Incorporating Tidal Uncertainty Into Probabilistic Tsunami Hazard Assessment (PTHA)

for Crescent City, CA

Tidal Report on Phase I – Journal Article Working Draft

October 6, 2013

Loyce M. Adams, Randall J. LeVeque, Frank I. González
University of Washington

Photo Source: http://tripsintohistory.com/wave-propagation-content/uploads/2012/03/800px-Crescent_City_California_harbor_aerial_viewUSArmyCorpofEngineeers1.jpg

This work was done as a Pilot Study funded by BakerAECOM.

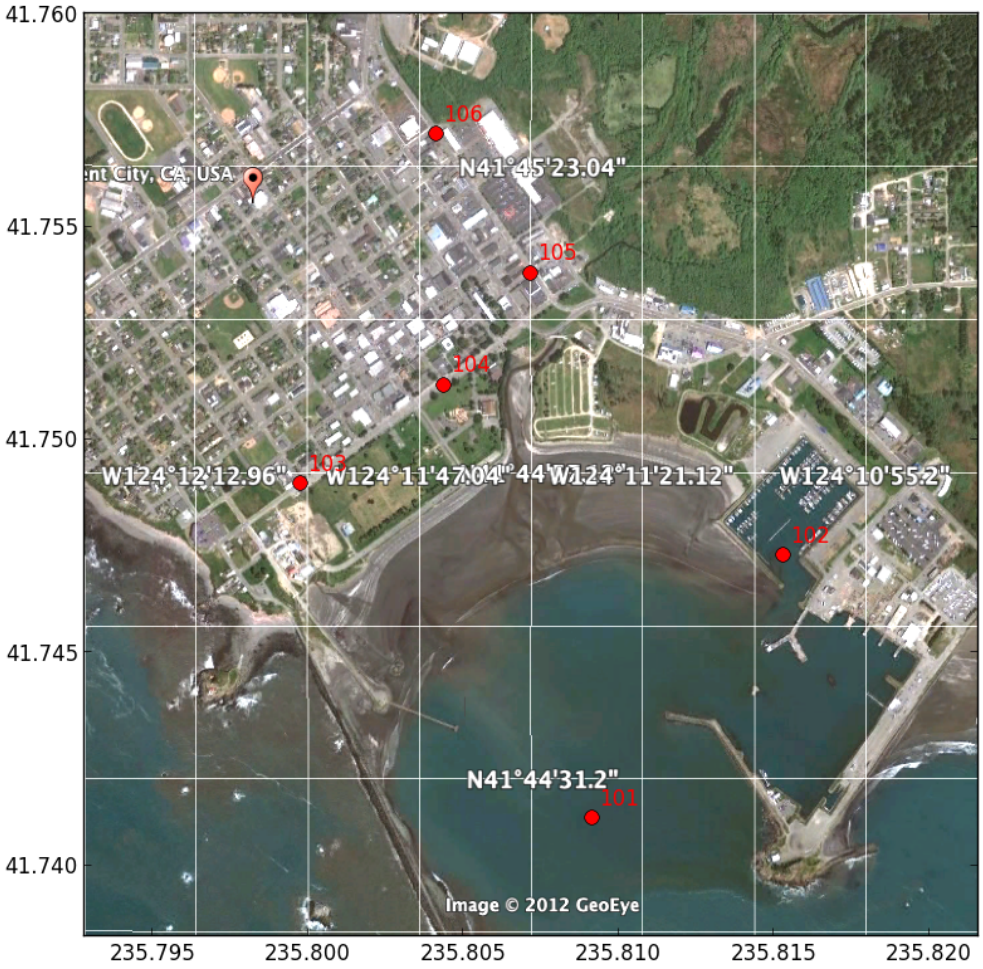


Figure 1: Zoom view of the harbor and downtown region. Gauges 101 through 106 in the computational results present below are indicated in this figure.

Contents

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 also made a series of improvements to the 2007 method of Mofjeld, et. al. [7] that was used for the 2009 Seaside, Oregon probabilistic study by González,et.al. [3]. We refer to these improved methods as the G2, G3, and G4 methods and that of [7] as the G method. All six methods are compared.

We believe 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 like Crescent City 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.

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, in 2010 Kowalik and Proshutinsky [6] 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, in 2011 a study of the 1964 Prince William Sound tsunami by Zhang et al. [10] 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.

In 1978, Houston and Garcia [5] 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 1970 work of Soloviev [9]. 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.

In 2007, Mofjeld, et. al. [7] developed a tidal uncertainty methodology that, unlike Houston and Garcia [5], 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 in 1984 by Van Dorn [1] for Pacific-wide tsunamis. As in Houston and Garcia [5], 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 is the method used in 2009 by González, et al. [3] in their 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 in 2012 as part of a PTHA study of Crescent City, California (González, LeVeque, and Adams [4]), a pilot stuy that was funded by BakerAECOM for the purpose of exploring 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 also describe a series of improvements to the method of Mofjeld, et. al. [7] which we refer to as the Gaussian or G-Method. We refer to these improved G-Methods as the G2-, G3-, and G4-Methods. All six methods are compared.

1.1 Notation and terminology

The following notation and terminology is used throughout this paper.

- 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 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. Some Cascadia Subduction Zone (CSZ) events cause a change in the bathymetry in the Crescent City region (uplift or more typically subsidence). We will sometimes use $\tilde{B}$ for the modified bathymetry, but normally B will refer to the pre-earthquake bathymetry.

- $\tilde{B} + h$ is the surface elevation relative to MHW output from GeoClaw.

- $\eta = \tilde{B} + h + (B - \tilde{B}) = h + B$ is the surface elevation relative to MHW plus the amount of subsidence calculated from GeoClaw output h.

- z will be used to denote the maximum observed GeoClaw value over the full time period of a tsunami of either h or $B + h$:

$$z = \begin{cases} h, & \text{the flow depth, in regions where } B > 0 \text{ (onshore)}, \\ B + h, & \text{the sea surface elevation plus subsidence, in regions where } B < 0. \end{cases} \tag{1}$$

- ζ will be used to denote the real value over the full time period of a tsunami of either d or $B + d$:

$$\zeta = \begin{cases} d, & \text{the flow depth, in regions where } B > 0 \text{ (onshore)}, \\ B + d, & \text{the sea surface elevation plus subsidence, in regions where } B < 0. \end{cases} \tag{2}$$

  GeoClaw inundation maps show z and hazard maps that include tidal variation show ζ. Sea surface elevation is relative to MHW.

- 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. This is important in the way we handle tidal uncertainty.

- Tsunami sources we used will generally be denoted in the form AASZe03r01, for example, which refers to realization number 1 of event number 3 on the Alaska Aleutian Subduction Zone. Many events have only one *realization* (model for how slip is distributed on the fault plane, and/or the resulting seafloor deformation). All of our 8 Alaskan Aleutian Subduction Zone events have only one realization, so we can shorten AASZe03r01 to simply AASZe03. Some events, e.g., a CSZ Mw 9.1 event, have multiple possible realizations, as will be evident in their naming convention. In the probabilistic modeling we typically assign a recurrence time to the event and then a conditional probability to each realization of the event.

- We use KmSZ, KrSZ, and SchSZ, to refer to the Kamchatka, Kuril, and South Chile Subduction Zones. TOH refers to Tohoku. CSZBe01r01-CSZBe01r15 refer to the Cascadia Subduction Zone Bandon sources of various sizes which we modelled as 15 realizations of a single 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 is $p = 1 - e^{-\nu}$ where the *rate* is $\nu = 1/T_M$. If ν is small then $p \approx \nu$ with an error that is $O(\nu^2)$. For example, if $T_M = 250$ then $\nu = 0.004$ and $p = 0.003992$. For larger T_M there is even less error. Since T_M is not accurately known, it is generally fine to assume $p = 1/T_M$.

When calculating the probability that one of several possible events might happen, some care is required. If two independent events are considered with annual probabilities p_1 and p_2 then the annual probability of at least one of them occurring is

$$p_{12} = 1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1 p_2.$$

If both probabilities are very small then $p_{12} \approx p_1 + p_2$ but for larger probabilities the more accurate expression must be used. Similarly, if we are interested in the probability of any one of N independent events occurring, the probability is

$$p_{1\ldots N} = 1 - (1 - p_1)(1 - p_2)\cdots(1 - p_N) \approx p_1 + p_2 + \cdots + p_N.$$

Again simply adding the probabilities is valid only if the result is much less than one, but otherwise not.

Note that when expressed in terms of Poisson rates, it is valid to add the rates: if $p_i = 1 - e^{-\nu_i}$ then $p_{1\ldots N}$ has rate $\nu_1 + \nu_2 + \cdots + \nu_N$ since

$$p_{1\ldots N} = 1 - e^{-\nu_1} e^{-\nu_2} \cdots e^{-\nu_N} = 1 - e^{-(\nu_1 + \cdots + \nu_N)}.$$

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. That is, the probability that E_j occurs is $P(E_j) = 1 - e^{-\nu_j}$. 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

$$P(\zeta > \zeta_i) = 1 \; - \; \prod_{j=1}^{J} \left(1 \; - \; (1 - e^{-\nu_j})\, P(\zeta > \zeta_i \,|\, E_j)\right) \; . \tag{3}$$

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 (3) gives

$$P(\zeta > \zeta_i) = 1 \,-\, \prod_{j=1}^{J} \left(1 \,-\, (1 - e^{-\nu_j}) \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk})\, P(E_{jk} \,|\, E_j) \right) . \tag{4}$$

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

$$\bar{\mu}_{ij} = (1 - e^{-\nu_j}) \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk})\, P(E_{jk} \,|\, E_j), \tag{5}$$

equation (4) can be written as

$$P(\zeta > \zeta_i) = 1 \,-\, \prod_{j=1}^{J} (1 - \bar{\mu}_{ij}) \tag{6}$$

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

$$P(\zeta > \zeta_i) \approx 1 \,-\, \prod_{j=1}^{J} e^{-\bar{\mu}_{ij}}. \tag{7}$$

If we again use the discussion in Section 1.2 to approximate $\bar{\mu}_{ij}$ in equation (5) by μ_{ij}, where

$$\mu_{ij} = \nu_j \sum_{k=1}^{k_j} P(\zeta > \zeta_i \,|\, E_{jk})\, P(E_{jk} \,|\, E_j), \tag{8}$$

we arrive at the expression for $P(\zeta > \zeta_i)$ that was used by González,et.al. [3] in the Seaside, Oregon study. That is,

$$P(\zeta > \zeta_i) \approx 1 \,-\, \prod_{j=1}^{J} e^{-\mu_{ij}}. \tag{9}$$

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 point of interest. We show how this is done in Section 2 when tidal uncertainty is ignored and three independent events with one realization each are used. That is, $k_j = 1$, for $j = 1 \ldots 3$, $E_{j1} = E_j$, $P(E_{j1} \,|\, E_j) = 1$, and $P(\zeta > \zeta_i \,|\, E_j)$ in equation (3) is either 1 or 0, depending on whether or not the inundation was above level ζ_i. With these assumptions, equation (3) is used to find the exceedance probability either using $P(E_j) = \nu_j$ or $P(E_j) = 1 - e^{-\nu_j}$. In Section 2, $P(E_j)$ is also denoted by p_j.

When tidal uncertainty is included, $P(\zeta > \zeta_i \,|\, E_{jk})$ in equation (8) is first found using one of the methods in Section 3, and then the known conditional probabilities $P(E_{jk} \,|\, E_j)$ and equation (8) are used to calculate the μ_{ij} that is used in (9) to find $P(\zeta > \zeta_i)$ for the hazard curve.

2 Generating hazard curves ignoring tidal variation

A key step in our probabilistic approach to producing hazard maps is the generation of a *hazard curve* at each point on a fine grid covering Crescent City and the surrounding area. 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 [3]. The hazard curve for inundation at a fixed (x, y) location (longitude and latitude) shows maximum depth of inundation ζ on the horizontal axis and probability of exceeding this value on the vertical axis. This is a cumulative probability function, but the fact that we consider the probability of *exceeding* each value means the probability does not generally approach 1 as $\zeta \to 0$. Instead, the value $P(\zeta > 0; x, y)$ is the probability of having any flooding $(\zeta > 0)$ at this point (x, y).

A sample hazard curve is shown in Figure 2 for the point $(235.80719, 41.75391)$, the location of Gauge 105 in Figure 1. This simple example was created by assuming that the only possible events are the first 3 Alaska Aleutian Subduction Zone (AASZ) characteristic tsunamis discussed in the recent report by González, LeVeque, Adams [4], and using the recurrence times from Table 1. We have also ignored tidal uncertainty to begin with and only consider inundation computed when the simulation is run at Mean Higher High Water (MHHW). The incorporation of tidal uncertainty is discussed in the next section.

By examining the synthetic tide gauge records at Gauge 105 for each of the three AASZ sources, the maximum depth of water, $\hat{\zeta}$, at this point can be determined for each scenario, with the maximum depths and annual rate of occurrence as given in Table 1.

Table 1: Three distinct AASZ events with the depth they inundate at one fixed (x, y) point.

j	E_j	**Max Inundation $\hat{\zeta}_j$ (m.)**	$P(E_j) = p_j \approx \nu_j$
1	AASZe01	$\hat{\zeta}_1 = 1.9$	$\nu_1 = 1/1313 = 0.000762$
2	AASZe02	$\hat{\zeta}_2 = 1.3$	$\nu_2 = 1/750 = 0.00133$
3	AASZe03	$\hat{\zeta}_3 = 4.0$	$\nu_3 = 1/750 = 0.00133$

Ignoring tides $(\zeta = z)$, the conditional probabilities $P(\zeta > \zeta_i \,|\, E_j)$ from equation (3) will either be 1 or 0, depending on whether the depth inundated above ζ_i or not. Since the recurrence times ν_j, $j = 1 \ldots 3$ are small, we estimate $P(E_j) = p_j = 1 - e^{-\nu_j}$ by ν_j. Then the probabilities $P(\zeta > \zeta_i)$ for different levels ζ_i can be computed from equation (3) as described below, and used to make a hazard curve.

We see from this data that there is 0 probability of exceeding depth $\zeta = 4$, while for any value of ζ between 1.9 and 4, the probability of exceedance is $p_3 = 0.00133$, since only one event AASZe03 inundates to this level.

On the other hand there are two events (AASZe03 and AASZe01) that exceed 1.3 m, and so for any ζ between 1.3 and 1.9 the probability of exceedance must be computed taking into account that either event might occur.

If these two events are independent with annual probabilities p_i and p_j then the probability of at least one event happening is

$$p_{ij} = 1 - (1 - p_i)(1 - p_j) = p_i + p_j - p_i p_j \approx p_i + p_j.$$

Hence the probability of event 1 or 3 happening is $p_{13} \approx 0.0021$, and this is the value displayed on the hazard curve for $1.3 < \zeta < 1.9$. Note that if $p_i = 1 - e^{-\nu_i}$ and $p_j = 1 - e^{-\nu_j}$ then $p_{ij} = 1 - e^{-\nu_i} e^{-\nu_j} = 1 - e^{-(\nu_i + \nu_j)}$, so the combined probability can also be computing by adding the Poisson rates $\nu_i = 1/T_i$.

Similarly, for $0 < \zeta < 1.3$ there are three possible events (AASZe03, AASZe01, or AASZe02) that give this level of inundation and so for any ζ in this range the probability of exceedance is $p_{134} = 0.0034$, computed via

$$p_{ijk} = 1 - (1 - p_i)(1 - p_j)(1 - p_k) = p_{ij} + p_k - p_{ij} p_k \approx p_i + p_j + p_k.$$

The latter approximation is valid if all probabilities are much less than 1. Similar formulas hold when more than 3 events are considered.

With only three possible discrete events to consider, the hazard curve is piecewise constant with jump discontinuities at the values $\zeta = \hat{\zeta}_2,\ \hat{\zeta}_1,\ \hat{\zeta}_3$ corresponding to the maximum inundation observed for each event. The magnitude of each jump is approximately equal to the probability of the corresponding event, as long as the sum of all probabilities of larger events is much less than 1.

Note that if an additional event were added that gave maximum inundation $\hat{\zeta}_4$ at the (x, y) point being studied, then a new jump discontinuity would be added to the hazard curve at the point $\hat{\zeta}_4$, with the portion of the curve to the left of $\hat{\zeta}_4$ shifted upwards by approximately p_4, the probability of occurance of this event. (More exactly by $1 - (1 - p_4) \prod_k (1 - p_k)$ where the product is over all events with inundation $\hat{\zeta}_k > \hat{\zeta}_4$.)

Note also that if there is uncertainty in the exact details of the slip pattern for one of these events, then we might replace the single realization, say AASZe01, by N slightly different realizations, which might give a range of inundations near $\hat{\zeta}_1$. If we assigned each a probability p_1/N, for example, then the hazard curve would be unchanged except in the vicinity of $\hat{\zeta}_1$, where the discontinuity of magnitude $p_{13} - p_3$ would be replaced by N discontinuities each with magnitude roughly $1/N$ as large, distributed near $\hat{\zeta}_1$ at points corresponding to the maximum inundation of each of these N realizations. If we considered a continuous distribution of possible realizations of AASZe01, then this would generally tend to smooth out the discontinuity into a continuous curve between the minimum and maximum inundations observed for different realizations, with the total drop in the exceedance probability over this interval remaining constant at $p_{13} - p_3$. If we do this for each of the events shown in Table 1 and Figure 2, we might get a hazard curve similar to what is shown in Figure 3. As explained in the next section, a similar smoothing of the hazard curve is observed when tidal uncertainty is incorporated, since this also increases the range of inundation values ζ that can be observed for each event.

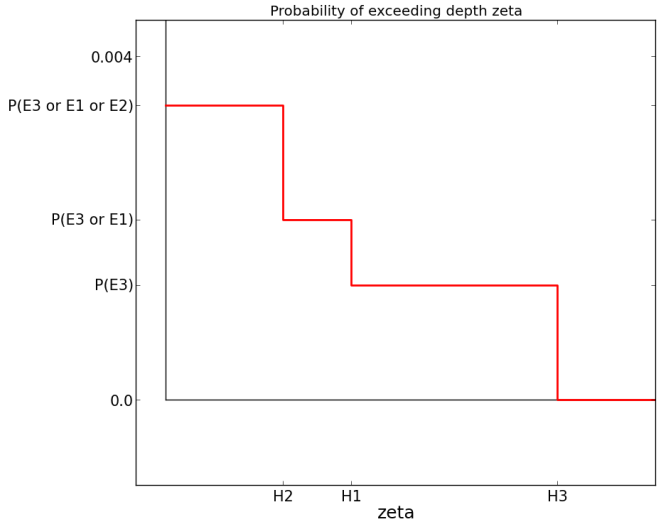


Figure 2: Sample hazard curve for a fixed (x, y) point when three distinct events are considered that give inundation to three depths. Using the data from Table 1.

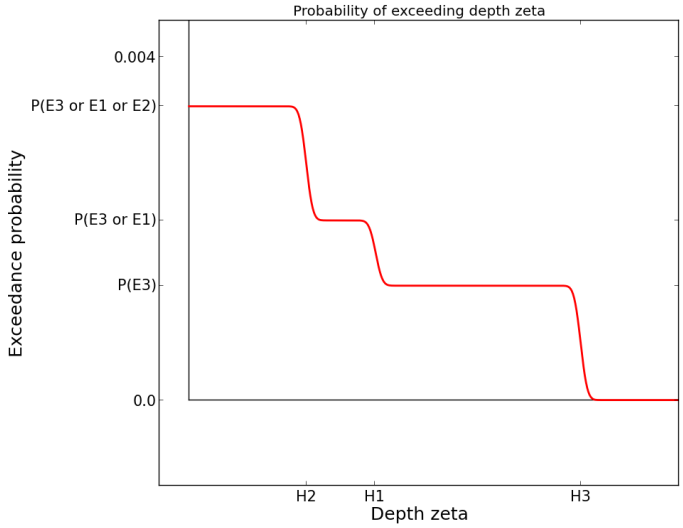


Figure 3: Sample hazard curve as in Figure 2, but with uncertainty in the amplitude added. The probability of each event is the same as before, but the extent of inundation may vary between different realizations, leading to a spreading of the jump discontinuity.

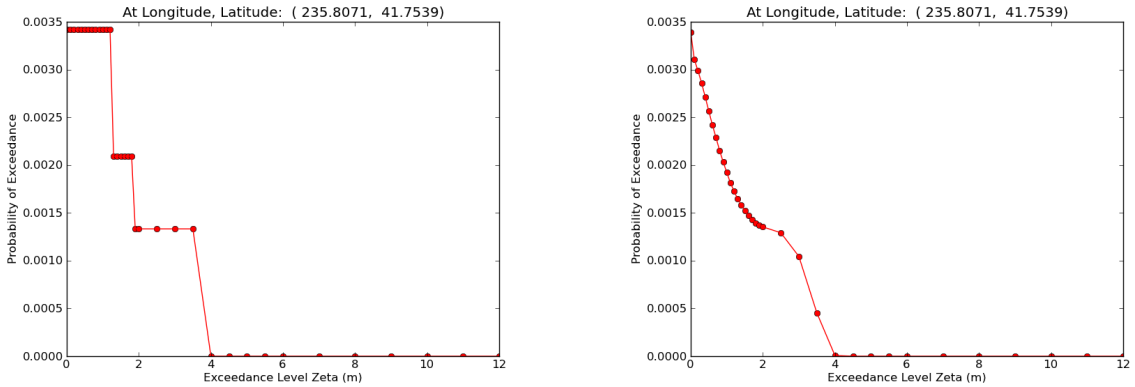


Figure 4: Sample hazard curve for a fixed (x, y) point when three distinct events are considered that give inundation to three depths. Left: Using the data from Table 1, as in Figure 2 but illustrating the use of a finite number of exceedance values ζ_k. Right: Corresponding curve for the same three events but when tidal uncertainty is included.

In practice we do not attempt to compute the hazard curve probabilities for all values of ζ at each (x, y). Instead we choose a finite set of exceedance values ζ_k and determine the probability of exceeding each ζ_k. We then approximate the hazard curve by a piecewise linear function that interpolates these values $(\zeta_k,\ P(\zeta > \zeta_k; x, y))$. The left plot in Figure 4 shows this approximation for the previous example. We do this because computing each value $P(\zeta > \zeta_k; x, y)$ requires combining information from all simulation runs together with tidal variation, as described in the next section, and is somewhat costly to perform. By choosing a finite set of ζ_k values the postprocessing is also simplified.

We have chosen the following exceedance values:

$$\begin{aligned} \zeta_k = 0,\ 0.1,\ 0.2,\ \ldots,\ 1.9,\ 2.0, \\ 2.5,\ 3.0,\ \ldots,\ 5.5,\ 6.0, \\ 7.0,\ ,8.0,\ \ldots,\ 12.0, \end{aligned} \tag{10}$$

which we believe is sufficiently dense to yield good approximations of the hazard curves in general. The curve determined from considering only the three events discussed above is shown in Figure 4.

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.

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 of seeing any flooding. The

$p = 0.01$ contour would again correspond to the inundation limit of the "100-year event".

When tidal uncertainty is included $P(\zeta > \zeta_i \,|\, E_{jk})$ will no longer be 1 or 0 depending on whether realization k of event E_j inundated beyond level ζ_i. In Section 3 we describe methods for calculating these probabilities including the effect of the tides. The probability $P(\zeta > \zeta_i)$ needed for the hazard curve is then gotten from equations (8) and (9).

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

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

$$P(\zeta > \zeta_i \,|\, E_{jk}). \tag{11}$$

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 the GeoClaw code is not modeling the tidal dynamics (i.e. the rise or fall of the tide within a single simulation, or how tidal currents could affect the inflow of the tsunami wave.) This limitation might be important to address and is a future research topic. In this section, we present new methods for incorporating the tidal effects as best we can given this limitation.

The GeoClaw information needs to be combined with tidal information at Crescent City to determine the probability in equation (11). We implemented six different methods for doing this and determined their relative merits. The six methods are referred to as the dt-Method, the Pattern-Method, the G-Method, the G2-Method, the G3-Method, and the G4-Method. The method developed by Mofjeld, et.al. [7] and used by González, et.al. [3] for the Seaside, Oregon study, will be referrred to as the G-Method. The G-Method is briefly described in Section 3.5 to fix ideas and to facilitate the understanding of the G2-Method as described in Section 3.6. The G3-Method is described in Section 3.7, and the G4-Method is described in Section 3.8. All six methods are compared in detail in Section 4. Results show that the dt-Method and the Pattern-Method give quite similar results for a properly chosen dt, that the G2-, G3-, and G4- Methods are successive improvements over the G-Method, and that 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 points.

The key ideas in the dt-Method, see Section 3.2, and the Pattern-Method, see Section 3.3 are summarized below:

- A tsunami wave that arrives at high tide will cause more flooding than the same wave arriving at low tide. But 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 3.2 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 3.3.

- 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.

We begin by giving the Crescent City tide gauge information and the probability density function and cumulative distribution function that we computed from this gauge data.

3.1 Crescent City tides

The tide gauge at Crescent City (Gauge No. 9419750) has the values shown in Table 2 for Mean Low Low Water (ξ_{MLLW}), Mean Low Water (ξ_{MLW}), Mean Sea Level (ξ_{MSL}), Mean High Water (ξ_{MHW}), and Mean High High Water (ξ_{MHHW}), respectively. In addition, we include the Lowest (ξ_{Lowest}) and Highest ($\xi_{Highest}$) water seen at the gauge in a year's data from July 2011 to July 2012. Unless explicitly stated, the tide levels we use are referenced to MSL.

Table 2: **Crescent City tide values**

Level	**Referenced to MSL**
ξ_{Lowest}	-1.83
ξ_{MLLW}	-1.13
ξ_{MLW}	-0.75
ξ_{MSL}	0.00
ξ_{MHW}	0.77
ξ_{MHHW}	0.97
$\xi_{Highest}$	1.50

A fixed number of bins is made from tide levels $-1.83 \leq \xi \leq 1.50$. Then the yearly tide data at Crescent City is associated with the appropriate bin to make the probability density function and associated cumulative distribution function shown in Figure 5. 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. The integration of the PDF function from a particular tidal level, $\xi = \tilde{\xi}$ to infinity, yields $P(\xi > \tilde{\xi})$, the value of the Cumulative Distribution Function at $\xi = \tilde{\xi}$.

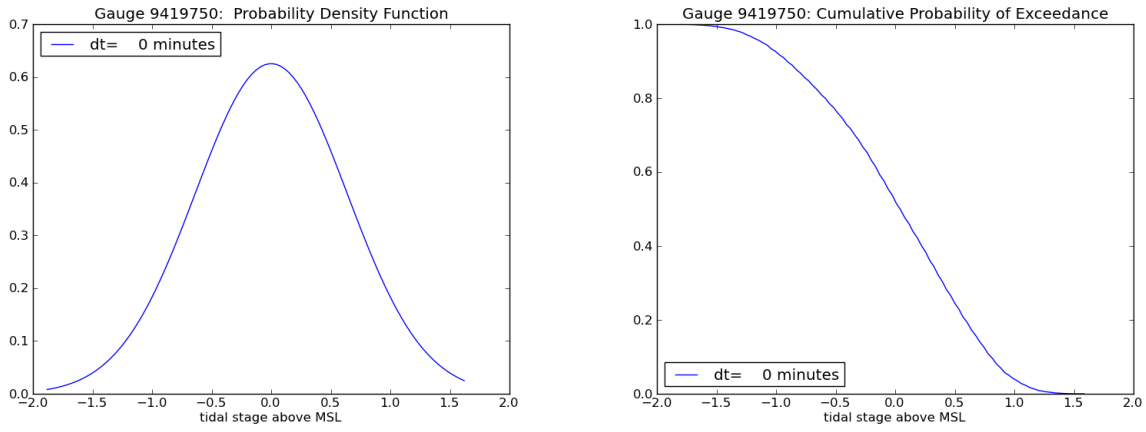


Figure 5: Crescent City Tidal Distributions Left: Probability Density Function (mean=0.0, σ = .638) Right: Cumulative Distribution Function

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. Since we need to model this uncertainty, we make use of the GeoClaw feature that permits simulations at any tide level. This gives additional information about the z acheived at different *static* tidal levels. We run each realization using at least three tide levels (usually Mean Low Water (MLW), Mean Sea Level (MSL) and Mean High High Water (MHHW)). We also use the famous Alaska 1964 event for intense analysis. For this event (one realization), we actually use 11 tide levels, ranging from Mean Low Low Water (MLLW) up to the maximum tide level seen at the Cresent City tide gauge.

3.2 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 upper plot of Figure 6 below. (This is illustrated for a case where 7 tide levels $\hat{\xi}_m$ were run.) 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.

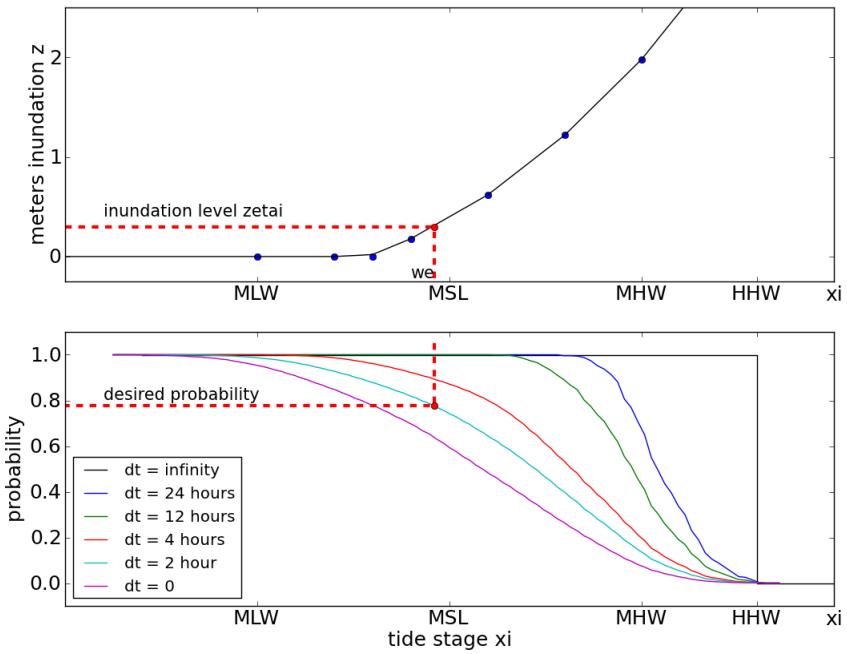


Figure 6: Finding $\hat{\xi} = w_e$ and $P(\xi > w_e \,|\, \mathrm{dt} = 2)$

The upper graph in Figure 6 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 as given below:

- If $\zeta = 0$ for all the data points:
  - Left and right extension: Use slope=0
- Else:
  - Left extension: Use slope=min(1, slope of first data segment)
  - Right extension: Use slope=max(1, slope of last data segment)

If the horizontal dashed line in the upper plot of Figure 6 (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 the conditional probability is taken as 1 for exceeding ζ_i ($P(\zeta > \zeta_i \,|\, E_{jk}) = 1$). Likewise, if the ζ_i inundation dashed line is above the entire graph, then ζ_i has probability of 0 of being exceeded and this conditional probablility is set to 0. If w_e is greater than or equal to the highest tide possible at the Crescent City gauge, ζ_i can never be exceeded and the conditional probability is set to 0. Finally, if w_e is less than the lowest tide possible at the Crescent City gauge, ζ_i is always exceeded and the conditional probability is set to 1.

Now, suppose for the moment that the GeoClaw tsunami of Figure 6 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 5 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

$$P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \mathrm{dt} = 0). \tag{12}$$

The cumulative probability distribution from Figure 5 is also shown as the curve labelled (dt=0) in the bottom plot in Figure 6. 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 6 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 $P(\xi > w_e \,|\, \mathrm{dt} = .25)$. Then, we have

$$P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \mathrm{dt} = .25)\,. \tag{13}$$

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 6 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 two 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 3.3, we see the dt-Method works remarkably well compared to the Pattern-Method for appropriately chosen dt. In particular, for all tsunamis in Table 5 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.

3.3 The Pattern-Method

This approach grew from the desire to automate the choice of dt in the dt-Method. Instead of achieving this automation of the dt-Method, 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. Our experience 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 and AASZe08, Figures 7 and 10), while for other tsunamis the probabilities are consistent with a varying dt (AASZe02 and KrSZe13, Figures 8 and 9). For some tsunamis, dt appears to have varied from as high as 4 down to 0 (KrSZe13, Figure 9), or from above 4 down to zero (AASZe02, Figure 8) as the tide level increases, while for other tsunamis such as the 1964 Alaska event (AASZe03, Figure 7), dt=1 closely approximates the pattern cumulative distribution. Choosing a fixed dt between 1 and 2 closely approximates the AASZe08 event as seen in Figure 10. To illustrate these findings, the figures below show the pattern cumulative distribution as a dotted line on the same graph as that for the dt-Method with varying dt.

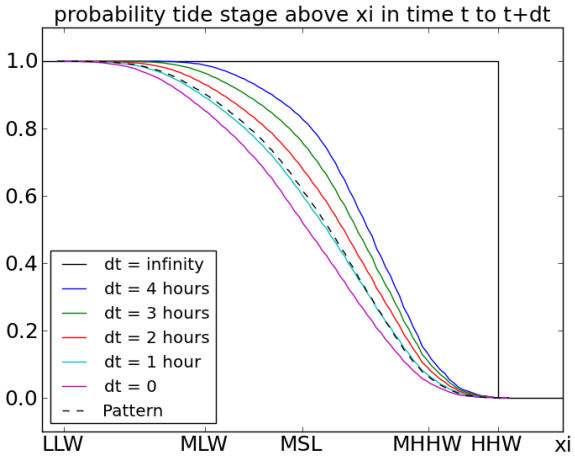


Figure 7: AASZe03: Pattern-Method to dt-Method Cumulative Distribution Comparison

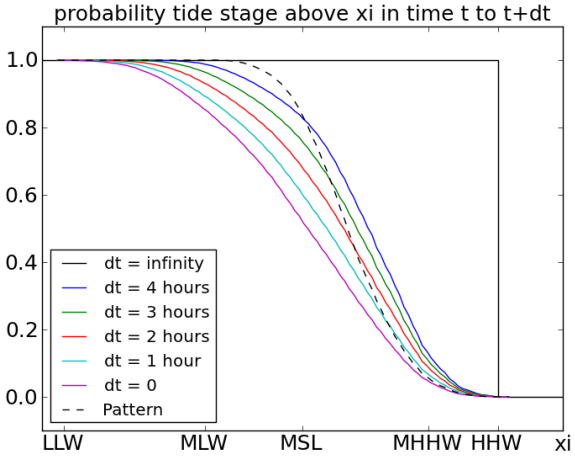


Figure 8: AASZe02: Pattern-Method to dt-Method Cumulative Distribution Comparison

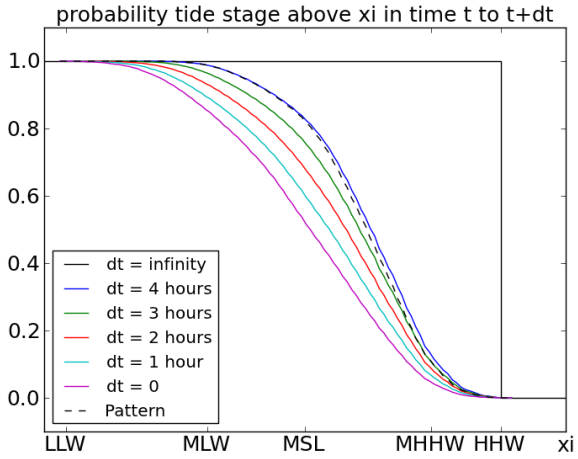


Figure 9: KrSZe13: Pattern-Method to dt-Method Cumulative Distribution Comparison

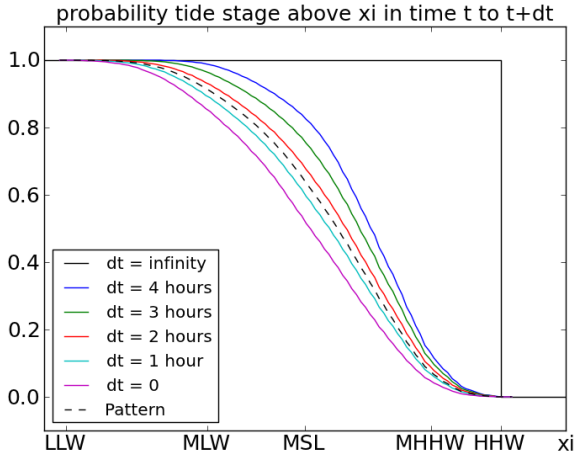


Figure 10: AASZe08: Pattern-Method to dt-Method Cumulative Distribution Comparison

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 Δ_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 terminating time of the K-th wave.

For example, if we have three waves, and the highest wave is W_2, the first wave is .5 meters lower than W_2 and the third wave is .8 meters lower than W_2, we would record $\Delta_1 = .5$, $\Delta_2 = 0$, and $\Delta_3 = .8$. If the first wave has a width of 15 minutes, the second a width of 20 minutes, and the third a width of 30 minutes, and the second wave appears 27 minutes after the first has completed, and the third appears 1 hour after the second has completed, we would record I_1=[0, 15], I_2=[42, 62], and I_3=[122,152]. The length of the pattern is $T_K = 152$ minutes.

In Figure 11, we show the GeoClaw tsunami for the AASZe02 event that was recorded at gauge 101 as the red graph and the pattern shown 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 3. 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 11.

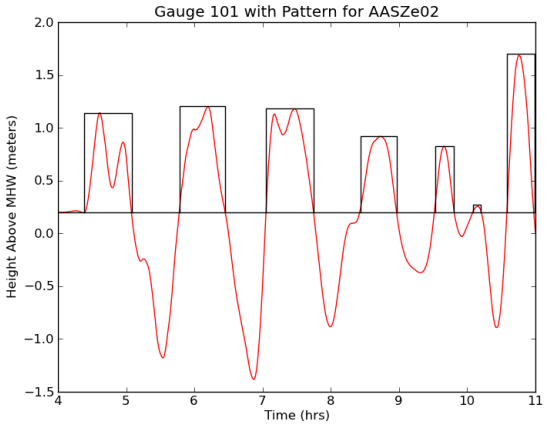


Figure 11: Pattern for AASZe02 at Gauge 101

Table 3: **Pattern Values for AASZe02**

Wave	**I=[S, T] (min)**	**Δ (m)**
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 $\bar{M}_k$: $\bar{M}_k = M_k - \Delta_k$.
- Compute $M_P = \max_k \bar{M}_k$.
- 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 the 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 for the method but is computed so comparisions can be made with the G method.

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 only sees the effect of the tsunami after nonlinear effects have been incorporated.

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 $P(\xi > w_e \,|\, \text{pattern})$ with the meaning

$$P(\xi > w_e \,|\, \text{pattern}) = P(\xi > w_e + \Delta_k \,\text{somewhere in}\, I_k \,\text{for some}\, k)\,. \tag{14}$$

Then, we have

$$P(\zeta > \zeta_i \,|\, E_{jk}) = P(\xi > w_e \,|\, \text{pattern})\,. \tag{15}$$

Advantages over the dt-Method

- By adjusting the M_k, we permit the possibility that a wave with amplitude Δ_k less than the maximum one seen at GeoClaw gauge 101 could also cause an inundation at gauge 101 of maximum z or higher if it occurred at a time when the tide level was at least Δ_k higher than that required of the maximum amplitude GeoClaw gauge 101 wave.
- By looking in each interval I_k, we take into account the actual width of each wave.
- We can ignore what the tide is doing when nothing is happening, and don't have to have a large dt to cover all the action. 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.
- Only gauge 101 needs to be examined.
- The procedure is automatic. Examining gauge 101 and running the code to generate the cumulative distribution is much faster than examining multiple gauges to make an informed decision about the choice of dt.

Possible limitations

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.

3.4 Mathematical description

In this section, we give a careful mathematical description of including tidal uncertainty that can be used to describe both the Pattern and dt-Methods and the G, G2, G3, and G4 methods of the next sections. We start by introducing the following notation:

- t is the time in minutes since the tsunami struck Crescent City's Gauge 101. It strikes Gauge 101 at time $t = 0$ and location P at time t_P.
- τ is the time in the Crescent City yearly tidal minute-record when the hypothetical tsunami strikes Gauge 101. T is the duration of the tsunami in minutes (assumed the same at all locations P).

- B is the original topography or bathmetry. If $B > 0$, the point is on land and B is the original height above ξ_{MHW}. If $B < 0$, the point is in the water and $-B$ its original distance below ξ_{MHW}. $\tilde{B}$ is the topography or bathmetry after subsidence (or uplift). $B - \tilde{B}$ is positive with subsidence and negative with uplift.

- $\xi(\tau + t)$ is the tidal stage at time t minutes after the tsunami struck Gauge 101 at time τ of the tidal record when the tidal stage was $\xi(\tau)$. This is assumed to be the same for all locations P. $\xi(\tau + t) - (B - \tilde{B})$ is the sealevel after subsidence (uplift) at time t.

- $d_P(t, \tau + t)$ is the flow depth above $\tilde{B}$ at time t corresponding to time $\tau + t$ of the tidal record at location P. $\zeta_P(t, \tau + t)$ is $d_P(t, \tau + t)$ for land and $d_P(t, \tau + t) + B$ for water points.

- $d_P(t, \tau + t) + \tilde{B} + \xi_{MHW}$ is the amount of the flow depth that is above ξ_{MSL}, and $d_P(t, \tau + t) + \tilde{B} + \xi_{MHW} - (\xi(t + \tau) - (B - \tilde{B}))$ is the amount that is above the new subsided (or uplifted) sealevel (tidal stage) which we will call the *amplitude*. Simplifying, we see the *amplitude* at location P is

$$d_P(t, \tau + t) + B + \xi_{MHW} - \xi(t + \tau). \tag{16}$$

  For water points, this amplitude is $\zeta_P(t, \tau + t) + \xi_{MHW} - \xi(t + \tau)$, and for land points it is $\zeta_P(t, \tau + t) + B + \xi_{MHW} - \xi(t + \tau)$.

- A GeoClaw simulation with sealevel set to $\hat{\xi}$ assumes a constant tidal record, $\tilde{\xi}(\tau) = \hat{\xi}$ for all τ, and produces a flow depth denoted by $h_P(t, \tau + t)$ instead of $d_P(t, \tau + t)$. The notation $z_P(t, \tau + t)$ instead of $\zeta_P(t, \tau + t)$ will mean $h_P(t, \tau + t)$ for land points and $h_P(t, \tau + t) + B$ for water points. We define $z_P(\hat{\xi}) = \max_{t_p \le t \le T + t_p} z_P(t, \tau + t)$, noting $\hat{\xi}$ is the static tide level for all $\tau + t$ in the GeoClaw simulation.

- We will refer to the upper graph in Figure 6 as a GeoClaw simulation curve. The points on this curve are $(\hat{\xi}, z_P(\hat{\xi}))$ associated with *static* GeoClaw runs at various tide levels $\hat{\xi}$. $A_P(t; \hat{\xi})$ is the amplitude of the GeoClaw tsunami at location P at time t minutes, which is the height of water above the subsided (or uplifted) initial background sealevel $\hat{\xi}$ of the GeoClaw simulation and from equation (16) is given by

$$A_P(t; \hat{\xi}) = h_P(t, \tau + t) + B + \xi_{MHW} - \hat{\xi}. \tag{17}$$

  We define $A_P(\hat{\xi}) = \max_{t_P \le t \le T + t_P} A_P(t; \hat{\xi})$.

Water points:

We first assume it is valid to equate the amplitudes in equations (16) and (17). This leads to $\zeta_P(t, \tau + t) + \xi_{MHW}$ being approximated by the original tidal stage (before subsidence or uplift) $\xi(\tau + t)$ plus the GeoClaw amplitude $A_P(t; \hat{\xi})$,

$$\zeta_P(t, \tau + t) + \xi_{MHW} \approx \xi(\tau + t) + A_P(t; \hat{\xi}), \tag{18}$$

and subtracting $A_P(\hat{\xi})$ from both sides gives

$$\zeta_P(t, \tau + t) + \xi_{MHW} - A_P(\hat{\xi}) \approx \xi(\tau + t) + A_P(t; \hat{\xi}) - A_P(\hat{\xi}) \,. \tag{19}$$

Taking the maximum of both sides of (19) over the duration of the tsunami gives

$$\zeta_P(\tau) + \xi_{MHW} - A_P(\hat{\xi}) \approx \bar{\xi}_P(\tau; \hat{\xi}) \tag{20}$$

where we have defined

$$\zeta_P(\tau) = \max_{t_P \le t \le T + t_P} \zeta_P(t, \tau + t) \tag{21}$$

and

$$\bar{\xi}_P(\tau; \hat{\xi}) = \max_{t_P \le t \le T + t_P} \Big(\xi(\tau + t) + A_P(t; \hat{\xi}) - A_P(\hat{\xi}) \Big) \,. \tag{22}$$

Equation (22) suggests that we can create a PDF at each location P for a random variable $\bar{\xi}_P(\hat{\xi})$ by sliding across the τ locations of the tidal record. For each τ, we simply increment a counter in the appropriate bin corresponding to the value in equation (22) and also in lower bins for the cumulative distribution function. Dividing by the number of slides gives the cumulative distribution function, and dividing by both the number of slides and the bin size gives the PDF. Equation (20) shows we can approximate random variable ζ_P by the relationship,

$$\zeta_P \approx A_P(\hat{\xi}) - \xi_{MHW} + \bar{\xi}_P(\hat{\xi}) \,. \tag{23}$$

This shows,

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(A_P(\hat{\xi}) - \xi_{MHW} + \bar{\xi}_P(\hat{\xi}) > \zeta_i) \tag{24}$$

which simplifies to

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(\bar{\xi}_P(\hat{\xi}) > \zeta_i - A_P(\hat{\xi}) + \xi_{MHW}) \,. \tag{25}$$

If we run GeoClaw at a particular static sealevel $\hat{\xi}$, then the maximum amplitude $A_P(\hat{\xi})$ for this simulation is

$$A_P(\hat{\xi}) = z_P(\hat{\xi}) + \xi_{MHW} - \hat{\xi} \tag{26}$$

and combining with equation (25), gives

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(\bar{\xi}_P(\hat{\xi}) > \zeta_i - z_P(\hat{\xi}) + \hat{\xi}) \tag{27}$$

and using $\zeta_i = z_P(w_e)$ gives

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(\bar{\xi}_P(\hat{\xi}) > z_P(w_e) - z_P(\hat{\xi}) + \hat{\xi}) \,. \tag{28}$$

Defining

$$\tilde{w}_e = z_P(w_e) - z_P(\hat{\xi}) + \hat{\xi} \tag{29}$$

gives

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(\bar{\xi}_P(\hat{\xi}) > \tilde{w}_e)\,. \tag{30}$$

It is clear that if the maximum amplitude $A_P(\hat{\xi})$ is computed from the static simulation with sealevel $\hat{\xi} = w_e$, then $\tilde{w}_e = w_e$. Also, if the GeoClaw curve at location P for $\hat{\xi}$ vs. $z_P(\hat{\xi})$ for multiple static simulations is linear with slope 1, then $\tilde{w}_e = w_e$; otherwise the values w_e and $\tilde{w}_e$ are different.

For a water point P, it is also likely the amplitude in (17) is independent of $\hat{\xi}$ and that $A_P(t;\hat{\xi}) = A_P(t)$ and $A_P(\hat{\xi}) = A_P$. This is because the GeoClaw flow depth $h_P(t, \tau + t)$ at time t is likely to be linear in the variable $\hat{\xi}$ with slope 1. This is important, since we can examine the time series $A_P(t)$ and its maximum A_P from only one GeoClaw run with any convenient $\hat{\xi}$, say $\hat{\xi} = \xi_{MHHW}$ to produce $\bar{\xi}_P(\tau)$ in equation (22) and the cumulative distribution for the random variable $\bar{\xi}_P$ in equation (23), which is now independent of $\hat{\xi}$.

Land points:

We also begin by equating the amplitudes in equations (16) and (17) to get

$$\zeta_P(t, \tau + t) + B + \xi_{MHW} \approx \xi(\tau + t) + A_P(t;\hat{\xi})\,. \tag{31}$$

Subtracting $A_P(\hat{\xi})$ from both sides gives

$$\zeta_P(t, \tau + t) + B + \xi_{MHW} - A_P(\hat{\xi}) \approx \xi(\tau + t) + A_P(t;\hat{\xi}) - A_P(\hat{\xi})\,. \tag{32}$$

Taking the maximum of both sides of (32) over the duration of the tsunami gives

$$\zeta_P(\tau) + B + \xi_{MHW} - A_P(\hat{\xi}) \approx \bar{\xi}_P(\tau;\hat{\xi}) \tag{33}$$

where again (21) and (22) apply. For a land point P, Equation (23) now becomes

$$\zeta_P \approx A_P(\hat{\xi}) - B - \xi_{MHW} + \bar{\xi}_P(\hat{\xi})\,, \tag{34}$$

and

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(A_P(\hat{\xi}) - B - \xi_{MHW} + \bar{\xi}_P(\hat{\xi}) > \zeta_i) \tag{35}$$

which simplifies to

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) \approx P(\bar{\xi}_P(\hat{\xi}) > \zeta_i - A_P(\hat{\xi}) + B + \xi_{MHW})\,. \tag{36}$$

If we run GeoClaw at a particular static sealevel $\hat{\xi}$ the maximum amplitude $A_P(\hat{\xi})$ for this simulation at land point P is

$$A_P(\hat{\xi}) = z_P(\hat{\xi}) + B + \xi_{MHW} - \hat{\xi} \tag{37}$$

and combining with equation (36) again gives equation (27) since B drops out, and using $\zeta_i = z_P(w_e)$ again gives equation (28). With the definition in (29), we get (30) and if the maximum amplitude $A_P(\hat{\xi})$ is computed from the static simulation with sealevel $\hat{\xi} = w_e$, then $\tilde{w}_e = w_e$. Also, if the GeoClaw curve at location P for $\hat{\xi}$ vs. $z_P(\hat{\xi})$ for multiple static simulations is linear with slope 1, then $\tilde{w}_e = w_e$; otherwise w_e and $\tilde{w}_e$ are different.

We also note that equations (23) and (34) show the random variable ζ_P at either water or land points is approximated by

$$\zeta_P \approx z_P(\hat{\xi}) - \hat{\xi} + \bar{\xi}_P(\hat{\xi})\,. \tag{38}$$

But now, for land points, we can not assume the GeoClaw amplitude is independent of $\hat{\xi}$ because the GeoClaw curve most likely will not have slope 1. Hence, the random variable $\bar{\xi}_P$ remains a function of $\hat{\xi}$, as does the right hand side of equation (38) which makes its use suspect for land points. Nevertheless, we will see in the next section that the G and G2 methods use this approach with $\hat{\xi} = \xi_{MHHW}$ and $\hat{\xi} = w_e$, respectively.

Simplifications:

For a computational grid of 500,000 grid points convering Crescent City and its harbor, we can not afford to make a PDF (and associated cumulative distribution) for each of these points as indicated by the subscripts of P in the previous discussion. This would require saving the time series for $A_P(t)$ at each of these points. There are two basic approaches to simplification that are used in the methods we developed and examined as described below.

1. Approximate the PDF for either the random variable $\bar{\xi}_P(\hat{\xi})$ (or equivalently the random variable ζ_P from equation (38)) by a normal distribution with appropriate mean and standard deviation, using the maximum amplitude $A_P(\hat{\xi})$ from a GeoClaw simulation at point P. This approach is taken by the G-Method and the G2-Method described in Sections 3.5 and 3.6, respectively. This is a cheap solution since it involves evaluating a single `erf` function using the known $A_P(\hat{\xi})$ value at each P.

2. Use point P's GeoClaw simulation curve to find w_e^P so that $\zeta_i = z_P(w_e^P)$ is exceeded on this curve whenever w_e^P is exceeded. Then use equation (22) only at the point corresponding to Gauge 101 in the water to create a PDF and cumulative distribution for $\bar{\xi}_{101}$ which is assumed independent of $\hat{\xi}$. The point on this cumulative graph corresponding to w_e^P would then give $P(\zeta_{101} > z_{101}(w_e^P) \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e^P)$, the probability that the maximum tide seen during the duration of the tsunami at Gauge 101 exceeds w_e^P (in the sense that this maximum tide would produce exceedance of the static Geoclaw value $z_{101}(w_e^P)$). Make the approximation that $\zeta_i = z_P(w_e^P)$ will be exceeded at point P with the same probability as $z_{101}(w_e^P)$ is exceeded at Gauge 101. We then set $P(\zeta_P > \zeta_i \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e^P)$ for both water and land points. By using this simplification, we only need to find one cumulative distribution from one time series (either by analytical estimation or by binning up data), and use this distribution to look up w_e^P associated with the ζ_i for point P. For a particular ζ_i, w_e^P varies across the locations P, but the distribution being used is the same. This approach avoids using an approximation formula for land points that involves an amplitude calculated on land where the GeoClaw simulation curve clearly does not have slope 1, and does not make

the assumption that tides and amplitudes add at any point other than at Gauge 101. This simplification is used by the dt-Method, the Pattern-Method, the G3-Method, and the G4-Method, and for these methods A_{101} is computed from the $A_{101}(t)$ times series of the sealevel $\hat{\xi} = \xi_{MHHW}$ simulation.

dt-Method Revisited

The dt-Method is modelling the GeoClaw tsunami as a contiguous square wave where $A_{101}(t) = A_{101}$ over a duration of `dt` minutes (starting at time $t = t_s$ when the largest waves start) during which the maximum value $\bar{\xi}_{101}(\tau)$ is

$$\max_{0\le t\le T}(\xi(\tau+t) + A_{101}(t) - A_{101}) = \max_{t_s\le t\le t_s+\texttt{dt}}(\xi(\tau+t)). \tag{39}$$

Hence,

$$P(\zeta_{101} > z_{101}(w_e^P) \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e^P) = P(\xi > w_e^P \,|\, \texttt{dt}) \tag{40}$$

where $P(\xi > w_e^P \,|\, \texttt{dt})$ is described for the dt-Method in Section 3.2. We also note that if the tsunami time series at any other water point P is also modelled as a square wave $A_P(t) = A_P$ of duration `dt`, then $\bar{\xi}_P = \bar{\xi}_{101}$ and we also get

$$P(\zeta_P > \zeta_i = z_P(w_e^P) \,|\, E_{jk}) = P(\bar{\xi}_P > w_e^P) = P(\bar{\xi}_{101} > w_e^P) = P(\xi > w_e^P \,|\, \texttt{dt}) \tag{41}$$

which is the same probability gotten by adopting the second simplification described above. When P is a land point, accepting the second simplification also leads to equation (41).

Pattern-Method Revisited

The Pattern-Method is modelling the GeoClaw tsunami at Gauge 101 by a piecewise constant function $A_{101}(t)$ that has K different square waves. Wave k is over time interval I_k and over this interval has an amplitude that is Δ_k less than the maximum amplitude A_{101}. Hence, the maximum value $\bar{\xi}_{101}(\tau)$ is

$$\max_{0\le t\le T}(\xi(\tau+t) + A_{101}(t) - A_{101}) = \max_k \max_{t\in I_k}(\xi(\tau+t) - \Delta_k), \tag{42}$$

and we see the cumulative distribution for $\bar{\xi}_{101}$ is that for the Pattern-method,

$$P(\bar{\xi}_{101} > w_e^P) = P(\xi > w_e^P \,|\, \texttt{pattern}) \tag{43}$$

as described in equations (14) and (15). For other water points P, we can expect $\bar{\xi}_P$ to be closely approximated by $\bar{\xi}_{101}$, but can not claim their cumulative distributions are equal as in the dt-Method. Hence for other water points and all land points, we simply accept the second simplication and choose

$$P(\zeta_P > \zeta_i = z_P(w_e^P) \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e^P) = P(\xi > w_e^P \,|\, \texttt{pattern}). \tag{44}$$

3.5 The G-Method

This method as described in Mofjeld, et.al. [7] was later used to incorporate tidal uncertainty in the Seaside study by Gonzalez, et.al. [3]. 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 [1], Rabinovich, et. al. [8], and Fine et. al. [2] for examples. 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, $\hat{\xi}$, either by using equation (26) or (37) for water or land points, respectively. This theoretical tsunami was then combined with local tidal information to develop the Gaussian probability density function and ultimately a cumulative distribution approximation described by the **erf** function. Mofjeld, et.al. [7] gives parameters for this method for a variety of locations. The parameters for Crescent City are:

- σ_0=0.638 is the standard deviation for the Crescent City tides.
- $\alpha' = 0.056$, $\beta' = 1.119$, C'=.707, α=0.17, β=.858, and C=1.044 are the regression parameters for the G-Method given in [7].

Mofjeld, et.al. [7] gave a formula for the standard deviation, σ, of the random variable ζ_P (and hence the random variable $\bar{\xi}_P(\hat{\xi})$) in equation (38) as a function of the amplitude at location P ($A_G = A_P(\hat{\xi})$) and the Crescent City parameters,

$$\sigma = \sigma_0 \left(1 - C' \, e^{-\alpha' \left(\frac{A_G}{\sigma_0} \right)^{\beta'}} \right) \tag{45}$$

and the mean of ζ_P, denoted ζ_0, as

$$\zeta_0 = z_P(\hat{\xi}) - \hat{\xi} + C(\xi_{MHHW}) \, e^{-\alpha \left(\frac{A_G}{\sigma_0} \right)^{\beta}}, \tag{46}$$

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

$$w_0 = C(\xi_{MHHW}) \, e^{-\alpha \left(\frac{A_G}{\sigma_0} \right)^{\beta}}. \tag{47}$$

We verified these six regression parameters are reasonable by performing our own regression analysis using our yearly tidal data. We assumed the form for σ in (45) and w_0 in (47) above and ran the Pattern method on ten 5-day proxy tsumanis with A_G taken as $\frac{1}{3}\sigma_0$, σ_0, $3\sigma_0$, $6\sigma_0$, $9\sigma_0$, $12\sigma_0$, $15\sigma_0$, $18\sigma_0$, $21\sigma_0$, and $30\sigma_0$. The results given by Matlab's **fminsearch** routine are close to those of [7] given above:

- $\alpha' = 0.0429$, $\beta' = 1.2189$, C'=0.7021
- α=0.1732, β=0.8538, and C=1.0646

It is interesting to note that the models above use the value of $C\xi_{MHHW}$ for w_0 when A_G approaches 0, and $(1-C')\sigma_0$ for σ. From equation (22), as $A(t)$ (and hence A) goes to zero, the mean of the random variable $\bar{\xi}_P(\hat{\xi})$ should be the mean of the dt-Method PDF when `dt=H` hours ($H = T/60$) denoted by $\mu(\texttt{dt-Method}\,|\,\texttt{dt=H})$. Likewise, the standard deviation approaches the standard deviation of the PDF for the dt-Method with `dt=H` as A approaches 0, denoted by $\sigma(\texttt{dt-Method}\,|\,\texttt{dt=H})$. Both these values can be calculated and compared to what the models above would give using the two different sets of regression parameters. We also give the model results when $A_G = 6.140125\,\sigma_0 = 3.9174$, the amplitude of the AASZe03 tsunami and compare these to the results of applying the Pattern method to the 5-day proxy tsunami.

- $\mu(\texttt{dt-Method}\,|\,\texttt{dt=H}) = 1.0803$. The approximation in [7] to this is $C\xi_{MHHW} = 1.0127$ and with our set of regression parameters, $C\xi_{MHHW} = 1.0646$.
- $\sigma(\texttt{dt-Method}\,|\,\texttt{dt=H}) = 0.1884$. The approximation in [7] is $(1-C')\sigma_0 = 0.1869$ and with our set of regression parameters, $(1-C')\sigma_0 = 0.1901$.
- Applying the Pattern method to the 5-day proxy tsunami when $A_G = 3.9174$ gave $w_0 = 0.4583$ and $\sigma = 0.3430$. The respective values using the parameters in [7] are $w_0 = 0.4520$ and $\sigma = 0.3436$, and with our set of regression parameters are $w_0 = 0.4569$ and $\sigma = 0.3354$.

These results show the parameters determined in [7] are consistent with the yearly tidal data we are using and with our own analysis. Hence, these will be the parameters we will use for the G-method.

The probability $P(\zeta_P > \zeta_i \,|\, E_{jk})$ is then calculated by the following formula given originally in Mofjeld [7] and used by González, et. al. [3] in the Seaside, Oregon study:

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\zeta_i - \zeta_0}{\sqrt{2}\,\sigma}\right)\right) \tag{48}$$

Equation (48) can also be written as

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\tilde{w}_e - C\,\xi_{MHHW}\,e^{-\alpha\left(\frac{A_G}{\sigma_0}\right)^{\beta}}}{\sqrt{2}\,\sigma}\right)\right) \tag{49}$$

where $\tilde{w}_e = \zeta_i - z_P(\xi_{MHHW}) + \xi_{MHHW}$. This is seen to be $P(\bar{\xi}_P(\xi_{MHHW}) > \tilde{w}_e)$, since from Section 3.4, $\zeta_P = \bar{\xi}_P(\hat{\xi}) + z_P(\hat{\xi}) - \hat{\xi}$ and for the G-Method, we take $\hat{\xi} = \xi_{MHHW}$ and $z_P(\hat{\xi}) = z_P(\xi_{MHHW})$. Since $\zeta_i = z_P(w_e)$, if the GeoClaw simulation curve of $\hat{\xi}$ vs $z_P(\hat{\xi})$ has slope 1, then $\tilde{w}_e = w_e$; otherwise w_e and $\tilde{w}_e$ are different.

3.6 The G2-Method

Motivation

When we implemented the G-Method, one GeoClaw simulation with tide level $\hat{\xi} = \xi_{MHHW}$ was used to produce $z_P(\xi_{MHHW})$, and these values were used to compute A_G in equations (26) and (37). The 2 in the G2-Method emphasizes that GeoClaw simulations with at least 2 different tide levels are used to calculate a more accurate amplitude.

Using only one simulation is appropriate whenever $z_P(\hat{\xi})$ and $\hat{\xi}$ are related by a linear relationship,

$$z_P = s\,\hat{\xi} \;+\; \bar{z} \tag{50}$$

and the slope $s = 1$. With these assumptions, it is easy to see that the values of A_G in (37) and (26) will be

$$(s-1)\,\hat{\xi} + \bar{z} + B + \xi_{MHW} \tag{51}$$

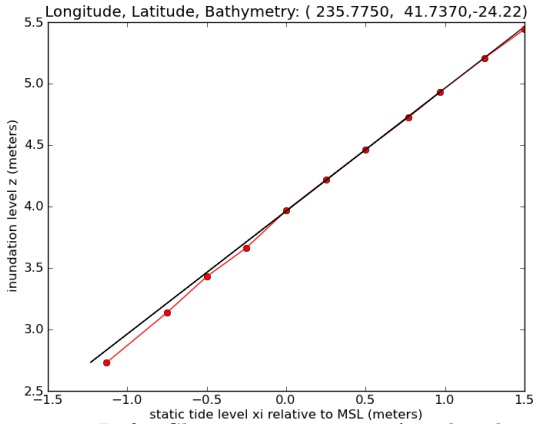
and

$$(s-1)\,\hat{\xi} + \bar{z} + \xi_{MHW} \tag{52}$$

respectively, and whenever $s = 1$ 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. For some locations, the relationship is fairly linear, but the slope is not 1. For other locations, it is best modelled with a piecewise linear function, and the slopes between segments allowed to be different. Figures 12 to 19 below illustrate this point for particular locations using 11 different tide levels for the Alaska 1964 event. The locations in Figures 12 to 16 are in the Crescent City harbor and are presented in the order of increasing bathmetry values. 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 each of these water points, the amplitudes calculated from the 11 runs were fairly close, as were the slopes. The line of slope 1 through $(\xi_{MHHW}, z(\xi_{MHHW}))$ is the black line in the figures.

The locations in Figures 17 to 19 are on land as indicated by the positive bathmetry values. These figures are again presented in order of increasing bathmetry values. Notice now, that for these positive bathmetry points, the slopes of these GeoClaw ξ vs. $z_P(\xi)$ curves are far from 1, and the amplitudes calculated for the 11 runs can be quite different. This indicates that the G-Method can be using the wrong amplitudes when calculating the probability of exceedance at ζ_i when ζ_i is far away from the value of $z(\xi_{MHHW})$ used for the one GeoClaw run required by the G-method. The grid locations in Figures 12, 17, 18, and 19 includes those where the dt-, G-, and G2- Methods differed the most with the Pattern method and those in Figure 16 (Left) and 17 (Left) are for Gauge 101 and Gauge 105, respectively. Gauge 101 is in the water, and Gauge 105 is on land, near a river, and has a high inundation level from most of our sources. Other locations are along transects encompassing the harbor.



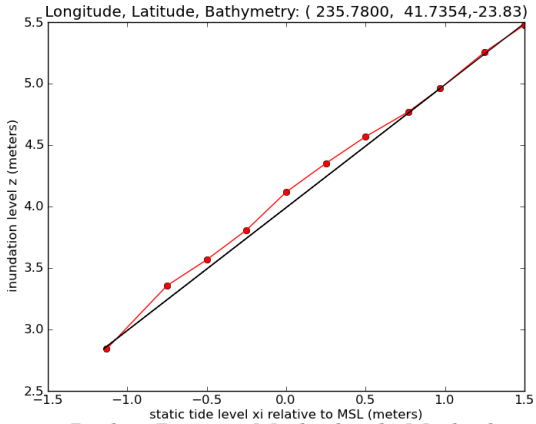
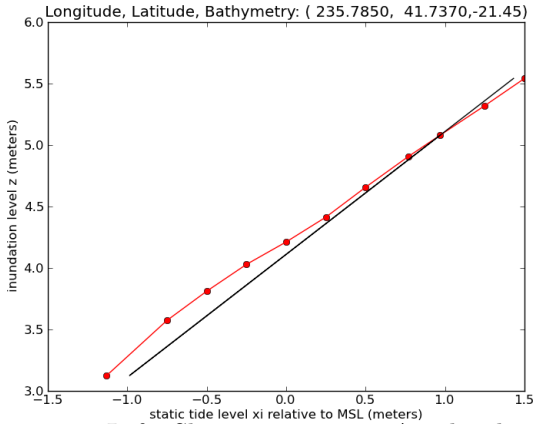


Figure 12: Left: Slopes .93 to 1.21, Amplitudes 4.63 to 4.74. Right: Pattern-Method - dt-Method is max. Slopes .75 to 1.36, Amplitudes 4.74 to 4.89



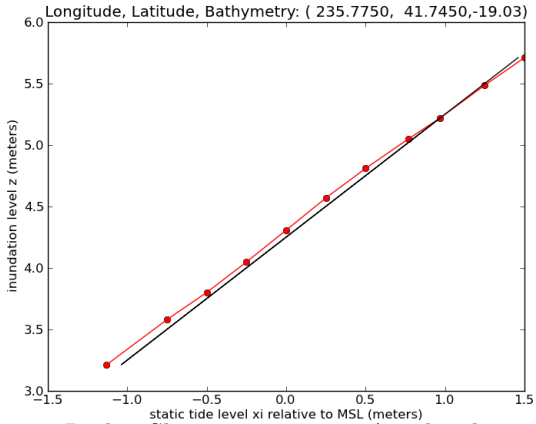
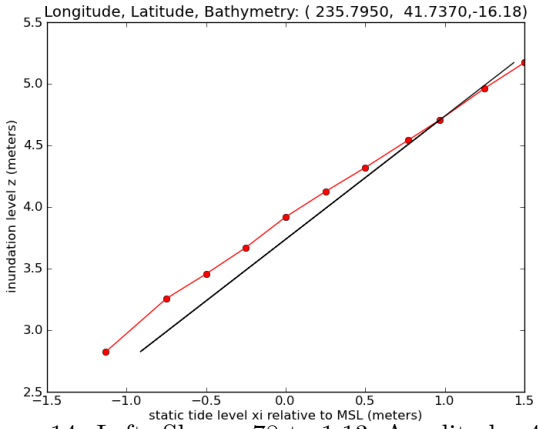


Figure 13: Left: Slopes .73 to 1.18, Amplitudes 4.84 to 5.09. Right: Slopes .85 to 1.05, Amplitudes 5.01 to 5.11



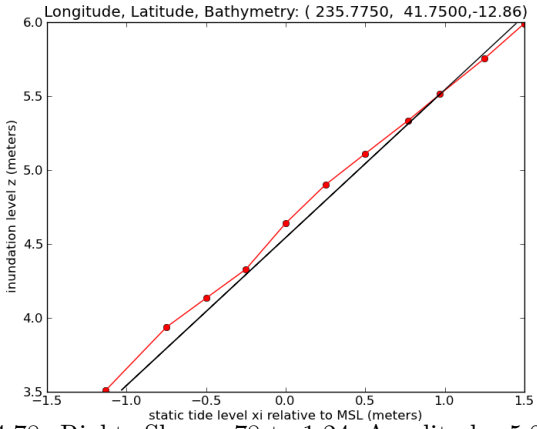
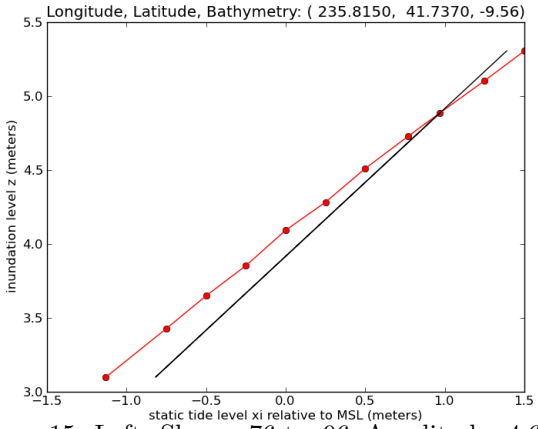


Figure 14: Left: Slopes .78 to 1.13, Amplitudes 4.48 to 4.78. Right: Slopes .78 to 1.24, Amplitudes 5.28 to 5.46



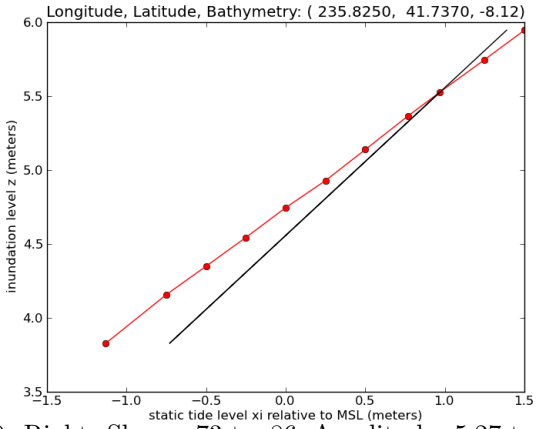
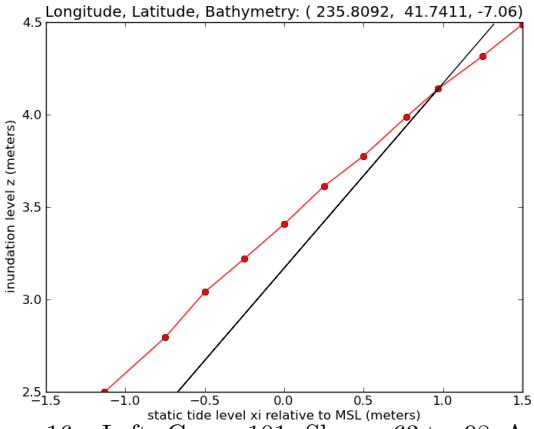


Figure 15: Left: Slopes .76 to .96, Amplitudes 4.62 to 5.0. Right: Slopes .73 to .86, Amplitudes 5.27 to 5.73



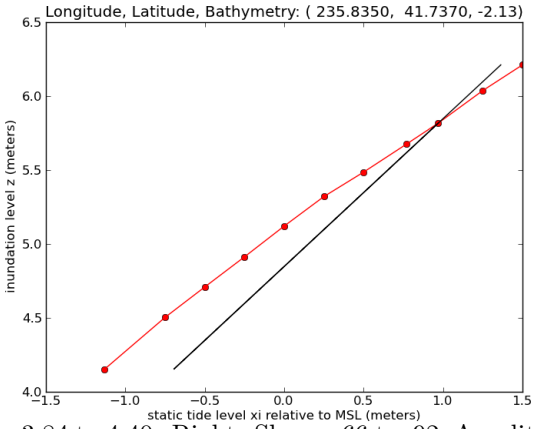
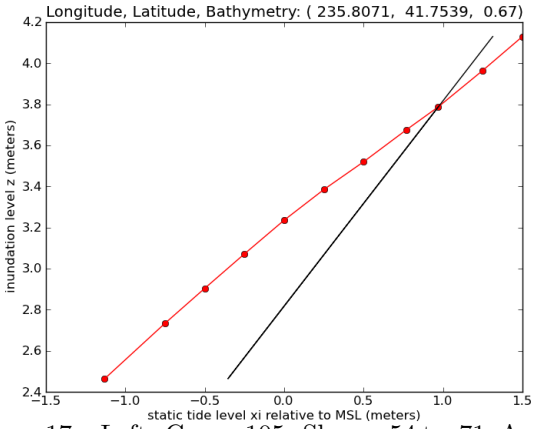


Figure 16: Left: Gauge 101: Slopes .63 to .98, Amplitudes 3.84 to 4.40. Right: Slopes .66 to .92, Amplitudes 5.56 to 6.05



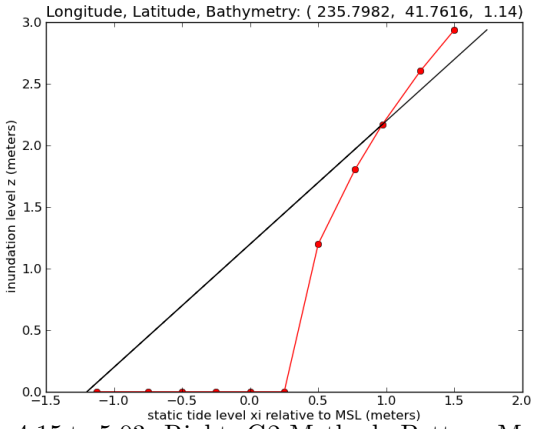
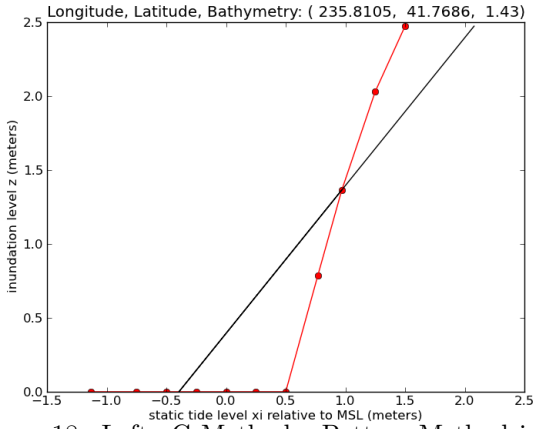


Figure 17: Left: Gauge 105: Slopes .54 to .71, Amplitudes 4.15 to 5.03. Right: G2-Method - Pattern-Method is max. Slopes 0 to 4.8, Amplitudes 1.91 to 3.26



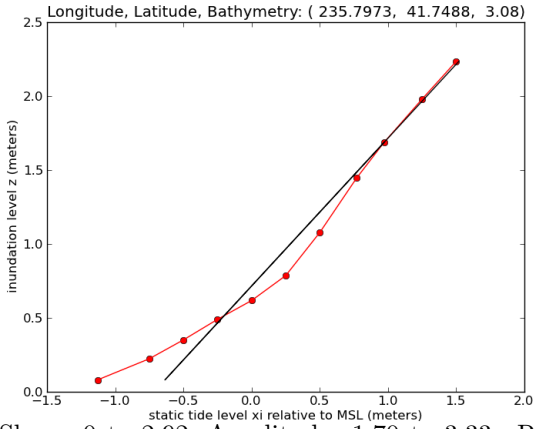
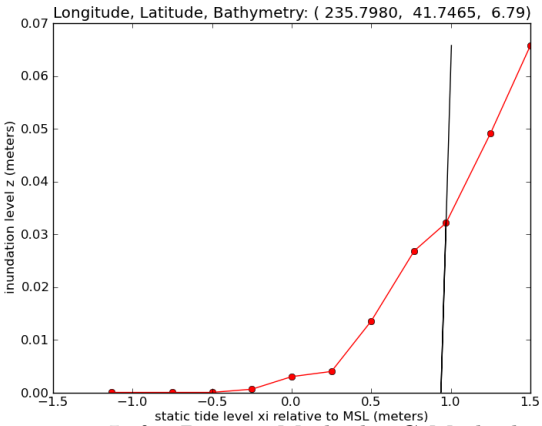


Figure 18: Left: G-Method - Pattern-Method is max. Slopes 0 to 2.92, Amplitudes 1.70 to 3.33. Right: dt-Method - Pattern-Method is max. Slopes .38 to 1.36, Amplitudes 4.39 to 5.06



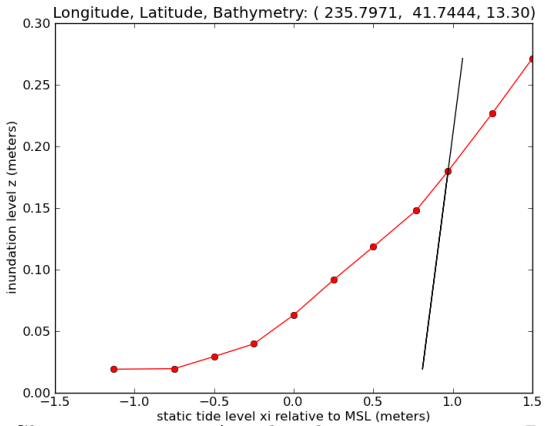


Figure 19: Left: Pattern-Method - G-Method is max. Slopes 0 to .07, Amplitudes 6.36 to 8.69. Right: Pattern-Method - G2-Method is max. Slopes .001 to .18, Amplitudes 13.1 to 15.2

G2-Method description

Since we have multiple GeoClaw z_P values associated with multiple static tide levels $\hat{\xi}$, as done in the dt-Method, we can find the minimum tide level, $\hat{\xi} = w_e$, that we could have used with GeoClaw to produce a maximum exceedance height $z_P(w_e)$ equal to the ζ_i of interest. Therefore, we use $\hat{\xi} = w_e$ to calculate the amplitude ($A_{G2} = A_P(\hat{\xi})$) in equation (37) as

$$A_{G2} = \zeta_i + B + (\xi_{MHW} - w_e) \text{ if } \ (B \geq 0) \tag{53}$$

and that in equation (26) by

$$A_{G2} = \zeta_i + (\xi_{MHW} - w_e) \text{ if } \ (B < 0)\,. \tag{54}$$

The G2-Method calculates

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{\zeta_\text{i} - \zeta_0}{\sqrt{2}\,\sigma}\right)\right) \tag{55}$$

where the mean of ζ_P is

$$\zeta_0 = \zeta_i - w_e + C\xi_{MHHW} e^{-\alpha\left(\frac{A_{G2}}{\sigma_0}\right)^\beta} \tag{56}$$

and the standard deviation, σ of ζ_P is

$$\sigma = \sigma_0\left(1 - C'\, e^{-\alpha'\left(\frac{A_{G2}}{\sigma_0}\right)^{\beta'}}\right). \tag{57}$$

Equation (55) then simplifies to be

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{w_e - C\,\xi_{MHHW}\, e^{-\alpha\left(\frac{A_{G2}}{\sigma_0}\right)^\beta}}{\sqrt{2}\,\sigma}\right)\right) \tag{58}$$

which is seen to be $P(\bar{\xi}_P(w_e) > w_e)$, since from Section 3.4, $\zeta_P = \bar{\xi}_P(\hat{\xi}) + z_P(\hat{\xi}) - \hat{\xi}$ and for the G2-Method, we take $\hat{\xi} = w_e$ and $z_P(\hat{\xi}) = \zeta_i$.

This simple change of amplitude makes a huge difference in the probabilities calculated for some grid locations as seen in Section 4.

3.7 The G3-Method

Since we have multiple GeoClaw simulation curve z_P values associated with multiple static tide levels $\hat{\xi}$, as done in the dt-Method, Pattern, and G2-Methods, we can find the minimum tide level, $\hat{\xi} = w_e$, that we could have used with GeoClaw to produce a maximum exceedance height $z_P(w_e)$ equal to the ζ_i of interest. Instead of using this information to calculate a better amplitude at point P, we use w_e to implement the second simplification by calculating

$P(\zeta_P > \zeta_i \,|\, E_{jk})$ using the cumulative distribution for $\bar{\xi}_{101}$ at Gauge 101 as approximated by the `erf` function assuming the 5-day proxy tsunami. That is,

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{w_e - w_0}{\sqrt{2}\,\sigma}\right)\right) \tag{59}$$

where the mean of $\bar{\xi}_{101}$ is

$$w_0 = C\,\xi_{MHHW} e^{-\alpha\left(\frac{A_{101}}{\sigma_0}\right)^{\beta}} \tag{60}$$

and the standard deviation σ of $\bar{\xi}_{101}$ is

$$\sigma = \sigma_0\left(1 - C'\,e^{-\alpha'\left(\frac{A_{101}}{\sigma_0}\right)^{\beta'}}\right). \tag{61}$$

Equation (59) becomes

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{w_e - C\,\xi_{MHHW}\,e^{-\alpha\left(\frac{A_{101}}{\sigma_0}\right)^{\beta}}}{\sqrt{2}\,\sigma}\right)\right) \tag{62}$$

which is seen to be $P(\bar{\xi}_{101} > w_e)$.

Using the cumulative distribution at Gauge 101 where tidal stages actually add with the tsunami amplitude gives an improvement over the G and G2 methods as seen in Section 4.

3.8 The G4-Method

Since we have multiple GeoClaw simulation curve z_P values associated with multiple static tide levels $\hat{\xi}$, as done in the dt, Pattern, G2, and G3 Methods, we can find the minimum tide level, $\hat{\xi} = w_e$, that we could have used with GeoClaw to produce a maximum exceedance height $z_P(w_e)$ equal to the ζ_i of interest. As in the G3-Method we use w_e to implement the second simplification by calculating $P(\zeta_P > \zeta_i \,|\, E_{jk})$ using the cumulative distribution for $\bar{\xi}_{101}$ at Gauge 101 as approximated by the `erf` function, but we assume a T minute proxy tsunami instead of a 5-day one. The values of T for all the sources in our study are given in Table 5. Mofjeld, et. al. [7] cautioned about using the 5-day approximations for small amplitude tsunamis, and the G4-Method is an attempt at an analytic approximation that uses the proper T. As in [7], we still use an e-folding time of 2 days.

We perform a regression analysis identical to that in Section 3.5 on data generated from the Pattern method applied to ten proxy tsunamis over the duration of T minutes (or $H = T/60$ hours) instead of 5 days to calculate a set of six new regression parameters. Using these parameters, we can check that the value of $C\xi_{MHHW}$ is close to the value of $\mu(\texttt{dt-Method}\,|\,\texttt{dt=H})$ which we calculate for verification. Likewise, we can check that the value of $(1 - C')\sigma_0$ is close to $\sigma(\texttt{dt-Method}\,|\,\texttt{dt=H})$. When $T = 267$ minutes (that for the AASZe03 tsunami) this procedure gives:

- $\alpha' = 0.001729$, $\beta' = 2.11217$, C'=0.28987
- α=0.0339, β=1.2961, and C=0.5000
- $\mu(\texttt{dt-Method} \,|\, \texttt{dt=H}) = 0.494185$. The regression parameters give $C\xi_{MHHW} = 0.484975$.
- $\sigma(\texttt{dt-Method} \,|\, \texttt{dt=H}) = 0.451629$. The regression parameters give $(1 - C')\sigma_0 = 0.45307$.
- Applying the Pattern method to the 267 minute proxy tsunami when $A_G = 3.9174$ gave $w_0 = 0.3370$ and $\sigma = 0.4678$. The set of regression parameters above gives $w_0 = 0.33956$ and $\sigma = 0.46727$, again verifying the model works well for this particular tsunami.

The G4-Method uses this new set of regression parameters to calculate

$$P(\zeta_P > \zeta_i \,|\, E_{jk}) = P(\bar{\xi}_{101} > w_e) = \frac{1}{2}\left(1 - \text{erf}\left(\frac{w_e - w_0}{\sqrt{2}\,\sigma}\right)\right) \tag{63}$$

where the mean of $\bar{\xi}_{101}$ is

$$w_0 = C\xi_{MHHW} e^{-\alpha\left(\frac{A_{101}}{\sigma_0}\right)^{\beta}} \tag{64}$$

and the standard deviation σ of $\bar{\xi}_{101}$ is

$$\sigma = \sigma_0 \left(1 - C'\, e^{-\alpha'\left(\frac{A_{101}}{\sigma_0}\right)^{\beta'}}\right). \tag{65}$$

This is the change to the G-Method that made the most improvement in the sense that the results were much closer to the Pattern method as seen in Section 4. This change is especially recommended for use with land points, since clearly a 5-day tsunami is not going to increase inundation levels on land when the major damage waves have passed in T minutes if T minutes is much less than 5 days.

3.9 Tidal methods summary

The table below summarizes the methods presented in the previous sections. The choices that distinguish the methods include the number of Geoclaw runs needed, the assumption made for the incident tsunami, the amplitude used for the tsunami, the cumulative distribution used to evaluate the conditional probability $P(\zeta > \zeta_i \,|\, E_{jk})$, and whether this distribution is found through a table lookup or an analytical approximation formula.

Table 4: Summary of Tidal Methods for finding $P(\zeta > \zeta_i \,|\, E_{jk})$ for a particular location categorized by longitude and latitude coordinates and known bathymetry B. $B < 0$ ($B \geq 0$) denotes water (land) points.

Method	GeoClaw runs	Tsunami	Amplitude	Cumulative Distribution	$P(\zeta > \zeta_i \,\|\, E_{jk})$
`dt`	multiple, use GeoClaw curve to find w_e	`dt` minutes of contiguous "damage" waves	N/A	Gauge 9419750 `dt` curve	$P(\xi > w_e \,\|\, \texttt{dt})$ (Table Lookup)
Pattern	multiple, use GeoClaw curve to find w_e	Gauge 101 time series for sealevel ξ_{MHHW} run with A_{101}	A_{101}	Convolution of Tsunami and Gauge 9419750 data using `pattern`	$P(\xi > w_e \,\|\, \texttt{pattern})$ (Table Lookup)
G4	multiple, use GeoClaw curve to find w_e	T-min Proxy, T from Table 5	A_{101}	Associated with T-Proxy PDF for $\bar{\xi}_{101}$ having mean and standard deviation given in (64) and (65)	$P(\xi > w_e \,\|\, A_{101})$ given in (63).
G3	multiple, use GeoClaw curve to find w_e	5-day Proxy Mofjeld Tsunami	A_{101}	Associated with Mofjeld PDF for $\bar{\xi}_{101}$ having mean and standard deviation given in (60) and (61)	$P(\xi > w_e \,\|\, A_{101})$ given in (59) and (62)
G2	multiple, use GeoClaw curve to find w_e	5-day Proxy Mofjeld Tsunami	Use w_e, ζ_i and (53) or (54) to get A_{G2}.	Associated with Mofjeld PDF for ζ_P having mean and standard deviation given in (56) and (57).	$P(\zeta > \zeta_i \,\|\, A_{G2})$ (55) or (58)
G	One run at ξ_{MHHW} to get $z_P(\xi_{MHHW})$	5-day Proxy Mofjeld Tsunami	Use ξ_{MHHW}, $z_P(\xi_{MHHW})$ and (26) or (37) to get A_G.	Associated with Mofjeld PDF for ζ_P having mean and standard deviation given in (46) and (45).	$P(\zeta > \zeta_i \,\|\, A_G)$ (48) or (49)

4 Method comparisons

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

In Table 5, we compare the probability density functions (PDFs) of the Mofjeld (G) and Pattern methods for each of the tsunamis considered in the Crescent City study. Table 5 shows there are huge differences between the G method (Mofjeld) and the Pattern method. Only for the four large amplitude tsunamis, such as 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. The first line in the table shows when we apply the Pattern method to the 5 day Proxy-tsunami that is assumed by the G method, the results are virtually identical to those of the G method. This is addressed in more detail in Section 4.2.1.

In the next section, we explain these differences and demonstrate that the G method is improved in a variety of ways by the G2, G3, and G4 methods. Recall, the G2 method still uses the Mofjeld formula (58) at land points, but chooses a better amplitude for the calculation of $P(\zeta > \zeta_i \,|\, E_{jk})$. The G3 method makes a bigger leap by rearranging the Mofjeld formula to seek the calculation of $P(\xi > w_e \,|\, A_{101})$ where the w_e is the water level at the point of interest that needs to be exceeded by the amplitude A_{101} tsunami at the water point of GeoClaw Gauge 101. This avoids the amplitude calculation used by the G method for land points. The G4 method improves the G3 method by using T in Table 5 instead of 5-days for the tsunami length. This choice of T is the crucial factor in addressing the differences in the means and standard deviations of G and Pattern methods, especially for small amplitude tsunamis.

Table 5: Mofjeld (G) and Pattern Method PDF comparisons. The length T in minutes and amplitudes A in meters (seen at Gauge 101) are given in columns 2 and 3 for all the tsunamis used in this study. Columns 4-7 give the mean ω_0 and standard deviation σ of the PDFs for $\bar{\xi}_{101}$ 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
AASZe05	281	1.98	.65	.27	.07	.61
AASZe06	91	.24	.94	.20	.15	.60
AASZe07	114	.24	.94	.20	.16	.60
AASZe08	114	.30	.93	.20	.18	.60
KmSZe01	308	.92	.80	.22	.15	.54
KmSZe02	312	.79	.83	.22	.36	.46
KrSZe01	275	.50	.88	.21	.22	.52
KrSZe02	246	.25	.94	.20	.35	.49
KrSZe03	255	.18	.96	.19	.42	.47
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
CSZBe01r05	160	11.30	.14	.53	.04	.63
CSZBe01r06	161	11.79	.13	.53	.04	.63
CSZBe01r07	160	7.78	.24	.46	.03	.63
CSZBe01r08	161	6.56	.29	.43	.03	.63
CSZBe01r09	160	6.72	.28	.43	.03	.63
CSZBe01r10	160	2.39	.60	.29	.03	.63
CSZBe01r11	163	4.79	.39	.37	.03	.63
CSZBe01r12	162	4.74	.39	.37	.03	.63
CSZBe01r13	160	3.20	.51	.32	.08	.61
CSZBe01r14	162	3.30	.50	.32	.05	.62
CSZBe01r15	160	3.25	.51	.32	.04	.63

4.2 AASZe03 comparisons

For the purposes of further comparing the G, G2, G3, G4, dt, and Pattern methods for including tidal uncertainty, we made GeoClaw runs of the Alaska 1964 event using 11 tide levels. These levels in meters referenced to MSL were -1.13, -0.75, -0.50, -0.25, 0.00, 0.25, 0.50, 0.77, 0.97, 1.25, and 1.5.

We considered 35 exceedances values, ζ_i, $i = 1 \ldots 35$. The first 20 range from $\zeta_1 = 0.0$ to $\zeta_{20} = 1.9$ in increments of 0.1. The next 9, ζ_{21} to ζ_{29}, are the values 2.0 to 6.0 in increments of .5, and finally ζ_{30} to ζ_{35} are the values 7.0 to 12.0 in increments of 1.0.

4.2.1 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 $\bar{\xi}$ 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 20 as the green and dashed red graphs. 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 20 by the differences between the black graph and the green (or dashed red) ones.

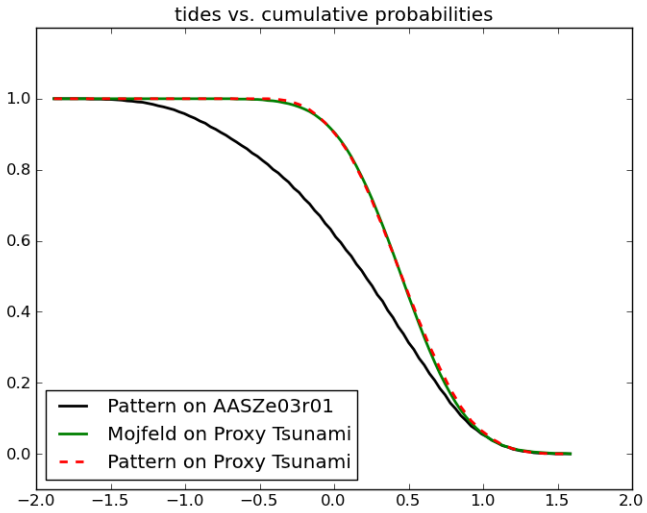


Figure 20: Pattern Method Validation

4.2.2 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 six methods for the Alaska 1964 tsunami (AASZe03). These are given in Table 6 below.

Based on these measures, the G4 method is closer to the Pattern method than the G3, G2, or G methods at Gauge 101. This indicates that changing the tsunami duration to T minutes instead of 5 days accounts for a huge part of the difference. The difference between the G4 and Pattern method is due to the differences between the actual GeoClaw tsunami pattern at Gauge 101 versus the assumed decaying Proxy tsunami.

Table 6: 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 six methods at the water point GeoClaw Gauge 101 to compute the Cumulative Distribution for the probability indicated in column 4 at this point. We note $A_{101,G2} = \zeta_i + \xi_{MHW} - w_e$ and $A_{101} = z(\xi_{MHHW}) + \xi_{MHW} - \xi_{MHHW}$. For all ζ_i, $A_{101,G2}$ and A_{101} are within .35m of each other as seen in Figure 16 (Left). The w_0 and σ for G2 are for when $\zeta_i = z(\xi_{MHHW})$. Recall $\tilde{w}_e = \zeta_i - z(\xi_{MHHW}) + \xi_{MHHW}$ and $\zeta_i = z(w_e)$.

Method	w_0	σ	$P(\zeta > \zeta_i \mid E_{jk})$
Pattern	.14	.60	$P(\xi > w_e \mid \texttt{pattern})$
dt	.12	.62	$P(\xi > w_e \mid \texttt{dt})$
G4	.34	.47	$P(\xi > w_e \mid A_{101}, T\,\texttt{min})$
G3	.45	.34	$P(\xi > w_e \mid A_{101}, 5\,\texttt{days})$
G2	.45	.34	$P(\xi > w_e \mid A_{101,G2}, 5\,\texttt{days})$= $P(\zeta > \zeta_i \mid A_{101,G2}, 5\,\texttt{days})$
G	.45	.34	$P(\xi > \tilde{w}_e \mid A_{101}, 5\,\texttt{days})$= $P(\zeta > \zeta_i \mid A_{101}, 5\,\texttt{days})$

4.2.3 Hazard and tide probability differences

The hazard probabilities ($P(\zeta > \zeta_i)$) given by (9) for the G, G2, G3, G4 and dt Methods could be compared to those of the Pattern-Method, but since the 1964 Alaska event is the only source for these comparisons, the differences will be essentially the same as the return rate $\nu_1 = 1.0/750. = 1.33e^{-03}$ multiplied by the tide probability ($P(\zeta > \zeta_i \mid E_{jk})$) differences. This is because $j = 1$, $k_j = 1$ and $P(E_{jk} \mid E_j) = 1$ in equation (8), and $1 - e^{-\nu_1}$ is well approximated by ν_1. Hence we only compare the differences in the tide probabilities. We also give the individual probability contour plots for the exceedance of $\zeta_i = 0$ meters and $\zeta_i = 2$ meters for each of these six methods in Section 4.2.4.

For each grid location, we compared the 35 probabilities $P(\zeta > \zeta_i \mid E_{jk})$ in equation (8) for the 1964 Alaska event. Again, $j = 1$, $k_j = 1$, and i ranges from 1 to 35. The numbers in Table

7 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.

Table 7: **Tide Probability** $P(\zeta > \zeta_i \,|\, E_{jk})$ **Differences**

	G-Pattern	**G2-Pattern**	**G3-Pattern**	**dt-Pattern**	**G4-Pattern**
max	+.747	+.493	+.290	+.006	+.158
min	-.936	-.090	-.0003	-.017	-8.5e-06

Indeed, differences close to 1 are observed in the first column. It is seen that the G3-Method and the G4-Method are much closer to the Pattern-Method (and hence the dt-Method) than is the G-Method or the G2-Method. This should be expected since now the G3-Method and the G4-Method, as well as the dt-Method and the Pattern-Method make use of more information about the relationship between the z a GeoClaw run produces and the tide level $\hat{\xi}$ that it used. The G3 and G4 Methods use the amplitude of the tsunami at Gauge 101 (instead of the amplitudes at the land points) as does the Pattern method. The major difference between the G3-Method and the Pattern method is that the G3-Method is using the 5-day proxy tsunami at Gauge 101 while the Pattern method is using the actual Geoclaw tsunami pattern of duration T minutes. The G4-Method improves the G3-Method by using a T-minute proxy tsunami at Gauge 101 instead of the 5-day one. Even though both the G4-Method and the Pattern methods use a T-minute tsunami, the T-minute proxy one gives a max difference of .158 in the conditional probabilites as compared to the actual GeoClaw Gauge 101 tsunami as seen in the last column of Table 7. This difference accounts for the different "patterns" of the two tsunamis and gives an indication of the magnitude of the error made by using the proxy one. Contour plots of the absolute value of these differences in the p-contours for exceeding $\zeta_i = 0$ and $\zeta_i = 2$ meters can be found in section 4.2.5.

For the 1964 Alaska event, dt=1 was used and it works surprising well compared with the Pattern-Method. For some sources in Table 5, we believe that choosing dt=2 or higher instead of dt=1 is warranted since waves of fairly equal magnitude persist for more than one *contiguous* hour. This would bring the results closer to those of the Pattern Method. However, the main idea of the Pattern Method was to avoid the art of "guessing" dt. The process of observing the wave pattern at gauge 101 (the gauge nearest the Crescent Tide Gauge) and estimating a *contiguous* time over which waves could equal or exceed a ζ_i value at the grid locations is cumbersome. It is also not as precise as automating the process by using the actual wave pattern as described in the Pattern-Method to create the cumulative probability distribution. However, it performs surprising well compared to the Pattern-Method for the sources in Table 5, and only requires one cumulative probability distribution to be found, instead of one for each specific GeoClaw tsunami pattern. More work needs to

be done to make a more accurate recommendation concerning the dt-Method and to find a more automated way of choosing dt.

We also expect the Pattern-Method to be more precise than the G, G2, G3, and G4 methods, since a theoretical decaying tsunami is not assumed in developing the Pattern-Method's cumulative probability distribution. In fact, 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 5. 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.

4.2.4 Probability contour plots

Figure 21 shows contour plots of exceeding $\zeta_i = 0$ meters (left) and $\zeta_i = 2$ meters (right) without the incorporation of tidal uncertainty. This plot is followed by the probability contour plots of exceeding $\zeta_i = 0$ meters using the dt, Pattern, G, G2, G3 and G4 methods in Figure 22. We organize the plots to allow a visual comparison over a map of Crescent City of the dt to the Pattern-Method, the G to the G2-Method, and the G3 to the G4-Method. At this level of detail, all six methods have the same qualitative detail for $\zeta_i = 0$, with the G method showing some discernable differences with the other methods.

The next set of contour plots in Figure 23 repeats the process for $\zeta_i = 2$. Now, it is clear there is a difference between the methods. The dt and Pattern-Methods look the same, as do the G2, G3 and G4-Methods. The G-Method is clearly different from other other 5. Also, the $\zeta_i = 2$ contours clearly show how the effect of the tides can cause inundation at a $\zeta_i > 0$ level that was not possible when no tidal uncertainty is included as shown in the right-side plot of Figure 21.

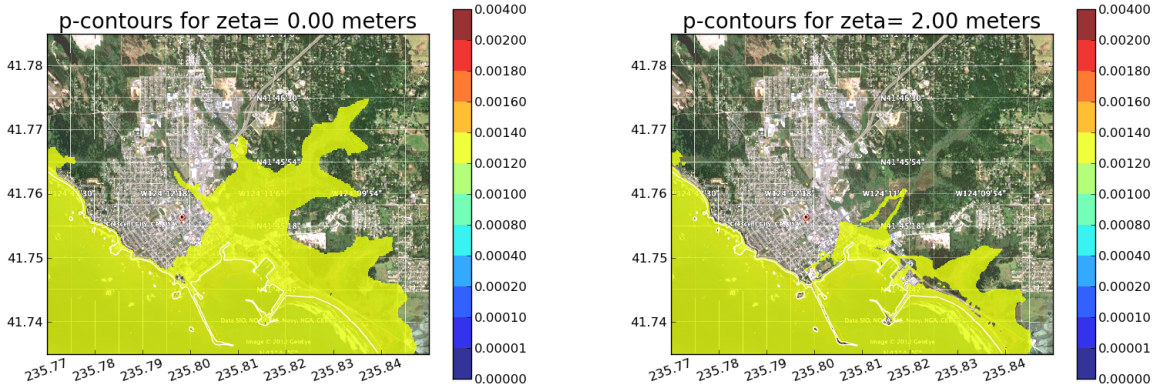


Figure 21: No Tides: Probability Contours for Exceeding $\zeta = 0$ m. (left) and $\zeta = 2$ m. (right)

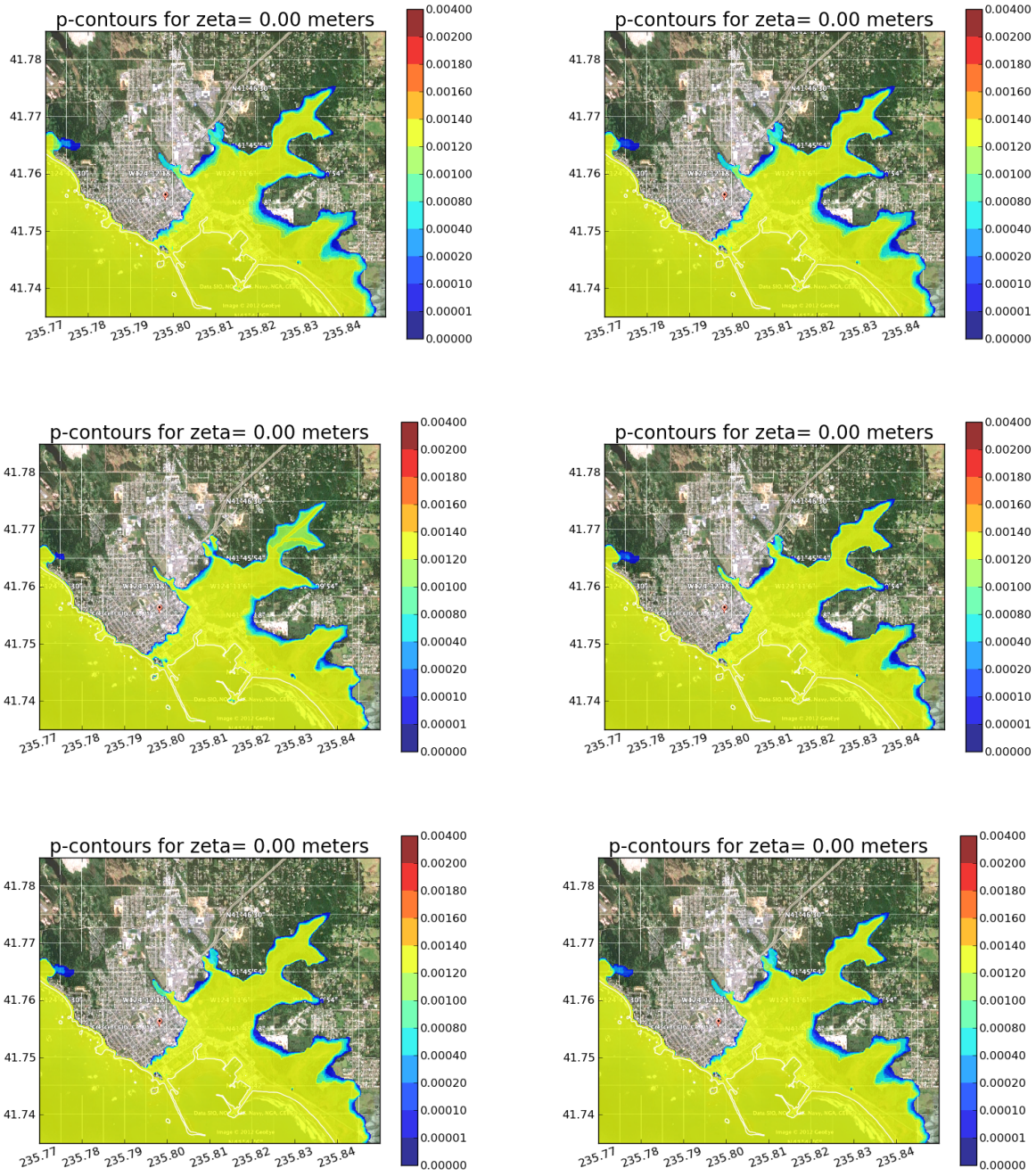


Figure 22: Probability Contours for Exceeding $\zeta = 0$ meters. Top (left: dt-Method, right: Pattern-Method), Middle (left: G, right: G2), Bottom (left: G3, right: G4)

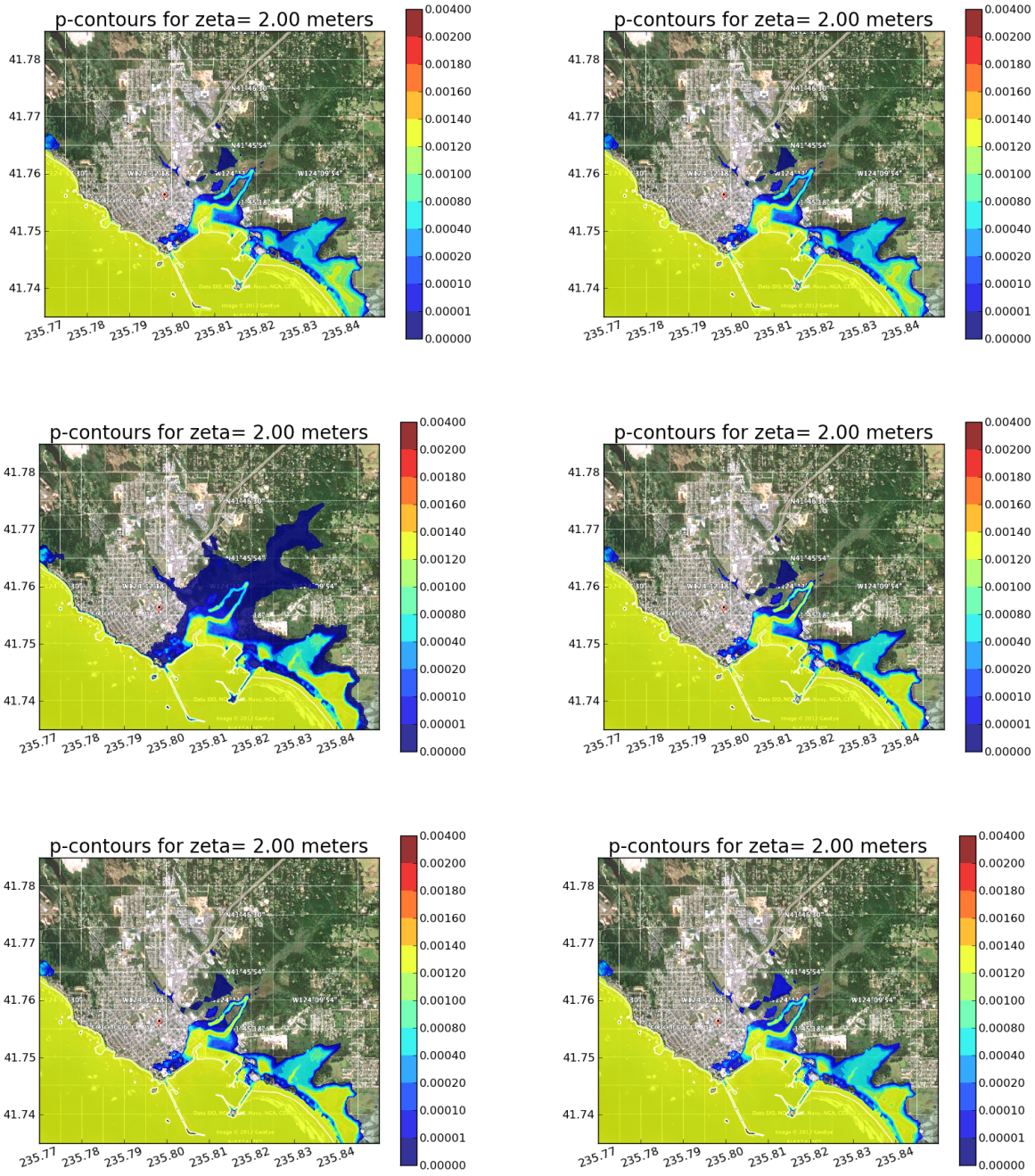


Figure 23: Probability Contours for Exceeding $\zeta = 2$ meters. Top (left: dt-Method, right: Pattern-Method), Middle (left: G, right: G2), Bottom (left: G3, right: G4)

4.2.5 Tide probability differences contour plots

In Figures 24, 25, 26, 27 and 28 we compare the dt-Method, the G, G2, G3, and G4 Methods 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 plots are organized to allow a visual comparison of the differences as ζ_i increases. Also, the colors in the plot give a quick comparison of each method's behavior compared to the Pattern-Method.

The magnitudes in the colorbar of Figure 24 for the dt-Method and Pattern-Method comparisons show their tidal probabilities differ by less than 2%. In fact, Table 7 gives the difference as less than 1.7%.

Figures 25 and 26 show the G2-Method is closer to the Pattern-Method than the G-Method along the boundaries of the inundated region as seen in less red colors in the $\zeta = 0$ contours and the smaller area of differences in the $\zeta = 2$ contours. In fact, the G-Method had the most visible differences with the Pattern-Method of all the methods, but it is encouraging to see that the major differences as indicated by the brighter colors are not nearly as prevalent as the smaller differences indicated by the blue and green colors. It is clear that selecting the amplitude in the G2-Method differently from that in the G-Method accounts for the observed improvement.

Comparing the right-side plots in Figures 25 and 26 with the left-side plots in Figures 27 and 28 shows the G3-Method gives less bright colors than the G2-Method in the $\zeta = 0$ contours, and is very similar to the G2-Method for the $\zeta = 2$ contours. We note that for $\zeta = .5$ and $\zeta = 1$ contours (not included), the G3-Method also shows improvement over the G2-Method in regions along the inundated boundary. These improvements by the G3-Method suggest it is better not to use amplitudes that are calculated at land-points for use with the proxy tsunami as does the G2-Method, but instead use the amplitudes of the Geoclaw tsunami at Gauge 101 and change the question from $P(\zeta > \zeta_i \,|\, E_{jk})$ at the land point to $P(\xi > w_e \,|\, A_{101})$ in the sense of the proxy tsunami at Gauge 101.

Figures 27 and 28 clearly show that the G4-Method has smaller differences with the Pattern-Method than does the G3-Method, both for the $\zeta = 0$ and $\zeta = 2$ contours. The G4-Method results show the biggest gain in fixing the G-Method to behave more like the Pattern method comes from using the proxy-tsunami duration to be T minutes instead of 5-days. This is crucial for land points as it eliminates the false result that for very small amplitudes, the mean of the distribution for $\bar{\xi}_{101}$ is around ξ_{MHHW}, instead of around the mean of the `dt`-Method when `dt` = `H`, where $H = T/60$ hours. Likewise, the standard deviation should be close to that of the `dt`-Method when `dt` = `H` for small amplitude tsunamis. The G-Method overestimates the probabilities for small amplitude tsunamis. We claim it is proper to use the Pattern Method, G4 method, or the `dt` method for these tsunamis.

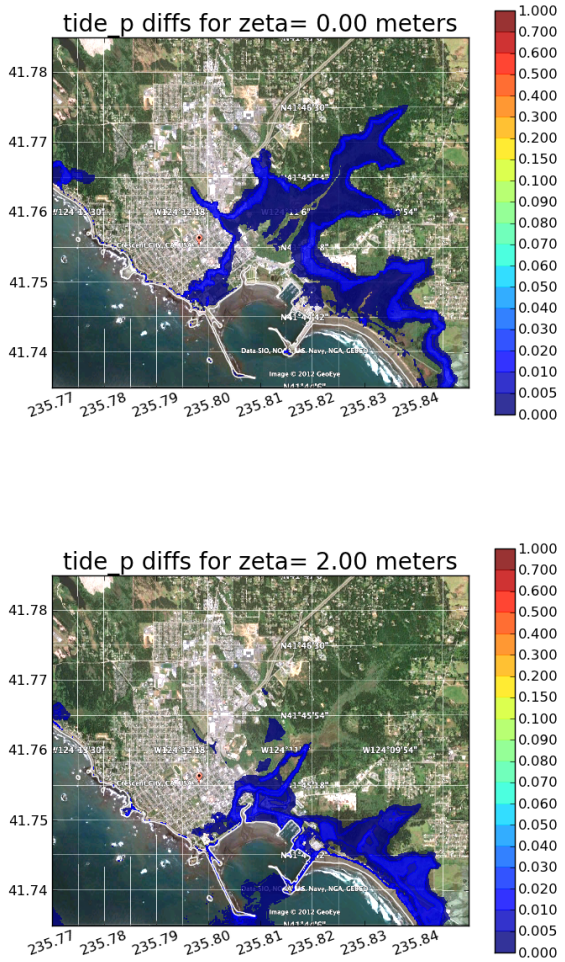


Figure 24: abs(dt-Pattern): Probability Difference Contours, Top: $\zeta = 0$ m., Bottom: $\zeta = 2$ m.

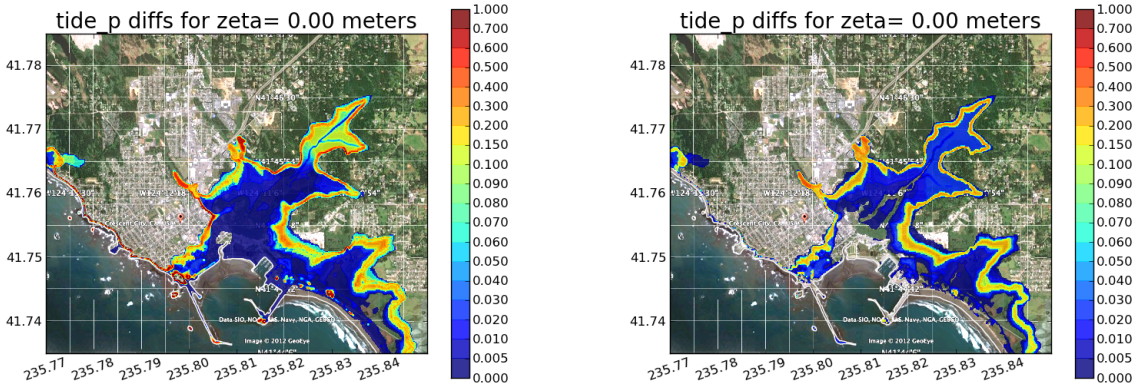


Figure 25: Probability Difference Contours for Exceeding $\zeta = 0$ meters, Left: abs(G-Pattern), Right: abs(G2-Pattern)

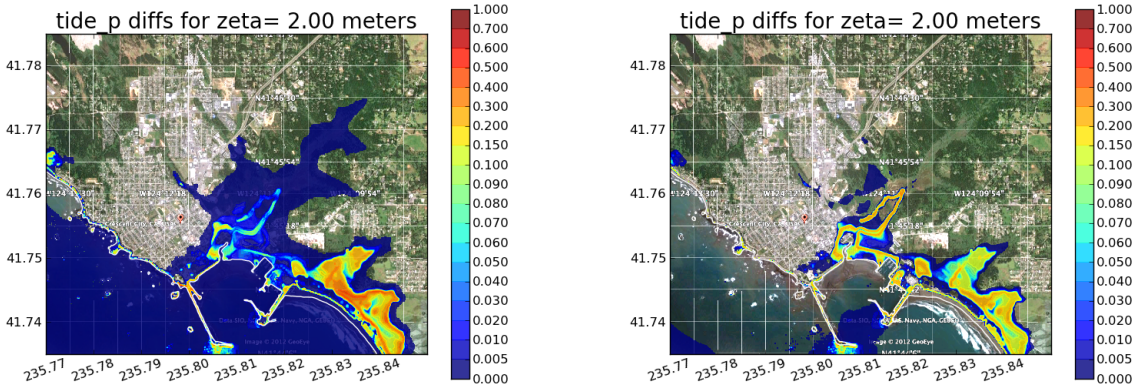


Figure 26: Probability Difference Contours for Exceeding $\zeta = 2$ meters, Left: abs(G-Pattern), Right: abs(G2-Pattern)

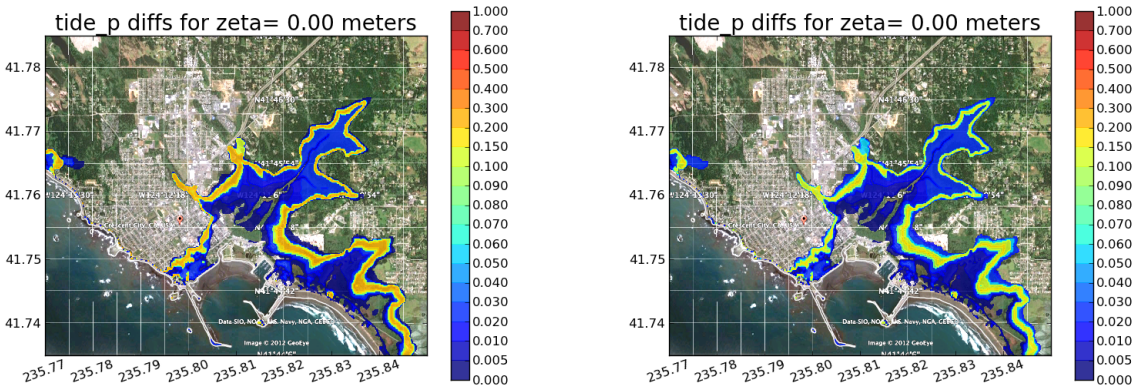


Figure 27: Probability Difference Contours for Exceeding $\zeta = 0$ meters, Left: abs(G3-Pattern), Right: abs(G4-Pattern)

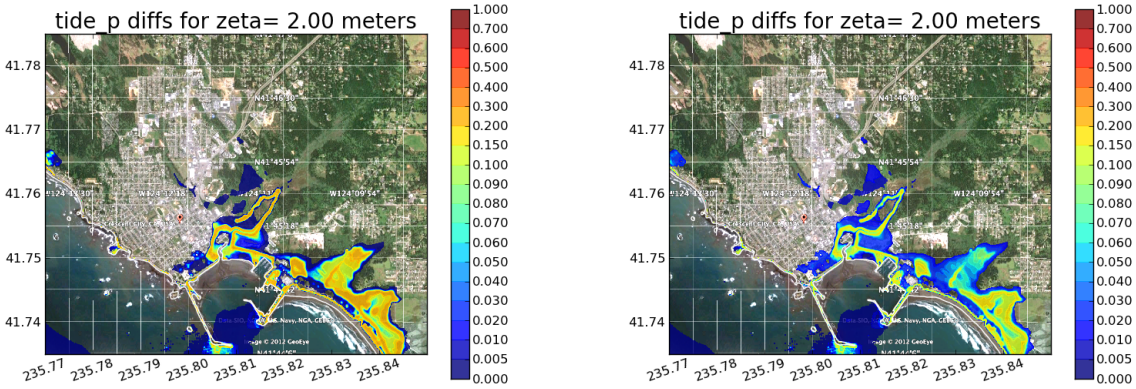


Figure 28: Probability Difference Contours for Exceeding $\zeta = 2$ meters, Left: abs(G3-Pattern), Right: abs(G4-Pattern)

5 Conclusions and open questions

This study has provided some advice to the community about modelling tidal uncertainty in PTHA.

- Multiple GeoClaw simulations at a given land point show that the maximum inundation is *not* a linear function of the tidal stage used for the simulation. This means there is *not* a single amplitude (height above the simulation sealevel parameter) for these points. This means using the G method for land points is not advised.
- The tsunami actual duration time T matters, and should be used in modelling land points. If instead one were interested in what goes on in the harbor after the tsunami waves have passed, increasing T to include the span of time (say 5 days) needed to see higher tide stages more frequently may be warranted for smaller amplitude tsunamis. It will depend on the questions one is trying to answer. The advice, though, is that land and water points are different for modelling purposes, and the results highly depend on T in either case.
- The Pattern-Method can be easily adjusted to have a different distribution for land as for water points if justified by the questions to be asked in a given study. For example, our next study will include currents in the harbor, and we might want T to be taken longer in the harbor than for points on land. The pattern is simply extended.
- For some earthquake sources, such as AASZe03, the proxy tsunami gives a reasonable approximation to the actual tsunami, but can be made even closer by fixing the proxy tsunami to have duration T as is done in the G4-Method. We are now looking at this question for all the sources in our study. So far, we can conclude that the most important difference is the time duration T.

Our attempts at modelling tidal uncertainty are not perfect. Here are some issues that warrant further consideration.

- We can not afford to save time series at every point of interest. The G method uses an analytic expression that depends on a single amplitude at each point, and hence is not appropriate at land points. The G2 and G3 methods improved upon this, but still didn't take account of the actual duration time of the tsunami. Even though the G4 method uses the actual duration to derive the mean and standard deviations in the analytic expression for a proxy T-minute tsunami, it still produced a $P(\zeta > \zeta_i \,|\, E_{jk})$ that was off by .158 from that of the Pattern-method, indicating the particular pattern of the actual tsunami does matter quite a bit. Also, the work to produce the distribution for the G4 method does not warrant its use over the Pattern method.
- The Pattern-Method is the best approach we investigated. It has the limitation that the distribution for $\bar{\xi}_{101}$ at Gauge 101 was used along with the GeoClaw simulation curve at each point P to infer probabilities $P(\zeta > \zeta_i \,|\, E_{jk})$ at each point P. Using more water gauges with GeoClaw simulation curves with slopes near 1 should be investigated. For Crescent City, most inundation of the city came from water passing through Gauge

101, but not all. The dt-Method is our simpliest approach because it only requires distributions to be formed by binning up the actual tide gauge data but so far we have no automatic way of choosing the value of `dt`.

- 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

[1] W.G. Van Dorn. Some tsunami characteristics deducible from tide records. *J. Physical Oceanography*, 14:353–363, 1984.

[2] Isaac V. Fine, Evgueni A. Kulikov, and Josef Y. Cherniawsky. Japan's 2011 tsunami: characteristics of wave propogation from observations and numerical modelling. *Pure and Applied Geophysics*, 2012.

[3] F. I. González, E L Geist, B. Jaffe, U Kanoglu, et al. Probabilistic tsunami hazard assessment at Seaside, Oregon, for near-and far-field seismic sources. *J. Geophys. Res.*, 114:C11023, Jan 2009.

[4] F. I. González, R. J. LeVeque, and L. Adams. Probabilistic tsunami hazard assessment (ptha) for Crescent City, CA. BakerAECOM Report for PTHA of Crescent City, CA, supported by FEMA Risk MAP Program, 2012.

[5] J. R. Houston and A. W. Garcia. Type 16 flood incurance study: Tsunami predictions for the West Coast of the continental United States, Crescent City, CA. USACE Waterways Experimental Station Tech. Report H-78-26, 1978.

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

[7] H.O. Mofjeld, F.I. Gonzáles, V.V. Titov, A.J. Venturato, and J.C. Newman. Effects of tides on maximum tsunami wave heights: Probability distributions. *J. Atmos. Ocean. Technol.*, 24(1):117–123, 2007.

[8] Alexander B. Rabinovich, Rogerio N. Candella, and Richard E. Thomson. Energy decay of the 2004 sumantra tsunami in the world ocean. *Pure and Applied Geophysics*, 168:1919–1950, 2011.

[9] S.L. Soloviev. Recurrence of tsunamis in the Pacific. Tsunamis in the Pacific Ocean, W.M. Adams, editor, Honolulu, East-West Center Press, 2011.

[10] Y. Zhang, R.C. Witter, and G.R. Priest. Tsunami-tide interaction in 1964 Prince William Sound tsunami. *Ocean Modeling*, 40:246–259, 2011.