



Fast evaluation of tsunami scenarios: uncertainty assessment for a Mediterranean Sea database

Irene Molinari^{1,2}, Roberto Tonini¹, Stefano Lorito¹, Alessio Piatanesi¹, Fabrizio Romano¹, Daniele Melini¹, Andreas Hoechner³, José M. González Vida⁴, Jorge Maciás⁵, Manuel J. Castro⁵, and Marc de la Asunción⁵

¹Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Roma 1, Via di Vigna Murata 605, 00143, Roma, Italy

²Department of Earth Sciences, Institute of Geophysics, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland

³GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany

⁴Dpto. Matemática Aplicada, Universidad de Málaga, EDANYA Group, 29071, Málaga, Spain

⁵Dpto. Análisis Matemático, Facultad de Ciencias, Universidad de Málaga, EDANYA Group, 29071, Málaga, Spain

Correspondence to: Irene Molinari (irene.molinari@erdw.ethz.ch) and Roberto Tonini (roberto.tonini@ingv.it)

Received: 27 April 2016 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 6 June 2016

Revised: 6 November 2016 – Accepted: 8 November 2016 – Published: 6 December 2016

Abstract. We present a database of pre-calculated tsunami waveforms for the entire Mediterranean Sea, obtained by numerical propagation of uniformly spaced Gaussian-shaped elementary sources for the sea level elevation. Based on any initial sea surface displacement, the database allows the fast calculation of full waveforms at the 50 m isobath offshore of coastal sites of interest by linear superposition. A computationally inexpensive procedure is set to estimate the coefficients for the linear superposition based on the potential energy of the initial elevation field. The elementary sources size and spacing is fine enough to satisfactorily reproduce the effects of $M \geq 6.0$ earthquakes. Tsunami propagation is modelled by using the Tsunami-HySEA code, a GPU finite volume solver for the non-linear shallow water equations. Like other existing methods based on the initial sea level elevation, the database is independent on the faulting geometry and mechanism, which makes it applicable in any tectonic environment. We model a large set of synthetic tsunami test scenarios, selected to explore the uncertainty introduced when approximating tsunami waveforms and their maxima by fast and simplified linear combination. This is the first time to our knowledge that the uncertainty associated to such a procedure is systematically analysed and that relatively small earthquakes are considered, which may be relevant in the near-field of the source in a complex tectonic setting. We find that non-linearity of tsunami evolution affects the reconstruction of the waveforms and of their maxima by introducing an almost unbiased (centred at zero) error distribution

of relatively modest extent. The uncertainty introduced by our approximation can be in principle propagated to forecast results. The resulting product then is suitable for different applications such as probabilistic tsunami hazard analysis, tsunami source inversions and tsunami warning systems.

1 Introduction

After the 2004 Indian Ocean tsunami, particular attention has been devoted to the improvement of tsunami warning systems (TWS) and probabilistic tsunami hazard analysis (PTHA), which currently represent two pillars in risk mitigation policies for the authorities of each country exposed to tsunami threat (Satake, 2014). At the same time, tsunami inversion techniques (Satake, 1987) have been greatly improved in the last decade (e.g. Lorito et al., 2016), characterised by a global surge of tsunamigenic earthquakes (Lay, 2015). Numerical modelling is nowadays a standard tool to accomplish all the above tasks.

However, the computational cost of numerical simulations still limits the feasibility for approaches which require (i) a very fast response and/or (ii) a massive amount of simulations, thus encouraging the development of efficient approximated solutions. Pre-calculated tsunami sources are commonly adopted by TWS to rapidly forecast tsunami effects which follow strong earthquakes. For example, stored scenarios are used in inversions of tsunami observations at DART

buoys and of seismic and geodetic data (e.g. NOAA/PMEL for the Pacific Ocean and GI-INA-TEWS project for Indonesia) or interpolated on the basis of real-time earthquake parameters (e.g. JMA for Japan and CENALT for France, for the north-eastern Atlantic Ocean and Western Mediterranean) (Bernard and Titov, 2015). Linear combinations (LCs) of elementary sources (ES) are also commonly used in earthquake source inversion (e.g. Yue et al., 2015), often jointly with other geophysical data (e.g. Romano et al., 2014). However, most of these databases only include large subduction earthquakes or other pre-defined faulting geometries and mechanisms, and they might be ineffective for areas characterised by complex tectonics as in the Mediterranean or Caribbean regions. Even along major megathrusts, TWS have been challenged several times in the last years by non-subduction earthquakes, such as outer-rise or strike-slip events: in these cases, approaches contemplating only the mapping of major subduction zones and megathrust (Gica et al., 2008) might lead to forecast failures. A similar argument holds for PTHA, in which a thorough exploration of the source variability in poorly mapped offshore source zones or around the major known faults is computationally very demanding (e.g. Geist and Parsons, 2006; Geist and Lynett, 2014; Lorito et al., 2015; Selva et al., 2016).

These limitations can be overcome by defining a database of ES for the sea level elevation that, properly queried and combined, is able to reproduce any tsunami initial condition and the corresponding tsunami impact, while significantly limiting the computational effort. No a priori assumptions about the seismic source geometry and the kind of tsunamigenic source are necessary, as long as the linear propagation of tsunami waves in deep water and the superposition principle hold. This methodology has been proposed in several studies, using Gaussian-shaped (e.g. Liu and Wang, 2008; Saito and Furumura, 2009; Saito et al., 2011; Tsushima et al., 2014; Mulia and Asano, 2015), pyramidal (Zaibo et al., 2003), conical or circular with positive elevation in the centre and negative at the edge (Choi et al., 2005; Zaitsev and Pelinovsky, 2011), rectangular prisms (Miranda et al., 2014) or cosine-tapered (Hossen et al., 2015) ES. The approach has been proposed in retrospective for rapid near-field forecasting of the Tohoku 2011 tsunami with tFISH/RAPiD (Tsushima et al., 2014), for PTHA (Selva et al., 2016) and for source inversion (Liu and Wang, 2008; Saito et al., 2011; Mulia and Asano, 2015; Hossen et al., 2015). However, such an approach has never been fully validated by a systematic assessment of the uncertainties that it introduces in the tsunami modelling or forecasting.

We present here a database of tsunami waveforms stored at densely spaced observation points (OPs) along the 50 m depth isobaths, obtained from a very large number of Gaussian-shaped tsunami ES covering the whole Mediterranean Sea. Given any static tsunami initial condition, the proposed procedure provides a rapid approximation of the corresponding full time history at any OP by LCs of the pre-

calculated waveforms associated to each selected ES. In addition to being independent of the source mechanism, the unit source size and density is suitable to satisfactorily reproduce not only the tsunamis generated by large earthquakes but also those generated by events as small as M6 earthquakes. The performance of this tool is analysed by quantifying its limits and errors in recovering an initial water displacement field and by assessing its usability in several different possible applications, such as probabilistic tsunami hazard analysis, tsunami source inversions and tsunami warning systems: for example, by propagating the estimated uncertainty in the probability distribution of the tsunami forecast (e.g. Annaka et al., 2007; Horspool et al., 2014).

2 Method and implementation

In this section we illustrate the approach followed to calculate the approximate tsunami waveforms generated by any given seismic source. The method is based on LCs of the contributions of elementary sea level displacement recorded at the 50 m isobath contour.

2.1 Elementary sources

The whole Mediterranean Sea is covered with a dense grid of $\sim 53\,000$ regularly spaced tsunami ES, placed at a distance of about 7 km in both north–south and west–east directions (Fig. 1). Each ES is described by a 2-D Gaussian function as follows:

$$\xi_i(x, y) = h e^{-\frac{(x-x_i)^2+(y-y_i)^2}{2\sigma^2}}, \quad (1)$$

where (x_i, y_i) is the centre of the i th ES, $h = 10$ m and $\sigma = 4$ km. The choice of these parameters is based on a trial and error procedure, during which different Gaussians sizes were tested. The chosen σ ensures reaching a compromise between the spatial resolution needed to approximate the relatively small-wavelength deformation field caused by earthquakes down to $M = 6.0$, while still having a sufficient number of grid points to represent the Gaussian field for unit source propagation.

2.2 Tsunami modelling

Numerical simulations have been performed using the Tsunami-HySEA code (de la Asunción et al., 2013) solving the non-linear shallow water equations in spherical or Cartesian coordinates using a hybrid numerical scheme that combines a finite difference (FD) two-step scheme similar to leap-frog for the propagation phase and a second-order finite volume (FV) flux-limiter TVD-weighted average flux (WAF) flux-limiter scheme for the inundation step. The combination of both schemes guarantees the mass conservation in the complete domain and prevents the generation of spurious high-frequency oscillations near discontinuities generated by

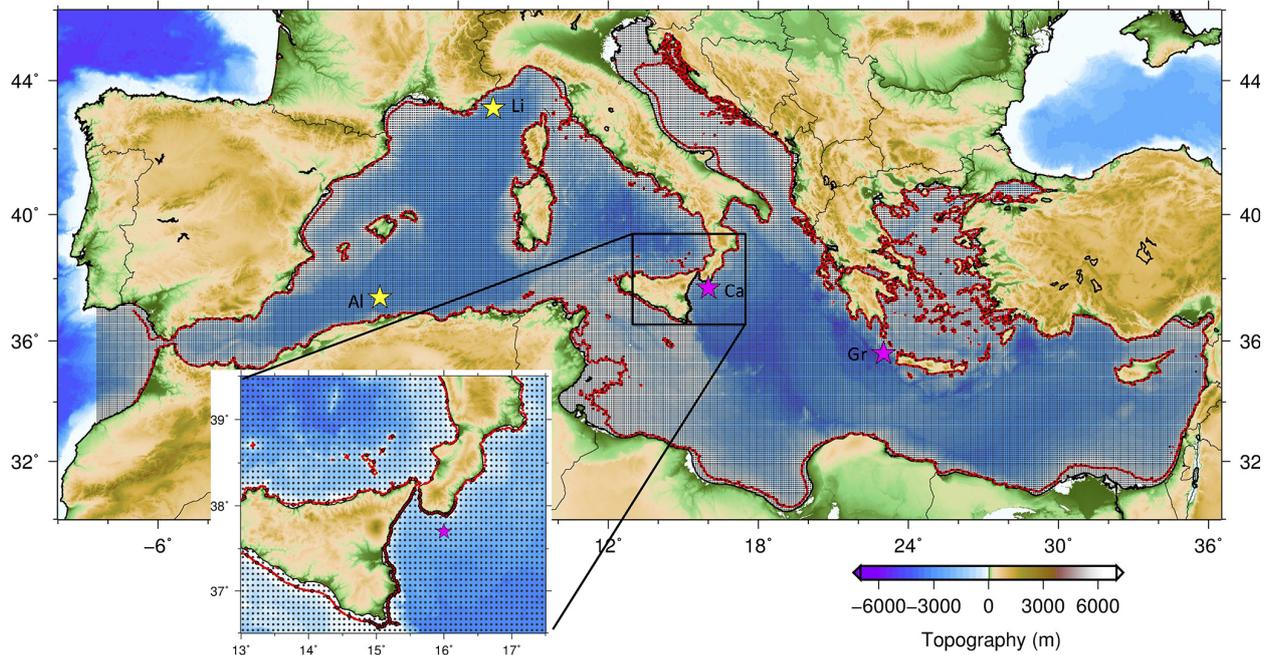


Figure 1. Spatial distribution of the Gaussian-shaped elementary sources (black dots) covering the Mediterranean Sea and position of tsunami receivers on the 50 m isobaths (red dots) where the pre-computed tsunami waveforms are evaluated. Yellow (maximum considered magnitude up to 8.0) and magenta (maximum considered magnitude up to 8.5) stars mark the epicentres used in our performance analysis presented in Sect. 3 (Al is Algeria, Li is Liguria, Ca is Calabria and Gr is Greece).

leap-frog type schemes. Moreover, this hybrid method provides a convenient scheme which uses FD in open sea, where the solution is smooth, and FV in coastal areas, where the solution can be characterised by strong discontinuities. The Tsunami-HySEA numerical model has undergone proper benchmarking (Macías et al., 2016) according to the community standards (e.g. Synolakis et al., 2008) within the framework of the US tsunami hazard program (<http://nws.weather.gov/nthmp/>). The code is implemented in CUDA (Compute Unified Device Architecture) and runs in multi-GPU architectures, yielding remarkable speedups in comparison with other CPU-based codes. Running the non-linear shallow water version of the code increases the simulation time by only $\sim 10\%$ with respect to the linear version. We simulate 8 h of propagation on the 30 arcsec bathymetric model SRTM30+ (http://topex.ucsd.edu/WWW_html/srtm30_plus.html), collecting the waveforms at the $\sim 13\,000$ OPs each ~ 2 km along the 50 m isobath. The waveforms are sampled each 30 s, a value that allows us to limit the final store size to ~ 5 TB while still sampling densely enough typical tsunami wavelengths.

2.3 Reconstruction and forecasting procedure

We follow two main steps to reproduce the tsunami generated by a given seismic source: (i) finding the coefficients for an approximated representation of the initial (I) water vertical displacement $Z_I(x, y)$ by LCs of a local ES subset and

(ii) combining accordingly the tsunami waveforms associated to each selected ES at each OP.

To find the coefficients, the ES whose centres fall in the area where $Z_I(x, y)$ is non-negligible ($> \frac{1}{100} \max(|Z_I|)$) are selected, and the values of $Z_I(x, y)$ at their centres $\alpha = \alpha_1, \dots, \alpha_{ng}$ are extracted, where ng is the total number of ES considered. The equivalent water displacement tsunami $Z_{EQ}(x, y)$ (EQ being equivalent) is then obtained by linearly combining the selected ES with weights α as

$$Z_{EQ}(x, y) = C_S \sum_{i=1}^{ng} \alpha_i \xi_i(x, y) = C_S Z_{SUM}(x, y), \quad (2)$$

where ξ_i is the i th ES defined in Eq. (1), Z_{SUM} is the weighted summation of all the selected ES and C_S is a scaling coefficient:

$$C_S = \frac{\max(Z_I(x, y)) - \min(Z_I(x, y))}{\max(Z_{SUM}(x, y)) - \min(Z_{SUM}(x, y))}. \quad (3)$$

For any tsunami scenario, C_S scales $Z_{EQ}(x, y)$ to the same maximum peak-to-trough distance of $Z_I(x, y)$ (see Fig. 2a to c). Finally, if $G_{i,m}(t)$ is the waveform generated by the i th ES at the location of the m th OP, an estimate of the tsunami time history $\zeta_m(t)$ at the m th OP may be obtained:

$$\zeta_m(t) = C_S \sum_{i=1}^{ng} \alpha_i G_{i,m}(t). \quad (4)$$

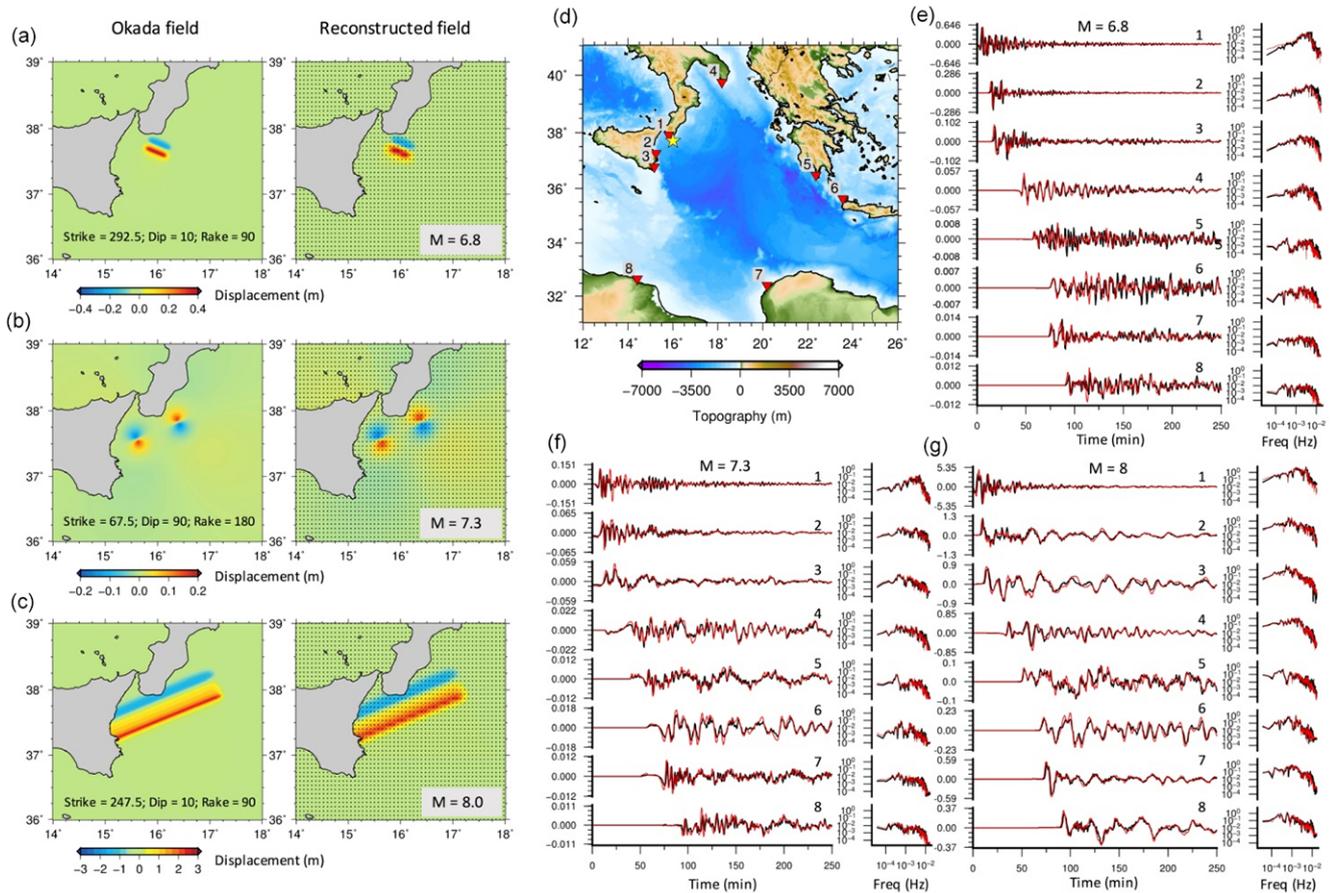


Figure 2. (a–c) Three examples of original initial conditions with different faulting parameters and earthquake magnitudes (left panels) and their reconstruction (right panels) obtained through the linear combination of the Gaussian-shaped sources. (d) Tsunami receivers (red triangles) and epicentres (yellow star) of scenarios in (a)–(c). (e–g) Comparison between simulated (black lines, NS in the main text) and linearly combined (red lines, LC in the main text) waveforms and frequency spectra, corresponding to the scenarios in (a)–(c). In panels (e) to (g) the vertical axes show the recorded wave heights expressed in metres.

The use of the scaling coefficient, C_S , also ensures a good corresponding maximum-to-minimum scaling of the combined waveforms, as shown in Fig. 2e and g. C_S coefficients could also be retrieved by linear inversions; however, in the present study, we seek for a balance between accuracy and an inexpensive and fast procedure.

In the next section, we test this method extensively to quantify the accompanying uncertainty, trace back the relative contributions of the different uncertainty sources and revise the method accordingly, in order to reduce the bias introduced by our approach.

3 Performance analysis

To test the proposed approach, we first visually compared the original and reconstructed initial conditions for several earthquake scenarios. Each scenario is represented with a rectangular fault with length and width assigned by the Wells and

Coppersmith (1994) empirical scaling relations. The seafloor displacement is calculated with Okada (1992) analytic expressions. Then, the corresponding tsunami waveforms obtained by LCs of the waveforms corresponding to the selected Gaussian ES are compared with those obtained by direct numerical simulations (labelled NS, hereafter) starting from the original initial condition. For NS, non-hydrostatic effects in the transfer of the displacement to the water column are approximated following Kajiura (1963), and the tsunami is then propagated using Tsunami-HySEA. All the waveforms with maximum amplitude less than 0.05 m are discarded as not significant for operational applications.

The initial conditions are qualitatively well reproduced in most of the cases; some examples are shown in Fig. 2. Small sources (Fig. 2a, b) and sharp changes in the initial field (Fig. 2c) are the most difficult to reproduce, because they contain features smaller than the resolution offered by the size and density of ES. The agreement between the corresponding tsunami waveforms is satisfactory in both time and

Table 1. Focal mechanisms and depths of the top of the faults considered testing the performances of the database in the Mediterranean Sea.

Fault mechanism	Strike (degree)	Dip (degree)	Rake (degree)
Normal	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	30	−90
	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	50	−90
	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	70	−90
Reverse	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	10	90
	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	30	90
	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	50	90
Strike-slip	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	70	0
	22.5, 67.5, 112.5, 157.5	90	0
	22.5, 67.5, 112.5, 157.5, 202.5, 247.5, 292.5, 337.5	70	180
	22.5, 67.5, 112.5, 157.5	90	180

frequency domain and for OPs in the near- and far-field of the source (Fig. 2e–g). However, we note small phase shifts and a slight amplitude overestimation.

A more thorough quantitative analysis is deemed necessary to assess limitations and uncertainties introduced by the method. Therefore, we test a large number of realistic earthquake scenarios with epicentres located in four areas where tsunamigenic earthquakes may occur (Fig. 1): offshore Algeria, Liguria and Calabria (Italy) and Crete (Greece). For each epicentre, we explore the dependence on the variation of the source parameters: the magnitude ($M = 6.0, 6.3, 6.8, 7.3, 7.7, 8.1$ and 8.5), the focal mechanism (72 combinations of strike, dip and rake; see Table 1) and the depth of the top of the fault (3 and 12 km). Approximately 4000 scenarios have been considered and the corresponding tsunami signals at ~ 1680 OPs (approximately each 8th point at an average distance of ~ 16 km), leading to a statistically robust amount of analysed waveforms (~ 6800000).

First, we analyse the misfit (Sect. 3.1) between LC and NS waveforms and then we also perform a comparison between the LC and NS maximum wave amplitudes (Sect. 3.2) in order to quantify the uncertainty related to different quantities, possibly required by different specific applications.

We argue that the main uncertainty sources are (i) the misfit between $Z_I(x, y)$ and $Z_{EQ}(x, y)$ and (ii) the linearity assumption and we analyse them separately to determine their relative importance (Sect. 3.3). Since we find that, between the two, the initial field reconstruction introduces a larger bias in the final result, an improvement of the reconstruction technique is proposed and verified in Sect. 3.4.

3.1 Prediction of the whole waveforms

The overall agreement between the waveforms predicted by LC and the corresponding ones obtained by NS is evaluated through the calculation of the misfit for each scenario and each OP. The misfit is defined through a cost function frequently used to compare tsunami signals in source inversions

(e.g. Romano et al., 2015):

$$E = 1 - \frac{2 \sum_{i=1}^{nt} h_{NS}(t_i) h_{LC}(t_i)}{\sum_{i=1}^{nt} h_{NS}^2(t_i) + \sum_{i=1}^{nt} h_{LC}^2(t_i)}, \quad (5)$$

where $h_{LC}(t)$ and $h_{NS}(t)$ are waveforms at a given OP obtained through LC and NS, respectively, and nt is the number of considered time steps. This cost function is computed considering the first 2 h of the tsunami waveform starting from the first-arrival time automatically detected as a change greater than the 2% of the absolute maximum/minimum of the whole waveform.

Overall, the waveforms are reproduced quite well, as shown by the rather narrowly peaked misfit distribution (Fig. 3a), whose median value (ν , in Fig. 3, i.e. the 50th percentile) is smaller than the mean value (μ , in Fig. 3). The analysis with respect to the earthquake magnitude, the faulting mechanism and the receiver location indicates that the misfit is most sensitive to earthquake magnitude (Fig. 3b–d) and to a lesser extent to the earthquake mechanism (not shown). The results have then been grouped into three classes depending on earthquake magnitude (i.e. strong, $M < 7$; major, $7 \leq M < 8$; great, $M \geq 8$). The misfit distribution is significantly wider for smaller magnitudes; this is explained by the inability (resolution) to represent the initial field in terms of Gaussian ES when the wavelength of the field is comparable to the size of the ES (Fig. 2b).

3.2 Prediction of maximum tsunami amplitudes

Maximum offshore tsunami amplitude is a widely used metric for both TWS and PTHA. For example, it is been used by Selva et al. (2016), who applied the same ES database and approach described in Sect. 2.1–2.3.

The differences between the maximum wave amplitudes predicted by the NS (H_{NS}) and by the LC (H_{LC}) are visualised in the scatter plots (H_{NS} vs. H_{LC} ; Fig. 3e–h) and in the histograms of the percentage error of H_{LC} , taking H_{NS} as the reference value, for all the OPs (Fig. 3i–l); the results are again grouped according to magnitude classes. The distribu-

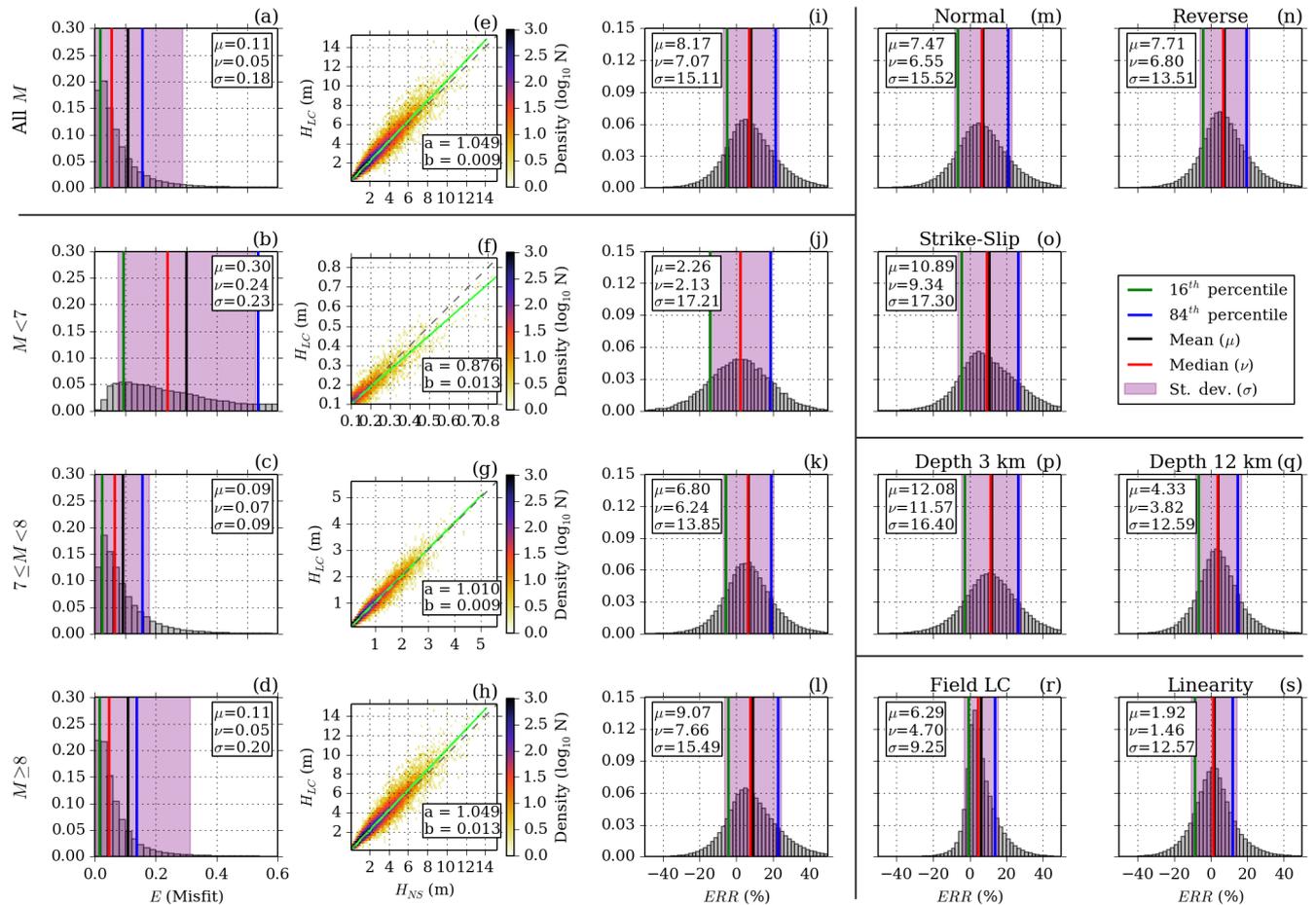


Figure 3. Validation results: (a) misfit between the considered LC and NS waveforms; (b–d) misfits grouped by earthquake magnitude; (e–h) scatter plots between LC and NS tsunami maxima; (i) LC percentage error with respect to NS; (j–l) percentage errors grouped by earthquake magnitude; (m–o) percentage errors grouped by faulting mechanisms; (p–q) percentage errors grouped by top-of-the-fault depths; (r) NL percentage error with respect to NS; (s) LC percentage error with respect to NL.

tion of points is well fitted by a line, with a smaller scatter corresponding to the highest tsunami amplitudes (Fig. 3e). The trend (green line) indicates that the LC slightly overestimates the target NS amplitudes. This slight overestimation occurs mostly at the highest magnitudes (Fig. 3h). A worse overall agreement is found for the lowest magnitudes, which show an opposite underestimation trend (Fig. 3f). The above described behaviour is illustrated also by the percentage error distributions. The group of lower magnitudes is characterised by the larger standard deviation but, conversely, its mean and median are smaller than the groups with larger magnitudes (Fig. 3j–l); that is they are characterised by a smaller bias. The overall mean and standard deviation are $\sim 8\%$ (overestimation) and $\sim 15\%$, respectively (Fig. 3i). In all cases, the 16th and 84th percentile are narrower than the standard deviation, indicating depleted tails; i.e. no excess of events with a poor matching is found compared to a Gaussian distribution.

The results have also been analysed by separating the scenarios according to earthquake faulting mechanisms

(Fig. 3m–o). The worst agreement occurs for strike-slip mechanisms, likely because of the relative complexity of their associated displacement fields with respect to the – most tsunamigenic – thrust and normal mechanisms.

When grouping the events according to fault depth (Fig. 3p, q), we find that shallower faults result in a greater overestimation with respect to the deeper ones, since their co-seismic fields are sharper. Again, in all cases, the tails are thinner than those of a Gaussian distribution.

3.3 Uncertainties in initial field reconstruction and validity of the linearity assumption

Since the LC procedure presented here results in an average overestimation with respect to the NS waveform maxima of $\sim 8\%$ (Fig. 3i), it is important to trace back the main causes of this bias. We here assume that the two main sources of uncertainty are (i) the misfit between the $Z_I(x, y)$ and $Z_{EQ}(x, y)$ and (ii) the linearity assumption. Thus, we simulate the tsunamis associated to the ~ 4000 reconstructed

tsunami initial conditions $Z_{EQ}(x, y)$ (hereafter labelled as NL, standing for non-linear). Then, we compute both the percentage error of NL (H_{NL}), taking NS (H_{NS}) as the reference value, and the percentage error of LC (H_{LC}), taking NL (H_{NL}) as the reference value (Fig. 3r and s). The differences between the maximum wave amplitudes resulting from the reconstructed (NL) and the reference displacement (NS) fields are, on the average, of $\sim 6.5\%$ (Fig. 3r). The very narrow distribution indicates a quite homogeneous overestimation, except for its negative skewness toward lower values. The uncertainty deriving from non-linear propagation is exemplified in Fig. 3s (mean $\sim 1.9\%$). That is, the linearity assumption introduces a very limited bias since, for example, the mean is much closer to zero in this case; conversely, the distribution of the values around the mean is slightly larger than in Fig. 3r. Hence, the proposed method reconstructs slightly amplified initial displacement fields mostly due to inaccurate reconstruction of the initial field, which is in turn likely due to the aforementioned effects from the summation of the tails of the Gaussians, whose support is not compact, and to resolution limits. Conversely, the linearity assumption holds quite well in the sense that the forecast is almost unbiased with reasonable dispersion around the central value.

3.4 Improvement of the initial field reconstruction

We then aim to reduce the bias introduced by the inaccurate reconstruction of the tsunami initial condition $Z_I(x, y)$ (see Sect. 2.3). We propose and test two method refinements to enhance the conformity of the initial condition and to improve the overall reconstruction of the initial displacement condition.

The first refinement serves to correct the uneven sampling of the Gaussian ES on a grid that has constant spacing of 30 arcsec in both north–south and east–west directions. The ES constant width ($\sigma = 4$ km, about 20 km base width) is approximated with a spacing of 4.5 arcmin in the longitude and 4 arcmin in the latitude direction. This results in a more pronounced overlap of the ES towards north. More precisely, if all the weights α from Sect. 2.3 are set to 1, in the north the total contribution of all ES at the centre of an ES amounts to 230 %, while in the south it is 190 %. Additionally, there is a further distortion effect caused by “cropping” of the ES falling inland. For instance, if an ES close to the coast has only one ES neighbour, e.g. because it is situated in narrow bay, their joint contribution amounts to $\sim 120\%$. In order to reduce these latitude/cropping distortion effects on the initial condition, we divide the weights α by the local corrections β before applying the global scaling factor, defined as follows:

$$\beta_j = \frac{1}{h} \sum_{j=1}^{ng} \xi_j(x_j, y_j), \quad (6)$$

where the symbols are the same as in Eq. (1). The effect of such correction is an improvement of $\sim 2\%$ towards a zero

mean (unbiased) distribution of the differences between the maximum wave amplitudes, while the associated uncertainties remain comparable to the previous result, as shown in Fig. 4a and c.

As a further step towards an improved representation of the initial water displacement $Z_I(x, y)$, we also introduce an alternative method to calculate the scaling factor C_S . Instead of preserving maximum peak-to-trough distance, we preserve the potential energy (C_S^{Enc}) of the tsunami initial condition, as defined in the following formula:

$$C_S^{Enc} = \frac{\int (Z_I)^2 dA}{\sum_{i=1}^{ng} \left(\frac{\alpha_i}{\beta_i}\right)^2 \cdot A_i}, \quad (7)$$

where A_i is the area of each element of the ES grid, which depends on the latitude.

The performances of the above corrections are tested against the direct fully non-linear simulation (NS) of the tsunamis generated by the target initial field; i.e. Fig. 4b and d should be compared to Fig. 3e and i. As shown both by the scatter plot and the histogram in Fig. 4b and d, the combined effect of the latitude–crop correction and of the potential energy preservation technique significantly improves the results. Indeed, now we obtain an almost centred distribution with $\mu = 0.45$; that is the corrections almost eliminate the systematic overestimation for amplitude preservation shown in Fig. 3i and Fig. 4c, providing an almost unbiased estimates, even lower than the one found in the linearity test (Fig. 3s). However, the total standard deviation is similar ($\sigma \sim 15$) and on the same order of magnitude as that due to non-linearity; hence, this effect is likely mainly due to local propagation effects around the receivers rather than to be related to the initial field reconstruction.

4 Discussion and conclusions

We present here a source mechanism-free tool to rapidly reconstruct the full waveform and the maximum wave heights predicted by any static tsunami initial water displacement, independently from any a priori assumptions on fault geometry. The reconstruction is obtained through a linear combination of a pre-computed database of tsunami waveforms generated using tsunami elementary sources.

For the first time, the validity of the method has been systematically tested against a wide range of realistic scenarios (~ 4000) by varying all the earthquake parameters, regarding its ability of forecasting full tsunami waveforms and maximum amplitudes at a very large number of forecast points placed along the 50 m isobath along the Mediterranean coasts. This analysis points out that the main source of bias is the amplitude-preserving reconstruction of the initial condition: adopting the simplest reconstruction strategy, we have seen that the tool provides a reasonable fit of the full waveforms and, on the average, the target tsunami elevation max-

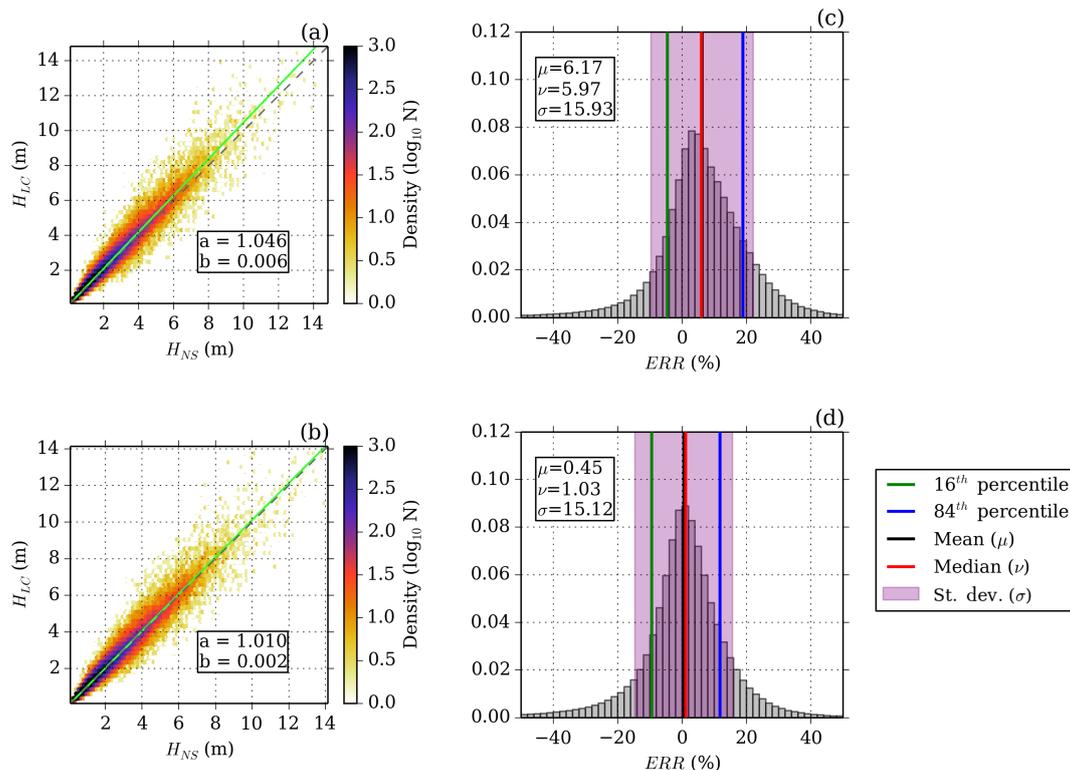


Figure 4. Validation results. Misfit between the considered LC and NS waveforms using (top) the latitude and coastal crop (lat/crop) correction, (bottom) lat/crop correction and potential energy preservation, for all magnitudes and mechanisms. (a, b) Scatter plots between LC and NS tsunami maxima and (c, d) LC percentage error with respect to NS for different method refinements.

ima are biased (overestimated) by $\sim 8\%$. The non-linearity of the propagation turns out to be weak enough for allowing an almost unbiased forecast. We also tested a different method for the estimation of the linear combination coefficients based on the potential energy, resulting as well in an overall almost unbiased ($\sim 0.5\%$) tsunami forecast. In all of the cases the standard deviation is $\sim 15\%$.

We point out that we have populated the ES database using non-linear shallow water equation simulations since performing NS with Tsunami-HySEA code increases the computational time by only $<10\%$ with respect to the linear scheme version; moreover, we were expecting only weak non-linearity, as the results confirmed, being most of the propagation in deep enough waters. Hence, we have not judged necessary to put efforts in switching off the non-linear terms in the code. Moreover, one future project we envisage is to investigate other reconstruction methods, such as the reduced base methods (Quarteroni and Rozza, 2014), that benefit from the (weak) non-linearity of the phenomenon to produce reconstructions of the initial waveforms suitable for closer to the coast predictions, while using the same database.

We consider that the results provided by our method are satisfactory for most of the practical applications such as probabilistic tsunami hazard analysis, tsunami source inver-

sion and tsunami warning systems. The present tool, in fact, has been already successfully used to develop an event-tree-based PTHA methodology which accounts for both aleatory and epistemic uncertainty (Selva et al., 2016), and it will be further applied, including the corrections introduced in this paper, to the first national PTHA in Italy and to the first homogeneous PTHA in the NEAM (north-eastern Atlantic, Mediterranean and connected seas) region (<http://www.tsumaps-neam.eu>).

Moreover, since they basically contain no bias, the uncertainty introduced by the approximations used can be propagated in a straightforward manner into the uncertainty associated to the final results, for example when defining the parameters of a log-normal distribution of the hazard impact metric, to be convolved with the probability density function (PDF) of representing different sources of aleatory uncertainty such as the natural variability of the earthquake source or the contribution of the tidal stage (see Annaka et al., 2007, and Horspool et al., 2014, for some examples).

An important advantage in TWS applications is that our database will allow managing the regime of large epistemic uncertainty concerning the faulting mechanism, when either fast moment tensors or direct tsunami measurements are not immediately available after a potentially tsunamigenic earthquake. This is almost always the case in the Mediterranean

or Caribbean seas, where, due to tectonic complexity combined to short tsunami arrival times at the coast, the faulting mechanism is highly unpredictable and its rapid estimation very challenging. The situation, however, applies to any potential source zone of large enough crustal earthquakes in the near-field of the coast.

A wide range of faulting mechanisms can be in fact readily explored using this database; the search can be guided by prior knowledge of the regional past seismicity and tectonic setting. For example, if an earthquake of a certain magnitude happens with a certain hypocentre, the PTHA event tree (Lorito et al., 2015; Selva et al., 2016) can be accessed for deriving a discrete PDF for the other earthquake parameters. Using the database, all the tsunami scenarios corresponding to this discrete PDF can be evaluated simultaneously at each given coastal point of interest; their values can be weighted with the values of the source PDF for finally obtaining a PDF for the tsunami forecast at the site. In other words, the (epistemic – as the event already happened and could, in principle, be measured) source uncertainty could be efficiently mapped in a probabilistic tsunami forecast through weighting according to the source mechanism probability of simultaneous evaluation of a number of tsunami scenarios.

In the presence of fast moment tensor solutions, the forecast uncertainty can be promptly reduced, while still incorporating errors in the real-time seismic solutions, by combining the latter with a priori assumptions on the source mechanism probability. These aspects will be better addressed in a future study that deals with the implementation of this tool for the Italian NEAMTWS Tsunami Service Provider (e.g. Bernardi et al., 2015).

5 Data availability

The underlying data, tsunami waveforms database and results are not available to the public. For scientific collaboration and data usage, interested researchers are invited to get in contact with the authors.

Author contributions. Irene Molinari, Roberto Tonini, Alessio Piatanesi and Stefano Lorito conceived the method and analysed the results; Irene Molinari, Roberto Tonini and Stefano Lorito wrote the manuscript; Irene Molinari and Roberto Tonini prepared the figures; Irene Molinari, Roberto Tonini and Andrea Hoechner performed numerical simulations and the performance analysis; Fabrizio Romano provided input and support to numerical simulations; Daniele Melini provided HPC support for numerical simulations; Jose M. González Vida, Jorge Macías, Manuel J. Castro and Marc de la Asunción provided the GPU code for numerical simulations. All authors reviewed the manuscript.

Acknowledgements. This work has been funded by: (i) the flagship project RITMARE funded by the Italian Ministry of Research and Education; (ii) the INGV-DPC Agreement Annex B2; (iii) the EU-project ASTARTE – Assessment, Strategy And Risk Reduction for Tsunamis in Europe – FP7-ENV2013 6.4-3, Grant 603839; (iv) the TSUMAPS-NEAM project, Agreement Number: ECHO/SUB/2015/718568/PREV26, co-financed by the European-Union Civil Protection Mechanism; and (v) the Spanish Government Research projects MTM2012-38383-C02-01 and MTM2015-70490-C2-1-R. We thank A. Armigliato and E. Pelinovsky for constructive comments, that helped to improve the manuscript. Figures have been prepared using the Generic Mapping Tools (Wessel and Smith, 1998), Numpy (van der Walt et al., 2011) and Matplotlib (Hunter, 2007) tools.

Edited by: I. Didenkulova

Reviewed by: A. Armigliato and E. Pelinovsky

References

- Annaka, T., Satake, K., Sakakiyama, T., Yanagisawa, K., and Shuto, N.: Logic-tree Approach for Probabilistic Tsunami Hazard Analysis and its Applications to the Japanese Coasts, *Pure Appl. Geophys.*, 164, 577–592, doi:10.1007/s00024-006-0174-3, 2007.
- Bernard, E. and Titov, V.: Evolution of tsunami warning systems and products, *Phil. Trans. R. Soc. A*, doi:10.1098/rsta.2014.0371, 2015.
- Bernardi, F., Lomax, A., Michelini, A., Lauciani, V., Piatanesi, A., and Lorito, S.: Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian Candidate Tsunami Service Provider, *Nat. Hazards Earth Syst. Sci.*, 15, 2019–2036, doi:10.5194/nhess-15-2019-2015, 2015.
- Choi, B. H., Pelinovsky, E., Lee, H. J., and Woo, S. B.: Estimates of tsunami risk zones on the coasts adjacent to the East (Japan) Sea based on the synthetic catalogue, *Nat. Hazards*, 36, 355–381, 2005.
- De la Asunción, M., Castro, M. J., Fernández-Nieto, E. D., Mantas, J. M., Ortega Acosta, S., and González Vida, J. M.: Efficient GPU implementation of a two waves TVD-WAF method for the two-dimensional one layer shallow water system on structured meshes, *Comput. Fluids*, 80, 441–452, 2013.
- Geist, E. L. and Parsons, T.: Probabilistic Analysis of Tsunami Hazards, *Nat. Hazards*, 37, 277–314, 2006.
- Geist, E. L. and Lynett, P. J.: Source processes for the probabilistic assessment of tsunami hazards, *Oceanography*, 27, 86–93, 2014.
- Gica, E., Spillane, M., Titov, V. V., Chamberlin, C., and Newman, J. C.: Development of the forecast propagation database for NOAA's Short-term Inundation Forecast for Tsunamis (SIFT), NOAA Tech. Memo. OAR PMEL-139, 89 pp., 2008.
- Hossen, M. J., Cummins, P. R., Dettmer, J., and Baba, T.: Time reverse imaging for far-field tsunami forecasting: 2011 Tohoku earthquake case study, *Geophys. Res. Lett.*, 42, 9906–9915, 2015.
- Horspool, N., Pranantyo, I., Griffin, J., Latief, H., Natawidjaja, D. H., Kongko, W., Cipta, A., Bustaman, B., Anugrah, S. D., and Thio, H. K.: A probabilistic tsunami hazard assessment for Indonesia, *Nat. Hazards Earth Syst. Sci.*, 14, 3105–3122, doi:10.5194/nhess-14-3105-2014, 2014.

- Hunter, J. D.: Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.*, 9, 90–95, 2007.
- Kajiura, K.: The leading wave of a tsunami, *Bull. Earthquake Res. Inst. Univ., Tokyo*, 41, 535–571, 1963.
- Lay, T.: The surge of great earthquakes from 2004 to 2014, *Earth Planet. Sc. Lett.*, 409, 133–146, 2015.
- Liu, L. F. and Wang, X.: Tsunami Source Region Parameter Identification and Tsunami Forecasting, *J. Earthq. Tsunami*, 02, 87, doi:10.1142/S179343110800027X, 2008.
- Lorito, S., Romano F., and Lay, T.: Tsunamigenic Major and Great Earthquakes (2004–2013): Source Processes Inverted from Seismic, Geodetic, and Sea-Level Data, in: *Encyclopedia of Complexity and Systems Science*, edited by: Meyers, R. A., Springer Science+Business Media New York, 2015, ISBN-13: 978-0-387-30440-3, 2016.
- Lorito, S., Selva J., Basili R., Romano, F., Tiberti, M. M., and Piatanesi, A.: Probabilistic Hazard for Seismically-Induced Tsunamis: Accuracy and Feasibility of Inundation Maps, *Geophys. J. Int.*, 200, 574–588, 2015.
- Maciás, J., Castro, M. J., Ortega, S., Escalante, C., and González-Vida, J. M.: Tsunami-HySEA Benchmark results, in: *NTHMP report for the MMS Benchmarking Workshop: Tsunami Currents*, 9 February 2015, doi:10.13140/RG.2.2.22999.47527, 2016.
- Miranda, J. M., Baptista, M. A., and Omira, R.: On the use of Green's summation for tsunami waveform estimation: a case study, *Geophys. J. Int.*, 199, 459–464, 2014.
- Mulia, I. E. and Asano, T.: Randomly distributed unit sources to enhance optimization in tsunami waveform inversion, *Nat. Hazards Earth Syst. Sci.*, 15, 187–196, doi:10.5194/nhess-15-187-2015, 2015.
- Okada, Y.: Internal deformation due to shear and tensile faults in a half-space, *Bull. Seism. Soc. Am.*, 82, 1018–1040, 1992.
- Quarteroni, A. and Rozza, G. (Eds.): *Reduced Order Methods for Modeling and Computational Reduction*, MS&A Series, Volume 9, Springer International Publishing, doi:10.1007/978-3-319-02090-7, 334 pp., 2014.
- Romano, F., Molinari, I., Lorito, S., and Piatanesi, A.: Source of the 6 February 2013 $M_w = 8.0$ Santa Cruz Islands Tsunami, *Nat. Hazards Earth Syst. Sci.*, 15, 1371–1379, doi:10.5194/nhess-15-1371-2015, 2015.
- Romano, F., Trasatti, E., Lorito, S., Piromallo, C., Piatanesi, A., Ito, Y., Zhao, D., Hirata, K., Lanucara, P., and Cocco, M.: Structural control on the Tohoku, earthquake rupture process investigated, by 3D FEM, tsunami and geodetic data, *Sci. Rep.*, 5631, 1–11, 2014.
- Saito, T. and Furumura, T.: Scattering of linear long-wave tsunamis due to randomly fluctuating sea-bottom topography: coda excitation and scattering attenuation, *Geophys. J. Int.*, 177, 958–965, 2009.
- Saito, T., Ito, Y., Inazu, D., and Hino, R.: Tsunami source of the 2011 Tohoku-Oki earthquake, Japan: Inversion analysis based on dispersive tsunami simulations, *Geophys. Res. Lett.*, 38, L00G19, doi:10.1029/2011GL049089, 2011.
- Satake, K.: Inversion of tsunami waveforms for the estimation of a fault heterogeneity: method and numerical experiments, *J. Phys. Earth*, 35, 241–254, 1987.
- Satake, K.: Advances in earthquake and tsunami sciences and disaster risk reduction since the 2004 Indian ocean tsunami, *Geoscience Letters*, 1, doi:10.1186/s40562-014-0015-7, 2014.
- Selva, J., Tonini, R., Molinari, I., Tiberti, M. M., Romano, F., Grezio, A., Melini, D., Piatanesi, A., Basili, R., and Lorito, S.: Quantification of source uncertainties in Seismic Probabilistic Tsunami Hazard Analysis (SPTHA), *Geophys. J. Int.*, 205, 1780–1803, doi:10.1093/gji/ggw107, 2016.
- Synolakis, C. E., Bernard, E. N., Titov, V. V., Kânuolu, U., and González, F. I.: Validation and verification of tsunami numerical models, *Pure Appl. Geophys.*, 165, 2197–228, 2008.
- Tsushima, H., Hino, R., Ohta, Y., Iinuma, T., and Miura, S.: tFISH/RAPiD: Rapid improvement of near-field tsunami forecasting based on offshore tsunami data by incorporating onshore GNSS data, *Geophys. Res. Lett.*, 41, 3390–3397, 2014.
- van der Walt, S., Colbert, S. C., and Varoquaux, G.: The NumPy Array: A Structure for Efficient Numerical Computation, *Comput. Sci. Eng.*, 13, 22–30, 2011.
- Wells, D. L. and Coppersmith, K. J.: New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seism. Soc. Am.*, 84, 974–1002, 1994.
- Wessel, P. and Smith, W. H. F.: New, improved version of the Generic Mapping Tools released, *EOS Trans. AGU*, 79, 579, doi:10.1029/98EO00426, 1998.
- Yue, H., Lay, T., Li, L., Yamazaki, Y., Cheung, K. F., Rivera, L., Hill, E. M., Sieh, K., Kongko, W., and Muhari, A.: Validation of linearity assumptions for using tsunami waveforms in joint inversion of kinematic rupture models: application to the 2010 Mentawai Mw 7.8 tsunami earthquake, *J. Geophys. Res.*, 120, 1728–1747, 2015.
- Zaibo, N., Pelinovsky, E., Kurkin, A., and Kozelkov, A.: Estimation of far-field tsunami potential for the Caribbean Coast based on numerical simulation, *Science Tsunami Hazards*, 21, 202–222, 2003.
- Zaitsev, A. I. and Pelinovsky, E. N.: Forecasting of tsunami wave heights at the Russian Coast of the Black Sea, *Oceanology*, 51, 907–915, 2011.