Natural Hazards manuscript No.
(will be inserted by the editor)

The Pattern-Method for Incorporating Tidal Uncertainty Into Probabilistic Tsunami Hazard Assessment (PTHA)

Loyce M. Adams · Randall J. LeVeque · Frank I. González

Received

Abstract In this paper we describe two new methods for incorporating tidal uncertainty into the probabalistic analysis of inundation caused by tsunamis. These two methods are referred to as the dt-Method and the Pattern-Method. We compare these methods to the method in [Mofjeld, et al(2007)] that was used for the 2009 Seaside, Oregon probabilistic study [González, et al(2009)]. We show the Pattern-Method is superior to past approaches because it takes advantage of our ability to run the tsunami simulation at multiple tide stages and uses the time history of flow depth at strategic gauge locations to infer the temporal pattern of waves that is unique to each tsunami source. Some sources give only one large wave, others give a sequence of equally dangerous waves spread over several hours. Combining these patterns with knowledge of the tide cycle at a particular location improves the ability to estimate the probability that a wave will arrive at a time when the tidal stage is sufficiently large that inundation above a level of interest occurs.

Keywords PTHA · hazard curves · 100-yr flood · GeoClaw · Poisson process · cumulative distribution

Mathematics Subject Classification (2000) 86-08

This work was done as a Pilot Study funded by BakerAECOM. Partial funding was also provided by NSF grant numbers DNS-0914942 and DNS-1216732 of the second author.

University of Washington
Dept. Applied Mathematics
E-mail: lma3@uw.edu

University of Washington
Dept. Applied Mathematics
E-mail: rjl@uw.edu

University of Washington
Dept. Earth & Space Sciences
E-mail: figonzal@uw.edu

1 Introduction

At Crescent City, the difference in tide level between mean lower low water (MLLW) and mean higher high water (MHHW) is about 2.1 meters. Coastal sites with such a significant tidal range experience tsunami/tide interactions that are an important factor in the degree of flooding. For example, [Kowalik and Proshutinsky(2010)] conducted a modeling study that focused on two sites, Anchorage and Anchor Point, in Cook Inlet, Alaska. They found tsunami/tide interactions to be very site-specific, with strong dependence on local bathymetry and coastal geometry, and concluded that the tide-induced change in water depth was the major factor in tsunami/tide interactions. Similarly, a study of the 1964 Prince William Sound tsunami [Zhang, et al(2011)] compared simulations conducted with and without tide/tsunami interactions. They also found large, site-specific differences and determined that tsunami/tide interactions can account for as much as 50% of the run-up and up to 100% of the inundation. Thus, probabilistic tsunami hazard assessment (PTHA) studies must account for the uncertainty in tidal stage during a tsunami event.

[Houston and Garcia(1978)] developed probabilistic tsunami inundation predictions that included tidal uncertainty for points along the US West Coast. The study was conducted for the Federal Insurance Agency, which needed such assessments to set federal flood insurance rates. They considered only far-field sources in the Alaska-Aleutian and Peru-Chile Subduction Zones, because local West Coast sources such as the CSZ (Cascadia Subduction Zone) and Southern California Bight landslides had not yet been discovered, and assigned probabilities to each source based on the work of [Soloviev(2011)]. Maximum runup estimates were made at 105 coastal sites rather than from actual inundation computations on land. The tidal uncertainty methodology began with a modeled 2-hour tsunami time series that was extended 24-hours by appending a sinusoidal wave with an amplitude that was 40% of the maximum modeled wave, to approximate the observed decay of West Coast tsunamis. This 24-hour tsunami time series was then added sequentially to 35,040 24-hour segments of a year-long record of the predicted tides, each segment being temporally displaced by 15 minutes. Determination of the maximum value in each 24-hour segment then yielded a year-long record of maximum combined tide and tsunami elevations, each associated with the probability assigned to the corresponding far-field source. Ordering the elevations and, starting with the largest elevations, summing elevations and probabilities to the desired levels of 0.01 and 0.002, produced the 100-year and 500-year elevations, respectively.

[Mofjeld, et al(2007)] developed a tidal uncertainty methodology that, unlike [Houston and Garcia(1978)], does not use modeled tsunami time series. Instead, a family of synthetic tsunami series are constructed, each with a period in the tsunami mid-range of 20 minutes and an initial amplitude ranging from 0.5 to 9.0 m that decreases exponentially with the decay time of 2.0 days, as estimated by [Van Dorn(1984)] for Pacific-wide tsunamis. As in [Houston and Garcia(1978)], linear superposition of tsunami and tide is assumed and the time series are added sequentially to a year-long record of

predicted tides at progressively later arrival times, in 15 minute increments. Direct computations are then made of the probability density function (PDF) of the maximum values of tsunami plus tide. The results are then approximated by a least squares fit Gaussian expression that is a function of known tidal constants for the area and the computed tsunami maximum. This expression provides a convenient means of estimating the tidal uncertainty, and was used by [González, et al(2009)] in the PTHA study of Seaside, OR.

In this paper we describe in detail two improved methods for incorporating tidal uncertainty into PTHA studies. The two new methods are referred to as the dt-Method and the Pattern-Method. They were developed as part of a PTHA study of Crescent City, CA in [González, et al(2012)], a pilot study funded by BakerAECOM to explore methods to improve products of the FEMA Risk Mapping, Assessment, and Planning (RiskMAP) Program.

Both the dt-Method and the Pattern-Method introduce major improvements to previous approaches. In both methods: (a) the assumption of linear superposition of the tide and tsunami waves is replaced by a methodology that utilizes multiple runs at different tidal stages; thereby introducing nonlinearities in the inundation process that are not accounted for in previous methods, and (b) synthetic time series are replaced by the actual time series computed by the inundation model. In addition, the Pattern-Method (c) takes account of temporal wave patterns that are unique to each tsunami source; for example, some sources produce only one large wave, others a sequence of equally dangerous waves that arrive over several hours. Combining these patterns with knowledge of the tide cycle at a particular location like Crescent City improves estimates of the probability that a wave will arrive at a time when the tidal stage is sufficiently large that inundation above a level of interest occurs. Finally, we compare these methods to that of [Mofjeld, et al(2007)] which we refer to as the Gaussian or G-Method.

1.1 Notation and terminology

- h refers to the water depth above topography or bathymetry. It is one of the primary variables of the shallow water equations that is output from a GeoClaw run at a static sealevel. The real water depth that includes the tsunami and the rising and falling of the tides is denoted d. B refers to the pre-earthquake topography or bathymetry as specified by the topography datasets, and is relative to Mean High Water (MHW) since that is the vertical datum of the fine scale Crescent City bathymetry. z will be used to denote the maximum observed GeoClaw value over the full time period of a tsunami of either h or $B + h$:

$$z = \begin{cases} h, & \text{the flow depth, where } B > 0 \text{ (onshore)}, \\ B + h, & \text{the sea surface elevation plus subsidence, where } B < 0. \end{cases}$$

- ζ will be used to denote the real value over the full time period of a tsunami of either d or $B + d$: GeoClaw inundation maps show z and hazard maps

that include tidal variation show ζ.

$$\zeta = \begin{cases} d, & \text{the flow depth, where } B > 0 \text{ (onshore)}, \\ B + d, & \text{the sea surface elevation plus subsidence, where } B < 0. \end{cases}$$

- ξ denotes the tide stage, relative to Mean Sea Level (MSL). With GeoClaw we can run the code with different sealevels $\hat{\xi}$, relative to MSL, that remain fixed over the tsunami duration.
- CSZ, AASZ, KmSZ, KrSZ, and SchSZ refer to the Cascadia, Alaskan Aleutian, Kamchatka, Kuril, and South Chile Subduction Zones. TOH refers to Tohoku. We denote tsunami sources in the form AASZe03, for example, event number 3 on the Alaska Aleutian Subduction Zone. Some events, e.g., a CSZ Mw 9.1 event, have multiple possible realizations. CSZBe01r01-CSZBe01r15 refers to the CSZ Bandon sources of various sizes of a single event modelled as 15 realizations. We assign a recurrence time to the event and then a conditional probability to each realization of the event.

1.2 Probabilities, rates, and recurrence times

By *probability of an event* we generally mean *annual probability of occurrence.* Specific earthquake events are often assumed to be governed by a Poisson process with some *mean recurrence time* T_M, in which case the annual probability of occurrence of event E_j is $P(E_j) = 1 - e^{-\nu}$ where the *rate* is $\nu = 1/T_M$. If ν is small then $P(E_j) \approx \nu$ with an error that is $O(\nu^2)$. For example, if $T_M = 250$ then $\nu = 0.004$ and $P(E_j) = 0.003992$. For larger T_M there is even less error. It is generally fine to assume $P(E_j) = 1/T_M$.

1.3 Probability of exceedance

We consider J tsunami events, with event E_j having a recurrence rate ν_j that obeys a Poisson process. We are interested in finding the probability that inundation height ζ exceeds level ζ_i at a grid location of interest. Typically, we are interested in all grid locations covering a fixed grid of the Crescent City area. The probability that E_j does not produce exceedance of ζ_i is

$$1 - (1 - e^{-\nu_j})\, P(\zeta > \zeta_i \,|\, E_j).$$

Then the probability that at least one event gives exceedance of ζ_i is

$$P(\zeta > \zeta_i) = 1 - \prod_{j=1}^{J} \left(1 - (1 - e^{-\nu_j})\, P(\zeta > \zeta_i \,|\, E_j)\right). \tag{1}$$

Furthermore, if event E_j is composed of k_j mutually exclusive realizations, so that when E_j occurs, exactly one of the realizations occurs, say E_{jk}, then

$$P(\zeta > \zeta_i \,|\, E_j) = \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk})\, P(E_{jk} \,|\, E_j)$$

where $\sum_{k=1}^{k_j} P(E_{jk} \,|\, E_j) = 1$. Substituting this into equation (1) gives

$$P(\zeta > \zeta_i) = 1 \; - \; \prod_{j=1}^{J} \left(1 \; - \; (1 - e^{-\nu_j}) \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk}) \, P(E_{jk} \,|\, E_j) \right) . \quad (2)$$

If we define $\bar{\mu}_{ij}$ as

$$\bar{\mu}_{ij} = (1 - e^{-\nu_j}) \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk}) \, P(E_{jk} \,|\, E_j), \quad (3)$$

equation (2) can be written as

$$P(\zeta > \zeta_i) = 1 \; - \; \prod_{j=1}^{J} (1 - \bar{\mu}_{ij}) \quad (4)$$

and following the discussion in Section 1.2, can be approximated as

$$P(\zeta > \zeta_i) \approx 1 \; - \; \prod_{j=1}^{J} e^{-\bar{\mu}_{ij}} . \quad (5)$$

If we approximate $\bar{\mu}_{ij}$ in equation (3) by μ_{ij}, where

$$\mu_{ij} = \nu_j \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk}) \, P(E_{jk} \,|\, E_j), \quad (6)$$

we arrive at the expression for $P(\zeta > \zeta_i)$ from [González, et al(2009)]:

$$P(\zeta > \zeta_i) \approx 1 \; - \; \prod_{j=1}^{J} e^{-\mu_{ij}} . \quad (7)$$

The procedure is to first find $P(\zeta > \zeta_i \,|\, E_{jk})$ in (6) using one of the methods in Section 2, and then use the known conditional probabilities $P(E_{jk} \,|\, E_j)$ and (6) to calculate the μ_{ij} that is used in (7) to find $P(\zeta > \zeta_i)$. By varying $i = 1 \ldots n_\zeta$ to cover more exceedance levels of interest, we can calculate the pairs $(\zeta_i, \, P(\zeta > \zeta_i))$, $i = 1 \ldots n_\zeta$ and construct a *hazard curve* for each fixed grid location (x, y) with the horizontal axis representing maximum depth of inundation ζ and the vertical axis the probability of exceeding this value. The terminology of hazard curves has been used for many years in probabilistic *seismic* hazard assessment (PSHA) and has been adopted in PTHA and used in past studies such as [González, et al(2009)].

In practice we use a finite set of exceedance values ζ_i and approximate the hazard curve by a piecewise linear function that interpolates the values $(\zeta_i, \; P(\zeta > \zeta_i; x, y))$. We do this because computing each $P(\zeta > \zeta_i; x, y)$ requires combining information from all simulation runs together with tidal variation and is somewhat costly to perform. We use the following $n_\zeta = 35$

exceedance values which we believe is sufficiently dense to yield good approximations:

$$\zeta_i = 0,\ 0.1,\ 0.2,\ \ldots,\ 1.9,\ 2.0,\ 2.5,\ \ldots,\ 5.5,\ 6.0,\ 7.0,\ \ldots,\ 12.0. \tag{8}$$

Figure 1 gives several hazard curves for one location in Crescent City, CA, showing each Subduction Zone's influence to the total hazard (in green).

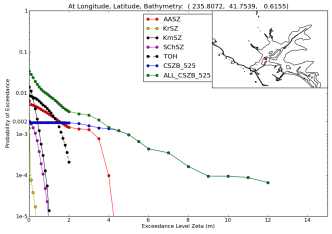

Fig. 1: Hazard Curves by Subduction Zone

Once the hazard curve at each (x, y) has been determined, the information contained in this curve can be used in two distinct ways. For a given probability such as $\bar{p} = 0.01$ it is possible to find the corresponding value ζ_{100} for which $P(\zeta > \zeta_{100}; x, y) = 0.01$. This could be interpreted as the depth of inundation expected in a "100-year event". By determining this for each (x, y) it is possible to plot the extent of inundation expected with probability $\bar{p}$ and the flow depth at each point inundated. [González, et al(2012)] shows the 100-yr flood for Crescent City, CA.

Conversely, one can choose a particular inundation level $\tilde{\zeta}$ and determine the probability of exceeding this value $P(\zeta > \tilde{\zeta}; x, y)$ at each point. A contour plot of this value over the spatial (x, y) domain then shows the probability of exceeding $\tilde{\zeta}$ at each point in the community. In particular, choosing $\tilde{\zeta} = 0$ would show probability contours (p-contours) of seeing any flooding. The $p = 0.01$ contour would again correspond to the inundation limit of the "100-year event". [González, et al(2012)] shows $\zeta = 0$ and $\zeta = 2$ meter p-contours for Crescent City, CA.

2 Methods for finding $P(\zeta > \zeta_i \,|\, E_{jk})$ including tidal variation

As outlined above, we need to find the probability that an inundation height ζ exceeds level ζ_i due to the k-th realization of source j whenever it occurs, denoted by

$$P(\zeta > \zeta_i \,|\, E_{jk}). \tag{9}$$

We want to find this probability at each location in a fixed grid covering the Crescent City area when the effect of the tides is taken into account. We note that GeoClaw code is not modeling the tidal dynamics, so GeoClaw information needs to be combined with tidal information at Crescent City to determine the probability in equation (9). We implemented three different

methods for doing this and determined their relative merits. The three methods are referred to as the dt-Method, the Pattern-Method, and the G-Method. All three methods are compared in detail in Section 3. The dt-Method and the Pattern-Method give quite similar results for a properly chosen dt but vary significantly from the G-Method, especially at land points. The Pattern-Method is a very robust method coupled to the wave pattern actually seen in GeoClaw for each individual tsunami. Furthermore, the Pattern-Method gives modelers a single method that can be used for both land and water locations. The G-Method [Mofjeld, et al(2007)] is briefly described in Section 2.4. The key ideas in the dt-Method, see Section 2.2, and the Pattern-Method, see Section 2.3 are summarized below:

A tsunami wave that arrives at high tide will cause more flooding than the same wave arriving at low tide. Nonlinearities in the governing equations mean that there will be nonlinearities in the tsunami-tide interaction. For example, if the tide stage is 1 meter higher, the resulting maximum flow depth at a point will not generally be exactly 1 meter higher, even at points that are inundated at both tide levels.

The GeoClaw code can easily be set to run with different (static) values of sea level in order to explore how the tide stage affects the level of inundation. The tide stage used for a run will be denoted by $\hat{\xi}$, relative to MSL.

For each exceedance level ζ_i and each grid point (x, y), we can use multiple GeoClaw runs to estimate how high the tide stage must be in order to observe a maximum GeoClaw flow depth above ζ_i at this point. This value of tide stage that must be exceeded will be denoted $\hat{\xi} = w_e$ below, the "water level to exceed". Note that w_e is different for each ζ_i at each (x, y) but in the discussion below we focus on a single point and exceedance level. We can then ask what the probability is that the tide stage at Crescent City will be above w_e when the tsunami arrives. If the tsunami consisted of a single wave of short duration, then the probability of exceeding ζ_i for this one realization would simply be the probability that the tide stage ξ is above w_e at one random instant of time i.e. a random point in the tide cycle. This can be estimated based on the past history of tides at Crescent City, as explained further below.

However, it is not this simple because most tsunamis consist of a sequence of waves that arrive over the course of several hours. During this time the tide may rise or fall considerably. If the tsunami consists of a sequence of closely spaced and equally large waves arriving over a period of Δt hours, then the better question to ask would be: what is the probability that the tide stage will be above w_e at *any* time between t_0 and $t_0 + \Delta t$, where t_0 is a random time. For fixed Δt this can also be determined from past tide tables. This approach is explained in Section 2.2 as the "dt-Method". Different events will require different choices of Δt. For example, a Cascadia Subduction Zone (CSZ) event typically gives one very large wave that causes most of the inundation. On the other hand farfield events may lead to a larger number of waves that arrive over many hours due to reflections from various distant points, any one of which could give flooding exceeding ζ_i if the tide stage is above w_e.

For some events, it may be that there are several such waves separated by many hours when no waves arrive that could cause the same level of flooding. In this case choosing a large Δt may overestimate the probability of inundation above ζ_i. Instead we might want to specify a *pattern* of times specific to one realization when the dangerous waves arrive. For example, if the tsunami consists of two large waves arriving 4 hours apart, the pattern might consist of a 1-hour window starting at time t_0 and another 1-hour window starting 4 hours later. We could then ask what the probability is that the tide stage will be above w_e at any time in this pair of windows, when t_0 is a random point in the tide cycle. This can also be determined based on the tide record and gives a smaller (and more accurate) probability than simply looking at a $\Delta t = 5$ hour window would. Similar questions can also be answered when the tsunami consists of multiple waves of different amplitudes. This is the basis of the "Pattern-Method" described in Section 2.3.

The dt and Pattern methods were designed to use GeoClaw simulation information at multiple *but static* tidal levels. These methods will work with other simulation codes that have the capability to produce similar results.

2.1 Crescent City tides

The tide gauge at Crescent City (Gauge No. 9419750) gives Mean Low Low Water ($\xi_{MLLW} = -1.13$), Mean Low Water ($\xi_{MLW} = -0.75$), Mean Sea Level ($\xi_{MSL} = 0.0$), Mean High Water ($\xi_{MHW} = 0.77$), and Mean High High Water ($\xi_{MHHW} = 0.97$). The lowest and highest water seen at the gauge in a year's data from July 2011 to July 2012 are $\xi_{Lowest} = -1.83$ and $\xi_{Highest} = 1.50$, respectively. The tide levels in meters are referenced to MSL. Figure 2 shows the probability density function and cumulative distribution for this yearly data. The horizontal axis represents tidal level and the vertical axis of the Cumulative Distribution Function represents the probability of exceedance of this level at any point in time.

Fig. 2: Crescent City Tidal Distributions Left: Probability Density Function (mean=0.0, $\sigma = .638$) Right: Cumulative Distribution Function

A GeoClaw simulation of the shallow water equations is done for each realization of each source. This gives a maximum inundation height z associated with the tide level set in GeoClaw, at each fixed grid location. Whether or not

this GeoClaw maximum z value is actually achieved or exceeded depends on the tidal levels at Crescent City during the tsunami. This represents aleatoric uncertainty, as we do know the tidal patterns at Crescent City, but we do not know when the tsunami will occur. To model this uncertainty, we make use of the GeoClaw feature that permits simulations at any *static* tide level.

2.2 The dt-Method

For each realization E_{jk}, we run GeoClaw simulations at multiple static tide levels, typically the three levels $\hat{\xi}_m$ for $m =$ MLW, MSL, and MHHW. We say tide level $\hat{\xi}_m$ produced the maximum GeoClaw inundation depth $z(\hat{\xi}_m)$ and plot the results with a piecewise linear function, as shown in the GeoClaw Simulation Curve in Figure 3. (This example used 7 tide levels.) The intersection of the vertical dashed line with the tide level axis will give the minimum *static* tide level $\hat{\xi} = w_e$ that could be used with GeoClaw to produce inundation depth $z = \zeta_i$. Hence, if tide level $\hat{\xi} > w_e$ were used for a GeoClaw run, we claim that $z > \zeta_i$ would result.

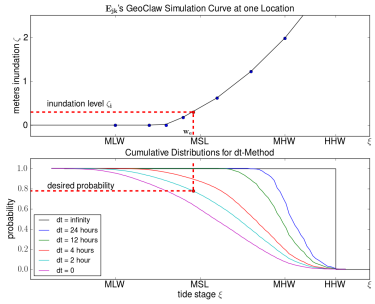

Fig. 3: Finding $\hat{\xi} = w_e$ and $P(\xi > w_e \,|\, \mathrm{dt} = 2)$

The upper graph in Figure 3 has been extended to the left and to the right of the data points (blue dots). We do this extension using a linear segment of appropriate slope. If $\zeta = 0$ for all the data points, we extend both left and right with slope=0; otherwise, we extend to the left using slope=min(1, slope of first data segment) and to the right using slope=max(1, slope of last data segment).

If the horizontal dashed line in the GeoClaw Simulation Curve at inundation level ζ_i intersects the extended graph in multiple places (as would happen for example if inundation does not occur until the tide reaches a particular level), we choose $\hat{\xi} = w_e$ to be the smallest tide level above which ζ_i is exceeded. As an example, if $\zeta_i = 0$, the graph shows we only exceed ζ_i if the tide is above the third blue dot, so the tide level associated with this point would be chosen for w_e. It could also happen that the ζ_i inundation dashed line falls below the entire graph (think of shifting the graph up by .5 meters and considering $\zeta_i = 0$). In this case $P(\zeta > \zeta_i \,|\, E_{jk}) = 1$. Likewise, if the ζ_i inundation

dashed line is above the entire graph, we set $P(\zeta > \zeta_i \,|\, E_{jk}) = 0$. If w_e is greater than or equal to the highest tide possible at the Crescent City gauge, we set $P(\zeta > \zeta_i \,|\, E_{jk}) = 0$. Finally, if w_e is less than the lowest tide possible at the Crescent City gauge, ζ_i is always exceeded and we set $P(\zeta > \zeta_i \,|\, E_{jk}) = 1$.

Now, suppose for the moment that the GeoClaw tsunami of Figure 3 consisted of only one wave with a very narrow width, say a spike even. If the tide level at Crescent City at the time this wave strikes exceeds w_e, then we say the conditional probability is 1 because both the tsunami wave and the Crescent City tide support this level of inundation. However, we don't know exactly when the tsunami will strike in the tidal cycle and the level might not exceed w_e. This aleatory uncertainty has been quantified in the cumulative probability distribution function in Figure 2 that gives the probability of exceedance of w_e at any particular instance of time. We refer to this time interval as $\mathrm{dt} = 0$, and say that $P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \mathrm{dt} = 0)$.

The cumulative probability distribution from Figure 2 is also shown as the curve labelled (dt=0) in the bottom plot in Figure 3. This graph shows how we would extract the desired probability by looking up w_e in the cumulative distribution table (when dt=0 we would drop the dotted vertical line to the bottom graph and then construct another dotted horizontal line to read off the desired probability). This is very convenient, since the question about $P(\zeta > \zeta_i \,|\, E_{jk})$ is changed to a simplier question about the tide levels at Crescent City and the *same* cumulative distribution table can be used for every grid point location in Crescent City (only w_e varies across the grid locations).

Next, suppose the tsunami represented in Figure 3 still consists of one wave, but a much wider one (say 15 minutes wide). It is convenient to think of a square wave with constant amplitude over this 15 minute interval. We still can find the constant GeoClaw tide level, w_e, that we need to exceed so that ζ will exceed ζ_i. The issue, though, is that the tide level at Crescent City will not remain constant during a 15 minute interval, although it changes by at most .18 meters. Do we need the Crescent City tide level ξ to exceed w_e during the entire 15 minutes to report exceedance of ζ_i? Would the same exceedance occur if $\xi > w_e$ for only 7.5 minutes while this square wave were passing into Crescent City? Taking this to the limit, would we still exceed ζ_i if $\xi > w_e$ at only one point in the 15 minute interval when the wave were coming into Crescent City? We don't know the answers to these questions, but choose to err on the side that would give the biggest probability. We will say exceedance of ζ_i occurs if the maximum value, $\bar{\xi}$, of ξ during the 15 minute (.25 hr) interval exceeds w_e, and denote the probability of any 15 minute interval as having a maximum value exceeding w_e as $P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \mathrm{dt} = .25)$.

We can of course consider the one-wave scenerio with waves wider than .25 hours since our experience shows that the wave width typically lasts between 5 and 45 minutes. The procedure is the same, and the requirement is that we are able to create a cumulative distribution table with columns corresponding to the size of dt, and rows corresponding to valid values of w_e. The bottom graph in Figure 3 illustrates several graphs of the columns of such a table, with one graph per column. The limiting case is considering an infinite dt which

would correspond to choosing the conditional probability to be 1 if the value of w_e is smaller than the largest tide level seen at Crescent City Gauge No. 9419750 and 0 otherwise.

The cumulative distribution corresponding to a finite dt is gotten as follows. We simply take a dt-window of time and slide it one minute at a time across a year's worth of Gauge 9419750 data. Each time the dt-slider window stops, we find the maximum tide level within the window. We increment a counter in the first bin whose right edge exceeds or equals this maximum (to create a histogram) and also in all lower bins (to create a cumulative histogram). Dividing by the number of times the dt-slider window stops gives us the probability mass function and cumulative distribution function, respectively. (The probablility density function is then obtained by dividing the probablility mass function by the binsize used.) A table is saved that records the cumulative probabilities for the valid tide levels with one column for each dt considered.

Tsunamis, however, consist of multiple waves of varying amplitudes and widths, and may have the biggest amplitudes spaced apart by hours during which the height of the tide alone will not change the maximum exceedance ζ value at a grid point location. Multiple waves of nearly or equal magnitude should increase the probability of exceedance of ζ_i since the time frame where w_e could be exceeded increases. Even waves with lesser magnitude than the largest one could produce exceedance of ζ_i if they came into Crescent City at a sufficiently higher tide level than w_e. Applying the dt-Method to these cases means finding a reasonable way to choose dt. We have the possibility in GeoClaw to record the time history for the tsunami wave (or its effect) at any computational location. Of course, doing this everywhere is prohibitive, but to assist this study, we place GeoClaw Gauge 101 at a location in the water near the Crescent City Gauge 9419750, and GeoClaw Gauge 105 at a point that usually inundates (near the river, but on land). We also have Gauge 33 near the shelf in deeper water, and GeoClaw Gauges 102, 103, and 104 on land. We record what we call the GeoClaw tsunami at Gauge 101, and its biggest effect is usually at Gauge 105. Examination of these computational gauges gives the time intervals and widths of the waves responsible for inundation. The width of the responsible wave of biggest amplitude certainly gives a minimum value for the *contiguous* dt interval, and we increase dt based on nearby potentially responsible waves.

In Section 2.3, we see the dt-Method works remarkably well compared to the Pattern-Method for appropriately chosen dt. In particular, for all tsunamis in Table 2 and [González, et al(2012)] the recommended values of dt can be given. We recommend dt=1 for the Kamchatka event KmSZe01 and dt=3 for KmSZe02. For the three Kuril events, we recommend dt=2 for KrSZe01, dt=3 for KrSZe02, and dt=4 for KrSZe03. For the Alaska events, we recommend dt=1 with the exception of dt=2 for AASZe02. The value dt=1 should be used for the Chilean event SChSZ01, the Tohoku event TOHe01, and the Cascadia Bandon CSZBe01r13 and CSZBe01r14 realizations. The value dt=0 should be used for the remaining Cascadia Bandon realizations, CSZBe01r01-

CSZBe01r12 and CSZBe01r15. We suspect that choosing dt beyond 4 will give overestimates of the probability as this points to a 4 hour *contiguous* interval.

2.3 The Pattern-Method

This approach grew from the desire to automate the choice of dt in the dt-Method. Instead of achieving this automation, we developed an even better method that is tailored to each realization's GeoClaw tsunami as seen at GeoClaw Gauge 101. The Pattern-Method uses the relative heights of the wave amplitudes seen at Gauge 101, their widths, and the times they occurred, (with the first wave starting at time 0), to first create a cumulative probability distribution (a table with one column) associated with this particular wave pattern. This is extra work, but the difference is that a fixed dt will not have to be chosen. Instead, the entire pattern will be taken into account to calculate the distribution. Figure 4 shows that for some tsunamis this new cumulative distribution when compared to the columns of the dt-Method's cumulative distribution gives probabilities similar to a fixed dt (AASZe03, dt=1), while for other tsunamis the probabilities are consistent with a varying dt (AASZe02). The pattern cumulative distribution is shown as a dotted line on the same graph as that for the dt-Method with varying dt.

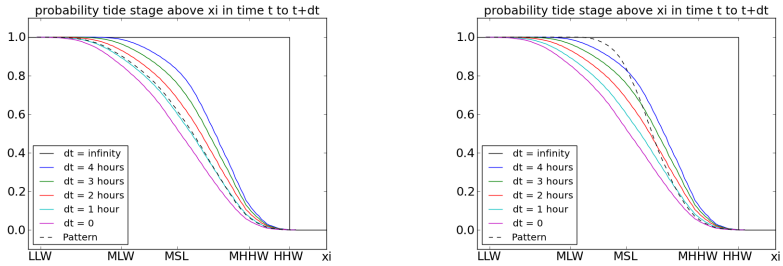

Fig. 4: Pattern to dt Comparison, Left: AASZe03, Right: AASZe02

Suppose Gauge 101 records K waves. We model wave W_k with a square wave and record the difference of its amplitude from that of the highest wave as D_k. We record the starting and terminating times of W_k as the interval $I_k = [S_k, T_k]$. These times are relative to the start of W_1, so we set $S_1 = 0$, and are recorded in minutes since our gauge 9419750 has minute data. The entire length of the pattern is then T_K minutes, the ending time of wave K.

In Figure 5, we show the GeoClaw tsunami for the AASZe02 event that was recorded at gauge 101 as the red graph and the pattern as the black graph. The first wave arrived at Crescent City 4 hours and 23 minutes after the earthquake and nothing significant was seen there after 11 hours. The pattern is well represented by the 7 waves shown. We are overestimating the probability a bit by using square waves, but we don't have to account for tides

during times that they can't possibly have any impact. A table showing the values that describe the pattern are given in Table 1. We note that the first wave began at 263 minutes after the earthquake and the amplitude of the largest wave W_7 was about 1.5 meters. The black horizontal line starts at .2 meters since the GeoClaw run was done at MHHW which is .2 meters above MHW, the zero level for the Gauge 101 plot in Figure 5.

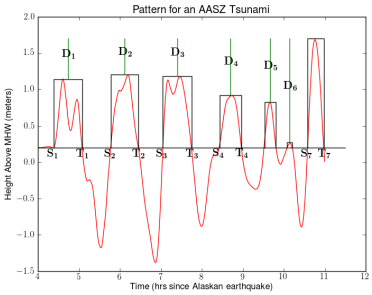

Fig. 5: Pattern for AASZe02

Table 1: Pattern Values

Wave W_k	$I_k = [S_k, T_k]$ Wave Interval (min since S_1)	D_k (meters) Difference to Tallest Wave
W_1	[000, 042]	0.561
W_2	[084, 124]	0.498
W_3	[160, 202]	0.517
W_4	[243, 275]	0.782
W_5	[309, 325]	0.876
W_6	[342, 349]	1.450
W_7	[372, 396]	0.000

As in the dt-Method, we take the valid values for the tide levels at Crescent City and put them into a fixed number of bins. But now we take our pattern-slider window that has length T_K and slide it one minute at a time across a year's worth of Gauge 9419750 data. Each time the pattern-slider window stops, we do the following:

- Find the maximum tide level, M_k associated with each I_k, $k = 1 \dots K$.
- Adjust M_k to get $\overline{M}_k$: $\overline{M}_k = M_k - D_k$.
- Compute $M_P = \max_k \overline{M}_k$.
- Increment a counter in the first bin whose right edge exceeds or equals M_P, the max for the pattern for this window stop, to create a histogram and also increment all lower bins to create a cumulative histogram.

Dividing the cumulative histogram by the number of times the pattern-slider window stops gives a cumulative distribution function for the probability of exceeding each valid tide level by a tsunami of this pattern. A table is saved that records the cumulative probabilities for all valid tide levels at the Crescent City gauge 9419750. The associated probability density function is not needed but is computed so comparisons can be made to the other methods.

After the pattern cumulative distribution is found, the method proceeds exactly as the dt-Method. We use the multiple GeoClaw simulations to find the minimum *static* tide level $\hat{\xi} = w_e$ that could be used with GeoClaw to produce inundation height $z = \zeta_i$. If we would make a GeoClaw simulation with tide level w_e (we don't do this), the thinking is that the resulting tsunami pattern values at GeoClaw gauge 101 would be the same as those obtained using any other tide level. This is because gauge 101 is in the water and records

the tsunami as it comes into Crescent City as opposed to being located at a land point that also feels the tsunami's nonlinear effects.

So, we need to find the probability that the Crescent City tide is sufficient for the tsunami pattern to exceed ζ_i by looking up w_e in our pattern cumulative distribution. We denote the probability that the tide exceeds w_e in the sense of the pattern as $P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \text{pattern})$ with the meaning

$$P(\xi > w_e \,|\, \text{pattern}) = P(\xi > w_e + D_k \text{ somewhere in } I_k \text{ for some } k). \quad (10)$$

The advantages over the dt-Method include:

- Only one synthetic gauge, GeoClaw gauge 101, needs to be examined.
- By adjusting the M_k, we permit the possibility that a wave with amplitude D_k less than the maximum one seen at GeoClaw gauge 101 could also cause an inundation at gauge 101 of ζ_i or higher if it occurred at a time when the tide level was at least D_k higher than that required of the maximum amplitude GeoClaw gauge 101 wave.
- By looking in each interval I_k, we take into account each wave's width. We only examine the tide during each interval I_k, not between. This allows a more accurate representation of tsunamis that have a longer duration.
- The procedure is automatic.

A possible limitation is that the Pattern-Method requires the simulation code to have GeoClaw's capability of a computational gauge. The dt-Method benefits from examining the gauges to determine dt, but if none were present, the recommended choice would be to use dt=1 for all near-field realizations and dt=2 for far-field events.

2.4 The G-Method

The G in the G-Method emphasizes that parameters are chosen to select a Gaussian probability density function for the maximum wave height of the tsunami and the tides. A 5-day theoretical tsunami with exponentially decaying amplitude having an e-folding time of 2 days was assumed at each grid location P. Other authors have used e-folding times to model the decay of tsunami wave energy, see [Van Dorn(1984)], [Rabinovich, et al(2011)], and [Fine, et al(2012)]. The tsunami's amplitude, A_G at location P is calculated using data at the grid location from one GeoClaw simulation using one tide level, $\hat{\xi}$. This theoretical tsunami was then combined with local tidal information and regression analysis used to develop an analytical expression for a Gaussian probability density function and a cumulative distribution described by the **erf** function.

[Mofjeld, et al(2007)] gives parameters for this method for a variety of locations. The parameters for Crescent City include $\sigma_0 = .638$, the standard deviation for the tides there, and the regression parameters $\alpha' = 0.056$, $\beta' = 1.119$, C'=.707, α=0.17, β=.858, and C=1.044.

The form assumed for the standard deviation, σ, of the random variable $\zeta_P(\hat{\xi})$ (and hence the random variable $\bar{\xi}_P(\hat{\xi}) = \zeta_P(\hat{\xi}) - z_P(\hat{\xi}) + \hat{\xi}$) is a function of the amplitude at location P, ($A_G = A_P(\hat{\xi})$), and the parameters:

$$\sigma = \sigma_0 \left(1 - C' \, e^{-\alpha' \left(\frac{A_G}{\sigma_0}\right)^{\beta'}}\right) \tag{11}$$

The form for the mean of $\zeta_P(\hat{\xi})$, denoted ζ_0, is

$$\zeta_0 = z_P(\hat{\xi}) - \hat{\xi} + C(\xi_{MHHW}) \, e^{-\alpha \left(\frac{A_G}{\sigma_0}\right)^{\beta}}, \tag{12}$$

and hence the mean of $\bar{\xi}_P(\hat{\xi})$, denoted w_0, is

$$w_0 = C(\xi_{MHHW}) \, e^{-\alpha \left(\frac{A_G}{\sigma_0}\right)^{\beta}}. \tag{13}$$

The G-Method calculates the probability as

$$P(\zeta_P(\hat{\xi}) > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\zeta_i - \zeta_0}{\sqrt{2}\,\sigma}\right)\right). \tag{14}$$

Equation (14) can also be written as

$$P(\zeta_P(\hat{\xi}) > \zeta_i \,|\, E_{jk}) = P(\bar{\xi}_P(\hat{\xi}) > \tilde{w}_e) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\tilde{w}_e - w_0}{\sqrt{2}\,\sigma}\right)\right) \tag{15}$$

where $\tilde{w}_e = z_P(w_e) - z_P(\hat{\xi}) + \hat{\xi}$, and we implement using the choice $\hat{\xi} = \xi_{MHHW}$. If the GeoClaw Simulation Curve $\hat{\xi}$ vs $z_P(\hat{\xi})$ has slope 1, $\tilde{w}_e = w_e$.

The G-Method has two major limitations. First, only one GeoClaw simulation with tide level $\hat{\xi} = \xi_{MHHW}$ is used to produce $z_P(\xi_{MHHW})$, and this value is used to compute A_G. Using only one simulation is appropriate whenever $z_P(\hat{\xi})$ and $\hat{\xi}$ are related by a linear relationship with slope 1, since then the value of A_G and the random variables ζ_P and $\bar{\xi}_P$ will be independent of the tide level $\hat{\xi}$ used for the GeoClaw run. Multiple GeoClaw Simulations Curves show that these assumptions are not true, especially for land locations. Figure 6 illustrates this point for one water and one land location, using 11 different tide levels for the Alaska 1964 event. The tsunami amplitudes were calculated using all 11 tide runs, and the range of amplitudes are given with the plot as well as the slopes of the piecewise linear segments in the graphs. Notice that for the water location, the amplitudes calculated from the 11 runs were fairly close, as were the slopes. The same is not true for the land location. The black line is the slope 1 through $(\xi_{MHHW}, z(\xi_{MHHW}))$.

The second limitation is the use of the same 5-day proxy tsunami (where the amplitude alone varies) as the pattern for modelling each tsunami in a PTHA study, especially when the major question being studied is the inundation at land points. As seen in Table 2, the duration of all the tsunamis studied that could impact the maximum inundation at a land point is much less than

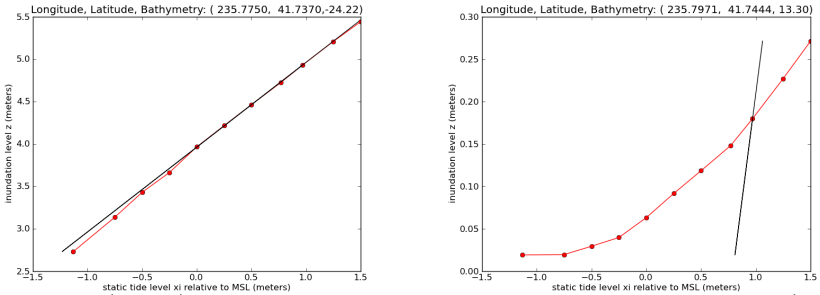

Fig. 6: Left (water): Slopes .93 to 1.21, Amplitudes 4.63 to 4.74. Right (land): Slopes .001 to .18, Amplitudes 13.1 to 15.2

5 days, and as seen by GeoClaw time series at Gauge 101, these tsunamis have patterns that are very different. A local tsunami from the Cascadia Subduction Zone will typically have only one or two waves occuring over a short time frame that are responsible for the maximum; whereas, far field events can have damaging waves occuring over a longer time frame, but still much shorter than 5 days with amplitudes that can increase over a time interval. It was not the first wave that was the largest in the real 1964 Alaska event, nor is it the first for many of the sources in Table 2. Also, for all these sources, it was rare to see the GeoClaw tsunami waves have non-increasing amplitudes in the first two hours after the arrival of the first wave. In fact, the first four Alaska sources did not have non-increasing amplitudes up through the first seven hours, and the second Kamchatka source did not have non-increasing amplitudes up through the first five hours. Such wave patterns pass through significant tidal variations.

3 Method comparisons

3.1 G and Pattern PDF Comparisons for $\bar{\xi}$ at Gauge 101 (multiple sources)

In Table 2, we compare the probability density functions (PDFs) of the Mofjeld (G) and Pattern methods for some of the tsunamis considered in the Crescent City study. Table 2 shows there are huge differences between the G-Method (Mofjeld) and the Pattern-Method. Only for the four large amplitude tsunamis CSZBe01r01, CSZBe01r02, CSZBe01r03, and CSZBe01r04 do the two methods have PDFs with similar means and standard deviations. For the other tsunamis in the table, the G-Method has a much higher mean and smaller standard deviation than the Pattern-Method.

3.2 AASZe03 comparisons

For the purposes of further comparing the methods, we made GeoClaw runs of the Alaska 1964 event using 11 tide levels. These levels referenced to MSL

Table 2: Mofjeld (G) and Pattern Method PDF comparisons. The length T in min. and amplitudes $A_G = A_{101} = z_{101}(\xi_{MHHW}) + \xi_{MHW} - \xi_{MHHW}$ in m. (seen at Gauge 101) are given in columns 2 and 3 for some tsunamis used in this study. Columns 4-7 give the mean ω_0 and standard deviation σ of the PDFs for $\bar{\xi}_{101} = \zeta_{101} - z_{101}(\xi_{MHHW}) + \xi_{MHHW}$ generated by the two methods.

Source Name	T (min)	A (m)	Mofjeld (G) w_0 (m)	Mofjeld (G) σ (m)	Pattern w_0 (m)	Pattern σ (m)
AASZe03-Proxy	7205	3.92	.45	.34	.46	.34
AASZe01	328	1.96	.65	.27	.12	.53
AASZe02	396	1.50	.71	.25	.36	.37
AASZe03	267	3.92	.45	.34	.14	.60
AASZe04	476	1.77	.67	.26	.18	.47
KmSZe01	308	.92	.80	.22	.15	.54
KrSZe01	275	.50	.88	.21	.22	.52
SChSZe01	106	.60	.86	.21	.16	.60
TOHe01	324	1.66	.69	.26	.07	.59
CSZBe01r01	329	14.18	.09	.56	.04	.63
CSZBe01r02	326	12.96	.11	.55	.04	.63
CSZBe01r03	326	13.31	.10	.55	.04	.63
CSZBe01r04	157	13.00	.11	.55	.04	.63
CSZBe01r09	160	6.72	.28	.43	.03	.63
CSZBe01r15	160	3.25	.51	.32	.04	.63

were -1.13, -0.75, -0.50, -0.25, 0.00, 0.25, 0.50, 0.77, 0.97, 1.25, and 1.5 meters. We also consider the 35 exceedance values given in (8).

3.2.1 G and Pattern cumulative comparisons for $\bar{\xi}$ at Gauge 101

We ran the Pattern-Method on the 5-day proxy tsunami that is assumed by the G-method and compared the resulting cumulative distributions for $\bar{\xi} = \zeta - z(\xi_{MHHW}) + \xi_{MHHW}$ at Gauge 101. The amplitude for the 5-day proxy tsunami was taken as that of the biggest wave seen at Gauge 101 for AASZe03. The two distributions when plotted are almost identical with values differing mostly less than 1% as seen in Figure 7 as the green and dashed red graphs and given in the first line of Table 2. The black graph is the distribution for $\bar{\xi}$ for the Pattern Method on the actual tsumani at Gauge 101 for which we used a $T = 267$ minute duration. This explains differences generated by the Mofjeld method (G-Method) and the Pattern-Method at the Gauge 101 for any real tsunami is not due to our methodology, but to the fact that the real tsunami is not well approximated by the proxy one. The Pattern-Method can capture the differences of each specific tsunami as seen in Figure 7 by the differences between the black graph and the green (or dashed red) ones.

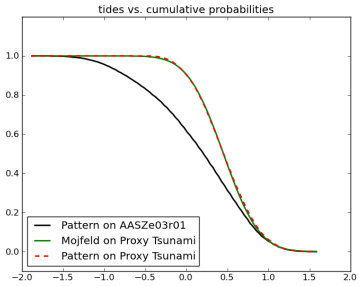

Fig. 7: Pattern Method Validation

3.2.2 All methods PDF comparisons for $\bar{\xi}$ at Gauge 101

Next, we computed the means and standard deviations of the PDFs for $\bar{\xi}$ at Gauge 101 using all three methods for the Alaska 1964 tsunami (AASZe03). These are given in Table 3 below.

Table 3: Method PDF comparisons for AASZe03 at Gauge 101. Columns 2 and 3 give the mean ω_0 and standard deviation σ of the PDFs for $\bar{\xi}$ used by the methods at GeoClaw Gauge 101 to compute the Cumulative Distribution for the probability indicated in column 4. We note $A_G = A_{101} = z_{101}(\xi_{MHHW}) + \xi_{MHW} - \xi_{MHHW}$, and $\tilde{w}_e = z_{101}(w_e) - z_{101}(\xi_{MHHW}) + \xi_{MHHW}$.

Method	w_0	σ	$P(\zeta > \zeta_i \,\vert\, E_{jk})$
Pattern	.14	.60	$P(\xi > w_e \,\vert\, \texttt{pattern}) = P(\bar{\xi} > w_e)$
`dt`	.12	.62	$P(\xi > w_e \,\vert\, \texttt{dt}) = P(\bar{\xi} > w_e)$
G	.45	.34	$P(\bar{\xi} > \tilde{w}_e)$ in (15)

3.2.3 Tide probability differences

For each grid location, we compared the 35 probabilities $P(\zeta > \zeta_i \,|\, E_{jk})$ in equation (6) for the 1964 Alaska event where $j = 1$, $k = k_j = 1$, and i ranges from 1 to 35. The numbers in Table 4 are over all the grid locations that cover the Crescent City area. The row labelled **max** is the maximum difference seen when the method being compared to the Pattern-Method gives the larger result, and the row labelled **min** is the difference seen when the Pattern-Method gives the larger result.

Indeed, differences close to 1 are observed in the first column and the second column shows that dt=1 for this particular tsunami gives results very close to those of the Pattern-Method. Both the dt and Pattern methods use

Table 4: **Tide Probability $P(\zeta > \zeta_i \,|\, E_{jk})$ Differences**

	G-Pattern	**dt-Pattern**
max	+.747	+.006
min	-.936	-.017

the amplitude of the tsunami at Gauge 101 (instead of the amplitudes at the land points) and assume its duration is T-minutes instead of 5 days. Further analysis given in [Adams, et al(2013)] shows that almost all of the -.936 is due to the G-Method's choice of 5 days, while all but .158 of the .747 is due to this choice. This remaining .158 difference is due to the use of a proxy decaying e-folding pattern of 2 days for the tsunami, rather than the observed pattern.

3.2.4 Tide probability differences contour plots

In Figure 8 we compare the dt-Method and the G-Method to the Pattern-Method by giving contour plots of the absolute value of the tide probability differences of exceeding $\zeta_i = 0$ meters and $\zeta_i = 2$ meters. The magnitudes in the colorbar show the tidal probabilities of the dt and Pattern methods differ by less than 2%. In fact, Table 4 gives the difference as less than 1.7%.

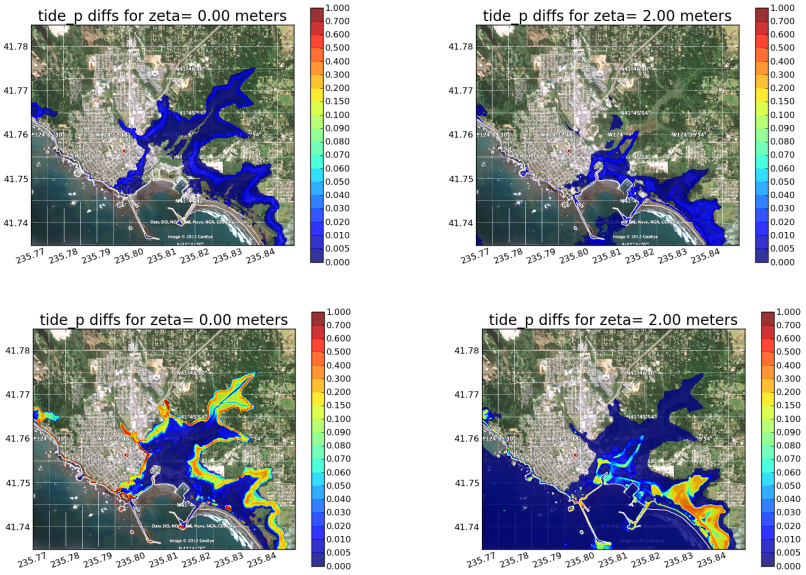

Fig. 8: Probability Difference Contours, Left: $\zeta = 0$ m., Right: $\zeta = 2$ m. Top: abs(dt-Pattern), Bottom: abs(G-Pattern)

4 Conclusions and open questions

This study has provided some advice to the PTHA community.

- The Pattern-Method is appropriate for both land and water locations.
- The G-Method should not be used for land points.
- Duration time T should be used for land points.

Issues that warrant further consideration are given below:

- The Pattern-Method can be extended to use different cumulative distributions for different groups of locations if warranted by a study's questions. For example, taking a larger T for water locations might be of interest.
- Using more computational gauges in addition to Gauge 101 could enhance the Pattern-Method.
- Finding an automatic way of choosing dt would enhance the dt-Method.
- We do not model the currents that are generated by the tide rising and falling. A tsunami wave arriving on top of an incoming tide could potentially inundate further than the same amplitude wave moving against the tidal current, even if the tide stage is the same. Modeling this is beyond the scope of current tsunami models.

References

Adams, et al(2013). Adams L, LeVeque R, González FI (2013) Incorporating tidal uncertainty into probabilistic tsunami hazard assessment (ptha) for Crescent City, CA. Final Tidal Report for Phase I PTHA of Crescent City, CA, supported by FEMA Risk MAP Program, 1-54

Fine, et al(2012). Fine IV, Kulikov EA, Cherniawsky JY (2012) Japan's 2011 tsunami: characteristics of wave propogation from observations and numerical modelling. Pure and Applied Geophysics DOI 10.1007/s00024-012-0555-8

González, et al(2009). González FI, Geist EL, Jaffe B, Kanoglu U, et al(2009) Probabilistic tsunami hazard assessment at Seaside, Oregon, for near-and far-field seismic sources. J Geophys Res 114:C11,023

González, et al(2012). González FI, LeVeque RJ, Adams L (2012) Probabilistic tsunami hazard assessment (ptha) for Crescent City, CA. BakerAECOM Report for PTHA of Crescent City, CA, supported by FEMA Risk MAP Program, 1-118

Houston and Garcia(1978). Houston JR, Garcia AW (1978) Type 16 flood insurance study: Tsunami predictions for the West Coast of the continental United States, Crescent City, CA. USACE Waterways Experimental Station Tech. Report H-78-26

Kowalik and Proshutinsky(2010). Kowalik Z, Proshutinsky A (2010) Tsunami-tide interactions: A Cook Inlet case study. Continental Shelf Research 30:633–642

Mofjeld, et al(2007). Mofjeld H, Gonzáles F, Titov V, Venturato A, Newman J (2007) Effects of tides on maximum tsunami wave heights: Probability distributions. J Atmos Ocean Technol 24(1):117–123

Rabinovich, et al(2011). Rabinovich AB, Candella RN, Thomson RE (2011) Energy decay of the 2004 sumantra tsunami in the world ocean. Pure and Applied Geophysics 168:1919–1950, DOI 10.1007/s00024-011-0279-1

Soloviev(2011). Soloviev S (2011) Recurrence of tsunamis in the Pacific. Tsunamis in the Pacific Ocean, W.M. Adams, editor, Honolulu, East-West Center Press, 149-163

Van Dorn(1984). Van Dorn W (1984) Some tsunami characteristics deducible from tide records. J Physical Oceanography 14:353–363

Zhang, et al(2011). Zhang Y, Witter R, Priest G (2011) Tsunami-tide interaction in 1964 Prince William Sound tsunami. Ocean Modeling 40:246–259