mu(m)_{\log(d_{t,k})}\right)^2}{2\sigma^2_{\log(d_{t,k})}} \right] \quad (5)$$

where N is the number of stations whose data are available at time t.

[33] Following *Iervolino et al.* [2006], the prior information on magnitude is given by the Gutenberg and Richter relationship for times before the first P arrival (t_{first}) is detected at the seismic network. As new data become available from the seismic network ($t > t_{first}$), the prior PDF must change according to the new available information on earthquake magnitude provided by the recorded waveform:

$$f_{t>t_{first}}(m) = f_{t-1}(m|d_{t-1}) \quad (6)$$

3.2. Magnitude Estimation From Peak Measurements

[34] At each time step after the first P detection, the seismic network outputs a set of peak amplitude measurements: P peak in windows of 2 and 4 s after first P; S peak in windows of 1 or 2 s after first S.

[35] While the P wave onset is identified by an automatic picking procedure [*Allen*, 1978], the S onset can be estimated from an automatic S picking or from a theoretical prediction based on the hypocentral distance given by the actual earthquake location.

[36] To verify the quality of the estimated S arrivals, on the basis of the hypocentral location, we extracted from each magnitude bin a set of events, with at least six records, and we located them using the 1-D velocity model by *Fukuyama et al.* [2003]. The obtained hypocentral estimates are then used to calculate the S arrival times which are compared to the manual picking. The distribution of the differences between the theoretical and the manual S picking is centered on zero and have a standard deviation of ±0.5 s.

[37] At a given time step after the first P wave detection at the network, progressively refined estimates of magnitude are obtained from P and S peak displacement data. These are preliminarily normalized to a reference distance R = 10 km:

$$\log(\mathrm{PD}_c^{10\,\mathrm{km}}) = \log(\mathrm{PD}_c^{R}) - C_c \log\left(\frac{R}{10}\right) \quad (7)$$

where the subscript c is for P and S phase, constant C_c is determined through the best fit regression expressed in equation (3), and R is the hypocentral distance, expressed in kilometers. Then the distance corrected peaks are correlated with the final magnitude through the relationship

$$\log(\mathrm{PD}_c^{10\,\mathrm{km}}) = A'_c + B'_c M \quad (8)$$

[38] Assuming a standard error of SE_c on peak displacements regression (8), the mean values and standard deviation of the quantity log (PD_c), can be written as

$$\begin{cases} \mu(m)_{\log(\mathrm{PD}_c)} = B'_c m + A'_c + C_c \log\left(\frac{R}{10}\right) \\ \sigma_{\log(\mathrm{PD}_c)} = \mathrm{SE}_c + \log\left(\frac{R}{10}\right)\Delta C_c + C_c \frac{1}{R}\Delta R \end{cases} \quad (9)$$

where ΔR and ΔC_c are the errors on the distance R and on the C_c coefficient in equation (7). In this framework it is possible to evaluate the PDF using several strategies including the use of 2-s P peak regression laws jointly with the S information.

[39] As discussed in section 3.3, the saturation effect does not prevent the use of the regression law for 2-s P data during the real-time magnitude estimation, so that the regression laws for both 2- and 4-s P data (in addition to 2-s S wave data) can be usefully included in a Bayesian probabilistic approach.

[40] By defining $F_{t,M}(m)$ as the cumulative PDF at time t, it is possible to estimate a range of magnitude variation [$M_{\min}$; $M_{\max}$], whose limits are based on the shape of the $F_{t,M}(M)$ function:

$$\begin{aligned} M_{\min} &: F_{t,M_{\min}}(m) = \int_{-\infty}^{M_{\min}} P_t(m|\underline{d}_t)dm = \alpha \\ M_{\max} &: F_{t,M_{\max}}(m) = \int_{-\infty}^{M_{\max}} P_t(m|\underline{d}_t)dm = 1-\alpha \end{aligned} \quad (10)$$

For example, if we assume α = 5%, then $M_{\min}$ and $M_{\max}$ will be the $F_t(m)$ evaluated at 0.05 and 0.95, respectively. Alternatively, the PDF can be used to predict the probability of exceeding a critical magnitude threshold (M_{cth}) given the estimated magnitude (M_{est}) defined as

$$P(M > M_{\mathrm{cth}}|M_{\mathrm{est}}) = \int_{M_{\mathrm{cth}}}^{+\infty} P_t(m|\underline{d}_t)dm \quad (11)$$

3.3. A Statistical Model to Account for the 2-s *P* Peak Saturation Effect

[41] In this section we discuss how to include in our Bayesian method the empirical correlation law for P peak displacement measured in a window of 2 s. The use of 2-s P data ($\mathrm{PD}_P^{2\,\mathrm{s}}$) for magnitude estimation is important for several reasons: (1) the $\mathrm{PD}_P^{2\,\mathrm{s}}$ regression law can predict the magnitude of events below the saturation threshold (M_{sat}), (2) $\mathrm{PD}_P^{2\,\mathrm{s}}$ data are the only P wave information from the near-source station with S-P time window shorter than 4 s, and (3) although more uncertain, the $\mathrm{PD}_P^{2\,\mathrm{s}}$ data can provide a preliminary magnitude evaluation, and its associated PDF can be used as prior information for further magnitude estimates.

[42] Let us hypothesize that the distribution for $\mathrm{PD}_P^{2\,\mathrm{s}}$ data follows the lognormal distribution for $M < M_{\mathrm{sat}}$ while it is

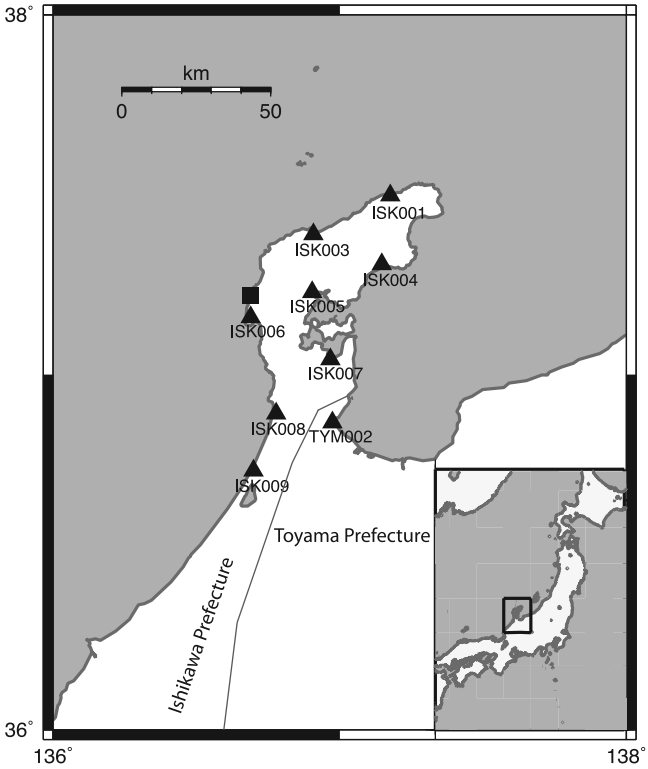

Figure 3. Epicentral position of 25 March 2007 Noto Hanto earthquake (black square). Since this event has a depth of 11 km, we select the K-Net stations which lie in a radius of 50 km from the epicentral position (black triangles).

constant for $M \geq M_{\text{sat}}$. This means that the $\text{PD}_P^{2\,\text{s}}$ data are predictive below the threshold magnitude, while the probability of occurrence for magnitude above the threshold is assumed to be constant and controlled by the prior PDF. In this case,

$$P\left(PD_P^{2\,\text{s}}|m\right) = \frac{1}{\sqrt{2\pi}\sigma_{\log\left(PD_P^{2\,\text{s}}\right)} PD_P^{2\,\text{s}}} \cdot \exp\left[-\frac{\left(\log\left(PD_P^{2\,\text{s}}\right) - \mu(m)_{\log\left(PD_P^{2\,\text{s}}\right)}\right)^2}{2\sigma^2_{\log\left(PD_P^{2\,\text{s}}\right)}}\right] \tag{12}$$

with

$$\begin{cases} \mu(m)_{\log\left(\text{PD}_P^{2\,\text{s}}\right)} = B_P^{\prime 2\,\text{s}} m + A_P^{\prime 2\,\text{s}} + C_P \log\left(\dfrac{R}{10}\right) & m < M_{\text{sat}} \\ \mu(m)_{\log\left(\text{PD}_P^{2\,\text{s}}\right)} = B_P^{\prime 2\,\text{s}} M_{\text{sat}} + A_P^{\prime 2\,\text{s}} + C_P \log\left(\dfrac{R}{10}\right) & m \geq M_{\text{sat}} \end{cases} \tag{13}$$

where $M_{\text{sat}} = 6.5$, $A_P{}'^{2\,\text{s}}$, $B_P{}'^{2\,\text{s}}$, and C_P are the coefficients in equation (8) and (7) retrieved for 2-s P peak.

4. Application to Japanese Earthquakes

[43] The real-time magnitude algorithm (RTMag) has to be conceived as part of a process chain, starting from the event detection at the seismic network, aimed at evaluating the seismic hazard in real time and the missed/false alarm probabilities [*Iervolino et al.*, 2006, 2007] for a given infrastructure target (hospital, nuclear/electric plants, schools, etc.). The input parameters for the RTMag algorithm are the event location coordinates with their uncertainties and P, S pickings. Hereinafter we assume that a real-time earthquake location algorithm is running and it is continuously delivering the earthquake location parameters, as a function of the time from the first P wave detection. Later, we discuss the effect of the event mislocation on the magnitude estimation.

[44] We assume that seismic records are streamed by the network as a multichannel, time synchronized data flow, with 1-s data packet length. At each time step, whose duration is equal to 1 s, the real-time magnitude estimation algorithm involves (1) signal processing, (2) P and S data acknowledgement, (3) P and S peaks measurement, (4) distance correction, (5) magnitude PDF estimation (equation (4)) at time t, and (6) magnitude prediction with associated uncertainty bounds.

[45] On strong motion data the signal processing involves, band pass filtering in a 0.075–3 Hz window, double integration, and ground displacement modulus evaluation. This is the same data processing as applied to retrieve the empirical prediction laws, described in section 3.

[46] The P and S data acknowledgement is a very critical step. At this stage several conditions are checked: (1) P phase duration: to avoid S phase overlap with the P phase information the S-P interval is measured. This check is also meant to verify if $\text{PD}_P^{2\,\text{s}}$ or $\text{PD}_P^{4\,\text{s}}$ peak will be readable on each station and (2) P or S peak availability: before performing the peak reading, the algorithm verifies that the whole 4-s P or 2-s S window has been acquired. The verification of these conditions makes it possible to read peaks in the acquired P or S phase window.

[47] Using the real-time location, the displacement peaks are corrected for distance attenuation and normalized to a distance of 10 km (equation (7)). The error introduced on peak correction for the effect of a wrong distance estimation is of the order of $C_c(R_{\text{true}}/R_{\text{est}})$ (equation (7)), where R_{true} and R_{est} are the actual and the estimated distance, respectively. If the error on the peaks amplitude correction is smaller than the standard error on the predictive regression laws (i.e., the peaks obtained with the right and the wrong correction lie between the dashed lines in Figure 2), the estimated magnitude is not changed significantly (see Appendix A). The information collected at each time step is used to evaluate the likelihood product (equation (5) and (9)) that will be combined with previous information via the Bayesian formulation expressed in equation (4). The most likely magnitude is given by the value corresponding to the maximum of the PDF of magnitude, while its associated uncertainty is given by the 95% confidence bound evaluated on cumulative PDF (see equation (10)).

[48] We applied the Bayesian method to estimate the final magnitude of the 25 March 2007 Noto Hanto earthquake (M_ω = 6.7 − M_{jma} = 6.9) and 16 January 1995, Kobe earthquake (M_ω = 6.9 − M_{jma} = 7.3). These events have not been included in the database used for retrieving the regression relationships of Figure 2. In order to simulate the real-time acquisition process, the earthquake signal has been packed in consecutive windows with duration of 1 s

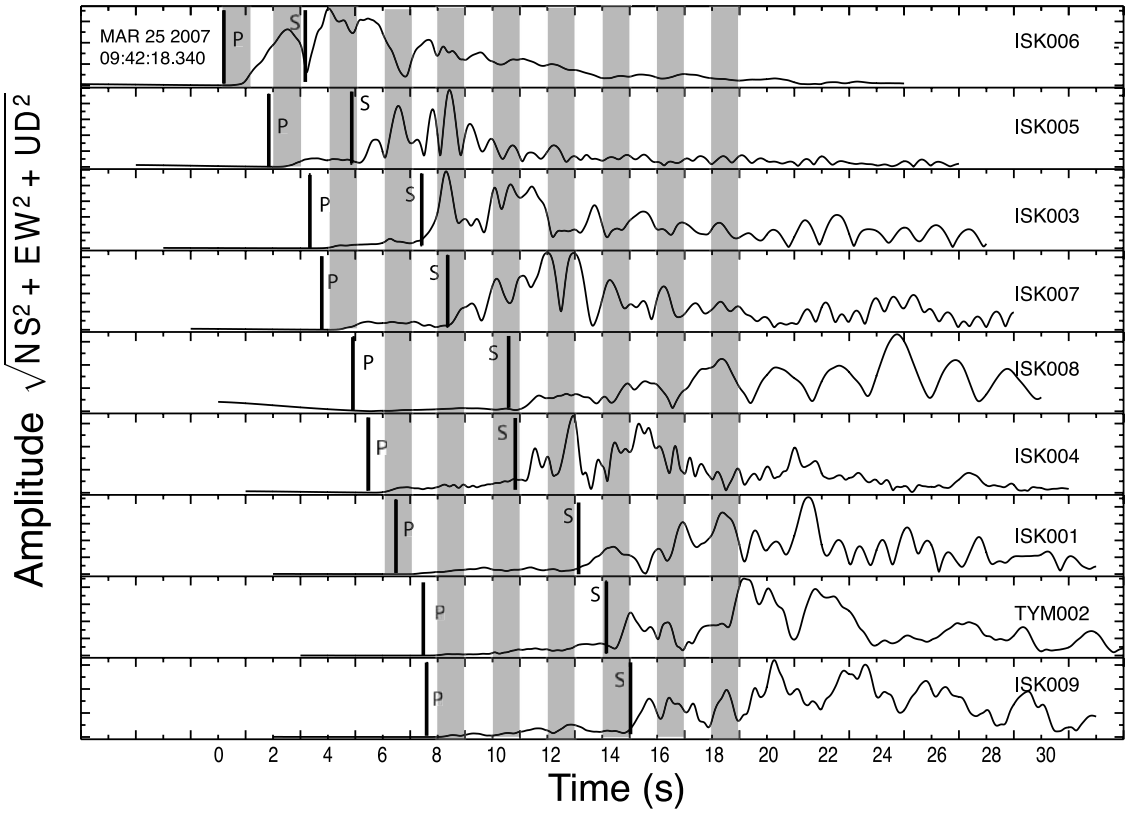

Figure 4. Modulus of displacement vector for the 25 March 2007 Noto Hanto event. The alternate white and grey stripes indicate the 1-s signal packets examined at each time step. The traces are normalized. This allows us to understand second by second which stations are recording the event and what sort of input (*P* or *S* peak) they are giving to the real-time system. For example, after 6 s from the first *P* phase picking, six station are recording the event; the 2-s *P* phase peak is available at ISK003 and ISK007 stations, while at the ISK006 station (nearest to the epicenter) the 2-s *S* phase peak is available.

starting with the first *P* pick detected at the network, as shown in Figure 4.

[49] In both the applications presented in this work we preliminarily use the location given in literature (Kyoshin Network Catalogue; *Ide et al.* [1996]). The first *S* arrival times have been estimated using a 1-D velocity model [*Fukuyama et al.*, 2003]. Also in this case, we found an average discrepancy of 0 s between the estimated and the observed *S* picking, with a standard error of about 0.5 s.

4.1. Application to 2007 Noto Hanto Earthquake (M_ω = 6.7 − M_{jma} = 6.9)

[50] The Noto Hanto earthquake (M_{jma} = 6.9), occurred on 25 March 2007 in the Ishikawa Prefecture, located about 300 km NW from Tokyo (Figure 3). The hypocenter was located by the JMA at 37.13°N, 136.41°E, at a depth of 11 km. According to the report from the Fire and Disaster Management Agency issued on April 2007, one person was killed, 297 persons were injured, 321 houses collapsed, and 357 houses were half destroyed. Most of damage was limited to wooden houses, and the lesson from 1995 (M_{jma} = 7.3) Kobe earthquake helped to reduce the triggered damaging effects, in particular those of fires.

[51] The following strategy was adopted to collect information available at several stations in a given time step: for near-source stations, with estimated *S-P* time smaller than 4 s, the PD_P^{2s} data are read, while for stations with estimated *S-P* time greater than 4 s we use only the PD_P^{4s} data. Given the about 0.5-s standard deviation on the difference between the manual and estimated picking of *S* phase, we use for magnitude estimation only the PD_S^{2s} data.

[52] In Figure 4 the modulus of ground displacement amplitude, obtained from the recorded accelerations, at several K-Net stations with hypocentral distances smaller than 60 km, is reported.

[53] Our strategy allows computation of the first magnitude estimation after 2 s from the first *P* trigger at the ISK006 using the PD_P^{2s} read at this station. Two seconds later the PD_P^{2s} at the ISK005 station is available.

[54] The first *S* wave information available is the PD_S^{2s} peak read on ISK006 station 6 s after the first *P* detection. Only 8 s after the first trigger, the PD_P^{4s} data become available at ISK003 and ISK007 stations.

[55] The results of the proposed strategy are illustrated in Figure 5. On the left side, the magnitude PDF is plotted as a function of time after the first *P* phase detection at the ISK006 stations. The shape of the magnitude probability density function, based on the PD_P^{2s} peak at ISK006 station (2-s curve on Figure 5), is controlled by the lognormal distribution for values smaller than *M* 6.5, while for larger values the magnitude PDF shape is defined by the prior information given by the Gutenberg-Richter (G-R) relation-

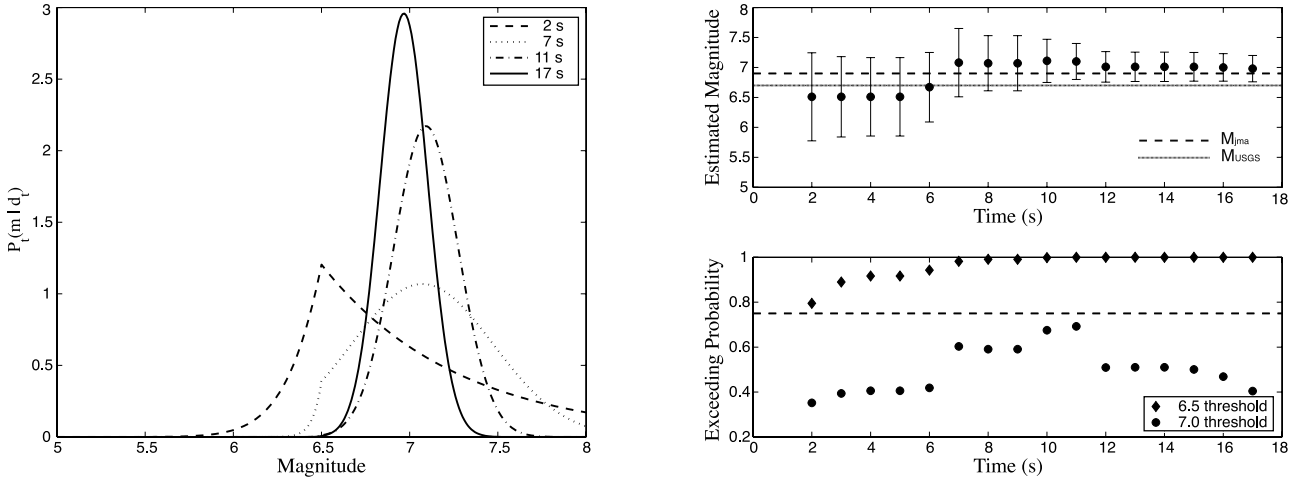

Figure 5. Application of RTMag to the Noto Hanto earthquake (25 March 2007). (left) PDF distribution at several time steps measured from the first *P* phase picking. (top right) Magnitude estimation with uncertainties as a function of time. (bottom right) Probability of exceeding magnitude 6.5 and magnitude 7.0 thresholds as a function of time.

ship. At this time the probability of exceeding *M* 6.5 is about 80% (see equation (11)), and the most likely magnitude is in the range [5.8–7.2].

[56] When the first $PD_S^{2\,s}$ peaks are available at ISK006 and ISK005 stations, the shape of magnitude PDF for $M > 6.5$ is no more controlled by the prior G-R distribution (7-s curve on Figure 5). The probability of exceeding magnitude 7 increases to 60%, the uncertainty starts to decrease and the most likely magnitude range is [6.5–7.5].

[57] As time passes by, the magnitude PDF shape becomes narrower and symmetrically distributed around the most likely magnitude value. The arrival of new readings in the subsequent seconds allows prediction, with a 95% of confidence, of the [6.8–7.0] magnitude range.

[58] An inaccurate hypocentral location will affect both the correction for the distance attenuation and the *S* pick measurement, the combination of which can influence the accuracy on magnitude estimation. To evaluate such effects, we performed a test hypothesizing a failure of the automatic *P* phase picker, indeed the event is located using few stations. In particular, we obtained 10 different hypocentral locations for the Noto Hanto event using different clusters of four stations. Figure 6 shows the inferred location based on the *P* phase automatically picked and the 1-D velocity model of *Fukuyama et al.* [2003]. The worst location is obtained using the four southern stations giving an epicenter shifted of 20 km in the northwest direction with an estimated depth of 5 km (the actual depth is 11 km).

[59] For each location, the magnitude values and their associated uncertainties are reported as a function of time using grey symbols in Figure 6, while the results obtained using the correct event location are drawn in black.

[60] Despite of the variability of the event location the magnitude estimates converge to a final value ranging from 6.5 to 7.0 with an average uncertainty of ±0.3. In the worst case the magnitude estimated is 6.5 one second after the first *P* arrival and it doesn't change significantly with the time. The shallowest locations (depth between 5 and 6 km) have a slower convergence and the final magnitude value (6.7) is reached 16 s after the first *P* arrival.

4.2. Application to 1995 Kobe Earthquake ($M_\omega = 6.9$ – $M_{jma} = 7.3$)

[61] A further event considered for testing the RTMag algorithm is the Kobe earthquake which occurred on 16 January 1995. This earthquake inflicted heavy damage in the Hanshin-Awaji region. Only in Kobe city the number of victims was 4571, and 14,678 people were injured; 67,421 structures fully collapsed and the great fire triggered by the event destroyed 6965 buildings (data extracted from *City of Kobe Office* [2008]).

[62] The complex rupture mechanism of this strike-slip earthquake has been widely investigated by numerous authors [*Wald*, 1996; *Yoshida et al.*, 1996; *Ide et al.*, 1996; *Sekiguchi et al.*, 1996, 2000], who locate the largest slip area on the fault segment extending SW of the hypocenter, beneath Awaji Island (Figure 7). In contrast, a smaller and deeper rupture segment occurred beneath the city of Kobe. Moreover, *Wald* [1996] observed that for all the accelerometers located nearby Kobe (which represent most of selected records) the strong motion is mainly influenced by the rupture segment beneath the city and only a small contribution comes from the SW distant rupture beneath Awaji Island.

[63] After this event, three dense seismic networks have been installed on the entire Japanese territory composed of high-sensitivity seismographs (Hi-Net), broadband sensors (F-Net), and strong motion accelerometers (K-Net). The strong motion records of Kobe earthquake used in the present work have been acquired by the JMA strong motion network operating at that time (Figure 7). The shape of the magnitude PDF and the consequent convergence to a final magnitude value, depend on several factors such as network density, source station geometry, presence of outliers (i.e., not operational stations, missed and/or false *P*, *S* picks). Figure 8 shows the modulus of the displacement vector for the Kobe, 1995 earthquake, evaluated in a set of stations

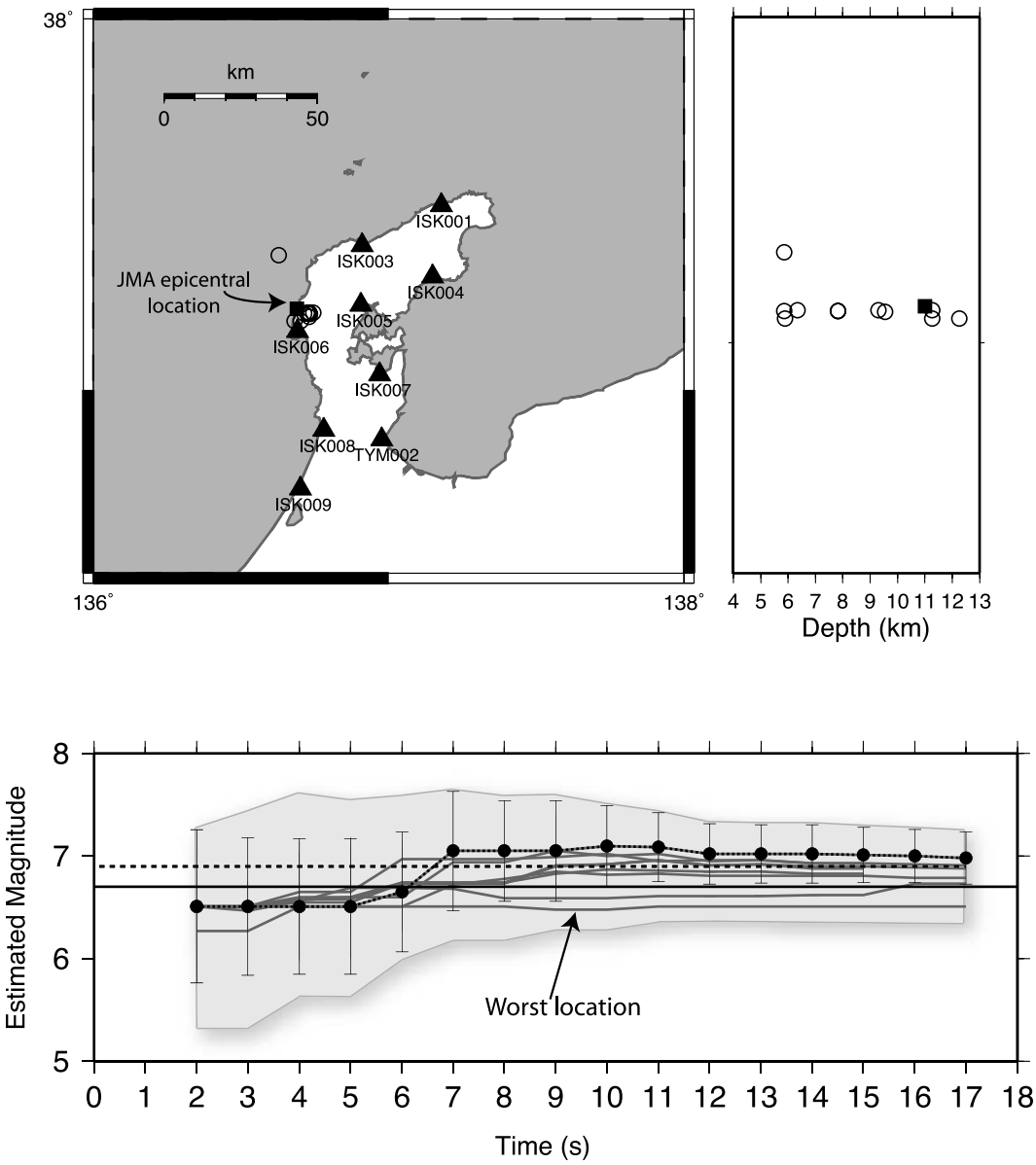

Figure 6. Effect of an inaccurate hypocentral location on real-time magnitude estimation. (top) Sketch of the epicentral area of the Noto Hanto earthquake. The black square is the JMA hypocentral location, the circles are the different hypocentral locations obtained using four stations among the lines reported on the map (black triangles). (bottom) Magnitude estimation as a function of time. The horizontal lines refer to the final magnitude values (black line, USGS M = 6.7 estimation; dashed line JMA M = 6.9 estimation). Black dots and error bars are the algorithm outputs corresponding to the JMA hypocentral location. The grey lines indicate the evolution of magnitude estimation for different mislocations. The grey shaded area is the envelop of the errors on each magnitude estimation.

belonging to the accelerographic network operating at the time. In Figure 9 we show the simulation of a Bayesian, evolutionary magnitude estimation.

[64] The magnitude PDF, 2 s after the first P trigger (based on P phase peak displacement read on a window of 2 s at NIS station see Figure 8), is centered at M 5.7 within an estimated uncertainty of ±0.7.

[65] The two critical thresholds of M = 6.5 and 7 lie on the tail of the PDF distribution; therefore, the initial probability to exceed M 6.5 and M 7 is very small. When the first 2-s S peak becomes available (4 s after the first P pick), the magnitude estimated is in the range [5.9–6.7]. Seven seconds after the first P pick, the magnitude PDF has been estimated using the PD_S^{2s} peaks from the KJM and PRI stations, jointly with the first PD_P^{4s} peak from the TDO station (see Figure 7 for stations distribution and Figure 8 for peak reading). Such a high information rate contributes to improve the accuracy of magnitude prediction providing the range [6.4–7.0]. As the time passes by, PD_P^{4s} and PD_S^{2s} readings are available from the northeast cluster of stations (see Figure 7 for stations distribution and Figure 8 for peak

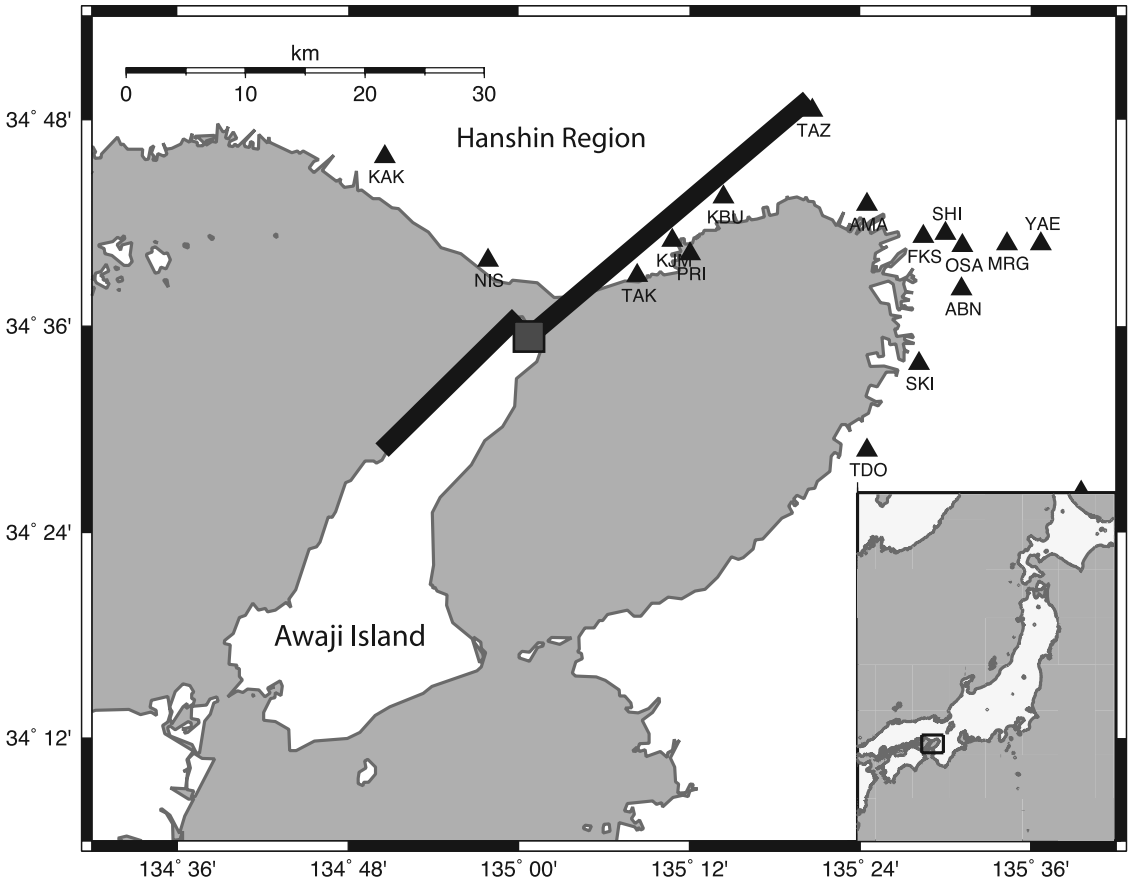

Figure 7. Location map of selected strong motion stations (black triangles) for the Kobe earthquake. The black lines present the surface projection of the fault plane segments, while the black square is the JMA epicentral location.

reading). The magnitude estimation converges to the value of 7.1 ± 0.2 (13 s after the first *P* detection).

5. Effect on PGA Prediction at the Target Site

[66] Since 2004, the Japanese Meteorological Agency has been working on developing an EEWS in Japan. From February 2004 to July 2006, JMA have sent out 855 earthquake early warnings, only 26 of which were recognized as false alarms [*Tsukada*, 2006; *Hoshiba et al.*, 2008].

[67] The dissemination of information is made in several steps, on the basis of a hybrid approach, in the sense that the system uses jointly the single station (on site) and the network based (regional) early warning techniques. The first warning is issued 2 s after the first *P* phase detection, if the maximum acceleration amplitude exceeds the threshold of 100 cm/s^2. The information on event location is not necessary at this stage for warning purposes. If there is no following *P* picking to the closer station, the "cancel" report is communicated, otherwise the warning is renovated with the evaluation of the earthquake location [*Horiuchi et al.*, 2005; *Odaka et al.*, 2003], magnitude [*Kamigaichi*, 2004] and expected intensity (defined following the JMA intensity scale) [*Midorikawa et al.*, 1999]. The thresholds for regional warning are magnitude 6 and level 5 of the JMA intensity scale. The warning is updated when subsequent hypocenter estimations vary by 0.2° for latitude or longitude, 20 km for hypocentral depth, 0.5 for magnitude and 1 for maximum seismic intensity. The final warning is transmitted when the estimated value converges to a stable value with time.

[68] Using the above criteria (0.5 variation on magnitude estimate) to set the early warning time, for Noto Hanto we achieve the following warning reports: the first warning (M = 6.5) is given 2 s after the first trigger; 6 s after the first trigger, the warning is updated (M = 7.0). After this time the estimated magnitude does not vary significantly (M = 6.94 is the final estimation).

[69] For Kobe earthquake the first warning time is issued 2 s after the first *P* trigger, but the estimated magnitude (M = 5.6) is below the warning threshold. Two seconds later (4 s after the first trigger), the magnitude reaches M = 6.3, and the first warning is issued. Ten seconds after the first *P* detection, the warning is updated (M = 6.8). The estimation gradually converges to M = 7.1 value but no warning update should be issued. We attribute such a slower convergence to the final magnitude value, with respect to the Noto Hanto earthquake simulation, to the sparse geometry of strong motion stations around the fault plane. Indeed, when the dense cluster of stations situated at Kobe city starts to give information to the system, the magnitude estimation improves and converges steadily to the final value. This condition is verified within 7–8 s from the first trigger.

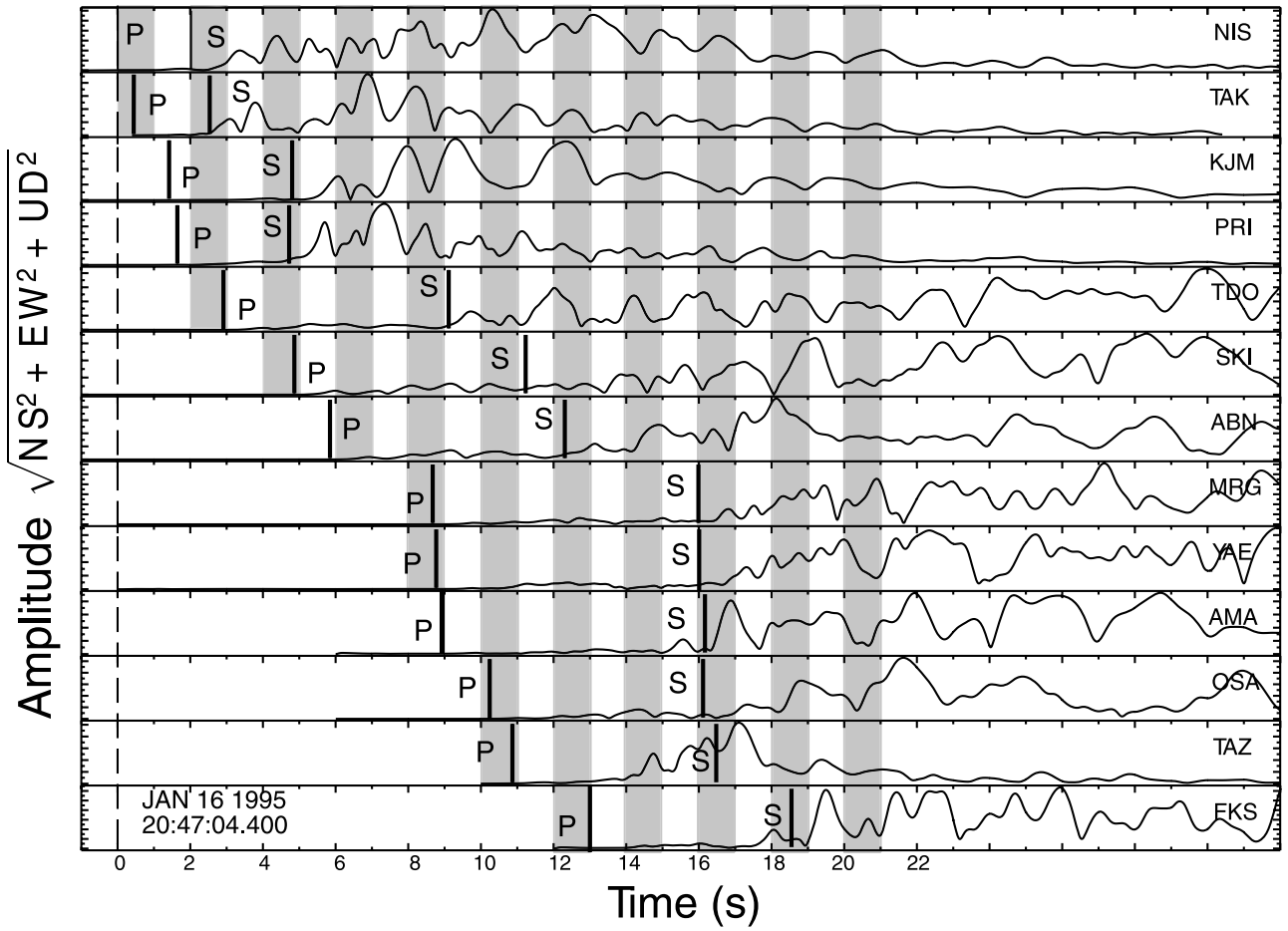

Figure 8. Modulus of displacement vector for the 16 January 1995 Kobe event. The alternate white and grey stripes indicate the 1-s signal packets examined at each time step. The zero on the time axes refers to the first event detection. To apply the evolutionary technique, it is important to synchronize records (usually with respect to the origin time). For Kobe event records, the absolute time of the first sample is not available. Therefore, we used the $t_{P\text{pick}} - t_{\text{First Sample}}$ range and the $t_{P\text{theo}} - t_{\text{Origin}}$, evaluated using P phase travel time, to mark the time of each station. The first S information is available after 4–5 s.

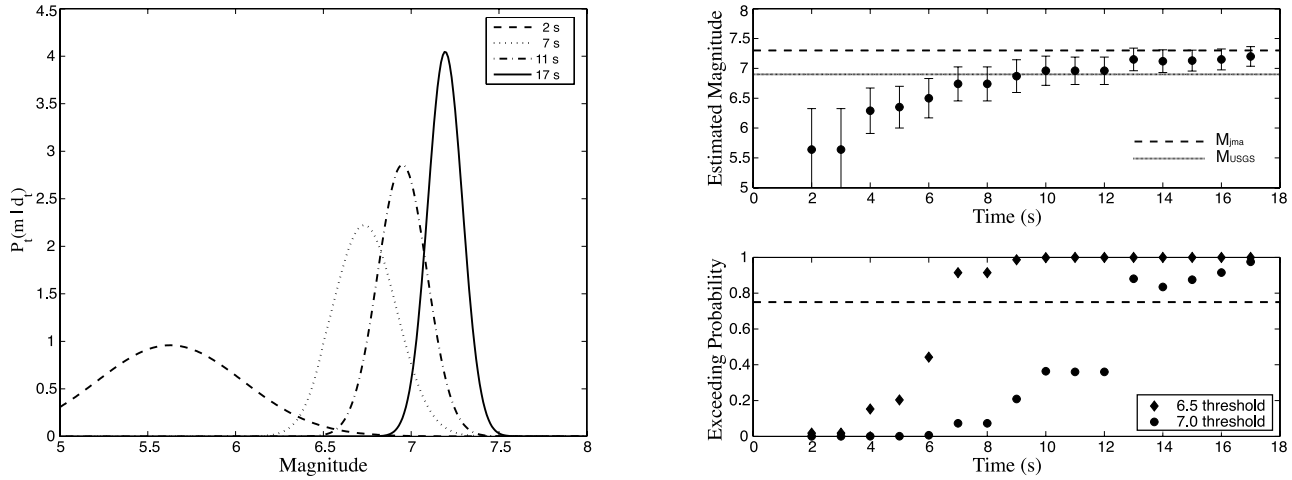

Figure 9. Application of RTMag to the Kobe earthquake (1995). (left) PDF distribution at several time steps measured from the first P phase picking. (top right) Magnitude estimation with uncertainties as a function of time. (bottom right) Probability to exceed magnitude 6.5 and magnitude 7.0 thresholds as a function of time.

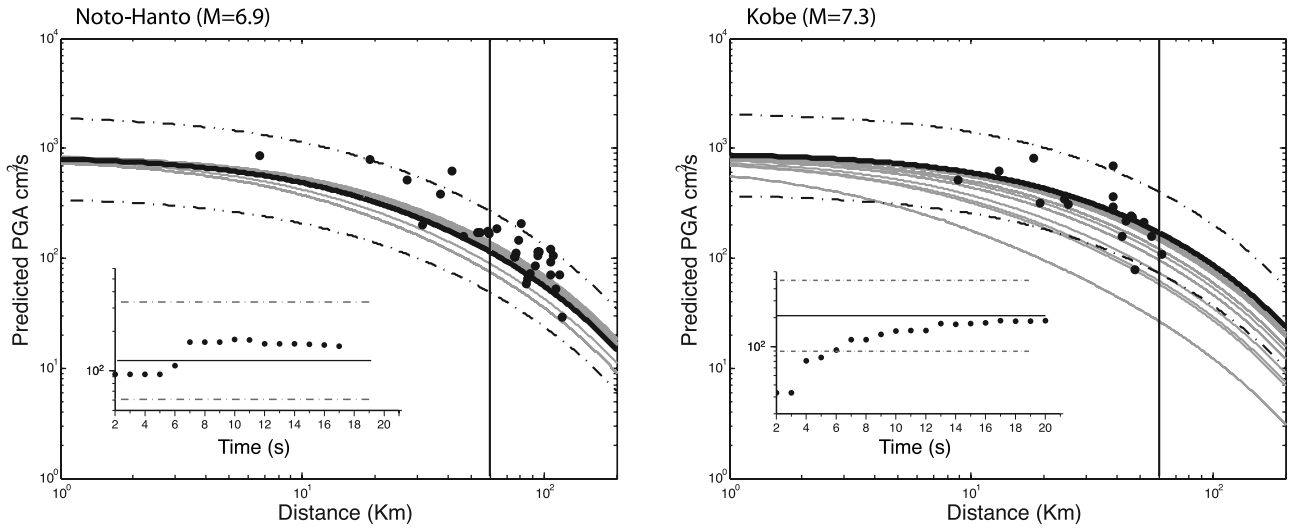

Figure 10. Comparison between *Kanno et al.*'s [2006] PGA attenuation laws evaluated off line and in real time for 2007 Noto Hanto (M = 6.9) and 1995 Kobe (M = 7.3) earthquakes. The black solid line corresponds to the off-line attenuation curve, here plotted with the associated error (dashed black lines). The grey lines are relative to the attenuation law, evaluated using the magnitude values estimated at each time step; the black dots are the observed PGAs. The inserts report the PGA prediction, as a function of time, for a site at 50 km from the epicentral area.

[70] Recalling that the real-time location and magnitude are the basic elements for the hazard assessment, it is interesting to understand what is the effect of magnitude estimation as a function of time on PGA prediction at a given target. With this intent, we compare for Noto Hanto and Kobe earthquakes, the PGA attenuation laws [*Kanno et al.*, 2006] evaluated for the actual magnitude values with relative errors, and the PGA attenuation laws evaluated step by step using the most likely magnitude value. We assume that the location is know rapidly with sufficient accuracy.

[71] Concerning the Noto Hanto earthquake, we observe that since the first estimation, the real-time attenuation curves lie in the error bounds of the final curve (Figure 10). This implies that the real-time PGA prediction is reliable in 2 s after the first P detection. Moreover the PGA prediction becomes stable in 5 s after the first P arrival.

[72] For the Kobe earthquake, the system is able to give in 6 s after the first P a PGA prediction included in the error bounds of the final curve. For a site distant 50 km from the epicenter, the PGA estimation becomes stable in 8 s after the first trigger.

[73] The time needed for predicting a peak ground motion parameter at a given target site will determine the "blind zone" radius and the lead time as a function of epicentral distance. The blind zone is the area where no early warning actions are possible, while the lead time is defined as the time interval between the arrival of S wave to the target and the first warning.

[74] In Figure 11 we draw a sketch of the epicentral area of the Noto Hanto earthquake with some critical facilities [*Hokuriku Electric Power Company*, 2006], along with the blind zone and lead time evaluated at several distances from the earthquake epicenter. We assume that the first warning has been issued in 2 s after the first P detection. The lead time is $\sim$10 s, at 60 km from the epicenter. This area corresponds to a part of Toyama and Ishikawa prefectures, where the water supply was damaged and some roads were partially destroyed but gas lines were not damaged.

[75] Several studies on early warning feasibility [*Goltz*, 2002; *Shresta*, 2001] indicate 10 s as a reasonable warning time to operate automatic actions such as sending alert to fire department command centers, making sure everyone is out of elevators, starting to move equipments out of buildings, activating data backup, evacuating bottom floors, stopping hazardous works, shutting down gas, alerting drivers, controlling traffic signals. The Noto Hanto scenario, based on the real-time algorithm RTMag, confirms its proficiency for the achievement of an effective regional EEWS, not only providing accurate and stable magnitude estimates and PGA predictions, but also ensuring reasonable lead times for targets at distances greater than 50 km from the epicentral area.

6. Discussion and Conclusions

[76] Real-time estimates of earthquake location and magnitude are necessary for regional warning systems [*Kanamori*, 2005]. However, the alarm decision needs a further step: the prediction, with quantified confidence, of a ground motion intensity at a distant target site. This problem can be approached following an evolutionary frame where probability density functions (PDFs) for earthquake location, magnitude and attenuation parameters are combined into a real-time probabilistic seismic hazard analysis [*Iervolino et al.*, 2006].

[77] Considering the peak displacement amplitude measured in the early portion of P and S waves, in this article we have shown that suitable PDFs for the earthquake magnitude can be evaluated. The PDFs of location and magnitude parameters will be used for the real-time probabilistic assessment of false alarms and loss estimation, which are

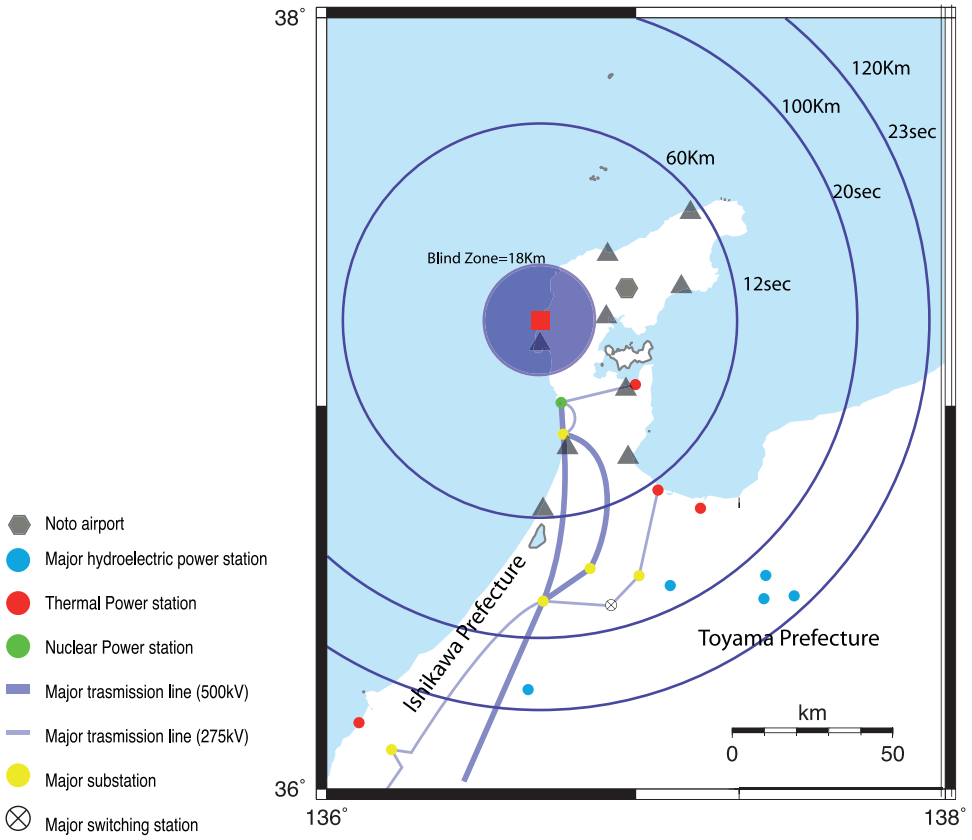

Figure 11. Sketch of the Noto Hanto earthquake area, with Hokuriku electric power stations. The circle refers to blind zone and lead time as a function of epicentral distance. The lead time is defined as the time interval between the arrival of *S* wave to the target and the first magnitude estimation. The Noto Airport runaway (the black hexagon at 25 km of epicentral distance in N–E direction) has been damaged, and all the airport operations were shut down the day of the earthquake. During the earthquake, the Nanao Ota thermal Power Plant Unit 2 (red circle near the ISK007 station) automatically shut down. High-voltage cable was disconnected at eight places, and 29 utility poles were damaged.

the key elements based on which automatic actions can be undertaken to mitigate the earthquake effects.

[78] From the analysis of acceleration records of Japanese earthquakes, we have found that a correlation exists between the event magnitude and the low-frequency peak displacement measured in the early portion of *P* and *S* signals. The 2-s *P* peak shows an excellent correlation with magnitude up to $M \sim 6.5$, while for higher magnitude values the peaks appear to scatter around a nearly flat trend. This effect is no more evident when considering a 4-s *P* window, while no saturation is observed for *S* wave data even though measured in a very narrow time window (1 s).

6.1. Implication on the Earthquake Rupture Physics

[79] A possible explanation of the different scaling of *P* and *S* peaks with magnitude, when both read in a time window of 2 s, and the observation of the saturation effect can be given using the kinematic concept of isochrone [*Spudich and Frazer*, 1984; *Bernard and Madariaga*, 1984]. The far-field radiation, recorded at a given station at a given time, can be expressed as a surface integral on the fault area delimited by the isochrone, defined as the set of points whose radiation arrives at a given station at a given time t. Assuming a constant rupture velocity, for a station P, placed at a distance *D* from the rupture front, whose radius is *r*, the isochrone formulation is

$$\tau(P,r) = \frac{r}{v} + \frac{D}{c} \tag{14}$$

where *v* is the rupture velocity and *c* is the *P* or *S* wave velocity. Note that the points on the isochrone do not irradiate simultaneously; the radiation emitted by the rupture front is delayed during the wave propagation, depending on the source-station geometry. Therefore, the rupture surface, where the first 2 s of *P* or *S* on strong motion records have been radiated from, is, in general, at least equal to, or bigger than, the fault surface which is expected to rupture in 2 s [*Festa*, 2007]. Moreover, as the *S* waves have slower velocity than *P* waves, the discrepancy between the fractured area and the surface imaged by the *S* isochrones of 2 s of duration will be much larger for the 2-s *P* isochrones.

[80] In a very recent work, *Murphy and Nielsen* [2008] argue that for magnitude below the saturation threshold (*M* 6.5), the average area imaged by the 2-s *P* isochrones is larger or comparable with the final fault size, thus explaining the observed correlation between PD and magnitude.

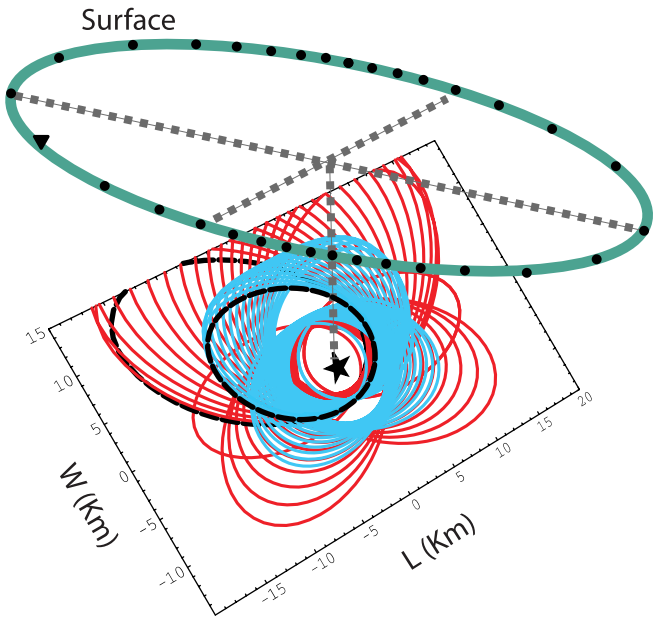

Figure 12. An example of *P* and *S* isochrones relative to a signal duration of 2 s. The *P* isochrones are in light blue, the *S* isochrones in red. The fault plane has a dip of 60°. A constant velocity rupture *v* of 3 km/s is assumed, and it is equal to 0.9 *S* wave velocity. The *P* and *S* velocity ratio is $\sqrt{3}$. The plotted isochrones are referred to 32 stations (black dots) covering the green ellipse with a constant azimuth step. We outline in black the *P* and the *S* isochrones relative to the station represented as a black triangle. The area imaged by the 2-s *P* isochrones roughly corresponds to the rupture surface of a magnitude 6.5, while the area sampled by the 2-s *S* isochrones, corresponds to a larger-magnitude event. The different size of the fault surface imaged by *P* and *S* isochrones can explain the saturation effect observed on the 2-s PD regression shown in Figure 2. Image generated and kindly provided to us by G. Festa.

This suggests that for higher magnitude values the saturation effect is due to an undersampling of the fault plane, when 2 s of *P* signal are used. Extending the *P* window to 4 s, or using 2 s of *S* window, larger fault surfaces are sampled, and the scaling extends over a larger magnitude range. In Figure 12 we show the 2-s *P* (light blue curves) and 2-s *S* isochrones (red curves) for a fault plane with a dip of 60°. An expanding circular fracture, with a constant velocity rupture *v* of 3 km/s, is modeled. And the ratio between the *P* and *S* wave velocity is assumed to be $\sqrt{3}$. The plotted isochrones are referred to 32 station covering the green ellipse with a constant azimuth step. The 2-s *P* isochrones encompass an average area roughly corresponding to the surface rupture of a magnitude 5.8–6.5 [*Wells and Coppersmith*, 1994], while a larger area is sampled by the 2-s *S* isochrones, corresponding to a magnitude of 6.5–7.0.

[81] The above argumentation is based on the simplistic assumption of neglecting the directivity effect. Nevertheless, as confirmed by our observations (Figure 2), this effect is averaged out when a large number of measurements is considered.

[82] The correlation between the initial peak displacement amplitude and the final magnitude, can be further explained in terms of basic earthquake source concepts. Given the explored range of distance and frequency, the initial *P* and *S* wave radiation can be assimilated in first approximation to the far-field effect of a point source, assuming that the rupture directivity effect is averaged out by the variable azimuthal position of the stations. In this case the initial *P* and *S* peak displacement amplitude, PD, at a given distance *r*, is expected to scale with moment rate $\dot{M}$ [*Aki and Richards*, 1980]:

$$\mathrm{PD} = \mathrm{const}\frac{1}{r}\dot{M}\left(t-\frac{r}{c}\right) = \mathrm{const}\frac{1}{r}\Delta\dot{u}\Sigma \qquad (15)$$

where *c* is the wave velocity, Δu is the average slip on the fault, and Σ is the active fault area during the initial stage of the rupture. According to theoretical models of rupture dynamics [*Kostrov*, 1964; *Scholz*, 1990], the slip rate amplitude, $\dot{u}$, scales linearly with dynamic stress drop $\Delta\sigma$. Therefore, an important implication of our experimental analysis of the Japanese data set, is that stress drop and/or active slip surface have to scale with magnitude in the initial stage of seismic ruptures, at least up to magnitudes of about 6.5–7. Assuming a constant stress drop model [*Keilis-Borok*, 1959], one can use equation (15) to derive the expected relationship between the initial *P* and *S* peak displacement, normalized at a reference distance, $\mathrm{PD}^{10\ \mathrm{km}}$ and magnitude *M*:

$$\log\left(\mathrm{PD}^{10\ \mathrm{km}}\right) = \mathrm{const} + \log(M_0) + \log(L) = \mathrm{const} + M \qquad (16)$$

where M_0 is the seismic moment and L is the fault length. Equation (16) is analogous to the original definition of Richter's magnitude, where peak *S* wave amplitudes are logarithmically related to the earthquake size as expressed by $\log(M_0)$ or magnitude. The logarithmic dependence of

peak displacement with magnitude/distance makes the proposed method robust and accurate, since moderate errors on amplitudes/distance lead to small errors on magnitude, as with all magnitude calculation procedures. We note that the scaling coefficient of peak displacement versus magnitude from our correlation analysis (Table 1) slightly underestimates the expected value (0.7–0.8 instead of 1, e.g., equation (16)) which is possibly due to data scatter and relatively poor data sampling for magnitude larger that M about 6.5.

6.2. Real-Time Magnitude Estimation Algorithm

[83] On the basis of the P and S peak displacement correlation laws, inferred for Japan earthquakes we implemented a probabilistic, evolutionary algorithm to estimate magnitude in real time. Its formulation, founded on the Bayes theorem, allows evaluation of the conditional PDF of magnitude, at any time step t after the first event detection. The information from different stations is expressed by the likelihood product, and then combined with prior information, given by the PDF retrieved in the previous time step and the Gutenberg-Richter distribution.

[84] The use of a statistical procedure allows (1) definition of the most likely magnitude as a function of time; (2) expression of its uncertainty in terms of confidence bound; and (3) evaluation of the probability of exceeding a fixed critical magnitude threshold. Moreover, the PDF of magnitude is one of the input functions needed for the real-time seismic hazard assessment.

[85] The Bayesian formulation allows the use of the regression law for 2-s P data, overcoming the observed saturation effect. Indeed, we hypothesize that the 2-s P peaks follow a lognormal probability distribution up to the magnitude saturation threshold, while for higher magnitude values the distribution is assumed to be uniform. This is equivalent to assume that the probability of occurrence of magnitude is controlled by the 2-s P information, up to the magnitude threshold, and by the ''a priori'' distribution for higher magnitude values.

[86] Since the final aim is to optimize the ratio between quantity and quality of the information, the advantage in using the 2-s P phase peaks is to collect a great number of readings in few seconds after the earthquake origin (depending on the network density in the epicentral area). Although this information may provide, for strong events, magnitude estimation with large uncertainties it can be used as ''prior'' information for subsequent estimates. On the other hand, the use of only 4-s P phase peaks and 2-s S phase peaks allows better constraint of the magnitude estimation, but such a strategy neglects the P contribution of nearest stations and reduces the information rate.

[87] The algorithm performance depends on the information rate: as more data are available at any second t the faster the algorithm will converge to the final magnitude value. A further factor that influences the accuracy of magnitude prediction is the real-time location, whose effect has been analyzed with specific tests. We make use of the hypocentral distance not only to normalize peak displacement amplitude, but also to estimate the S phase arrival time.

[88] An example of how real-time earthquake location can be influenced by station failures is given by *Satriano et al.* [2008]. They show that for a working seismic network, only in the very first seconds after the event, with a few (0 to 3) recorded arrivals, a non triggering (or mistriggering) station can bias the hypocentral estimation. However, the location improves as soon as additional stations trigger, thus reducing and, eventually, eliminating the bias. This would imply that for early warning applications, the magnitude estimation should rely on an unbiased location as soon as at least ~4 stations correctly trigger, thus ensuring the convergence to a more accurate magnitude value.

6.3. Results From the Application to Japanese Earthquakes

[89] The real-time magnitude algorithm (RTMag) has been applied to the 2007 Noto Hanto (M_{jma} = 6.9) and to the 1995 Kobe (M_{jma} = 7.3) Japanese earthquakes. In both applications we selected strong motion records with hypocentral distance less than 60 km. For Noto Hanto event, we used records from the K-Net and KiK-Net stations evenly deployed along the Noto Peninsula, with an average distance of 5 km. In contrast, the Kobe earthquake has been recorded by the JMA strong motion network operating before the K-Net and KiK-net installation. Even if the average distance between stations is of the order of 5–6 km, the source-station geometry is more irregular and characterized by a dense cluster of stations in the city of Kobe.

[90] We find that for the Noto Hanto earthquake, the first warning is given within 2 s from the first P detection, and it is improved 4 s later. On the other hand, for Kobe earthquake, the first warning is issued 4 s after the first trigger and is updated 6 s later.

[91] During the Noto Hanto event, the JMA system successfully issued the first warning 5 s after the first P detection. The magnitude estimation algorithm, implemented in the JMA system, is based on a relationship between magnitude, hypocentral distance and maximum displacement amplitude recorded on P phase and S phase [*Kamigaichi*, 2004]. The magnitude is given by the average of estimates on different stations.

[92] Compared with the JMA algorithm, the main advantages of the RTMag procedure are the capability to estimate magnitude from the early portion of P and S phases, and the probabilistic foundation. Indeed, the algorithm combines rigorously the information coming from different stations in different time steps, allowing estimates of the uncertainty on the prediction. These characteristics make the magnitude estimation fast, reliable and robust.

[93] A main conclusion of our work is the feasibility of regional early warning using the S wave data from near-source accelerographic stations. We demonstrate that for dense networks it is possible to estimate the magnitude using S waves within 4–5 s from the first P wave detection, so that the estimated lead time is still useful for sending an alert to distant target sites.

[94] A further result discussed in this paper concerns the capability to predict in real time, using the Japanese attenuation law and time-dependent magnitude estimates, PGA values comparable to those evaluated using the final event magnitude. Considering the standard errors on empirical attenuation relationships, even rather large uncertainties on earthquake location and magnitude can provide

useful PGA predictions at a target sites, 50–100 km distant from the source. The application of our method to Japanese earthquakes shows that reliable predictions of peak ground motion can be obtained a few seconds after the first *P* arrival at the network, despite of a significant uncertainty in the initial magnitude estimates. This result supports the feasibility of earthquake early warning downstream of its engineering application, having as a final aim the probabilistic assessment of expected loss reduction through a triggered security action.

Appendix A: Effect of Real-Time Location on Peak Amplitude Correction

[95] Let us define for such station of the EEWS network the quantities: (1) R_{true} True hypocentral distance, i.e., the distance estimated in real time after 4–5 triggers or the distance estimated off line and (2) R_{est} The hypocentral distance estimated in real time, used for the peak amplitude correction (see equation (3)). The "true" correction on peak amplitude logarithm is

$$\log\left(\text{PD}^{10}_{\text{Off}-\text{Line}}\right) = \log\left(\text{PD}^{R}\right) - C\log\left(\frac{R_{\text{true}}}{10}\right) \tag{A1}$$

while the real-time correction is given by

$$\log\left(\text{PD}^{10}_{\text{Real}-\text{Time}}\right) = \log\left(\text{PD}^{R}\right) - C\log\left(\frac{R_{\text{est}}}{10}\right) \tag{A2}$$

We express the effect on amplitude correction due to the error on distance estimation as

$$\begin{aligned}\log\left(\text{PD}^{10}_{\text{Real}-\text{Time}}\right) - \log\left(\text{PD}^{10}_{\text{Off}-\text{Line}}\right) &= -C\log\left(\frac{R_{\text{est}}}{10}\right) + C\log\left(\frac{R_{\text{true}}}{10}\right) \\ &= C\log\left(\frac{R_{\text{true}}}{R_{\text{est}}}\right)\end{aligned} \tag{A3}$$

If such a difference is less than the standard error on regression, then the magnitude estimation does not change significantly. Assuming that the regressions have an average standard error of 0.3 (see Table 1) and the C coefficient is ≈ -1, the condition $\log(\text{PD}^{10}_{\text{Real-Time}}) - \log(\text{PD}^{10}_{\text{Off-Line}}) = \text{SE}$ is reached for $R_{\text{true}} = 2\ R_{\text{est}}$.

[96] **Acknowledgments.** We wish to thank the Associate Editor R. Console and the two anonymous reviewers for their comments and suggestions helpful to the extension and clarification of several key concepts of this work. We are grateful to Gaetano Festa and Shane Murphy for their discussions on seismic fracture modeling and to Claudio Satriano for his comments on the real-time location. This work was partially funded by AMRA scarl through the EU-SAFER project (contract 036935). The earthquake records used in this study are recorded at KiK-Net and K-NET by NIED, downloaded through their Web sites.

References

ABS Consulting (2007), Flash report on the Noto Hanto earthquake, technical report, Houston, Tex.

Aki, K., and P. G. Richards (1980), *Quantitative Seismology: Theory and Methods*, W. H. Freeman, San Francisco, Calif.

Allen, R. (2007), The ElarmS earthquake early warning methodology and application across California, in *Earthquake Early Warning System*, edited by P. Gasparini, G. Manfredi, and J. Zhao, pp. 21–43, Springer, Heidelberg, Germany.

Allen, R. M., and H. Kanamori (2003), The potential for earthquake early warning in southern California, *Science*, *300*, 786–789.

Allen, R. V. (1978), Tiltmeter observations near a large earthquake, *Bull. Sesimol. Soc. Am.*, *68*(3), 855–857.

Bernard, P., and R. Madariaga (1984), A new asymptotic method for the modeling of near-field accelerograms, *Bull. Seismol. Soc. Am.*, *74*(2), 539–557.

Campbell, K. W. (1985), Strong motion attenuation relations: A ten-year perspective, *Earthquake Spectra*, *1*, 759–804.

City of Kobe Office (2008), The great Hanshin-Awaji earthquake statistics and restoration progress, technical report, Kobe, Japan.

Cua, G., and T. Heaton (2007), The Virtual Seismologist (VS) method: A Bayesian approach to earthquake early warning, in *Earthquake Early Warning System*, edited by P. Gasparini, G. Manfredi, and J. Zschau, pp. 97–132, Springer, Heidelberg, Germany.

Dziewonski, A. M., T. A. Chou, and J. H. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, *86*, 2825–2852.

Festa, G. (2007), Does the slip in the early steps of the rupture scale with the final magnitude of the event?, *Eos Trans. AGU*, *88*(52), Fall Meet. Suppl., Abstract S24B-04.

Font, Y., H. Kao, S. Lallemand, C.-S. Liu, and L.-Y. Chiao (2004), Hypocentre determination offshore of eastern Taiwan using the maximum intersection method, *Geophys. J. Int.*, *158*(2), 655–675, doi:10.1111/j.1365-246X.2004.02317.x.

Fukuyama, E., W. L. Ellsworth, F. Waldhauser, and A. Kubo (2003), Detailed fault structure of the 2000 western Tottori, Japan, earthquake sequence, *Bull. Seismol. Soc. Am.*, *93*(4), 1468–1478.

Goltz, J. D. (2002), Introducing earthquake early warning in california: A summary of social science and public policy issues, technical report, Governor's Off. of Emergency Serv., Pasadena, Calif.

Hellemans, A. (2006), Do early tremors give sneak preview of quake's power?, *Scienze*, *314*, 1670, doi:10.1126/science.314.5806.1670.

Hokuriku Electric Power Company (2006), Annual report, technical report, Toyama City, Japan.

Horiuchi, S., H. Negishi, K. Abe, A. Kamimura, and Y. Fujinawa (2005), An automatic processing system for broadcasting earthquake alarms, *Bull. Seismol. Soc. Am.*, *95*(2), 708–718.

Hoshiba, M., O. Kamigaichi, M. Saito, S. Tsukada, and N. Hamada (2008), Earthquake early warning starts nationwide in Japan, *Eos Trans. AGU*, *89*(8), 73.

Ide, S., M. Takeo, and Y. Yoshida (1996), Source process of the 1995 Kobe earthquake: Determination of spatio-temporal slip distribution by Bayesian modeling, *Bull. Seismol. Soc. Am.*, *86*, 547–566.

Iervolino, I., V. Convertito, M. Giorgio, G. Manfredi, and A. Zollo (2006), Real-time risk analysis in hybrid earthquake early warning systems, *J. Earthquake Eng.*, *10*, 867–885.

Iervolino, I., M. Giorgio, and G. Manfredi (2007), Expected loss-based alarm threshold set for earthquake early warning systems, *Earthquake Eng. Struct. Dyn.*, *36*, 1151–1168, doi:10.1002/eqe.675.

Kamigaichi, O. (2004), JMA earthquake early warning, *J. Jpn. Assoc. Earthquake Eng.*, *3*, 134–137.

Kanamori, H. (2005), Real-time seismology and earthquake damage mitigation, *Annu. Rev. Earth Planet. Sci.*, *33*, 195–214.

Kanno, T., A. Narita, N. Morikawa, H. Fujiwara, and Y. Fukushima (2006), A new attenuation relation for strong ground motion in Japan based on recorded data, *Bull. Seismol. Soc. Am.*, *96*(3), 879–897.

Katsumata, A. (1996), Comparison of magnitudes estimated by the Japan Meteorological Agency with moment magnitudes for intermediate and deep earthquakes, *Bull. Seismol. Soc. Am.*, *86*(3), 832–842.

Keilis-Borok, V. (1959), On estimation of the displacement in an earthquake source and of source dimensions, *Ann. Geofis.*, *12*, 205–214.

Kostrov, B. V. (1964), Selfsimilar problems of propagation of shear cracks, *Appl. Math. Mech.*, *28*(5), 1077–1087.

Midorikawa, S., K. Fujimoto, and I. Muramatsu (1999), Correlation of new JMA instrumental seismic intensity with former JMA seismic intensity and ground motion parameters, *J. Inst. Soc. Safety Sci.*, *1*, 51–56.

Murphy, S., and S. Nielsen (2008), Dynamic and kinematic investigation of earthquake magnitude determination by first few seconds of waveforms, *Bull. Seismol. Soc. Am.*, in press.

Nakamura, Y. (1988), On the urgent earthquake detection and alarm system (Uredas), in *Proceedings of the 9th World Conference on Earthquake Engineering*, vol. VII, pp. 673–678, Univ. do Porto, Fac. de Eng., Porto, Portugal.

Odaka, T., K. Ashiya, S. Tsukada, S. Sato, K. Ohtake, and D. Nozaka (2003), A new method of quickly estimating epicentral distance and magnitude from a single seismic record, *Bull. Seismol. Soc. Am.*, *93*(1), 526–532.

Olson, E. L., and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature*, *438–10*, 212–215.

Rydelek, P., and S. Horiuchi (2006), Is the earthquake rupture deterministic?, *Nature*, *442–20*, E5–E6.

Rydelek, P., C. Wu, and S. Horiuchi (2007), Comment on "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on stron motion records" by Aldo Zollo, Maria Lancieri, and Stefan Nielsen, *Geophys. Res. Lett.*, *34*, L20302, doi:10.1029/2007GL029387.

Satriano, C., A. Lomax, and A. Zollo (2008), Real-time evolutionary earthquake location for seismic early warning, *Bull. Seismol. Soc. Am.*, *98*(3), 1482–1494, doi:10.1785/0120060159.

Scholz, C. H. (1990), *The Mechanics of Earthquakes and Faulting*, Cambridge Univ. Press, New York.

Sekiguchi, H., K. Irikura, T. Iwata, Y. Kakehi, and M. Hoshiba (1996), Minute locating of faulting beneath Kobe and the waveform inversion of the process during the 1995 Hyogo-Ken Nambu, Japan earthquake using strong ground motion records, *J. Phys. Earthquake*, *44*, 473–487.

Sekiguchi, H., K. Irikura, and T. Iwata (2000), Fault geometry at the rupture termination of the 1995 Hyogo-Ken Nanbu earthquake, *Bull. Seismol. Soc. Am.*, *90*(1), 133–177.

Shresta, B. K. (2001), Disaster reduction and response preparedness in Japan: A Hyogo approach, paper presented at CDC 2001: The Second Tampere Conference on Disaster Communication, Motorola, Tampere, Finland.

Spudich, P., and L. N. Frazer (1984), Use of ray theory to calculate high-frequency radiation from earthquake sources having spatially variable rupture velocity and stress drop, *Bull. Seismol. Soc. Am.*, *74*, 2061–2082.

Tsuboi, C. (1954), Determination of the Gutenberg-Richter' s magnitude of earthquakes occurring in and near Japan (in Japanese with English abstract), *J. Seismol. Soc. Jpn.*, *7*, 185–193.

Tsukada, S. (2006), Earthquake early warning system in Japan, paper presented at 6th Joint Meeting of UJNR Panel on Earthquake Research, Coop. Program in Nat. Resour., Tokushima, Japan.

Wald, D. J. (1996), Slip history of the 1995 Kobe, Japan, earthquake determined from strong motion, teleseismc, and geodetic data, *J. Phys. Earth*, *44*, 489–503.

Wells, D., and K. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, *84*(4), 974–1002.

Wu, Y. M., and H. Kanamori (2005a), Rapid assessment of damage potential of earthquakes in Taiwan from the beginning of *P* waves, *Bull. Seismol. Soc. Am.*, *95*(3), 1181–1185.

Wu, Y. M., and H. Kanamori (2005b), Experiment on an onsite early warning method for the Taiwan early warning system, *Bull. Seismol. Soc. Am.*, *95*(1), 347–353.

Wu, Y.-M., and L. Zhao (2006), Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning, *Geophys. Res. Lett.*, *33*, L16312, doi:10.1029/2006GL026871.

Wu, Y.-M., H.-Y. Yen, L. Zhao, B.-S. Huang, and W.-T. Liang (2006), Magnitude determination using initial P waves: A single-station approach, *Geophys. Res. Lett.*, *33*, L05306, doi:10.1029/2005GL025395.

Yoshida, S., K. Koketsu, B. Shibazaki, T. Sagiya, and Y. Yoshida (1996), Joint inversion of near- and far-field waveforms and geodetic data for the rupture process of the 1995 Kobe earthquake, *J. Phys. Earth*, *44*, 437–454.

Zhao, D., F. Ochi, A. Hasegawa, and A. Yamamoto (2000), Evidence for the location and cause of large crustal earthquakes in Japan, *J. Geophys. Res.*, *105*, 13,579–13,594.

Zollo, A., M. Lancieri, and S. Nielsen (2006), Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records, *Geophys. Res. Lett.*, *33*, L23312, doi:10.1029/2006GL027795.

Zollo, A., M. Lancieri, and S. Nielsen (2007), Reply to comment by P. Rydelek et al. on "Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records," *Geophys. Res. Lett.*, *34*, L20303, doi:10.1029/2007GL030560.

M. Lancieri, Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Napoli, 328 Via Diocleziano, I-80124 Napoli, Italy. (lancieri@ov.ingv.it)

A. Zollo, Dipartimento di Scienze Fisiche, Università di Napoli Federico II, I-80125 Naples, Italy.