Forecasting tsunami inundation with convolutional neural networks for a potential Cascadia Subduction Zone rupture

David Grzan[1], John B Rundle[2,3,4], Geoffry C Fox[5,6], and Andrea Donnellan[4]

[1]Affiliation not available
[2]Physics and Astronomy, University of California
[3]Santa Fe Institute
[4]Jet Propulsion Laboratory, California Institute of Technology
[5]Biocomplexity Institute and Initiative, University of Virginia
[6]Computer Science Department, University of Virginia

February 9, 2023

Forecasting tsunami inundation with convolutional neural networks for a potential Cascadia Subduction Zone rupture

David P. Grzan[a]*, John B. Rundle[a,b,c], Geoffry C. Fox[d,e], Andrea Donnellan[c]

[a]Physics and Astronomy, University of California, Davis, CA, USA
[b]Santa Fe Institute, Santa Fe, NM, USA
[c]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[d]Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, USA
[e]Computer Science Department, University of Virginia, Charlottesville, VA, USA
*Corresponding Author, email: grzanprojects@gmail.com

August 2, 2022

Abstract

Tsunamis in the last two decades have resulted in the loss of life of over 200,000 people and have caused billions of dollars in damage. There is therefore great motivation for the development and improvement of current tsunami warning systems. The work presented here represents advancements made towards the creation of a neural network-based tsunami warning system which can produce fast inundation forecasts with high accuracy. This was done by first improving the waveform resolution and accuracy of Tsunami Squares, an efficient cellular automata approach to wave simulation. It was then used to create a database of precomputed tsunamis in the event of a magnitude 9+ rupture of the Cascadia Subduction Zone, located only ∼100 km off the coast of Oregon, US. Our approach utilized a convolutional neural network which took wave height data from buoys as input and proved successful as maps of maximum inundation could be predicted for the town of Seaside, OR with a median error of ∼0.5 m.

1 Keywords

- Tsunami Simulation
- Machine Learning
- Tsunami Early Warning
- Cascadia Subduction Zone
- Tsunami Squares

2 Introduction

The Cascadia Subduction Zone stretches from Southern Canada down to Northern California and has a history of large, submarine earthquakes. According to records, the average recurrence rate is about 500 years with the last event occurring 300 years ago in which the entire 1000 km length of the subduction zone ruptured and produced a tsunami causing destruction reaching as far as Japan (Clague, 1997; Heaton & Hartzell, 1987; Atwater et al., 1995). It is proposed that either a large magnitude 9+ earthquake or several smaller magnitude 8+ earthquakes are estimated to occur some time in the future (Heaton & Hartzell, 1987). Because the Cascadia Subduction Zone is located only ∼100 km from the coast, tsunamis generated here are particularly hazardous because of how little time it takes to reach the coast. For this reason, the

rapid forecasting of these tsunamis is crucial for the dissemination of warnings which can drastically reduce casualties.

In recent history, similar subduction zones have triggered large tsunamis, causing massive loss of life. The most recent example is the 2011 Tohoku tsunami, where Japan suffered a loss of life exceeding 19,500 citizens (Imamura & Anawat, 2011). In this case, the initial tsunami warning was based on an initial underestimate of the earthquake magnitude (M7.9), where the actual magnitude was much larger (M9.0) (Satake, 2014). This lead to a tsunami wave height estimate which was much lower, causing residents to stay in dangerous areas due to a misunderstanding of the risks (Ando et al., 2011). Several studies have been carried out since, aiming to produce faster and more reliable tsunami forecasts. These include the development of a system to produce real time high resolution tsunami inundation simulations using supercomputers (Musa et al., 2018), near-field tsunami forecasting based on rapid estimations of tsunami source functions (Tsushima et al., 2012, 2014), and an algorithm which finds the best matching precomputed waveforms from virtual observation points in order to directly obtain the corresponding inundation map (Gusman et al., 2014). However, the earthquake source inversions these methods rely on potentially take tens of minutes (Tsushima et al., 2011, 2012), are prone to varying results due to slightly incorrect input parameters (Mai et al., 2016), and are subject to error due to ground sensor rotation or tilt (Kubota et al., 2018). In addition, forward-modeling approaches such as the one developed by Oishi et al. (2015) was able to produce an inundation forecast in under 1.5 min, but required nearly 10,000 cores to do so and is therefore inaccessible to areas with more limited resources. From this, several challenges need to be met in order to increase the speed and accuracy of near-field tsunami inundation forecasts in the event of a Cascadia Subduction Zone rupture.

To respond to these challenges, several studies have since expanded on work done by Gusman et al. (2014) by using increasingly efficient selection techniques on precomputed tsunami databases. Mulia et al. (2018) improves on their method of finding the best fitting scenario based on waveforms from virtual observation points, and instead uses an interpolation algorithm on a map of low resolution maximum tsunami heights to produce a unique inundation map. Following this study, Mulia et al. (2020) replaces the interpolation algorithm with a two-stage deep neural network in order to gain additional efficiency. Similar to this approach, Fauzi & Mizutani (2020) uses a comparable database of low and high resolution simulations to train a convolutional neural network (CNN) and a multilayer perceptron (MLP) find the best match in the database or to produce a unique solution. In a different approach, Makinoshima et al. (2021) utilizes a CNN to forecast inundation waveforms at specific locations along the Japan coast, using wave height time series data from many buoys as input to the network.

In this study, we employ a CNN to forecast inundation resulting from a given rupture of the Cascadia Subduction Zone based on currently existing and hypothetical configurations of wave height measurement buoys. This method avoids the inherent inaccuracies and delays associated with earthquake source inversion techniques by directly using wave height time series buoy data similar to Makinoshima et al. (2021). In addition, we forecast the entire maximum inundation map as was done by Mulia et al. (2020); Fauzi & Mizutani (2020). Dissimilar to the studies mentioned, the earthquake database here is created using techniques similar to Goda et al. (2018), where correlations are introduced between slip values on the fault, creating localized asperities which heavily influence the location of affected coastal areas (Mueller et al., 2015). For this study, Seaside, Oregon is chosen as the site of interest because of its low elevation, proximity to the coast, and overall susceptibility to a future tsunami disaster (González et al., 2009; Dominey-Howes et al., 2010; Park & Cox, 2016; Park et al., 2019).

2.1 Study Site

The city chosen for this study is Seaside, located on the coast of Oregon, USA. It is a relatively small city with a population of around 6,500 according to the 2010 census (Bureau, 2022). It was chosen because it is only 130 km away from the Cascadia Subduction Zone (CSZ), a fault which runs parallel to the North American West Coast stretching from Vancouver to Northern California. The CSZ is positioned between the Juan de Fuca and North American plate, where the heavier Juan de Fuca plate slowly slides beneath the lighter North American Plate at a shallow angle. At shallower depths (<30 km), the plate is locked by frictional forces. However at deeper depths, the edge of the sinking plate becomes molten under intense heat and stress, allowing large slipping between the plates to occur, resulting in large earthquakes (magnitude 8.0+).

Because of the massive length of this fault, it is capable of producing earthquakes that can exceed magnitude 9.0. The average interval between large earthquakes (M8.0+) is approximately 500 years, but historical records show that the interval can vary anywhere from 210 to 910 years (Atwater et al., 2011).

There have been several studies detailing the probabilistic tsunami hazard assessment for Seaside, OR (González et al., 2009; Dominey-Howes et al., 2010; Park & Cox, 2016; Park et al., 2019). In these studies, maximum tsunami amplitudes for different time spans such as 100 or 500 years are first calculated by looking at historical records and determining the largest earthquake most likely to occur in this range of time. From this, the corresponding maximum tsunami amplitude is calculated and fed into an inundation simulator to produce the inundation map. Historic studies suggest Seaside has been struck by several tsunamis in the past, either from CSZ ruptures or from other far-field tsunamis (Fiedorowicz & Peterson, 2002). It is known that on average once every 500 years, the Cascadia Subduction Zone ruptures, causing a massive tsunami that strikes the coast of North America (Peterson et al., 2008). However, other tsunamis which have origins elsewhere have caused flooding in Seaside. These would notably be the 1946 Alaska event, the 1960 Chile event, and the 1964 Alaska event which caused significant property damage (Lander et al., 1993).

Seaside, Oregon is particularly vulnerable as it sits directly on a gradually sloping beach, providing little to no barrier for any large wave that comes its way. Figure 1 shows two aerial views in addition to a map in order to highlight the city's susceptibility to a large wave.

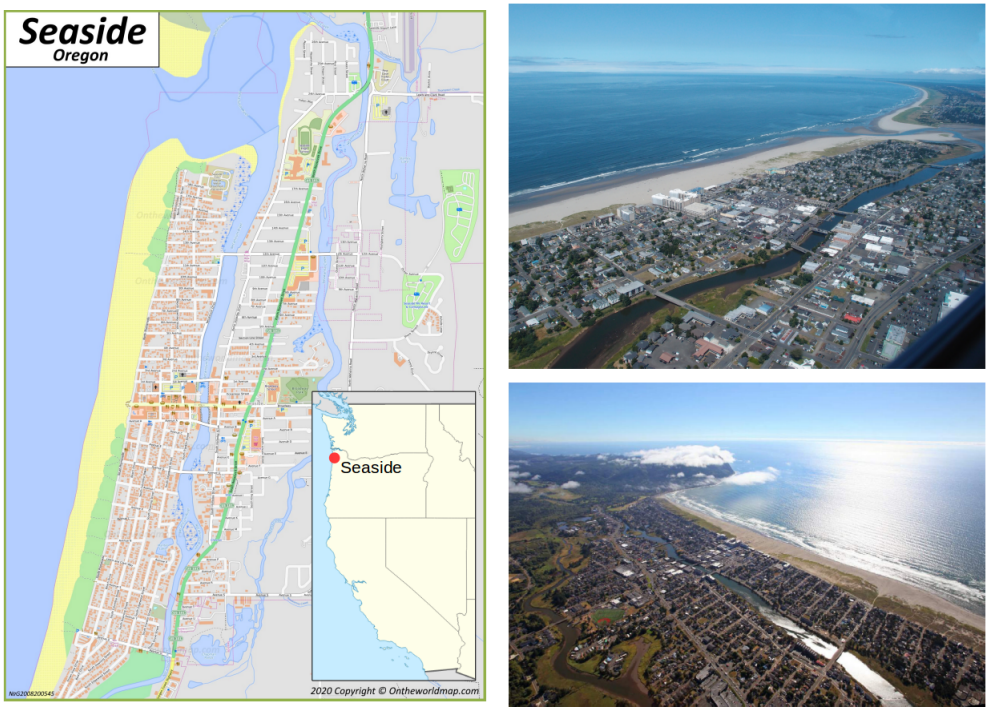

Figure 1: Left: Map of Seaside showing the buildings and city layout (*Seaside Map*, 2020). Top right: Aerial photo of the city which particularly highlights its susceptibility to an incoming tsunami (F., 2011). Bottom right: Another aerial photo looking south (Frank, n.d.).

3 Database Development

3.1 Earthquake Rupture Scenarios

Generating earthquake rupture scenarios for a tsunami database is crucial because it forms the initial conditions from which all subsequent simulations begin. If the CSZ ruptures in a way which is contained within the basis defined by this database, then the neural network should be able to make an accurate prediction for inundation. In an attempt to span a large amount possible earthquakes caused by the Cascadia Subduction Zone within the Oregon region, a database of 3000 earthquakes was made.

When developing such a collection of earthquakes, each one must remain sufficiently unique such that there are no substantially overlapping scenarios in the database. To accomplish this, previous studies involving machine learning on a database of tsunamis either take each subfault and randomly assign a slip value (Makinoshima et al., 2021), or systematically vary the position and size of the fault over a region (Mulia et al., 2020; Fauzi & Mizutani, 2020). However in this study, a procedure inspired by Goda et al. (2018) was used.

Shown in Figure 2 are five stochastically generated slip distributions used in the database created by a process similar to Goda et al. (2018).

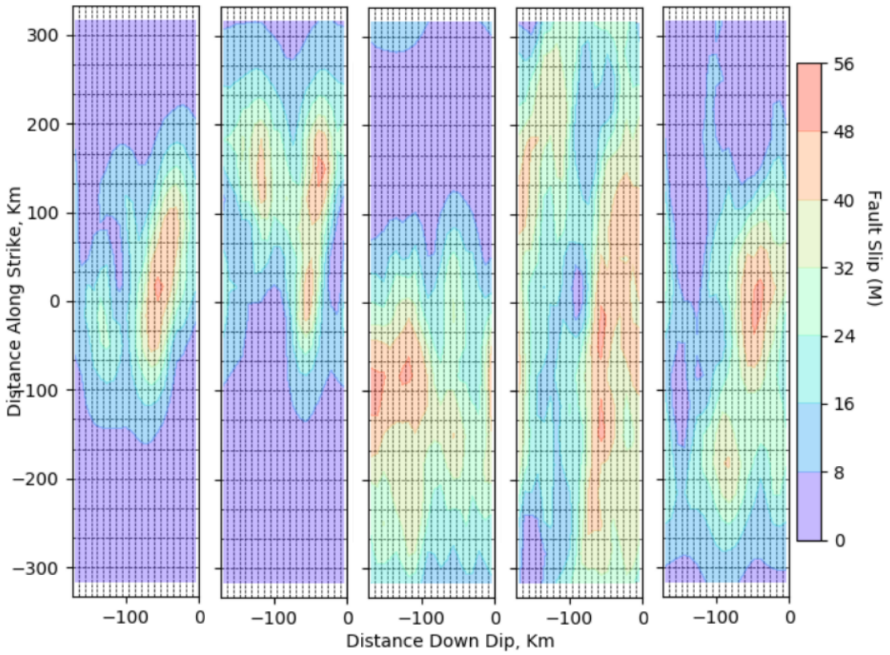

Figure 2: Five randomly generated slip distributions used in this study made by introducing correlations, setting maximum slip targets, and applying cutoffs.

The procedure used to create slip distributions such as the ones depicted in Figure 4.5 begins with the definition of the rectangular fault itself. The rectangular fault defined here stretches ∼660 km from Northern Washington to Southern Oregon with a width of about ∼170 km. The upper most edge of the fault follows the Cascadia fault line and extends downward at a shallow angle of ten degrees towards the East where it reaches a depth of 30k. The fault is split up into a 20x20 grid of rectangular subfaults, totaling 400. Each rectangular subfault is given its own individual slip value, but the direction of the slipping is unidirectional.

The process of assigning slip values to each of the 400 subfaults is as follows. First, each subfault is given a value from a Gaussian distribution of mean 10 m and standard deviation 60 m. Negative slip values will be generated from this distribution but they will be handled later. Then, a portion of the fault is chosen along the latitudinal direction. The slip values within this range are then multiplied by an arbitrarily large factor such as 20 as used here. This is meant to prioritize a randomly selected region of the fault, as one cannot assume the entire fault will rupture at once. For example, in Figure 2, the left most slip distribution was generated by amplifying the slip values from -150 km to 200 km.

After the rupture is localized, correlations are introduced to the 20x20 grid of subfaults by way of the Fourier filter technique, which takes a grid of random values and transforms them into the frequency domain where the slope of the power spectral decay is manually set by choosing a Hurst number. This technique has also been implemented by Løvholt et al. (2012) where 500 heterogeneous slip realizations were generated for an analysis on he variability of runup.

The process of introducing correlations is as follows. After the grid has been initialized by some set of semi-randomly generated values h_{ij}, a fast Fourier transform of the 2D grid is taken, giving H_{ij}. First, the radial wave number k_{ij} is defined as

$$k_{ij} = \sqrt{1 + i^2 + j^2} \tag{1}$$

Then, the Fourier transformed values H_{ij} are scaled by the factor $k_{ij}^{-\beta/2}$. This is written as

$$H'_{ij} = H_{ij}/k_{ij}^{\beta/2} \tag{2}$$

where β depends on the Hurst parameter H and is defined as

$$\beta = 2(H + 1) \tag{3}$$

The Hurst parameter is set to one in this study, indicating maximal correlation for this grid.

An inverse fast Fourier transform is then applied to obtain the new correlated slip values h'_{ij}. Afterwards, the current slip distribution is scaled such that its peak reaches a target maximum slip of 50 m. The negative slip values are then set to 0 to prevent slip in the opposite direction.

In order to obtain the surface displacement of the seafloor due to the slipping of the rectangular subfaults, analytical solutions of dislocation in an elastic half-space are used (Okada, 1985). Here, only the vertical component of displacement is considered for the tsunami initial condition.

The distribution of the moment magnitudes M_W produced by the procedure above for all 3000 earthquakes can be seen in Figure 3.

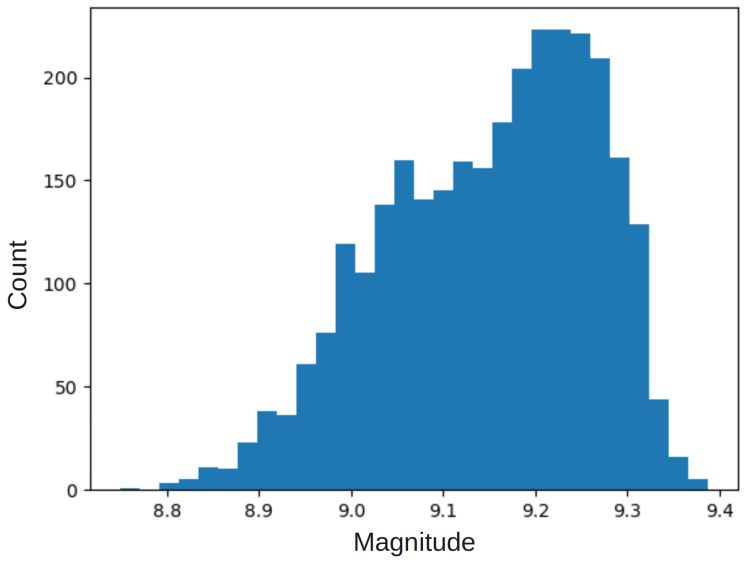

Figure 3: Distribution of magnitudes for all 3000 earthquakes in the earthquake database. Magnitudes range from ∼8.8 to ∼9.4.

Since the mean slip values were not adjusted to be log-normal, the distribution of magnitudes is not expected to be normal. The magnitudes range from 8.8 to nearly 9.4, representing a large Cascadia Subduction Zone rupture. The machine learning algorithm's predictability will therefore be restricted to earthquakes of magnitudes within this range.

3.2 Tsunami and Inundation Simulations

The tsunami and inundation simulations were both carried out by Tsunami Squares (TS), a wave and inundation simulator which uses a cellular automata approach to equivalently solve non-linear wave equations, making it computationally efficient. It has been tested against numerous historic tsunamis and landslides as well as other proven simulators to verify its accuracy (Xiao et al., 2015; Wang et al., 2019; Wilson et al., 2020). In addition, the recent improvements made to TS increase the resolution of the waveform, allowing more information to be preserved from the initial condition. For the purposes of creating a database of tsunami

and inundation simulations, TS is a highly suitable as it is computationally fast and now has sufficient resolution to produce distinguishable tsunami simulations with similar initial conditions. Distinguishability is very important in the creation of such a database because it ensures every unique earthquake event is paired uniquely with its own unique tsunami, decreasing the number of overlapping events (Mulia et al., 2018, 2020). Computational speed is also important in the creation of a database as it increases the number of total simulations which can be used to train the neural network. For this study, computational power was limited to a single Intel Core i7-8700 CPU (3.20GHz) node with 12 cores, making computational efficiency a necessity. Similar studies involving a database of inundation simulations are usually limited to a few hundred events (Mulia et al., 2020; Fauzi & Mizutani, 2020), as inundation simulations can be very costly to simulate since they rely on solving non-linear equations in order to propagate waves and runup.

The wave simulation is split up into two separate simulations: a large, low resolution simulation of the entire region and a small, high resolution simulation of the coast. It is typical to first simulate the tsunami wave in a low resolution bathymetry grid, then switch to a high resolution bathymetry grid for the coast where small details in the topography are more important for runup. This is to save considerable computation time and resources, as using high resolution bathymetry for the entire tsunami event would be very inefficient.

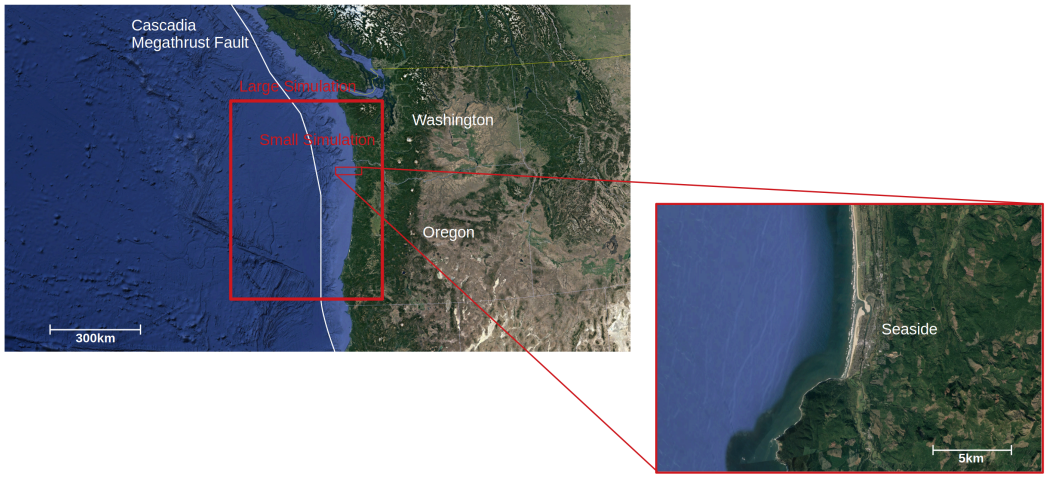

Figure 4: Map of the Cascadia Subduction Fault showing the simulated regions. A large, low resolution simulation is used to simulate the tsunami wave until it nears the coast. A smaller, high resolution simulation is then used for inundation. (Google, 2022).

Figure 4 shows how the CSZ is partitioned for the purpose of simulating. The larger simulation nearly spans the length of the Washington and Oregon coast and extends ∼500 km into the sea. Since it is so large, a 60 arc-second resolution bathymetry was used (N. N. G. D. Center, 2009). Also, to minimize computation time, a time step of four seconds was chosen which is the largest time step TS can allow given the constraints of the simulator. The wave was allowed to propagate for 200 time steps (13 minutes) where most waves are ∼50 km away from the coast. At this point, the last time step of the large simulation is used as the initial condition for the smaller simulation of the coast. For this, a high resolution bathymetry of the U.S. Northwest coast was used (N. N. G. D. Center, 2003). It is at a considerably higher resolution of three arc-seconds compared to 60 arc-seconds for the larger simulation and provides enough resolution for hazard assessment along the coast. As for the parameters of the coastal simulation, 1000 time steps (50 min) were simulated with a time step of three seconds, providing enough time for a given wave to reach the shore and drain back into the ocean. For both the low resolution and high resolution simulations, the computation time required to produce each simulation was around 20 minutes.

Figure 5 shows sequential frames for one of the high resolution inundation simulations contained in the database. In this example, inundation begins as soon as ∼25 minutes after the earthquake occurs which is followed by the main wave, striking ∼10 minutes later. An analysis of the database-average waveform confirms Figure 5 is a representative case as most tsunamis inundate the coast ∼25 minutes after the initial

earthquake.

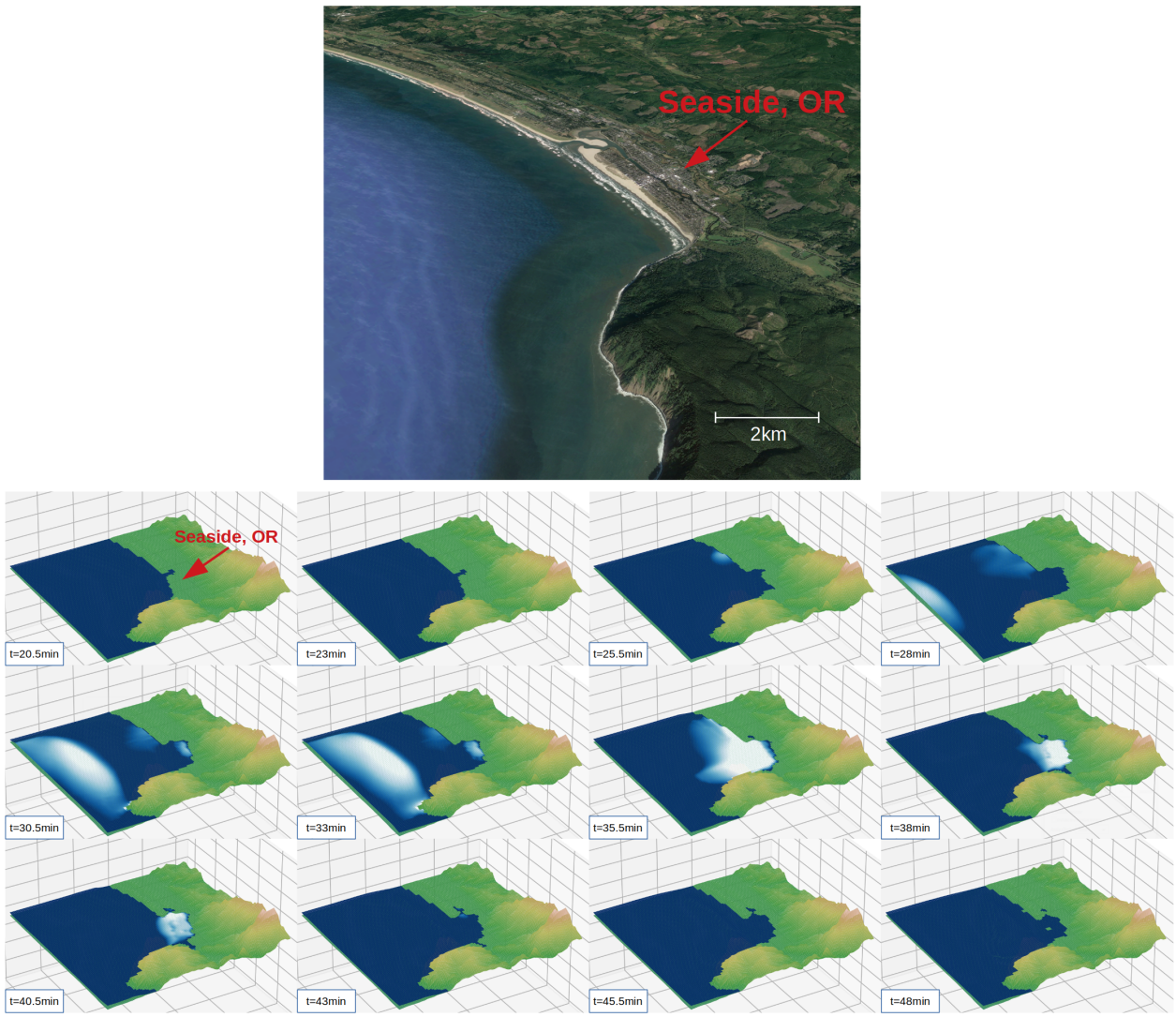

Figure 5: Simulation of the coast of Seaside, OR, USA for a randomly selected event in the database. Inundation begins ∼25 min after the earthquake has occurred, whereas the main wave strikes 10 min later. (Google, 2022).

Once the database of the high resolution simulations of the coast have been established, a map of the maximum flow depth is created for each simulation. Here, the flow depth is defined as the height of the water measured from the surface of the ground.

3.3 Buoy Data

Tsunami detection buoys are designed to measure tsunami wave heights in the open ocean generated by underwater earthquakes. They are deployed specifically for the purpose of early detection of tsunamis to aid in the issuing of an accurate warning. In its standard mode of operation, wave height data is recorded every 15 minutes. However, when a large seismic event occurs, the buoys switch into event mode where wave height data is recorded every 15 seconds. The simulated buoy data used here assumes event mode.

For a given tsunami in the database, each buoy can be represented by a time series consisting of 50 data points, each separated by 15 seconds totaling ∼13 minutes. Each buoy's time series is then stacked to form the input that the neural network will see. More on the data formatting will be discussed in further sections.

NOAA's National Data Buoy Center has a map of all the buoys currently in service for the West Coast of the United States. These locations were referenced to test how well a neural network could forecast inundation based on existing infrastructure. The 14 buoys which were referenced are shown in Figure 4.16. These do not represent all of the available buoys in the simulated area. Some were removed based on their distance from the site of interest. In addition, any buoys which were very close to each other were seen as duplicates and were similarly omitted.

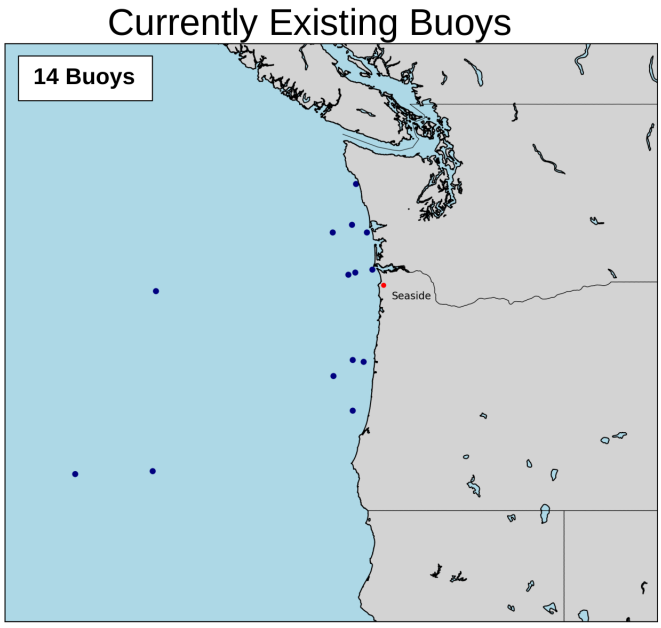

Figure 6: Currently operational wave height measurement buoys near the site of interest, Seaside, Oregon. (N. N. D. B. Center, 2022)

To test the capabilities of the inundation forecasting method, different numbers of hypothetical sensors in various configurations were used. Square grids with various separation distances were tested to explore how inundation prediction error changes with increasingly tight grids. The spacing of the tightest grid configuration was chosen to be 18 km, with every subsequent grid spacing expanding by 9 km. In total, 18 grid spacings were tested ranging from 18 km to 173 km with the total number of buoys ranging from $\sim$10 to $\sim$800. Figure 4.17 shows four examples of the differently spaced grids used in this study.

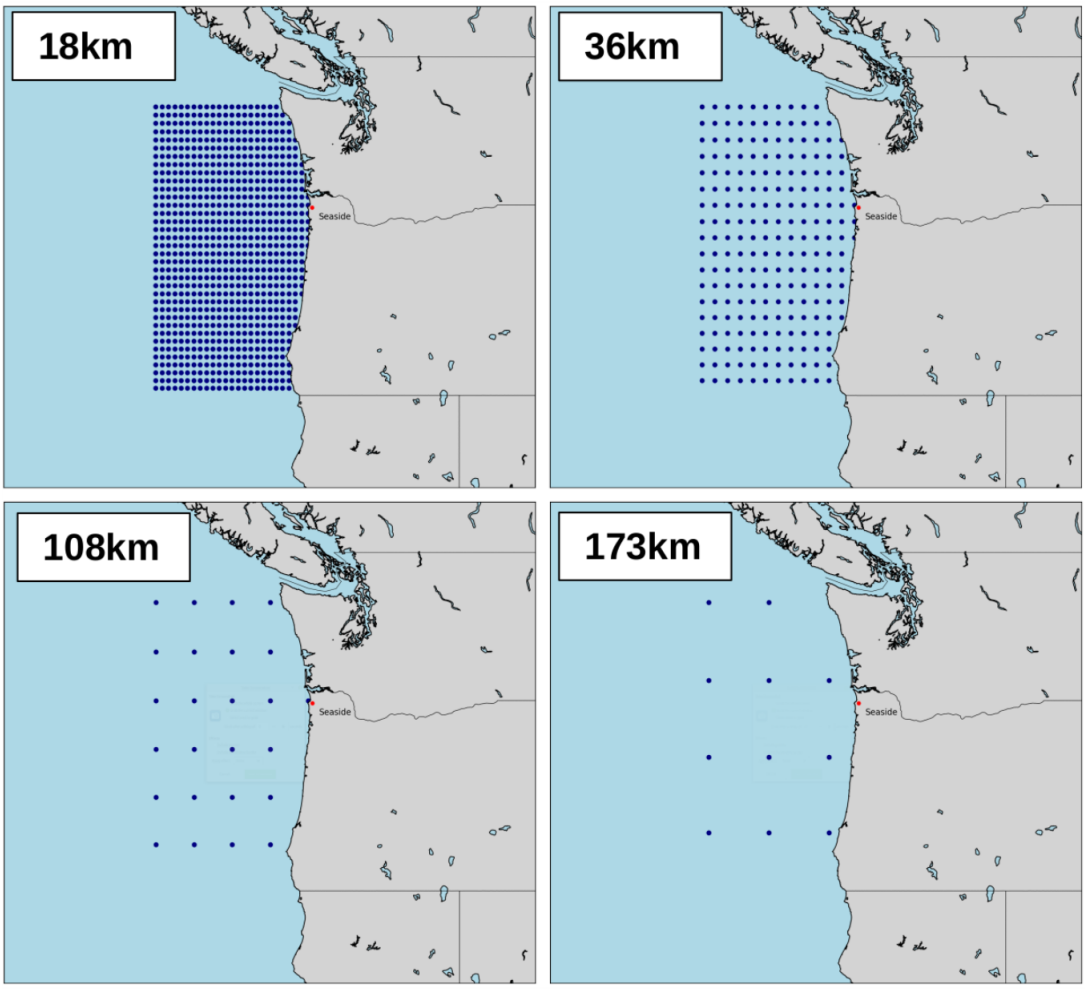

Figure 7: 4 of the 18 hypothetical buoy configurations used to create input data for the neural network. Grid spacings range from 18 km to 173 km.

3.4 Convolutional Neural Network

In order to convert the buoy data into an inundation forecast, a convolutional neural network (CNN) was utilized and implemented in PyTorch (Paszke et al., 2017). CNNs mainly specialize in pattern recognition and image classification (Lo et al., 1995; Lawrence et al., 1997; Ciregan et al., 2012; Garcia & Delakis, 2004; Sermanet et al., 2013; Russakovsky et al., 2015). It has seen uses in facial recognition, natural language processing, X-ray image analysis, object detection, and many other areas. A CNN was chosen mainly for its potential to detect patterns in the time series arrays containing wave height data (buoy data) used in this study.

Typical CNN models consist of layers constructed in the following order: an input layer, one or more convolutional layers with pooling layers mixed in, two or more fully connected layers, an output layer. The CNN architecture developed for this study follows this general structure and can be seen below in Figure 8.

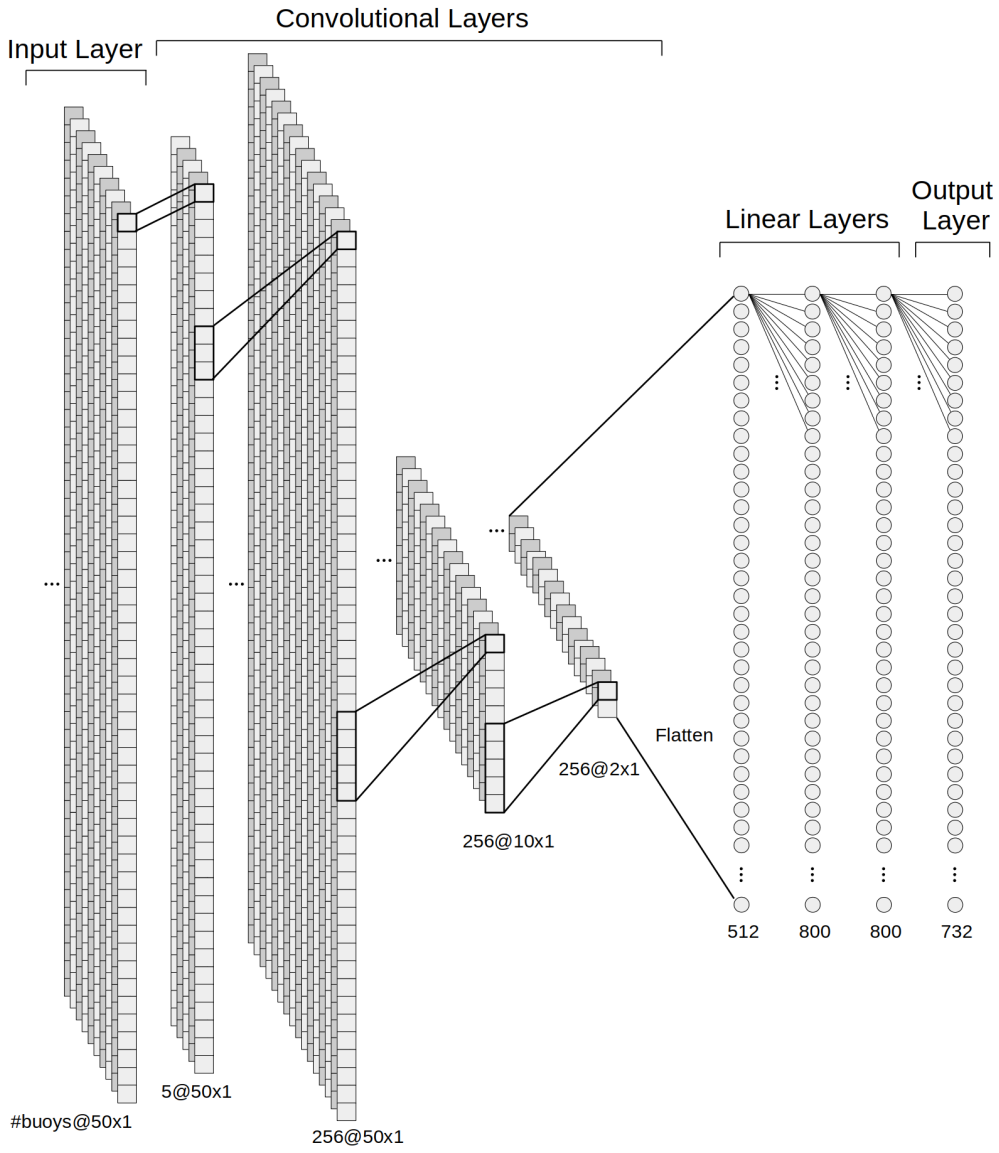

Figure 8: The CNN architecture. Each column in the input layer consists of the time series from one buoy, where the depth (or number of channels) equals the number of buoys. Data is passed through the subsequent convolutional layers where it is reduced to a flattened array of length 512. This is then passed on to two fully connected layers where a flattened inundation map is produced.

Layer	**Kernel Size**	**Stride**	**Padding**	**Input Channels**	**Output Channels**
Conv	1	1	0	# Buoys	5
Conv	3	1	1	5	256
Conv	5	5	0	256	256
Conv	5	5	0	256	256
Linear	None	None	None	512	800
Linear	None	None	None	800	800
Output	None	None	None	800	732

Table 1: The parameters of each layer for the CNN presented in Figure 8. In between each layer is a ReLU function.

The CNN architecture in Figure 8 was inspired by the neural networks found in Makinoshima et al. (2021) and Toledo-Marín et al. (2021), but was ultimately determined by trial and error.

The CNN in Figure 8 consists of four convolutional layers and two fully connected layers. The first convolutional layer has a kernel size of 1x1 and is meant to act as a filter across channels (Lin et al., 2013). Since the kernel extends down through all channels, the weights assigned to the 1x1 kernel are equivalent to a weight assigned to the channel itself. If the optimization algorithm decides one channel (or one buoy) is particularly important, it will give that channel a high weight. In essence, the 1x1 convolutional layer used here takes the number of input channels (or input buoys) and distills them to five channels. Since the number of buoys varies from 10 to 800 depending on which grid is used (Figure 7), the addition of this type of convolutional layer was critical to ensure that the following layers will receive a consistent amount of parameters. This removes the neural network structure as a confounding variable in the comparison of accuracy between differently sized inputs.

The following layer uses a kernel size of three with a padding value of one to preserve the length of each channel. This layer outputs 256 channels of length 50. The following two layers are convolutional pooling layers and act to downsample the data. To achieve this pooling effect, the kernel and stride of the filters were both set to five to downsample the length from 50 to 10 in the first layer, then from 10 to 2 in the second layer. The convolutional pooling layer approach to downsampling was chosen over the traditional maximum pooling operation solely due to an empirical improvement in performance. The 256 channels of length two are then flattened into a 1D array of 512 elements and fed into the two linear layers of length 800, where it is finally transformed into a 732 pixel inundation map by the output layer. To prevent overfitting, dropout with a probability of 0.001 was implemented directly before the output layer. Dropout refers to the process of temporarily ignoring, or "dropping out" different nodes in a given layer during each epoch of the training process. The network is therefore training on a different randomly generated thinned version of the network at every epoch. This has the effect of combining many different neural network architectures together to produce an averaged output, allowing the network to generalize well as it approximates an ensemble approach (Srivastava et al., 2014).

Regarding the hyperparameters of the CNN, a learning rate of 0.0005 was chosen and trained for a total of 800 epochs. At the end of the 800 epochs, the weights from the epoch with the best validation accuracy was chosen. A mean squared error cost function was chosen as it performed better compared to the least absolute error cost function. The Adam Optimizer Algorithm (Kingma & Ba, 2014) was used for gradient descent.

4 Results and Discussion

4.1 Various Grid Configurations

Here we discuss the inundation forecasting results from the buoy data used as input for the CNN. In addition to using existing buoys located off the coast of Oregon, several hypothetical arrangements of buoys were also used to test the capabilities of this method. The hypothetical arrangements consist of grids of buoys spanning the simulated region, each with their own buoy-to-buoy separation distance. In total, 18 grids were tested with separation distances ranging from 18 km to 171 km as shown in Figure 7.

Figure 9 shows three plots containing results for the average absolute error, the average percent error, and the percent difference in coverage. For these measures of error, both predicted and expected inundation heights below 15 cm are ignored. Otherwise, a measure of the average error will likely return falsely accurate since the calculation will be diluted by a large number of closely matching predicted and expected heights of around zero.

The equation used to calculate average event error for each event j is given as

$$E_j = \frac{1}{N_j} \sum_{i=0}^{N_j} |h_i - \hat{h}_i| \qquad \text{if } \hat{h}_i > 0.15 \text{ m or } h_i > 0.15 \text{ m} \tag{4}$$

where E_j is the average error for the j^{th} event in the testing set, $\hat{h}_i$ is the expected inundation height of the i^{th} pixel, h_i is the predicted height of the i^{th} pixel, and N_j is the total number of pixels in event j which satisfy the Boolean expression.

The percent error for each event j is defined by

$$\text{Percent E}_j = \frac{100}{N} \sum_{i=0}^{N} \frac{|h_i - \hat{h}_i|}{\hat{h}_i} \qquad \text{if } h_i > 0.15 \text{ m} \tag{5}$$

where $\hat{h}_i$ is the expected inundation height of the i^{th} pixel, h_i is the predicted height of the i^{th} pixel, and N_j is the total number of pixels in event j which satisfy the Boolean expression.

The percent difference in coverage for each event j is defined by

$$\text{Percent Difference}_j = 100 \cdot \frac{|n - \hat{n}|}{\hat{n}} \tag{6}$$

where $\hat{n}$ represents the total number of inundated pixels in the expected inundation map and n represents the total number of inundated pixels in the predicted inundation map for a given event j in the testing set. An inundated pixel is defined as one with a maximum height greater than 0.15 m.

To display the error for all 300 events in the testing set, the 10^{th}, 50^{th}, and 90^{th} percentiles were calculated to give an accurate representation of the error distribution. For reference and comparison purposes, the results using the existing buoys are shown in the form of black horizontal lines.

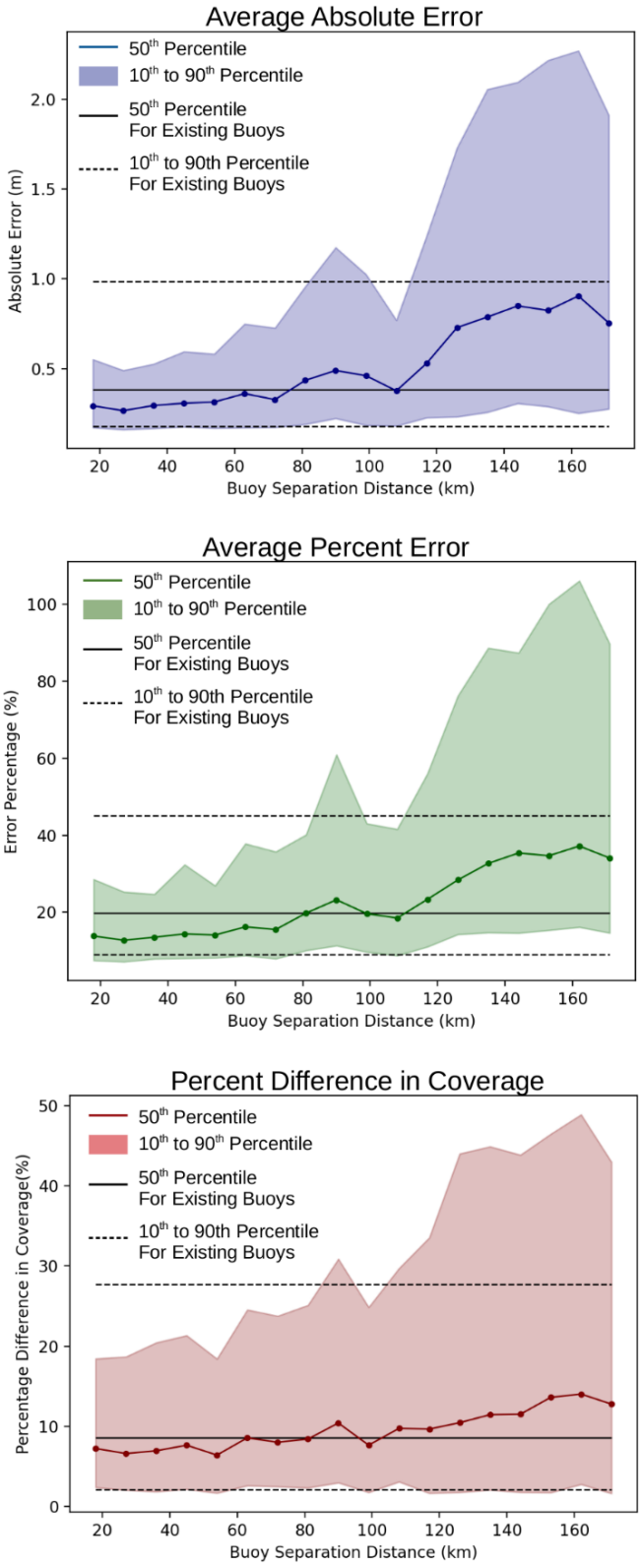

Figure 9: Measures of error as a function of buoy separation distance (increasing sparseness). Percentiles for the results from existing buoy locations are indicated by solid and dotted black lines. Results most notably indicate 90^{th} percentiles increase significantly with increasing buoy sparseness.

The average absolute error plot shown in Figure 9 shows an upwards trend in the 50^{th} percentile. This is expected since as the buoy separation distance increases, the total number of buoys decreases, providing the CNN with less data from which to make a prediction. A tight grid of buoys (separation distance 20-40 km) produces a median error of ∼0.25 m and the existing buoys produce a slightly higher median error of ∼0.35 m. In fact, the median error of the grids outperform existing buoys until a separation distance of ∼80 km. However, the difference between 0.25 m and 0.35 m error is not very significant. Focus is therefore given to the 90^{th} percentile, which is ∼0.5 for tight grids and ∼1.0 for existing buoys. For larger buoy separation distances, the 90^{th} percentile increases dramatically to over 2 m. This means for configurations with sparse grids (existing buoys and 100-180 km separation distances), the chances of a high error prediction are significant.

As with the average absolute error plot, the percent error results show a steadily increasing trend in the 50^{th} percentile. A tight grid of buoys produce a median percent error of ∼15% while the existing buoys produce ∼20%. However, although the median percent errors are similar, the 90^{th} percentile of the existing buoys is significantly higher at ∼45% compared to the tight grid configurations at 35%. This indicates that a tight grid of buoys will ensure accurate predictions for almost all events in the testing set while the existing buoys have the potential to produce a prediction with nearly 50% error.

The percent difference in coverage is mainly consistent for the 50^{th} percentile at ∼10%. The only significant trend can seen in the 90^{th} percentile where it generally increases from 20% to 50%.

An interesting feature of plots in Figure 9 can be seen in the form of dips in the error present at separation distances of 110 km and 170 km. This is due to some grid configurations containing buoys positioned very near to the point of interest. These buoys offer critical information the CNN needs in order to make a more accurate inundation prediction, thus decreasing the error for grid configurations with these buoys.

Figures 10, 11, and 12 show forecasts for six randomly selected events from the test set for three buoy configurations: existing buoys, a grid with an 18 km separation distance, and a grid with a 162 km separation distance.

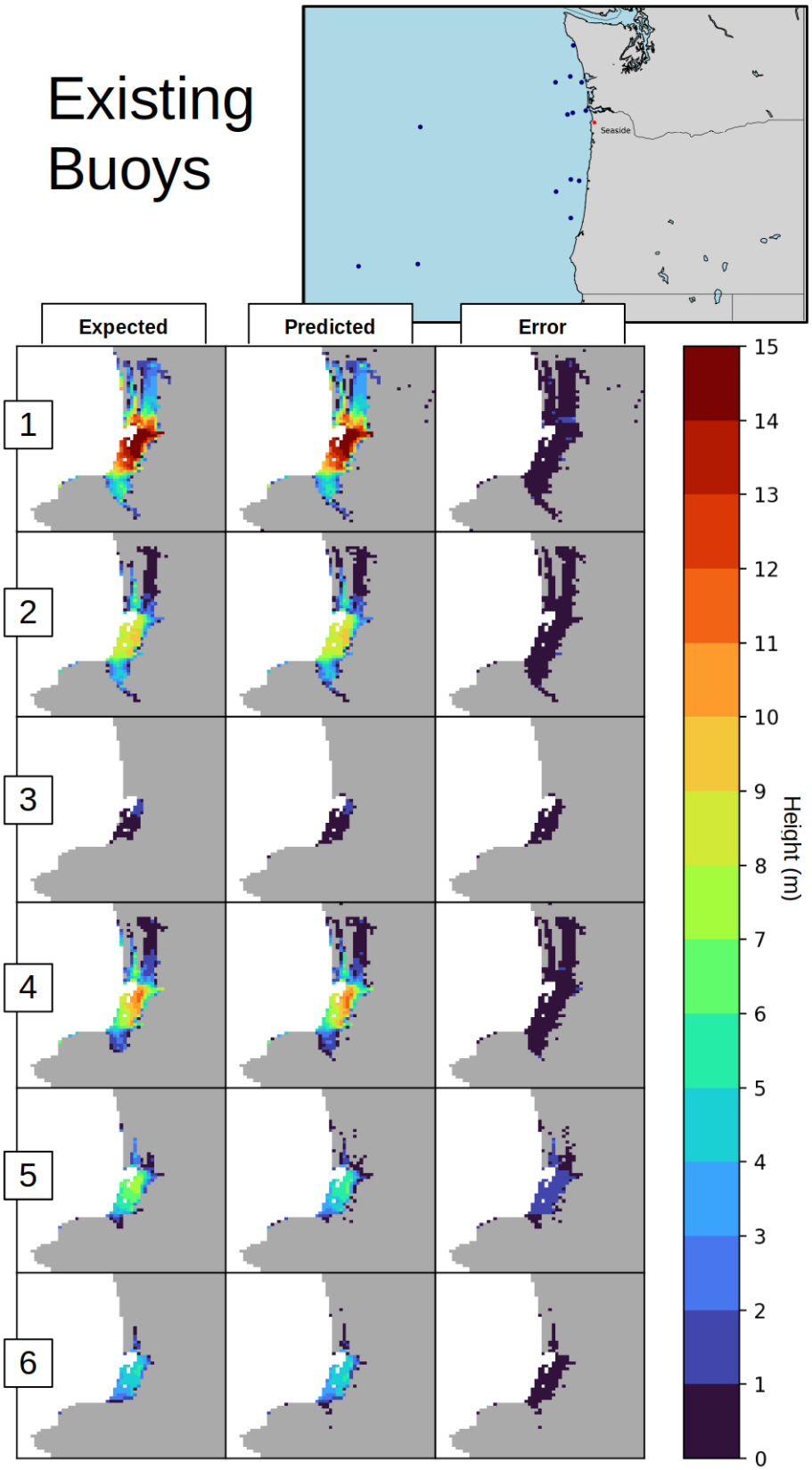

Figure 10: Results for six randomly selected events in the testing set using existing buoys.

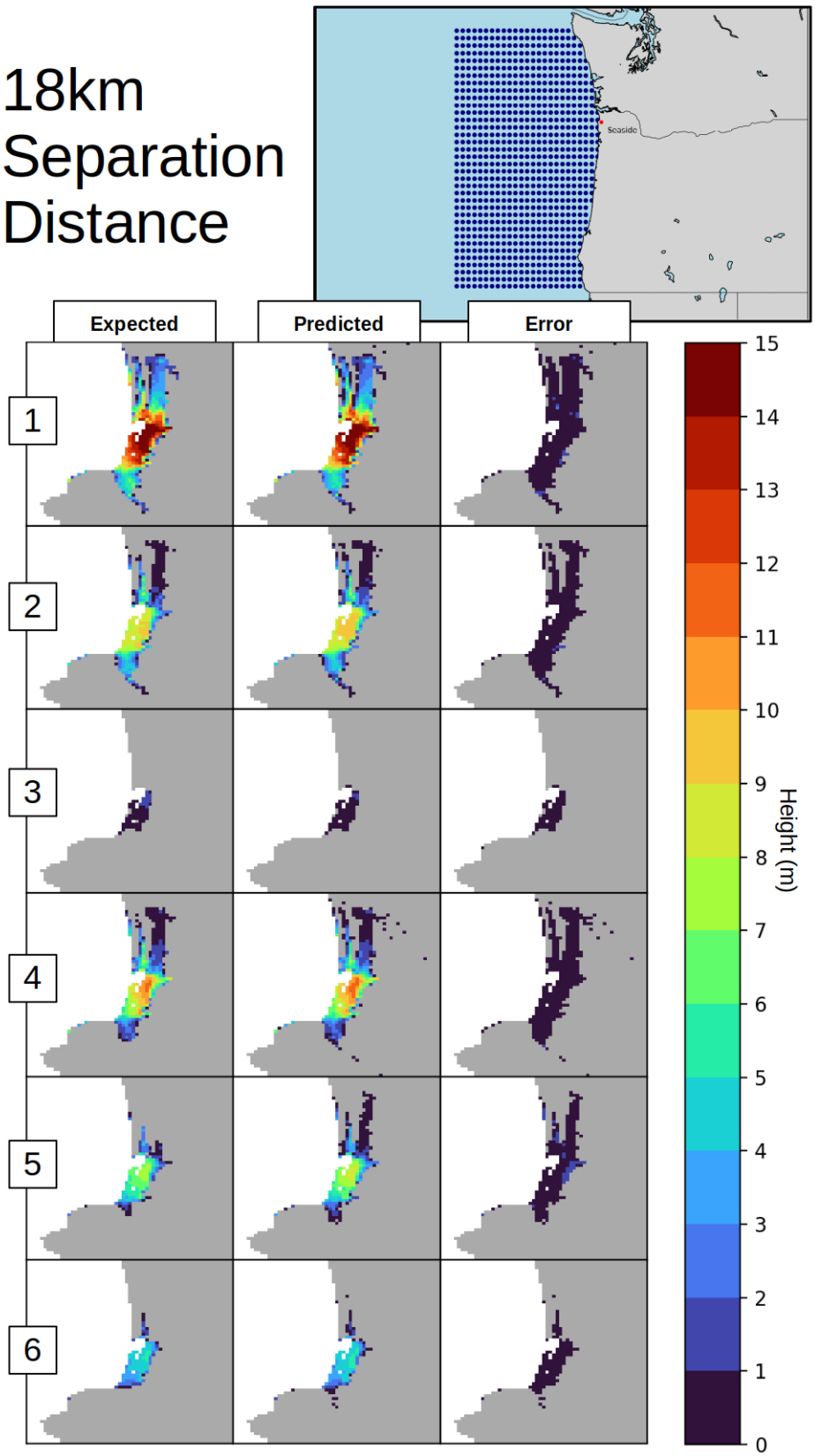

Figure 11: Results for six randomly selected events in the testing set using a hypothetical grid of buoys with 18 km separation distance.

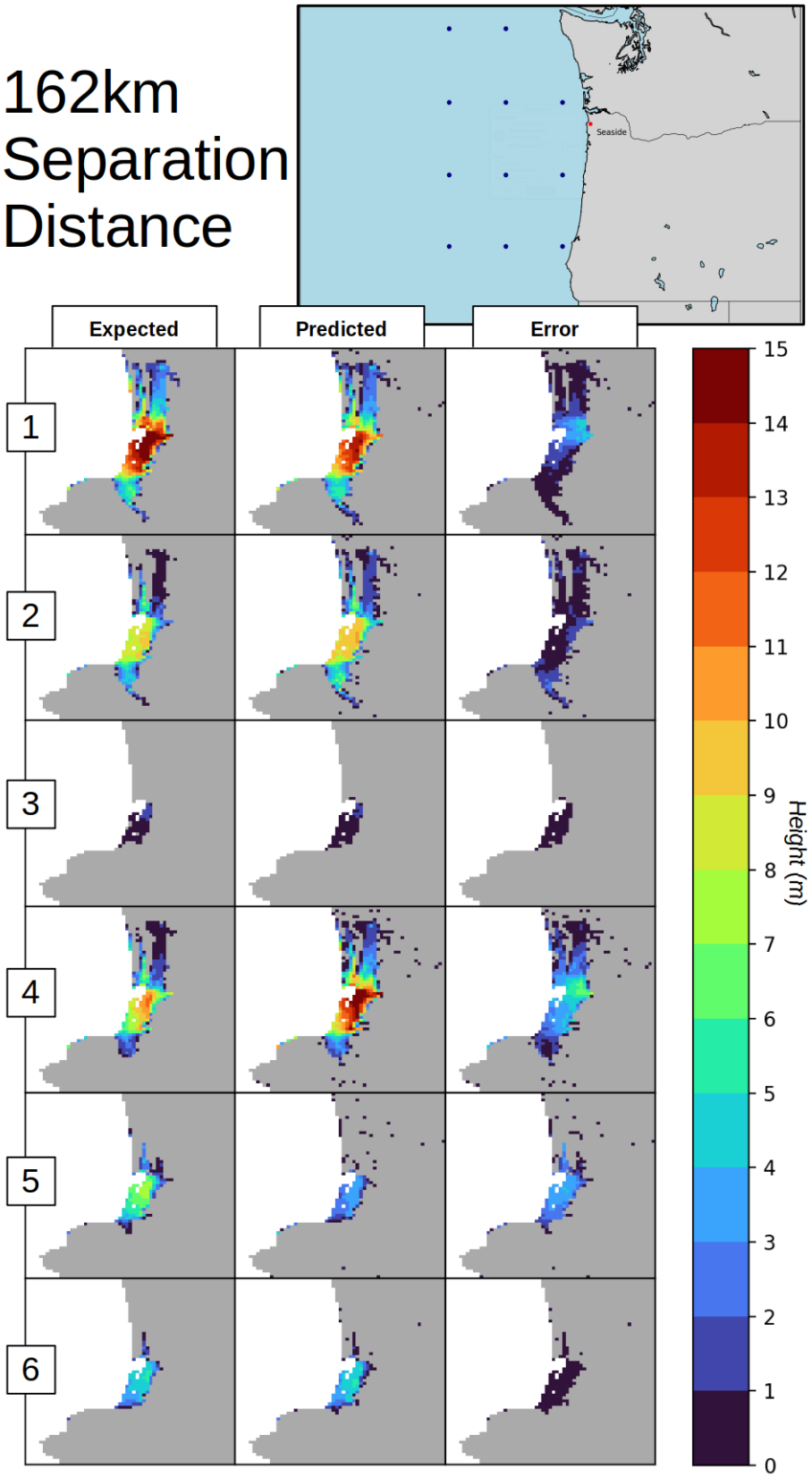

Figure 12: Results for six randomly selected events in the testing set using a hypothetical grid of buoys with 162 km separation distance, the arrangement with the highest error.

Overall, the results from the six randomly selected events using the existing buoys (Figure 10) seem acceptable. The error remains under 1 m for the most part with the exception of event #5 which exhibits error in the 1-2 m range. The results from the 18 km separation distance grid (Figure 11) show more consistency as the errors for all events shown are below 1 m. We begin to see erroneous predictions when viewing the results from the 162 km separation distance grid (Figure 12). With the exception of two out of the six events shown, the remaining four events display error above 2 m, with one event reaching up to 8 m.

Since a random selection of events does not reveal the true accuracy of these models, is valuable to analyze the predictions with the highest absolute error and the highest percent error. The results can be seen in Figure 4.34.

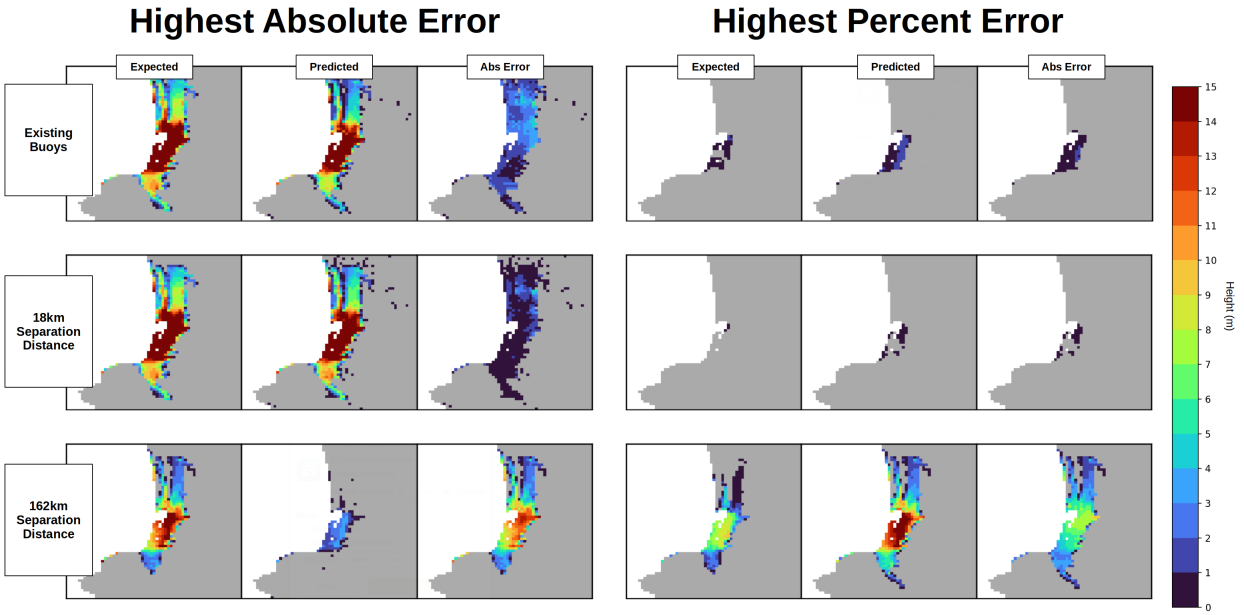

Figure 13: The events with the highest absolute error and the highest percent error for the models using existing buoys, an 18 km separation distance grid, and a 162 km separation distance grid.

It is clear from Figure 13 that both the highest absolute error prediction and the highest percent error prediction using the 162 km grid display significant deviation from the expected. This means given a particular event, there is a chance for a severely inaccurate prediction. Both the existing buoys and the tight arrangement of buoys perform moderately well in their respective worst case scenarios. However, the tight arrangement of buoys perform slightly better in its worst case. This is in line with Figure 9 which shows a lower 90^{th} percentile for the tight grid of buoys compared to the existing buoys. It can also again be concluded that although the existing buoys perform similarly to a tight grid of buoys for a random selection of events, the severity of the most erroneous predictions decrease with a tight grid.

It is important to note that by the nature of a percent error formula, high percent error events will naturally tend towards those with relatively low inundation heights. Using the same logic, events with high absolute error tend towards those with already high inundation heights. This tendency can be seen for the worst cases of both the existing buoys and the 18 km grid. However, the 162 km grid does not follow this tendency because the errors are so high.

4.2 Sensitivity Analysis

In the tests below, we aim to find an arrangement of the minimal amount of buoys required to achieve results equivalent to that of a dense network of tightly spaced buoys. The number of buoys contained in the grids tested above range from 10 for the sparsest arrangement to over 800 for the tightest arrangement. However, not every buoy is equally utilized by the CNN when predicting inundation. The CNN naturally prioritizes buoys which contain more predictive power and as a result, only a few are necessary. This

prioritization can be measured in terms of sensitivity, a measure of how much of an effect one input variable has on the accuracy of the output.

To explore the sensitivity of the CNN with respect to its input buoy data, an occlusion test was performed (Makinoshima et al., 2021; Zeiler & Fergus, 2014; Nielsen & Voigt, 2018). It works by omitting one input parameter at a time and observing its effect on the accuracy. In this case, we first take the trained CNN for a particular grid configuration and run it without any modifications on the test set to obtain the base average error. Then, the values of a selected buoy's time series are set to zero before giving it to the CNN to obtain a value of the modified average error. The ratio of the modified average error to the base average error is then calculated and scaled to a range of one to two. A sensitivity value of one would indicate the buoy has no effect on the prediction, while a value of two would indicate the buoy has strong influence on the prediction. This process is repeated for every buoy in the grid to obtain a map displaying relative sensitivity. The sensitivity maps for the eight tightest grid arrangements are shown in Figure 14. To help aid in the visualization of the placement of the buoys, the peak of the database-averaged tsunami waveform at the end of the buoy observational period were overlaid (13 min after the start of the simulation).

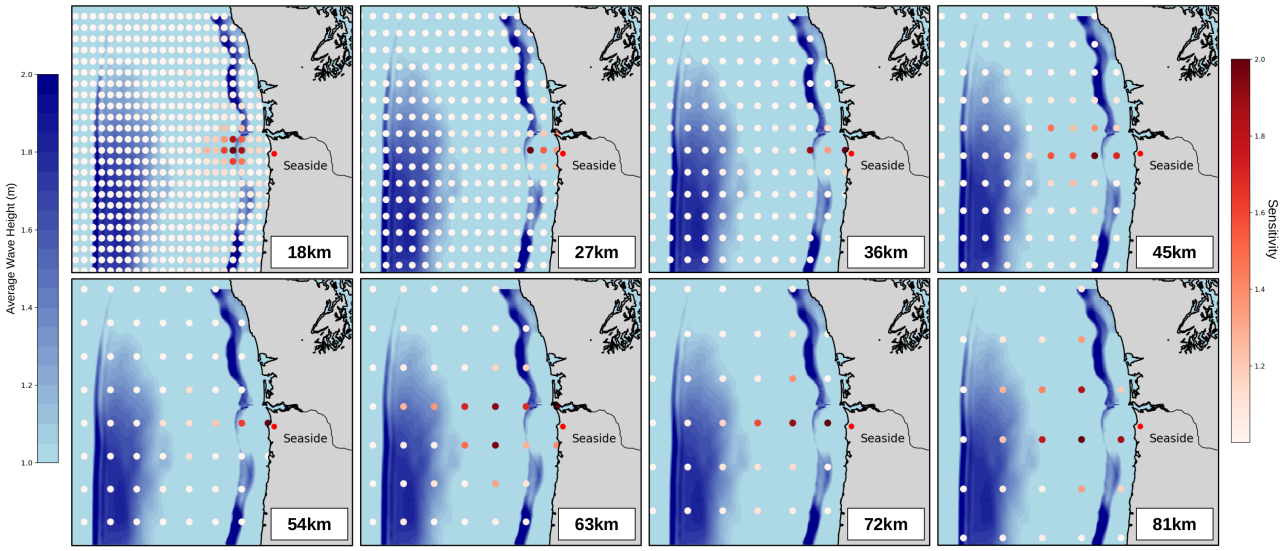

Figure 14: Buoy sensitivity tests for the eight tightest grids. Overlaid is the database-averaged waveform at the end of the buoy observational period (13 min). The CNN is most sensitive to buoys near the peak of the waveform at 13 min or buoys near the site of interest.

An interesting result can be seen in the upper left hand plot in Figure 14 showing the sensitivity map with the tightest buoy spacing. It indicates that if given access to an extremely dense grid of wave height time series data, the CNN will prioritize buoys at the location where the wave is at its peak (at the time the buoys stop collecting data). This sensitive region generally persists with increasing buoy separation with the exception of 36 km and 54 km, where the most sensitive buoys are the ones positioned nearest the site of interest. This is an important result as it indicates the CNN can gain similar accuracy by either prioritizing near the site of interest, or near the peak of the waveform. The positioning of the buoys relative to the site of interest is also important. From buoy spacings of 27 km, 36 km 45 km, 54 km, and 72 km, the most sensitive buoys were ones in a line extending perpendicularly out from the site of interest. The two grid spacings which did not share this characteristic, 63 km and 81 km, split the sensitivity between two parallel lines adjacent to the site of interest.

From this, further testing was done to find a configuration which uses a minimal amount of buoys while still maintaining the accuracy of a tight grid of buoys. The decided upon configuration can be seen in Figure 15. It consists of four buoys, spaced 27 km apart, arranged in a line extending out to sea from the point of interest.

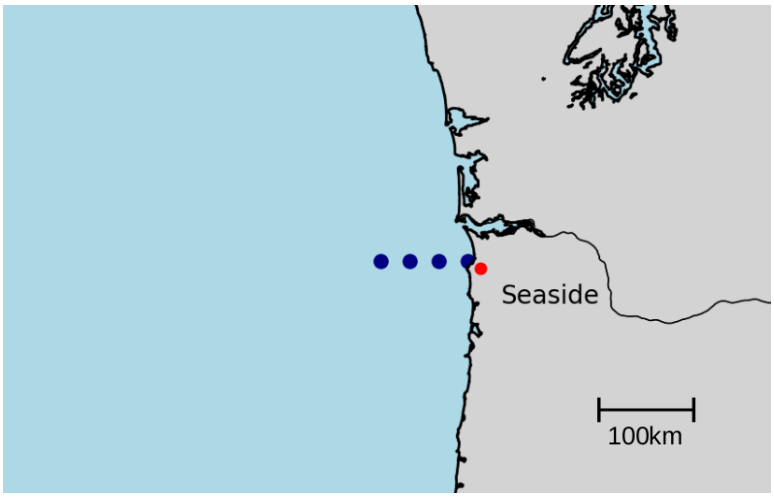

Figure 15: A minimal arrangement of buoys tested with aims to achieve similarly accurate results as a tight grid. Chosen based on sensitivity measurements (Figure 4.35) are four buoys spaced 27 km apart extending out to sea from the site of interest.

The results for the four buoy configuration in Figure 15 can be seen below in a layout similar to that of Figure 9. However, instead of the black horizontal lines representing existing buoys, they now represent the results of the four buoys arranged in a line extending out to sea.

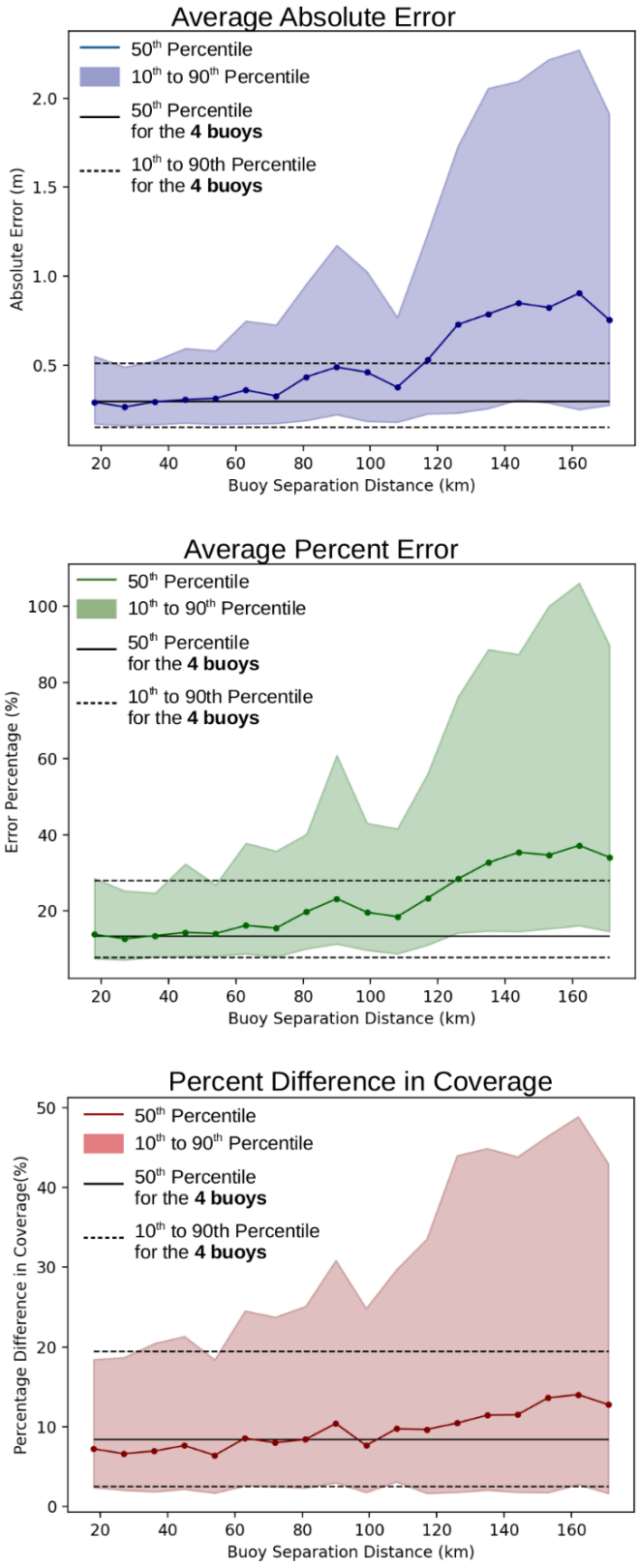

Figure 16: Plots identical to that of Figure 9 except with black lines now indicating results from the arrangement of four buoys (Figure 15). The four buoy arrangement achieves equally accurate results as the tightest grid spacings.

Results indicate that the accuracy using the four buoy arrangement for all three metrics are similar to that of a tight grid of buoys (separation distance <50 km). These results show that although a tight grid of buoys will produce inundation forecasts which are both accurate and consistent, the CNN does require all buoys in order to obtain an equally accurate forecast.

However, the findings here do not necessarily indicate that only four buoys are needed for any CSZ rupture scenario. Results are heavily confined by the locations and the magnitudes of the earthquake scenarios present in the database, namely earthquakes of magnitude ∼9 with epicenters very near the site of interest. Earthquakes outside these constraints may heavily affect which buoys are necessary. For example, if an earthquake with an epicenter far north of the site of interest were to occur, the wave would travel mainly south until reaching the site of interest. In this case, a line of buoys extending East to West would not capture the incoming waveform, thus leading to an erroneous or even missed prediction. In addition, the four buoy arrangement may not produce similar results given a different buoy observational window. In this study, 13 min of buoy data was collected before using it as input for the CNN. For shorter observational windows, buoys placed farther out to sea may be prioritized instead.

5 Conclusion

In this study, we presented a machine learning model (CNN) trained on buoy wave height data in order to forecast tsunami inundation due to a potential rupture of the Cascadia Subduction Zone (CSZ) for the town of Seaside, Oregon. Due to the short distance between the CSZ and the coast of Oregon, an early and accurate tsunami warning is crucial. This method introduces an alternative to existing direct non-linear forward modeling in real-time to obtain inundation forecasts, where rapid tsunami source estimations and significant computational resources are needed to simulate nonlinear tsunami propagation.

Work done here modifies the CNN and input data structure of Makinoshima et al. (2021) to forecast maximum runup height maps similar to that of Fauzi & Mizutani (2020) and Mulia et al. (2020). Instead of confining the experiments to existing buoys, various hypothetical buoy configurations were tested to explore the predictive power of this method. The study here also produces earthquakes using a procedure similar to that of Goda et al. (2018), where correlations and target maximums were introduced to the slip map to produce more realistic rupture scenarios.

Tests using currently operating buoy stations were found to produce accurate inundation forecasts with a median error of ∼0.5 m. When using a hypothetical dense grid of buoys as input to the CNN, it was found to decrease the magnitude and frequency of large error predictions. It was also found through a sensitivity analysis that only a small number of buoys were needed by the CNN to produce accurate forecasts. Results indicate that a configuration of four strategically placed buoys were sufficient to obtain results as accurate as a dense grid of buoys for earthquakes (M∼9.0) localized near the Oregon coast for the city of Seaside.

The proposed CNN method using time series buoy data shows promise to produce fast and accurate inundation predictions. In addition, the method of determining the minimal number of necessary buoys through a sensitivity measure proved to be successful. In the future, we aim to expand the database to include a wider range of earthquake magnitudes and epicenter locations to more accurately cover the total possible rupture scenarios for the Cascadia Subduction Zone. The length of the buoy observational period will also be varied to determine its effect on accuracy, exploring the possibility of a continuous prediction based on the amount of data collected.

6 Acknowledgments

Research by DPG and JBR was supported under NASA grant NNX12AM22G, and by DoE grant DE-SC0017324 to the University of California, Davis. Research by GCF was supported by grant NSF NSF2212550 and is gratefully acknowledged. Portions of this research were also carried out at the Jet Propulsion Laboratory, California Institute of Technology under contract with NASA.

References

Ando, M., Ishida, M., Hayashi, Y., & Mizuki, C. (2011). Interviews with survivors of tohoku earthquake provide insights into fatality rate. *Eos, Transactions American Geophysical Union*, *92*(46), 411–412.

Atwater, B. F., Musumi-Rokkaku, S., Satake, K., Tsuji, Y., & Yamaguchi, D. K. (2011). *The orphan tsunami of 1700: Japanese clues to a parent earthquake in north america*. University of Washington Press.

Atwater, B. F., Nelson, A. R., Clague, J. J., Carver, G. A., Yamaguchi, D. K., Bobrowsky, P. T., ... others (1995). Summary of coastal geologic evidence for past great earthquakes at the cascadia subduction zone. *Earthquake spectra*, *11*(1), 1–18.

Bureau, U. S. C. (2022). *U.s. census website.* `https://www.census.gov/`. (date accessed: 2022-05-15)

Center, N. N. D. B. (2022). *Station 46089 - tillamook, or.* `https://www.ndbc.noaa.gov/station_page.php?station=46089`. (date accessed: 2022-04-07)

Center, N. N. G. D. (Ed.). (2003). *U.s. coastal relief model vol.8 - northwest pacific.* NOAA National Centers for Environmental Information. (date accessed: 2021-1-05) doi: 10.7289/V5H12ZXJ

Center, N. N. G. D. (Ed.). (2009). *Etopo1 1 arc-minute global relief model.* NOAA National Centers for Environmental Information. (date accessed: 2021-1-05) doi: 10.7289/V5C8276M

Ciregan, D., Meier, U., & Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. In *2012 ieee conference on computer vision and pattern recognition* (pp. 3642–3649).

Clague, J. J. (1997). Evidence for large earthquakes at the cascadia subduction zone. *Reviews of Geophysics*, *35*(4), 439–460.

Dominey-Howes, D., Dunbar, P., Varner, J., & Papathoma-Köhle, M. (2010). Estimating probable maximum loss from a cascadia tsunami. *Natural hazards*, *53*(1), 43–61.

F., D. (2011). `https://www.web.archive.org/web/20161023052841/http://www.panoramio.com/photo/62067486`. (date accessed: 2022-04-02)

Fauzi, A., & Mizutani, N. (2020). Machine learning algorithms for real-time tsunami inundation forecasting: a case study in nankai region. *Pure and Applied Geophysics*, *177*(3), 1437–1450.

Fiedorowicz, B., & Peterson, C. (2002). Tsunami deposit mapping at seaside, oregon, usa. *Geoenvironmental Mapping*, *1*, 630–648.

Frank, D. (n.d.). `http://www.donfrankphotography.com/`. (date accessed: 2022-06-02)

Garcia, C., & Delakis, M. (2004). Convolutional face finder: A neural architecture for fast and robust face detection. *IEEE Transactions on pattern analysis and machine intelligence*, *26*(11), 1408–1423.

Goda, K., Yasuda, T., Mai, P. M., Maruyama, T., & Mori, N. (2018). Tsunami simulations of mega-thrust earthquakes in the nankai–tonankai trough (japan) based on stochastic rupture scenarios. *Geological Society, London, Special Publications*, *456*(1), 55–74.

González, F., Geist, E. L., Jaffe, B., Kânoğlu, U., Mofjeld, H., Synolakis, C., ... others (2009). Probabilistic tsunami hazard assessment at seaside, oregon, for near-and far-field seismic sources. *Journal of Geophysical Research: Oceans*, *114*(C11).

Google. (2022). *Google earth.* `http://www.google.com/earth/index.html`. (date accessed: 2022-03-20)

Gusman, A. R., Tanioka, Y., MacInnes, B. T., & Tsushima, H. (2014). A methodology for near-field tsunami inundation forecasting: Application to the 2011 tohoku tsunami. *Journal of Geophysical Research: Solid Earth*, *119*(11), 8186–8206.

Heaton, T. H., & Hartzell, S. H. (1987). Earthquake hazards on the cascadia subduction zone. *Science*, *236*(4798), 162–168.

Imamura, F., & Anawat, S. (2011). Damage due to the 2011 tohoku earthquake tsunami and its lessons for future mitigation. In *Proceedings of the international symposium on engineering lessons learned from the* (pp. 1–4).

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kubota, T., Suzuki, W., Nakamura, T., Chikasada, N. Y., Aoi, S., Takahashi, N., & Hino, R. (2018). Tsunami source inversion using time-derivative waveform of offshore pressure records to reduce effects of non-tsunami components. *Geophysical Journal International*, *215*(2), 1200–1214.

Lander, J. F., Lockridge, P. A., & Kozuch, M. J. (1993). *Tsunamis affecting the west coast of the united states, 1806-1992* (No. 29). US Department of Commerce, National Oceanic and Atmospheric Administration .

Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, *8*(1), 98–113.

Lin, M., Chen, Q., & Yan, S. (2013). Network in network. *arXiv preprint arXiv:1312.4400*.

Lo, S.-C. B., Chan, H.-P., Lin, J.-S., Li, H., Freedman, M. T., & Mun, S. K. (1995). Artificial convolution neural network for medical image pattern recognition. *Neural networks*, *8*(7-8), 1201–1214.

Løvholt, F., Pedersen, G., Bazin, S., Kühn, D., Bredesen, R. E., & Harbitz, C. (2012). Stochastic analysis of tsunami runup due to heterogeneous coseismic slip and dispersion. *Journal of Geophysical Research: Oceans*, *117*(C3).

Mai, P. M., Schorlemmer, D., Page, M., Ampuero, J.-P., Asano, K., Causse, M., ... others (2016). The earthquake-source inversion validation (siv) project. *Seismological Research Letters*, *87*(3), 690–708.

Makinoshima, F., Oishi, Y., Yamazaki, T., Furumura, T., & Imamura, F. (2021). Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks. *Nature communications*, *12*(1), 1–10.

Mueller, C., Power, W., Fraser, S., & Wang, X. (2015). Effects of rupture complexity on local tsunami inundation: Implications for probabilistic tsunami hazard assessment by example. *Journal of Geophysical Research: Solid Earth*, *120*(1), 488–502.

Mulia, I. E., Gusman, A. R., & Satake, K. (2018). Alternative to non-linear model for simulating tsunami inundation in real-time. *Geophysical Journal International*, *214*(3), 2002–2013.

Mulia, I. E., Gusman, A. R., & Satake, K. (2020). Applying a deep learning algorithm to tsunami inundation database of megathrust earthquakes. *Journal of Geophysical Research: Solid Earth*, *125*(9), e2020JB019690.

Musa, A., Watanabe, O., Matsuoka, H., Hokari, H., Inoue, T., Murashima, Y., ... Kobayashi, H. (2018). Real-time tsunami inundation forecast system for tsunami disaster prevention and mitigation. *The Journal of Supercomputing*, *74*(7), 3093–3113.

Nielsen, A. A., & Voigt, C. A. (2018). Deep learning to predict the lab-of-origin of engineered dna. *Nature communications*, *9*(1), 1–10.

Oishi, Y., Imamura, F., & Sugawara, D. (2015). Near-field tsunami inundation forecast using the parallel tunami-n2 model: Application to the 2011 tohoku-oki earthquake combined with source inversions. *Geophysical Research Letters*, *42*(4), 1083–1091.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, *75*(4), 1135–1154.

Park, H., Alam, M. S., Cox, D. T., Barbosa, A. R., & van de Lindt, J. W. (2019). Probabilistic seismic and tsunami damage analysis (pstda) of the cascadia subduction zone applied to seaside, oregon. *International Journal of Disaster Risk Reduction*, *35*, 101076.

Park, H., & Cox, D. T. (2016). Probabilistic assessment of near-field tsunami hazards: Inundation depth, velocity, momentum flux, arrival time, and duration applied to seaside, oregon. *Coastal Engineering*, *117*, 79–96.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... Lerer, A. (2017). Automatic differentiation in pytorch.

Peterson, C. D., Cruikshank, K. M., Jol, H. M., & Schlichting, R. B. (2008). Minimum runup heights of paleotsunami from evidence of sand ridge overtopping at cannon beach, oregon, central cascadia margin, usa. *Journal of Sedimentary Research*, *78*(6), 390–409.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... others (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), 211–252.

Satake, K. (2014). The 2011 tohoku, japan, earthquake and tsunami. *Extreme Natural Hazards, Disaster Risks and Societal Implications*, *1*, 310–321.

Seaside map. (2020). `https://www.ontheworldmap.com/usa/city/seaside/`. (date accessed: 2022-04-01)

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929–1958.

Toledo-Marín, J. Q., Fox, G., Sluka, J. P., & Glazier, J. A. (2021). Deep learning approaches to surrogates for solving the diffusion equation for mechanistic real-world simulations. *Frontiers in Physiology*, *12*.

Tsushima, H., Hino, R., Ohta, Y., Iinuma, T., & Miura, S. (2014). tfish/rapid: Rapid improvement of near-field tsunami forecasting based on offshore tsunami data by incorporating onshore gnss data. *Geophysical Research Letters*, *41*(10), 3390–3397.

Tsushima, H., Hino, R., Tanioka, Y., Imamura, F., & Fujimoto, H. (2012). Tsunami waveform inversion incorporating permanent seafloor deformation and its application to tsunami forecasting. *Journal of Geophysical Research: Solid Earth*, *117*(B3).

Tsushima, H., Hirata, K., Hayashi, Y., Tanioka, Y., Kimura, K., Sakai, S., ... Maeda, K. (2011). Near-field tsunami forecasting using offshore tsunami data from the 2011 off the pacific coast of tohoku earthquake. *Earth, planets and space*, *63*(7), 821–826.

Wang, J., Ward, S. N., & Xiao, L. (2019). Tsunami squares modelling of the 2015 june 24 hongyanzi landslide generated river tsunami in three gorges reservoir, china. *Geophysical Journal International*, *216*(1), 287–295.

Wilson, J. M., Schultz, K. W., Grzan, D., Rundle, J. B., Ward, S. N., Bhaskar, R., ... Kaushal, H. (2020). Tsunami squares simulation of megathrust-generated waves: Application to the 2011 tohoku tsunami. *Progress in Disaster Science*, *5*, 100063.

Xiao, L., Ward, S. N., & Wang, J. (2015). Tsunami squares approach to landslide-generated waves: application to gongjiafang landslide, three gorges reservoir, china. *Pure and Applied Geophysics*, *172*(12), 3639–3654.

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818–833).