The effect of uncertainty in earthquake fault parameters on the maximum wave height from a tsunami propagation model

D. Burbidge\textsuperscript{1,*}, C. Mueller\textsuperscript{2}, and W. Power\textsuperscript{2}

\textsuperscript{1}Geoscience Australia, Canberra, Australia
\textsuperscript{2}GNS Science, Lower Hutt, New Zealand
*now at: GNS Science, Lower Hutt, New Zealand

Received: 26 March 2015 – Accepted: 27 April 2015 – Published: 22 May 2015

Correspondence to: D. Burbidge (d.burbidge@gns.cri.nz)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Over the last decade tsunami propagation models have been used extensively for both tsunami forecasting and hazard and risk assessment. However, the effect of uncertainty in the earthquake source parameters on the results of the tsunami model has not always been examined in great detail. Here we have undertaken a systematic study of the uncertainty in the maximum wave height of a tsunami ($h_{\text{max}}$) as a function of the uncertainty in the rupture parameters of the earthquake that generates it (specifically the strike, dip, rake, depth and magnitude). We have shown that even for the simple case of a tsunami propagating over flat bathymetry, the Coefficient of Variation (CoV) and skewness of the distribution of $h_{\text{max}}$ was a complex function of the choice of rupture parameter, distance and azimuth. The relationships became even more complex as the bathymetry used became more realistic. This has major potential implications for both how warning centres operate in the future and how the uncertainty in parameters describing the source should be incorporated into future probabilistic tsunami hazard assessments.

1 Introduction

Since the 2004 Indian Ocean tsunami, there has been a major increase globally in tsunami propagation modelling for use in both tsunami warning and hazard and risk assessment. Probabilistic tsunami hazard assessments (PTHAs) have been created for the US (Geist and Parsons, 2006; González et al., 2009), Australia (Burbidge et al., 2008, 2009), New Zealand (Power et al., 2007; Power, 2013), the Mediterranean (Sørensen et al., 2012), the Northwest Indian Ocean (Heidarzadeh and Kijko, 2011), Indonesia (Horspool et al., 2014) and the even the entire globe (Løvholt et al., 2014). At the same time, hundreds to thousands of tsunami propagation models have been created to inform real-time tsunami forecasting and alerts (e.g., Greenslade et al., 2007, 2013).
Various physical parameters influence the tsunami wave field, such as tides (Weisz and Winter, 2005), dispersion of wave propagation (Glimsdal et al., 2014), Coriolis Force (Shuto, 1991), effects of friction (Myers and Baptista, 2001) and source effects. The accuracy of tsunami simulation not only depends on the consideration of these factors in the numerical implementation, but also on the variability and uncertainties associated with them. Dao and Tkalich (2007) reviewed the numerical effects of dispersion, Coriolis Force, coordinate systems (Cartesian or spherical), bottom friction, tides and wave equation used (Boussinesq-type vs. Non-linear Shallow Water Equations). They found that astronomic tides and bottom friction have large impacts in shallow water, whereas dispersion only has a considerable effect on waves travelling over long distances. The particular type of numerical implementation, e.g. the choice of wave equation (linear and non-linear shallow water wave equations, Boussinesq-type or full Navier-Stokes equations) and their corresponding capacity to incorporate the factors mentioned above also has an influence on the accuracy of the simulation. Other studies, such as Løvholt et al. (2012) or Davies et al. (2015), have investigated the effect of non-uniform slip on the near shore maximum tsunami height. However, these studies have generally focused on one particular location and thus on a limited range of distances and azimuths. In Geist (2002), the effect of non-uniform slip on the far field was stated to be “less than 10 %” but the exact azimuth and distance at that point was not discussed.

One aspect that has not been studied in great detail is the sensitivity, or uncertainty, in the maximum tsunami wave height due to uncertainty in the earthquake’s geometrical source parameters such as strike, dip and rake. Here we present a systematic study of this issue, starting with simple source models in a flat ocean and then moving on to three examples which use a more realistic bathymetry.

Having a better understanding of this problem has the potential to be important for not only for future tsunami hazard assessments but also tsunami forecasting and source inversions. In PTHAs it might be possible to treat the uncertainty in source parameters as an aleatory, rather than epistemic, uncertainty and include it in a prob-
Probabilistic assessment as discussed by Geist and Parsons (2006). This could reduce the number of propagation models required as in Thio et al. (2010). Outside of hazard assessment, another potential use would be to know how close to a warning threshold a modelled tsunami wave height from an event must be in order for the difference to be “insignificant” given the current uncertainty in the source’s rupture parameters. This could then have been used to inform the resulting warning given to the public. The uncertainty in source parameters could affect the reliability of assumptions made inverting for the source using tsunami mareogram data.

For the purposes of this paper, we have focused on trying to characterise the uncertainty in the maximum offshore tsunami wave height at a particular point ($h_{\text{max}}$). The reason we have selected this particular model output is that this is the one most commonly used for both PTHAs and for tsunami alert threshold levels in warning systems.

The level of sensitivity in $h_{\text{max}}$ to variations, or uncertainties, in source properties can be measured in a variety of ways. Here we have quantitatively estimated this sensitivity by calculating the Coefficient of Variation (CoV) of $h_{\text{max}}$. CoV was defined here to be equal to $\sigma_{\text{max}}/\mu_{\text{max}}$ where $\sigma_{\text{max}}$ was the corrected sample SD and $\mu_{\text{max}}$ was the mean value of $h_{\text{max}}$ at a particular location. Other metrics, such as $\sigma$ itself, could be used but CoV has the advantage of being both a dimensionless and reasonably common metric for estimating the dispersion of a distribution. We have estimated the CoV by running $N$ tsunami propagation models each with a different value of a particular source parameter selected from a normal distribution with a SD centred at the parameter’s mean. $\sigma_{\text{max}}$ at a given point is then:

$$
\sigma_{\text{max}}(x,y) = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (h_{\text{max}}^{i}(x,y) - \mu_{\text{max}}(x,y))^2}
$$  

where $h_{\text{max}}^{i}(x,y)$ was the maximum tsunami wave height at a particular location for the $i$th model run.
To the authors’ knowledge, previous studies into tsunami sensitivity or uncertainty with respect to variations in the source parameters have typically been very location and event specific and/or have only considered a few parameters with a handful of models (i.e. $N$ in the equation above was small). For example, Titov et al. (1999) and Gica et al. (2007) studied the effect on offshore tsunami maximum wave heights near Hawaii from earthquakes from the Aleutian Islands, Chile and Japan for various combinations of dip, strike, rupture dimensions, hypocentre, slip displacement and rake angle. While this was a comprehensive list of parameters, the limited number of sources and target sites make it difficult to know to what degree the results of that study can be applied to other areas. For the location and events that they did study, they found that the tsunami wave height is mostly effected by changes in fault dimensions, strike angle, and slip displacement but not as much by rake, dip, epicenter location and focal depth. This sensitivity was not effected by distance, even in the far field, thus their conclusion was that the earthquake could not be treated as a point source.

Okal and Synolakis (2008) have also performed a few tests of the effect of shifting the epicentre and rake on maximum tsunami wave heights predicted from their numerical model. Again, this was only for a few examples and therefore cannot be used to calculate the CoV. However, they did conclude from their study that the far field pattern was robust to these variations. Okal and Synolakis (2004) have also examined the effect of varying a range of rupture parameters on the maximum runup from near shore events. However, they did not examine the CoV on the maximum runup nor the effect on the maximum offshore wave height from more distant events. In addition, Løvholt et al. (2012) have examined the effect of changing the dip and depth on the CoV from a set of heterogeneous slip events using a plane wave tsunami model with idealised bathymetry. They found that the CoV of the maximum runup from varying the slip decreased when the depth of the fault was increased but was unchanged when the dip was varied. However, they did not specifically look at the CoV from varying the bulk rupture parameters (e.g. strike, dip or depth) nor did they examine the effect of changing the distance to the rupture or the azimuth. Finally, Xing et al. (2014) examined the effect
of strike, rake, dip and magnitude on the maximum tsunami wave height at locations off the eastern coast of Australia for two tsunami sources, one on the New Hebrides trench and one on the Puysegur trench. They found that $h_{\text{max}}$ was changed when any of these parameters were varied. However, these authors only studied five cases per parameter and again the study was specific to a particular set of locations and sources.

Here we have looked at how the CoV changes for a range of azimuths and distances for a given uncertainty in a particular rupture parameter using one particular tsunami propagation model, EasyWave (Babeyko, 2012). The rupture parameters chosen for this study were: strike, dip, rake, magnitude and depth. The slip on all the models shown was uniform. We have done this in order to answer the following simple questions:

– Does the CoV vary with distance, azimuth or magnitude and how is it affected by bathymetry?

– Is $h_{\text{max}}$ normally distributed? If not, is the shape of the distribution also a function of distance, azimuth, magnitude or bathymetry?

2 Method

The method used here for assessing the uncertainty in $h_{\text{max}}$ was conceptually simple, if computationally intensive. It consisted of the following steps:

1. Choose a bathymetry.

2. Select a standard (or reference) set of rupture parameters.

3. Choose a random number from a normal distribution with a width given by the SD in the parameter to be studied ($\sigma_{\text{strike}}, \sigma_{\text{dip}}, \sigma_{\text{rake}}$ or $\sigma_{\text{depth}}$ depending on the parameter).
4. Run the tsunami propagation model with this parameter and then save the maximum wave height at all points in the model’s domain.

5. Repeat steps 3 and 4 for \( N \) iterations, each with a different, randomly generated, value of the parameter to be studied. The maximum wave heights at all points in the model was then saved for each iteration.

6. Use these models to calculate \( \mu_{\text{max}} \) across the model domain.

7. Use Eq. (1) was used to calculate \( \sigma_{\text{max}} \) for all points in the model domain.

8. Calculate the ratio of \( \sigma_{\text{max}} \) and \( \mu_{\text{max}} \) to calculate the CoV at every point in the domain.

9. Map the resulting CoV values.

In addition to calculating the CoV, we have also binned the \( N \) models in order to examine the shape of the resulting distribution.

Finally, we have also mapped the sample skewness in order to provide a more quantitative measure of the shape of the distribution of \( h_{\text{max}} \) across the model’s domain. The sample skewness, \( S \) was given by (Mantalos, 2010):

\[
S = \frac{1}{N} \sum_{i=1}^{N} \left( h_{\text{max}}^i - \mu_{\text{max}} \right)^3 \left( \frac{1}{N} \sum_{i=1}^{N} \left( h_{\text{max}}^i - \mu_{\text{max}} \right)^2 \right)^{3/2} \tag{2}
\]

When \( S > 0 \) the distribution is usually skewed to the right, i.e. it has a large (or heavy) tail above the mean. Log-normal distributions are an example of this type of distribution. When \( S < 0 \) the distribution is skewed to the left and the heavy tail is below the mean. If \( S = 0 \) then the distribution is evenly distributed around the mean (as it would be for the normal or uniform distribution for example). Maps of \( S \) allowed us to see whether the shape of distribution of \( h_{\text{max}} \) changed with azimuth or distance.
For the studies shown here, we have set \( N = 100 \). To test that this was adequate, we ran one set of runs with \( N = 50 \) and found that this changed the maximum CoV in a test model by 11%. However, when we ran one set of models with \( N = 200 \) and we found that this changed the maximum CoV observed in the model by less than 1%. Therefore \( N = 100 \) was chosen as a reasonable balance between accuracy in the maps and computational speed.

### 2.1 Bathymetry

For this study we used three increasingly complex bathymetry data sets. The first was a 80° × 80° bathymetry model with a constant depth of 3678 m (the average depth of the ocean Charette and Smith, 2010). The second was an 80° latitude by 42° longitude bathymetry model with a constant depth of 3678 m to the west of 181.25° longitude and a “step” up to a constant elevation of 100 m above mean sea level to the east of that point. Both models were calculated on a 4 arc minute grid. The first bathymetry can be viewed as a simplified version of the bathymetry near an oceanic subduction zone and the second for a (highly) simplified continental subduction zone. Having uniform bathymetry removes bathymetry variations from the problem and allows us to understand the patterns in CoV better. The stepped model could be made more like an actual continental margin by (for example) including a sloped ramp up to 100 m. However, the main aim of this bathymetry is just to demonstrate the effect of a basic process, in this case a simple reflection, rather than be a demonstration of the effect of a continental margin on CoV or skewness.

Some models were also run with both a 2 arc minute and a 8 arc minute grid and in both cases there was only a small change in the maximum CoV observed in the model (less than 5%). However, there were some minor changes in the pattern. Therefore some of the details in the maps shown later could be influenced by the numerical resolution of the grid. This could be due to the different levels of numerical dispersion in the models with different grid resolutions or because some of the details of the intial
deformation pattern and subsequent waves were missed for the coarser resolutions. However, the overall pattern appeared to be independent of the model resolution.

The final bathymetry model used were two subsections of ETOP02 global elevation model (NOAA, 2006). Both were calculated on a 2 arc minute grid. ETOP02 is one of the standard bathymetry models commonly used for tsunami propagation calculations. These models illustrated the effect of realistic bathymetry on the CoV and S maps.

2.2 Reference fault parameters

The uniform and stepped bathymetry tsunami runs used a uniform slip model with a set of “standard” values. The standard values were:

- Dip = 20°
- Strike = 0°
- Rake = 90°
- Depth to the top edge of the rupture = 10 km

The dip was chosen to be 20° as this is a typical value for the average dip of the seismic part of a subduction zone (e.g. the average dip of the interface in Slab 1.0 of Hayes et al., 2012 varies from 8 to 30° depending on the zone). Since we are mainly interested in tsunamigenic earthquakes the rake was set to be pure thrust and the depth of the top edge was kept shallow. The strike was chosen to propagate the tsunami along the equator and minimize any distortions simply due to the map projection. It also made the north and south parts of the domain symmetric which is a check that the sampling is adequate. When a parameter was varied randomly the standard value listed above was the random distribution’s mean.

The sea floor deformation created by these sources was calculated on 10 km × 10 km patches using the Okada equations as implemented in EasyWave and then summed. The fault’s centroid was positioned at 180° W, 0° N and dipped to the east. The scaling
relationship used to estimate the rupture dimensions and slip for a particular magnitude were based on the scaling relations of Abe (1975). These relations hold for magnitudes up to that at which the rupture width extends for the full width of the fault plane (to the brittle ductile boundary where fault-locking no longer occurs). For the purposes of this study, the maximum width was assumed to be 150 km. Above this magnitude it was assumed that the displacement continues to scale proportionately with the fault length; this is known as the L-model (Scholz, 1982; Hanks and Bakun, 2002). Although other, more recent, scaling relations could have been used, this particular set of scaling relations had the advantage of already being implemented in the computational framework (see Sect. 2.3) and the effect of different scaling relations on $h_{\text{max}}$ and CoV was not intended to be the main focus of this study.

For the models which were run using the ETOPO2 bathymetry, the epicentre and mean strike used were those of the 2011 Kermadec earthquake, 2007 Solomon Islands earthquake or the 2006 Java earthquake. The strike of these models was set to be equal to that for one of the nodal planes for these events calculated by the USGS. Note that only the epicentre and strike were based on these events, the dip and rake were kept the same as the flat and stepped bathymetry examples in order to aid comparison. These three locations were chosen as they are good examples of subduction zones from different tectonic environments and thus bathymetries. The Kermadec zone is on a fairly typical oceanic subduction zone, the Solomon Islands zone is an example of a zone with complex bathymetry in the source region (i.e. multiple small islands) and Java is a typical example of a continental subduction zone.

2.3 Numerical models and scripts

The tsunami propagation model used here was EasyWave (Babeyko, 2012). The main reason this particular model was chosen was that it has been optimised for computational speed and is open source (http://trac.gfz-potsdam.de/easywave). The time step for all the models shown here was the one calculated automatically by EasyWave. For the purpose of the uncertainty calculations the Coriolis effect was not included.
A robust and efficient framework was required to manage the large amount of data and the simulated scenarios this method produced. A Python-based object oriented Application Programming Interface (API) was developed that augments and drives the EasyWave tsunami simulation program. The API allowed us to automate this parameter study using EasyWave as the tsunami simulation kernel. All the source models investigated in this study were created and managed with this API. The API also managed simulation on GNS Science’s cluster used for the computations, i.e. it farmed simulation scenarios out to cluster and collected data after simulation completion.

All the subsequent post-processing, including the map generation, were calculated by a set of post-processing GMT scripts (Wessel et al., 2013). The output grids from EasyWave and for the statistics were a mix of Golden Software format NetCDF files produced by EasyWave and NetCDF format files created by GMT.

3 Results

3.1 EasyWave and COMCOT comparison and validation against Chile 2010

As there has not been a great deal of published validation of EasyWave other than Greenslade et al. (2013), the results from an EasyWave tsunami propagation calculation were compared to those from COMCOT (Liu et al., 1995; Wang and Liu, 2007; Wang et al., 2008; Wang and Power, 2011) for a historical event. Figure 1 shows the maximum tsunami wave height generated by the Chile 2010 earthquake calculated with COMCOT. The parameters used to model this event were from Power (2013). Also shown in the figure are the location of the DART gauges at the time of the event. Figure 2 compares the de-tided mareograms at the DART gauges for this event (black curves) and the results from a COMCOT model (red curves) and an EasyWave model (blue curves) of the tsunami produced by this earthquake. For the example shown in the figure, the optional Coriolis term was not included in the EasyWave calculation. Apart from that difference, the two numerical models were solving the same set of lin-
ear shallow water wave equations. Both of the numerical models used identical source parameters and bathymetry grids. Generally speaking the maximum wave height at the gauges was similar for all three curves. The average difference between the maximum wave height at the DART locations calculated with COMCOT to that calculated with EasyWave was 10\%.

When the Coriolis term was included in EasyWave this was reduced to 4\% (figure not shown). This reduction in misfit was mostly achieved by an improved fit between the maximum wave height calculated with EasyWave and that with COMCOT at the more distant DART locations (e.g. at DARTs 52403 or 21413).

As can be seen from Fig. 2 they were slightly too high when the Coriolis term is neglected. Other than these minor differences, the agreement between the two models and with observational data was excellent, particularly for DARTs close to the source (e.g. DART 32412). This suggests that EasyWave produces comparable results to COMCOT, a long established and widely used tsunami model, and that it produces results consistent with observations given the right source and bathymetric model. It also suggested that the Coriolis effect can be safely neglected for distances up to 40° in this particular case. For other cases, for example for earthquakes located at a different latitude or tsunamis propagating in a different direction, the effect could be more important.

3.2 Uniform, flat bathymetry

3.2.1 Strike

Figure 3 shows the mean maximum wave height ($h_{\text{max}}$) over 100 uniform slip models with $\sigma_{\text{strike}}$ of 10° and a magnitude of $M_W$ 9.5. This level of uncertainty would be typical for a well constrained earthquake focal mechanism. The total rotational uncertainty (strike and dip) for a focal mechanism is typically between 5 and 20° (Kagan, 2003). The other parameters are at their standard values (see Sect. 2.2). The bathymetry was flat with a uniform depth. As one might expect, the tsunami propagated as two “beams”; one going to the east and and one to the west of the earthquake rupture's
initial location. The effect of averaging over 100 models with varying strike was that these beams become more "smeared" at their edges than they would be if only one model was simulated.

Figure 4a–c shows maps of the CoV of $h_{\text{max}}$ from three sets of earthquakes with magnitudes $M_W$ 7.5, 8.5 and 9.5 respectively. The sets of earthquakes varied in the strike by 10° but were otherwise the same. As can be seen, the largest values of CoV were on each side of the two tsunami beams (Fig. 3). The CoV was higher and more focused on the footwall (left) beam than for the hanging wall (right) beam for $M_W$ 7.5 (Fig. 4a) but not for the higher magnitudes. The variance was always at a minimum along strike (i.e. due north and south along the 180°W line of longitude). The range of CoV values went from 0.7 to 0.001 for $M_W$ 9.5. For $M_W$ 7.5 it ranged from 0.3 to 0.001 but the bulk of the region was well below 0.2. For $M_W$ 9.5 the most common CoVs across the model domain were between 0.05 and 0.1 but many of the points were between 0.5 and 0.7 (histogram not shown, but this can be seen from the range of colour values in Fig. 4c). Overall, the pattern was symmetric between the northern and southern halves of the model domain. However, on close inspection some minor asymmetries were seen between the northern half of the model and the southern half, particularly due north and south of the earthquake’s epicentre.

Figure 4d shows the CoV map when $\sigma_{\text{strike}}$ is reduced to 5° for a magnitude of $M_W$ 9.5. In this case, reducing $\sigma_{\text{strike}}$ did not change the pattern very much but reduced the amplitude and concentrated the larger CoV values into a smaller area. The maximum value of CoV reduced to 0.5 (down from 0.7). Also, the maximum effects of strike uncertainty occurred further from the source.

Figure 5a–c shows the histogram of the maximum wave heights at locations A–C on Fig. 5d for a $M_W$ 9.5 event. As can be seen, the distribution of maximum wave heights was far from being normally distributed at the locations shown. For location A the distribution had a strong negative skew, while for C it had a strong positive skew. At B the distribution was not significantly skewed, but was close to uniform. This was despite $\sigma_{\text{strike}}$ having a Gaussian distribution. Figure 5d also shows the skewness values across the
whole region more generally. The $h_{\text{max}}$ distributions were generally positively skewed, except for points in the beam close to the source where they were negatively skewed.

### 3.2.2 Dip

Figure 6a–c shows the CoV maps when $\sigma_{\text{dip}}$ is $5^\circ$ for a $M_W$ 7.5, 8.5 and 9.5 event respectively. Different plate models typically differ by between 1 to 10$^\circ$ in their estimates of the average dip of an interface (e.g. see Table 1 in Hayes et al., 2012), so 5$^\circ$ was chosen as a reasonable estimate of the typical uncertainty in dip. The other parameters were held at their reference values.

Unlike the previous example, the maximum CoV was along strike (i.e. to the north and south of the epicentre) rather than to each side of the tsunami beam. For the $M_W$ 7.5 example it was also higher in the hanging wall direction (to the right of the figure) rather than the footwall direction. This difference became less strong (more concentrated into a smaller region) as the magnitude increased. The CoV ranged from 0.4 to 0.01 for $M_W$ 9.5 and from 0.2 to 0.002 for $M_W$ 7.5.

Figure 6d shows the skew pattern for the $M_W$ 9.5 case. In this case the distributions were not as skewed as they were in the previous example. The skew was generally small and negative except immediately above the rupture where it was either strongly positive or negative.

### 3.2.3 Rake

Figure 7a–c shows the CoV maps when $\sigma_{\text{rake}}$ is set equal to 20$^\circ$ for a $M_W$ 7.5, 8.5 and 9.5 event respectively. Again, this is typical uncertainty in a well constrained focal mechanism (e.g. see Shaw and Jackson, 2010). For the $M_W$ 7.5 case the regions of maximum CoV were on the hanging wall of the fault and on the footwall side for $M_W$ 8.5 and 9.5. Unlike the previous two examples, the range of the CoV for the $M_W$ 9.5 and $M_W$ 7.5 events was essentially identical. It ranged from 0.3 to 0.005 for $M_W$ 9.5 and from 0.3 to 0.04 for $M_W$ 7.5.
The skewness for the $M_W$ 9.5 set of events is shown in Fig. 7d. To the north and south of the rupture the distributions were had a positive or had zero skew. Everywhere else the distributions were negatively skewed.

### 3.3 Depth

Figure 8a–c shows the CoV maps when the depth to the top edge of the fault was varied by a $\sigma_{\text{depth}}$ of 2.5 km for sets of $M_W$ 7.5, 8.5 and 9.5 events. Depths above 0 km were rejected to prevent “air quakes”. This uncertainty is fairly low for a typical earthquake, but was chosen to ensure that the distribution in depth is still approximately Gaussian after removing the “air quakes”.

These CoV maps had higher values on the hanging wall side than the footwall side for the $M_W$ 7.5 case (Fig. 8a). The CoV was also generally higher for $M_W$ 7.5 than for the other magnitudes. However, overall the CoV was much lower than for strike variations. The range of CoV went from 0.2 to 0.002 for the $M_W$ 9.5 set of events and from 0.2 to 0.03 for $M_W$ 7.5 set. However, the area covered with a high CoV was much larger for the $M_W$ 7.5 case than is was for the $M_W$ 9.5 case.

The skewness was negative on the footwall (left) side, postive to the north and south of the rupture and mostly near zero on the hanging wall side for the $M_W$ 9.5 case (Fig. 8d).

#### 3.3.1 Multiple parameters

The final example for the uniform bathymetry case we show here is one where all the parameters are allowed to vary around their reference values. In the example shown in Fig. 9, $\sigma_{\text{strike}} = 10^\circ$, $\sigma_{\text{dip}} = 5^\circ$, $\sigma_{\text{rake}} = 20^\circ$ and $\sigma_{\text{depth}} = 2.5$ km. The magnitudes varied from $M_W$ 7.5, 8.5 and 9.5. As one might expect this pattern was a combination of all the previous patterns, in this case dominated by the strike pattern, particularly for the higher magnitudes. The CoV ranged for $M_W$ 9.5 is 0.6 to 0.1 and for $M_W$ 7.5 is 0.4 to
0.1. For $M_W$ 9.5, skewness varied from strongly positive to weakly negative depending on the azimuth (Fig. 9d).

### 3.4 Stepped bathymetry

Figure 10a shows what happens to a CoV map if a reflecting barrier is placed due east of the fault. In this particular example, we show the effect on the CoV for a $M_W$ 9.5 uniform slip rupture with $\sigma_{\text{strike}} = 10^\circ$. As described in Sect. 2.1 the bathymetry increased east of 182.5° to 100 m a.s.l. Everywhere else it was the same as in the previous examples. As can be seen, in this example the beam and the regions just to each side of the beam had the highest CoV. The CoV also increased as one moved further away from the source. The range of the CoV for this example was from 0.5 to 0.001 just to the west of the epicentre.

Figure 10b shows the effect of this on the skewness field (cf Fig. 5b). The areas to both sides of the beam were strongly positively skewed, those elsewhere were negatively skewed. The pattern also changed immediately above and to the west of the fault, where it was mostly positively skewed but with some areas which had a strong negative skewness.

### 3.5 Realistic bathymetry

The previous examples all used flat or stepped bathymetric models. While this is very useful for determining basic patterns, in real cases the bathymetry is highly non-uniform. Here we repeat some of the above experiments for hypothetical earthquakes on the Kermadec, Solomon Islands and Java subduction zones. The bathymetry used for these three examples is shown in Fig. 11.

#### 3.5.1 Kermadec event

Figure 12a–c shows the CoV maps for (a) $\sigma_{\text{strike}}$ of 10°, (b) $\sigma_{\text{dip}}$ of 5° and (c) all the rupture parameters being allowed to vary. The epicentre of the event was at 183.762° E, 3384
The mean strike was 205°. The other parameters were the same as those list in Sect. 2.2. The magnitude is $M_W 8.5$ for all cases shown. For the last case, where multiple parameters were being varied, the $\sigma$ values were the same as those used in the example discussed in Sect. 3.3.1. As can be seen, the patterns in this case were broadly similar to the uniform bathymetry case (e.g. Fig. 4b) to the southeast of the epicentre but the non-uniform bathymetry to the northwest of the epicentre greatly increased the complexity of the pattern. Further away from the epicentre, there were patches with a particularly high CoV, such as those to the southeast of New Zealand. These appeared to be in areas of shallow bathymetry or in areas with some shallow bathymetry between them and the source. The CoV in Fig. 12 ranges from 0.5 to 0.04.

Figure 12d shows the skewness values for the last case (where all parameters allowed to vary). Skewness was generally non-zero and was neither consistently positive nor negative but rather varied across the region.

### 3.5.2 Java event

Figure 13a–c shows the CoV maps for an event off Java with the same set of $\sigma$ values used in the previous example. The epicentre of the event was at 107.33° E, 9.32° S. The mean strike was 295°. Again the broad pattern was similar to that found from the uniform or stepped bathymetry models, but this ceased as soon as any complex bathymetry was reached. For example, the lines of maximum CoV split as the tsunami went around the southwest corner of Western Australia. The CoV range for Fig. 13c was 0.5 to 0.05.

Figure 13d shows the skewness pattern for this case. Again the pattern was only similar to the one found with flat bathymetry until the wave reached complex bathymetry.

### 3.5.3 Solomon Islands event

Finally we show an example of the CoV where the bathymetry is complex in the source region, in this case the Solomon Islands subduction zone (Fig. 14). The epicentre of
the event was at 157.06° E, 8.43° S. The mean strike was 333°. The basic patterns of flat bathymetry examples can now barely be seen, if at all. The highly complex bathymetry in the source made predicting the CoV pattern at a given location difficult, if not impossible. The skewness map (Fig. 14d) is similarly complex. The highest CoV and skewness value were, in this case, due north of the earthquake’s epicentre. The CoV ranged from 0.8 to 0.04 (Fig. 14c). The maximum CoV was significantly higher in this case than for the other two examples, even though the σ values for the various parameters were the same. However, the area of extremely high CoV values were very small.

4 Discussion

A few general conclusions can be drawn from these results. Firstly, the exact value of CoV and the skewness of its distribution was a very strong function of the choice rupture parameter, azimuth, distance and bathymetry. In some cases the reasons for this can be fairly easily understood. For example, consider a set of MW 9.5 earthquakes with varying strike. In the direction of the beam any change in strike, positive or negative, will always act to reduce the $h_{\text{max}}$ for points in the beam and increase it for points just outside the beam. Thus the distribution was negatively skewed in the beam and positively skewed outside of it (Fig. 5). The magnitude of this effect will be at its greatest for points just outside the beam since they can go from being entirely inside the beam to entirely outside of it with just a small change in strike. Thus the CoV was at a maximum there (Fig. 4).

However, the patterns in the other cases are not as intuitive. Having the maximum CoV along strike when the dip was varied is probably due to the way the dip changes the initial crustal deformation pattern by bring the line of displacement closer to the trench. In a similar way, the other changes in CoV are probably due to the way changes in other parameters affect the initial deformation pattern.
The fact that the patterns and values change with magnitude suggests strongly that these effects are also linked to the changing dimensions and aspect ratio of the source region. At lower magnitudes ($M_W$ 7.5 in our examples) the source appears to be “point-like” except in the near-field. At intermediate magnitudes ($M_W$ 8.5 in our examples) the source dimensions mean that the source is more “area-like”. At large magnitudes ($M_W$ 9.5 in our example) the aspect ratio changes such that the length becomes much greater than the width, and the source becomes “line-like” in the far-field.

The reduced sensitivity to uncertainty in depth between $M_W$ 7.5 and $M_W$ 8.5 can be understood in this context. At $M_W$ 7.5 the rupture width was small and therefore occupies a small range of depths, so uncertainty in the depth of the top edge made a significant change to the overall deformation pattern and subsequent tsunami. However at $M_W$ 8.5 the larger rupture surface already occupies a wide range of depths, so uncertainty in the depth of the top edge made proportionately less difference overall.

A result of this complexity is that it is very difficult to make general statements about the level of uncertainty in $h_{\text{max}}$ given an uncertainty in any of the source parameters. For some particular locations or azimuths a small uncertainty in strike made very little difference to the result (i.e. less than 10 %), in other locations it changed $h_{\text{max}}$ by 20 % or even by more than 50 % (Fig. 4). It all depended on the azimuth, and for the latter examples, the bathymetry between the source and the location. This is broadly consistent with Gica et al. (2007) where the same $10^\circ$ change in strike could change the wave height measured at Hawaii by between 12 and 84 % depending on the location of the earthquake relative to the island. When the bathymetry in the source location was complex, such as in the Solomon Islands case (Fig. 14), the CoV and skewness maps became impossible to distinguish from noise and only general statements about the maximum upper bound on the CoV or $S$ can really be made.

Initially, the authors assumed that it might be possible to treat the uncertainty in the source parameters as an aleatory, rather than epistemic, uncertainty in PTHAs as discussed in the Sect. 1. However, our study shows that including $\sigma$ uncertainties in PTHAs as aleatory uncertainty can only ever be very approximate. It will always be
difficult to be sure whether the values used aren’t over- or under-estimating the hazard at a particular location given the highly, and inconsistently, skewed distributions of $h_{\text{max}}$. Since the skew can change from positive to negative over very short length scales these issues cannot be simply solved by using a different type of $\sigma$ (e.g. a log normal $\sigma$). The ideal solution clearly has to be to run a large number of models to try to ensure that the hazard from the events in any tails of these skewed distributions are included in the assessment. However, this can become very computationally challenging for larger assessments.

Similarly, this also means that tsunami databases for tsunami forecasting and warning systems need to be very large. Current warning and forecasting systems still rely on the use of a limited set of precalculated scenarios and do not currently include any assessment of the spatial distribution of the CoV for each scenario, although the CoV between different scenarios has been calculated (Greenslade et al., 2013). Given the large number of possible events, calculating the CoV for each event would be an extremely large computational task. It was also noticeable that the CoV distribution for events involving realistic bathymetry tends to become very scattered in shallow coastal areas outside of the near field, possibly the result of complex interference patterns involving multiple waves. Another explanation might be that the resolution of the grids used in our study was not high enough in the coastal regime to properly assess the CoV. In either case, many forecast methodologies rely on warning zones, which are sections of coast in which a warning threshold is crossed once a particular proportion of maximum wave heights (e.g. the 95th percentile) exceeds a specified level (e.g., Uslu and Greenslade, 2013). An area for further study is to see to what extent thresholds defined in these aggregate terms are sensitive to uncertainties in source parameters. If this does effect the reliability of thresholds, it therefore seems advisable to move away from precalculated tsunami databases and use fast tsunami simulation programs instead that allow for the calculation of both the CoV and $\mu_{\text{max}}$ as the tsunami event unfolds (i.e. ensemble forecasting). The CoV and the shape of the distribution of $h_{\text{max}}$ from this ensemble of models can then inform about the reliability of the tsunami...
forecast (and thus the warning) for any given point of interest. It also suggests that warnings should try to move towards taking uncertainties in the source into account more directly in the warning, for example using Bayesian Networks (Blaser et al., 2011, 2012).

In addition to the potential issues just discussed, our results also indicate that the inversion of the tsunami source based on DART buoy information will be affected by the relative positions of the source and the DART buoys. If a DART buoy happens to be located in an area that has a low CoV for a particular fault parameter, we would expect the resulting inversion for that parameter to be poorly constrained. In other words, the inverted source has the potential to be non-unique. For example, if the inversion only has DARTs close to a $M_W$ 9.5 earthquake, the maximum wave height at those DARTs will not be significantly effected by an error in strike unless one of them happens to be close to the edge of the rupture (see Fig. 4c and d). Thus the strike may not be well constrained. However, the same error in strike could make a large difference in the observed maximum wave height further from the source (i.e. in the red areas in Fig. 4c and d). This is consistent with the observation of Wang (2008) that gauges off the centre line of the tsunami propagation are more useful for constraining the source than those in tsunami beam itself. Ultimately, this sort of effect will create more uncertainty in the predicted wave fields. Reducing the uncertainty ideally requires techniques which can measure the tsunami wave height over broad areas (e.g. using remote sensing) or include additional types of data (e.g. seismic or geodetic). Also many inversion algorithms assume a Gaussian or log normal distribution of misfits, as can be seen from the maps of skewness this is not always the case. The effect of this on tsunami inversion assumptions is also be another potential area of future research.

5 Conclusions

The main conclusion of this study is that “the uncertainty in the maximum wave height of a tsunami is a complex function of our uncertainty in the source parameters and
bathymetry”. Even for the case of a completely flat bathymetry, complex patterns of CoV and skewness were seen. These patterns became even more complex when realistic bathymetries were used. While the specifics of these CoV maps may be influenced by the particular choice of numerical and bathymetry models used here, the overall patterns are probably not. For example, the high CoV lobes to each side of the beam when the strike was varied appear to be a function of the beam-like nature of tsunami propagation. Thus any model or bathymetry is likely to have a broadly similar CoV map even if the details may be different.

Given the complexity of CoV (and thus $\sigma$) simplified methods of taking earthquake uncertainty into account in PTHAs have the potential to be quite inaccurate. Depending on the way $\sigma$ or CoV is chosen, they will over-estimate the hazard in some locations and under-estimate it in others. Also, $\sigma$ does not follow a simple normal or log normal distribution as shown by the fact that the skewness also changes with distance and azimuth. This suggest that the best way to incorporate uncertainty in earthquake parameters in future PTHAs is still to model all reasonably possible earthquake ruptures. Similarly, these results give further impetus towards using real-time ensemble tsunami propagation models for warnings, rather than relying on limited catalogues of possible future tsunamis. The already substantial computational task of both activities will thus likely need to grow even further in future in order to take uncertainties such as these into account.

**Author contributions.** DB ran the models, wrote the scripts for analysing the results, prepared the figures and wrote the bulk of the text. CM wrote the Python API which managed the simulations. Both CM and WP contributed text to the final version of the manuscript, mostly to the Introduction, Method and Discussion sections.

**Acknowledgements.** The authors would like to thank A. Babeyko for assisting the authors with the EasyWave tsunami propagation model. The authors would also like to thank Xiaoming Wang for providing the Chile 2010 tsunami source model, the de-tided DART gauge data, the COMCOT input file and the maximum wave height figure (Fig. 1). All the other figures in this paper were created using GMT (Wessel et al., 2013). Finally, the authors would also like to
thank Xiaoming Wang, Nick Horspool and Jane Sexton for reviewing this document. The paper is published with the permission of the CEO of Geoscience Australia.

References


NOAA: 2-minute Gridded Global Relief Data (ETOPO2v2), US Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, available at: http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html (last access: 8 October 2009), 2006. 3377


Figure 1. The maximum wave height from the Chile 2010 tsunami calculated with the COMCOT model. Also shown are the locations of the DART gauges at the time of the event. From Power (2013).
Figure 2. The de-tided mareograms at selected DART stations from the 2010 Chile tsunami (black). Also shown are the tsunami waveforms calculated using the COMCOT (red) and the EasyWave models (blue) from a simplified source model for this event. The y axis shows the wave height relative to mean sea level in m, while the x axis shows the number of hours after the earthquake occurred.
Figure 3. The mean of the maximum tsunami wave height \( \mu_{\text{max}} \) in m for a \( M_w \) 9.5 event with a \( \sigma_{\text{strike}} = 10^\circ \). The bathymetry is uniform and completely flat. The earthquake’s rupture has a 20° dip to the right, it’s mean strike is north-south (i.e. 0°), the depth to the top edge is 10 km and it is centred at 180°W, 0°S. The slip along the rupture plane is uniform and has a 90° rake (i.e. it is a pure thrust earthquake). The black box on the figure shows the surface projection of the mean rupture plane. The solid line on the left is the top edge of the plane and the dashed line on the right is bottom edge.
Figure 4. The effect of $\sigma_{\text{strike}}$ on the co-efficient of variation (CoV) of $h_{\text{max}}$. (a) shows the CoV map from a set of 100 tsunami generated by $M_W$ 7.5 earthquakes which vary only in strike. (b, c) shows the effect on CoV when the magnitude of these events are increased to $M_W$ 8.5 and $M_W$ 9.5. For (a–c) $\sigma_{\text{strike}} = 10^\circ$. The bathymetry and other parameters are the same as in Fig. 3. Note that values above 0.4 are all shaded the same colour. 0.4 was chosen to be the maximum for these figures in order to be consistent with later figures. (d) shows the CoV map from a set of $M_W$ 9.5 events with a $\sigma_{\text{strike}} = 5^\circ$. 

3398
Figure 5. Example histograms for three locations surrounding a $M_W 9.5$ event with a $\sigma_{\text{strike}} = 10^\circ$. (a) is at 160° E, 0° and is directly in the mean path of the beam. This is an example of a location where the distribution has a strong negative skew. (b) is an example at the edge of the beam (160° E, 5° S). It is not skewed, but nor is it normally or log normally distributed. (c) is at 160° E, 10° S and is thus just off the beam. This is an example of a location where the distribution of maximum wave heights has a strong positive skew. (d) shows the location of the histograms in (a–c) relative to the rupture. The background image shows the skewness (see Eq. 2) of the maximum tsunami wave height distribution across the domain.
Figure 6. The effect of $\sigma_{\text{dip}}$ on the CoV and skewness of $h_{\text{max}}$. The magnitude of the earthquakes for each set were (a) $M_W$ 7.5, (b) $M_W$ 8.5 or (c) $M_W$ 9.5. $\sigma_{\text{dip}} = 5^\circ$ for all the sets shown here. The mean dip is 20$^\circ$. The strike is fixed to be due north-south. The bathymetry and other parameters are otherwise the same as in Fig. 4 and are held constant for all 100 iterations. (d) The skewness of the distribution of the maximum tsunami wave heights for a $M_W$ 9.5 and $\sigma_{\text{dip}} = 5^\circ$. 
Figure 7. The effect of $\sigma_{rake}$ on the CoV and skewness of $h_{max}$. The magnitude of the earthquakes for each set were (a) $M_W 7.5$, (b) $M_W 8.5$ or (c) $M_W 9.5$. $\sigma_{rake} = 20^\circ$ for all the sets shown here. The mean rake is $90^\circ$. The bathymetry and other parameters are otherwise the same as in the previous examples. (d) The skewness of the distribution of the maximum tsunami wave heights for a $M_W 9.5$ and $\sigma_{rake} = 20^\circ$. 

3401
Figure 8. The effect of $\sigma_{\text{depth}}$ (depth to the top edge of the rupture) on the CoV and skewness of $h_{\text{max}}$. The magnitude of the earthquakes for each set were (a) $M_W$ 7.5, (b) $M_W$ 8.5 or (c) $M_W$ 9.5. $\sigma_{\text{depth}} = 2.5$ km for all the figures shown here. The mean depth to the top edge was 10 km. The bathymetry and other parameters are otherwise the same as in the previous examples. (d) The skewness of the distribution of the maximum tsunami wave heights for a $M_W$ 9.5 and a $\sigma_{\text{depth}} = 2.5$ km.
Figure 9. The effect of varying multiple fault parameters on the CoV and skewness. The magnitude of the earthquakes for each set were again (a) $M_w 7.5$, (b) $M_w 8.5$ or (c) $M_w 9.5$. $\sigma_{\text{strike}} = 10^\circ$, $\sigma_{\text{dip}} = 5^\circ$, $\sigma_{\text{rake}} = 20^\circ$ and $\sigma_{\text{depth}} = 2.5$ km. The bathymetry and other parameters are otherwise the same as in the previous examples. (d) The skewness of the distribution of the maximum tsunami wave heights for a $M_w 9.5$ and these $\sigma$ values.
Figure 10. The effect of having a step in the bathymetry on the CoV and skewness of $h_{\text{max}}$. (a) shows the CoV from a set of $M_w$ 9.5 events with a $\sigma_{\text{strike}} = 10^\circ$. The elevation has a step at 182.5° where it suddenly increases to 100 m above mean seal level. The other parameters are otherwise the same as in Fig. 4b. (b) The effect of the step on the skewness of the distribution of the maximum tsunami wave heights (cf. Fig. 5b).
Figure 11. The bathymetry models used for the (a) Kermadec and Solomon Islands scenarios and (b) the Java scenario. The red symbols show the location of the epicentres for each scenario; a circle for the Kermadec scenario, a square for the Solomon Islands scenario and a triangle for the Java scenario.
**Figure 12.** The CoV and skewness maps of $h_{\text{max}}$ from a set of earthquakes on the Kermadec trench. Each set consisted of $M_w 8.5$ earthquakes where either one parameter was allowed to vary (cases a or b) or all five were allowed to vary (case c). For (a) $\sigma_{\text{strike}} = 10^\circ$, for (b) $\sigma_{\text{dip}} = 5.0^\circ$ and for (c) $\sigma_{\text{strike}} = 10^\circ$, $\sigma_{\text{dip}} = 5^\circ$, $\sigma_{\text{rake}} = 20^\circ$ and $\sigma_{\text{depth}} = 2.5$ km. (d) The skewness map for case (c).
Figure 13. The CoV and skewness maps of $h_{\text{max}}$ from a set of earthquakes on the Java subduction zone. Each set consisted of $M_w$ 8.5 earthquakes where either one parameter was allowed to vary (cases a or b) or all five were allowed to vary (case c). For (a) $\sigma_{\text{strike}} = 10^\circ$, for (b) $\sigma_{\text{dip}} = 5.0^\circ$ and for (c) $\sigma_{\text{strike}} = 10^\circ$, $\sigma_{\text{dip}} = 5^\circ$, $\sigma_{\text{rake}} = 20^\circ$ and $\sigma_{\text{depth}} = 2.5$ km. (d) The skewness map for case (c).
Figure 14. The CoV and skewness maps of $h_{\text{max}}$ from a set of earthquakes on the Solomon Islands subduction zone. Each set consisted of $M_W$ 8.5 earthquakes where either one parameter was allowed to vary (cases a or b) or all five were allowed to vary (case c). For (a) $\sigma_{\text{strike}} = 10^\circ$, for (b) $\sigma_{\text{dip}} = 5.0^\circ$ and for (c) $\sigma_{\text{strike}} = 10^\circ$, $\sigma_{\text{dip}} = 5^\circ$, $\sigma_{\text{rake}} = 20^\circ$ and $\sigma_{\text{depth}} = 2.5$ km. (d) The skewness map for case (c).