DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

Applying Sampling Methods to Parameter Reduced Bathymetry Data in Tsunami Simulation

Daniel Baur

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Bachelor's Thesis in Informatics

Applying Sampling Methods to Parameter Reduced Bathymetry Data in Tsunami Simulation

Anwendung von Samplingverfahren auf Parameterreduzierte Bathymetriedaten in Tsunami-Simulation

Author:	Daniel Baur
Supervisor:	Michael Bader
Advisor:	Anne Reinarz
Submission Date:	16.11.2020

I confirm that this bachelor's thesis in informatics is my own work and I have documented all sources and material used.

Munich, 16.11.2020 Daniel Baur

Abstract

The simulation of natural phenomena like tsunamis is a time- and resource-intensive task, in no small part due to the size of the required input parameter space. It is in our interest to reduce said parameter space while retaining a reasonable degree of simulation accuracy. Building on results gained in [Wal20], we examine reconstructions produced by a deep feature consistent variational autoencoder. We compare cutouts of the GEBCO 2020 dataset as well as the 2011 Tohoku dataset against their reconstructions based on a multiple criteria: Lowest points, mean and standard deviation as well as different error metrics and their relation to the bathymetric gradient of our datasets. To further test the reconstruction accuracy for actual application we run a solitary wave simulation on our datasets and again compare the results by plotting water height at a buoy point. We use both a visual plot and the time frequency misfit method to determine the accuracy of our reconstructed datasets for tsunami simulation. Since our simulations produce reasonable results we conclude that the model proposed in [Wal20] provides us with sufficiently accurate reconstructions of our datasets for the purpose of tsunami simulation.

Keywords: *Parameter Reduction, Bathymetry, Tsunami Simulation, GEBCO, ExaHyPE, Time Frequency Misfit, Variational Autoencoder*

Contents

Abstract **iii**

1 Introduction **1**
1.1 Impetus 1
1.2 Scope 2

2 Tools **3**
2.1 ExaHyPE 3
2.2 Sebastian Walter's Master Thesis 5
2.3 Time Frequency Misfit 7
2.4 GEBCO 9

3 Initial Evaluation of Output Data **10**
3.1 The GEBCO 2020 Dataset 10
3.1.1 Lowest Points 10
3.1.2 Mean and Standard Deviation 11
3.1.3 Error Metrics and Gradient 14
3.2 The Tohoku 2011 Dataset 17

4 Application of Sampling Methods to Bathymetric Data **18**
4.1 The Tohoku 2011 Tsunami Simulation 18
4.2 Solitary Gaussian Wave Simulation 22

5 Conclusion **29**

List of Figures **30**

List of Tables **32**

Bibliography **33**

1 Introduction

The study of natural phenomena like earthquakes and tsunamis has always been a highly relevant topic of human observation. For people living in areas affected by such catastrophes analyzing their behaviour and effects is quite literally a matter of life and death. With the rise of modern computational technology new avenues have been opened, especially in the realm of accurate simulations to help understand and predict these phenomena.
This is not, however, without its problems. Chief among these is the fact that the amount of input parameters required for accurate results is usually staggeringly high, making simulation of these phenomena a time- and resource-intensive task [Sul15]. This of course poses a challenge in terms of computation and in turn creates the impulse to find ways to reduce the input parameter space to something more manageable. Such reductions do however carry the risk of distorting our datasets, making it impossible to draw accurate conclusions. The outcome of simulations is generally heavily affected by bathymetric differences if said differences occur in coastal areas with strong bathymetric gradients, while divergences around the source of disturbance (in case of e.g. a tsunami) are less impactful but can still be significant [TS96].
It is therefore of paramount importance that we test our models to ensure that simulations run on the reduced parameters are still representative of our original data or rather, how big the effect of divergences between original and reconstruction are.

1.1 Impetus

This thesis is based on the Master's thesis of Sebastian Walter [Wal20], which evaluated the usage of deep learning for parameter reduction in bathymetric data. While the method proposed in Mr. Walter's thesis showed promise in terms of accuracy, an in-depth evaluation of the results did not take place at the time. It falls to us to examine whether the presented model holds up when tested using a larger selection of datasets and simulations.

1.2 Scope

This thesis aims to further evaluate Mr. Walter's results. To do this we will first analyze cutouts selected from the GEBCO dataset in regards to their differences. We then run a number of wave simulations on the same datasets to determine the actual impact of the difference between original and reconstruction.
In Chapter 2 we give an overview of the tools used for this thesis, which include brief summaries of Sebastian Walter's Master's thesis [Wal20], the time frequency misfit method and the ExaHyPE engine.
In Chapter 3 we discuss possible pitfalls in regards to the bathymetric data we are working with. We also carry out a rudimentary evaluation of the reconstruction algorithm to get an initial picture of the reconstructed datasets' accuracy in simple simulations. For this purpose we analyze our datasets along the lines of mean, standard deviation as well as different error metrics and how they correspond to the bathymetric gradient of said datasets.
Finally, in Chapter 4 we reconstruct the results of the tsunami simulation described in [Wal20] and apply solitary wave simulations to our previously defined bathymetry. We evaluate the results between original and reconstructed datasets to ascertain whether the reduced data is sufficiently accurate as to not negatively affect our simulations.

2 Tools

This chapter's aim is to give a brief overview of the tools that were used in the writing of this thesis. In practice this means a short introduction to ExaHyPE, which was used to apply sampling methods to our bathymetry data as well as a brief description of Sebastian Walter's Master Thesis, the practical aspect of which this thesis is meant to evaluate. Additionally we give a brief overview of the time frequency misfit criteria for quantitive comparison of time signals described by Kristeková et al [KKM09].

2.1 ExaHyPE

ExaHyPE (**E**xascale **H**yperbolic **P**DE **E**ngine) is an open source engine that is, as the name suggests, used to solve first-order hyperbolic partial differential equations.

ExaHyPE offers a comprehensive set of key features, including High-order ADER-DG [TMN01], dynamic mesh refinement on cartesian grids and addition of user-provided code. It also offers a suite of post-processing and plotting options, such as support for different output formats, to allow quick plotting in a dedicated software such as ParaView [Rei+19].
A particular benefit of ExaHyPE is that it allows the user to largely avoid having to interact with the underlying solvers. Instead parameters are set in a specification file which is then used by ExaHyPE to generate the necessary glue code.
ExaHyPE solves equations of the following form:

$$\frac{\partial}{\partial t}\mathbf{Q} + \nabla \cdot \mathbf{F}(\mathbf{Q}, \nabla \mathbf{Q}) + \mathbf{B}(\mathbf{Q}) \cdot \nabla \mathbf{Q} = \mathbf{S}(\mathbf{Q}) + \sum_{i=1}^{n_{ps}} \delta_i$$

- $\mathbf{Q}$: State vector
- $\mathbf{F}$: Flux tensor
- $\mathbf{B}(\mathbf{Q})$: Non-conservative flux
- $\mathbf{S}(\mathbf{Q}$: Sources
- $\sum_{i=1}^{n_{ps}} \delta_i$: Point sources

This formulation is quite flexible and allows to model a wide range of wave-based applications, but we only care about the implementation of shallow water equations in this case. The SWE can be written as

$$\frac{\partial}{\partial t}\begin{pmatrix} h \\ hu \\ hv \\ b \end{pmatrix} + \frac{\partial}{\partial x}\begin{pmatrix} hu \\ hu^2 + 0.5gh^2 \\ huv \\ 0 \end{pmatrix} + \frac{\partial}{\partial y}\begin{pmatrix} hv \\ huv \\ hv^2 + 0.5gh^2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ hg \cdot b_x \\ hg \cdot b_y \\ 0 \end{pmatrix} = 0$$

- **h**: height of the water column
- **u, v**: horizontal flow velocity
- **b**: bathymetry
- **g**: gravity
- **x, y**: partial differentiation

We generally summarize (h,hu,hv,b) as vector **Q**.
For our application of sampling methods, we make use of both ADER-DG and Finite Volume (FV) Solvers. To avoid numerical oscillations we apply A-Posteriori limiting as described by Reinarz et al in [Rei+19]. If a solution computed by the ADER-DG solver is deemed troubled by the limiter during a given timestep, it is reevaluated by the FV solver using values from earlier steps. Due to the implementation of its components the state of our SWE solver is well-balanced.
Of particular interest to us is that ExaHyPE offers functionality for working with NetCDF files via a number of external libraries. Through simple yaml configuration files we gain the ability to extract bathymetric data from our original files and make said data available for our solvers.
To top it off ExaHyPE already comes prepared with a number of simple scenarios for the SWE solvers such as a solitary wave or wetting-and-drying problem. This in combination with NetCDF support will allow us to rapidly test simulations on a variety of datasets [Sch+15].

2.2 Sebastian Walter's Master Thesis

Mr. Walter's thesis [Wal20] explores multiple deep learning approaches to reduce the parameter space of the GEBCO dataset while maintaining the ability to reconstruct the values in the original space. To achieve this it uses autoencoders (AEs) as discussed by Rumelhart et al [RHW86]. AEs are neural networks that compress their original input into a lower dimensional space and try to reconstruct it with as little loss of information as possible. This process aims to keep only the most important features of the input data.
One of the approaches Walter explores in his thesis is the use of a variational autoencoder (VAE) as described in [KW14]. The VAE is meant to reduce the bottleneck of regular AEs. Instead of simply scaling back up from the point of most reduction, the VAE uses variational inference to encode input as distribution over a latent space. Points from this latent space can then be sampled and scaled back up, making the VAE much more robust.
To further improve the VAE deep feature consistency can be used. A framework for this is proposed by Hou et al in [Hou+16]. A VAE is combined with a pre-trained deep convolutional neural network (CNN). The CNNs hidden features are used to define a feature conceptual loss for training of the VAE, meaning that the VAE is penalized by differences between it's own input and output. Hou et al conclude that this will force the VAE to achieve a more accurate reconstruction.

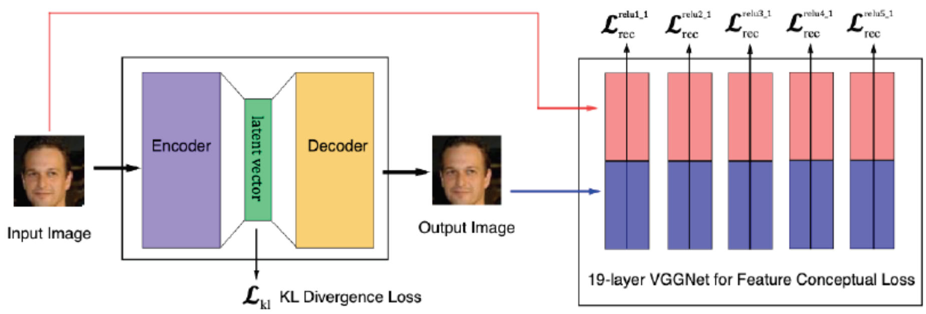

Figure 2.1: Deep feature consistent variational autoencoder, image taken from [Hou+16]

Walter comes to the conclusion that a variational autoencoder with deep feature consistency (DFC VAE) does in fact produce the most accurate results.

To affirm this in the context of bathymetric simulations he evaluates a simulation of the 2011 Tohoku tsunami dataset on the original bathymetric data and a reconstruction of said data. To compare the datasets he plots the water level at a buoy position over 7500 seconds (shown in 2.2). A simulation over flat bathymetry is also included to showcase the effects of drastically different bathymetry on the results. While we can see differences especially as time goes on, the reconstruction is accurate in the higher amplitudes and the differences later on are still only within the range of about one meter. Based on these results Walter concludes to use the DFC VAE model for further analysis.

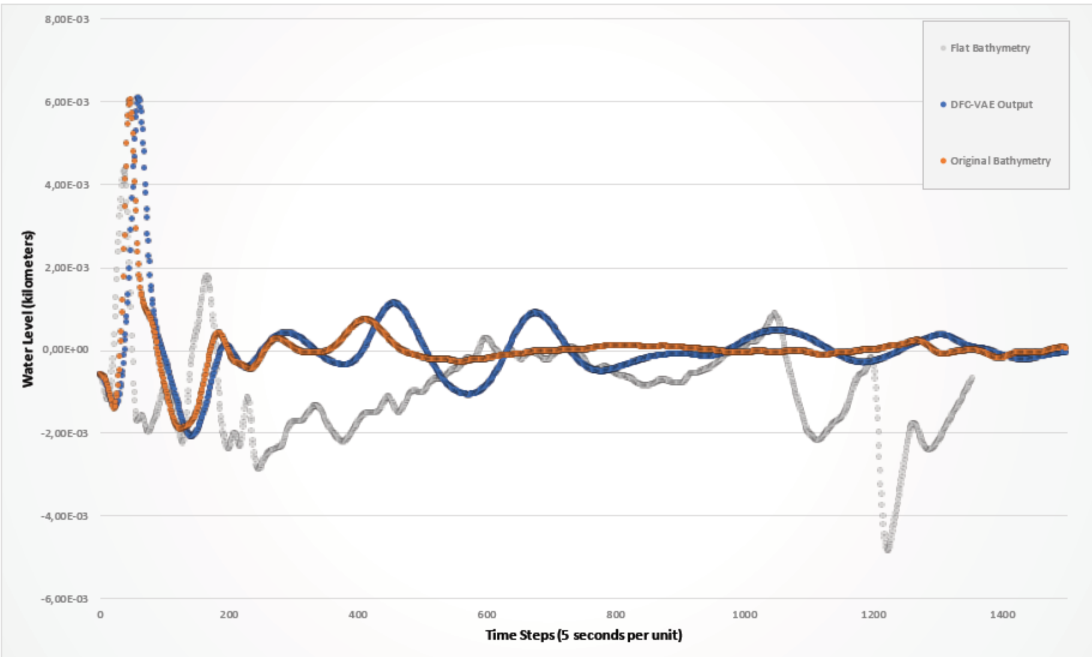

Figure 2.2: Water height at fixed coordinate resulting from tsunami simulations [Wal20]

2.3 Time Frequency Misfit

To get a better view of whether the signals measured for a given pair of original and reconstructed datasets are similar and whether their differences matter for us we use the Time Frequency Misfit comparison for time signals described by Kristeková et al in [KKM09].
For all further descriptions we define $s(t), sr(t)$ as signal and reference signal respectively, with t denoting time. The time-frequency representation of a given signal is used to obtain its time evolution at any frequency. We can obtain it via the continuous wavelet transform, which is defined for s(t) as follows:

$$CWT_{(a,b)}\{s(t)\} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} s(t)\psi * (\frac{t-b}{a})dt \tag{2.1}$$

- **t**: Time
- **a**: Scale parameter
- **b**: Translational parameter
- ψ: Analyzing wavelet
- *: Complex conjugate function

The TF representation of s(t) is defined as W(t,f) = $CWT_{(f,t)}\{s(t)\}$. With this representation a signal $\phi(t,f)$ and envelope $A(t,f)$ of $s(t)$ are defined as

$$\phi(t,f) = Arg[W(t,f)], A(t,f) = |W(t,f)| \tag{2.2}$$

Furthermore, the difference between two envelopes/phases at a (t,f) point is defined as

$$\Delta A(t,f) = A(t,f) - Ar(t,f) = |W(t,f)| - |Wr(t,f)| \tag{2.3}$$

and

$$\Delta\phi(t,f) = Arg\left[\frac{W(t,f)}{Wr(t,f)}\right] \tag{2.4}$$

TF Misfit criteria are defined to quantify the relative difference between envelopes or phases at any given point. The envelope difference is normalized through the representation of its reference signal. The phase criterion is already correctly quantified, we can, however, divide (10) by π if we want to choose the range (-1,1) $(-\pi, \pi)$. We get

$$TFEM_{LOC}(t,f) = \frac{\Delta A(t,f)}{Ar(t,f)} \tag{2.5}$$

$$TFPM_{LOC}(t,f) = \frac{\Delta\phi(t,f)}{\pi} \tag{2.6}$$

Last but not least, global misfit criteria are defined to describe the difference at a certain point as well as said point's significance in regards to the maximum envelope/phase of the reference signal.

$$TFEM_{GLOB}(t,f) = \frac{Ar(t,f}{max_{t,f}\{Ar(t,f)\}} TFEM_{LOC}(t,f) \tag{2.7}$$

$$TFPM_{GLOB}(t,f) = \frac{Ar(t,f}{max_{t,f}\{Ar(t,f)\}} TFPM_{LOC}(t,f) \tag{2.8}$$

This last part is crucial for our purposes, since the global normalization allows us to gauge accuracy for the higher amplitudes of our signals, by giving less significance to differences in the smaller amplitudes. This is acceptable to us, since these differences have a comparatively smaller impact on the accuracy of our simulations.
Kristekova et al also include a scale for goodness of fit based on [And04]. We use these for evaluation since they are already tuned for earthquake-engineering applications. EM and PM are mapped to a rising scale from 0 to 10.

Misfit Envelope	Misfit Phase	Goodness-of-Fit Numerical value	Verbal value
± 0.00	± 0.0	10	
± 0.11	± 0.1	9	excellent
± 0.22	± 0.2	8	
± 0.36	± 0.3	7	good
± 0.51	± 0.4	6	
± 0.69	± 0.5	5	fair
± 0.92	± 0.6	4	
± 1.20	± 0.7	3	
± 1.61	± 0.8	2	poor
± 2.30	± 0.9	1	
± ∞	± 1.0	0	

Figure 2.3: Discrete goodness of fit values against misfit values, taken from [KKM09]

2.4 GEBCO

This thesis uses the General Bathymetric Chart of the Oceans (GEBCO) gridded bathymetric data set [20220]. The 2020 version is used for the bulk of this paper, though we also make use of the 2011 version for its dataset of the Tohoku tsunami. It has to be noted that, due to different projection methods, the two datasets differ from each other in major ways. This can be confirmed even by the naked eye as seen in figure 2.4. Furthermore, there is a clear improvement in quality between the reconstructions of the 2011 dataset and its 2020 equivalent.

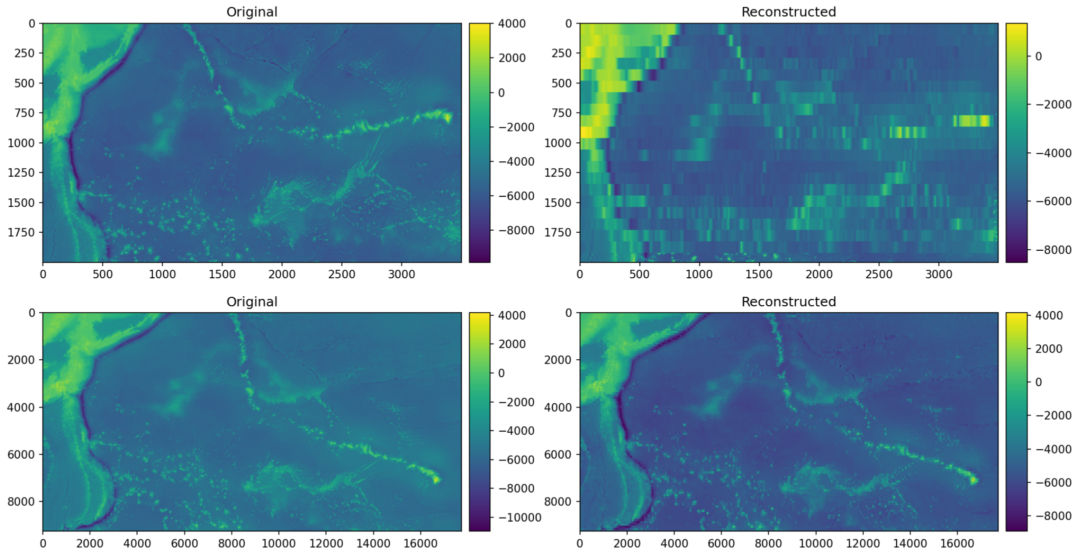

Figure 2.4: Comparison between cutouts of the Tohoku area from 2011 (top) and 2020 (bottom) as well as their reconstructions

The grid is available in the public domain and free to use. It provides a combination of land and seafloor topography. The land imagery is taken from the NASA Blue Marble dataset, while the bathymetric data is based on ship-track soundings. Interpolation between soundings is guided by satellite-derived gravity data. In case of availability, datasets with higher accuracy are included in the GEBCO dataset, which in turn means that measurement quality is not uniform for the entire set [Oce].

3 Initial Evaluation of Output Data

Our first step is to do a brief examination of the output data provided by the reconstruction algorithm. One key factor that has to be considered is the reconstructed output and its relation, or rather its lack thereof, to the real world. Our inputs consist of data points that provide a clear and meaningful context to real world phenomena which is essential for any simulations we might want to run, but the output is essentially just numbers without meaning.

Per Walter's thesis [Wal20] the reconstruction is meaningfully similar to the original dataset, though stark bathymetric features (e.g. cliffs, narrow canyons etc) can be eroded, meaning that the implementation can have difficulties with rapid changes in relative depth.

3.1 The GEBCO 2020 Dataset

The GEBCO dataset is both too large to work on as a whole and is in part unsuited for our purposes: Since we only want to look at bathymetric data, we have no need for data of the continental landmasses. To make it easier on us while still being able to draw accurate conclusions from our data, we define 100 small (in relation to the full dataset) cutouts of predominantly coastal and marine areas.

3.1.1 Lowest Points

To evaluate the accuracy of our output we used a python notebook with pre-loaded weights to generate 100 reconstructed data sets from randomly chosen selections of the GEBCO dataset.

In our first step we compare the lowest points of the original and reconstructed datasets. While this metric is not very meaningful on its own, it will allow us to reinforce conclusions drawn from other metrics like standard deviation, mean and mean squared error, later on.

Figure 3.1 shows our comparison of lowest points for all 100 datasets. In all cases there is at least some loss of depth at the lowest point. We can recognize that almost all datasets experience a loss of depth during the reconstruction process. In most of our datasets this loss at the lowest point lies in the 1 to 2 kilometer range with some outliers

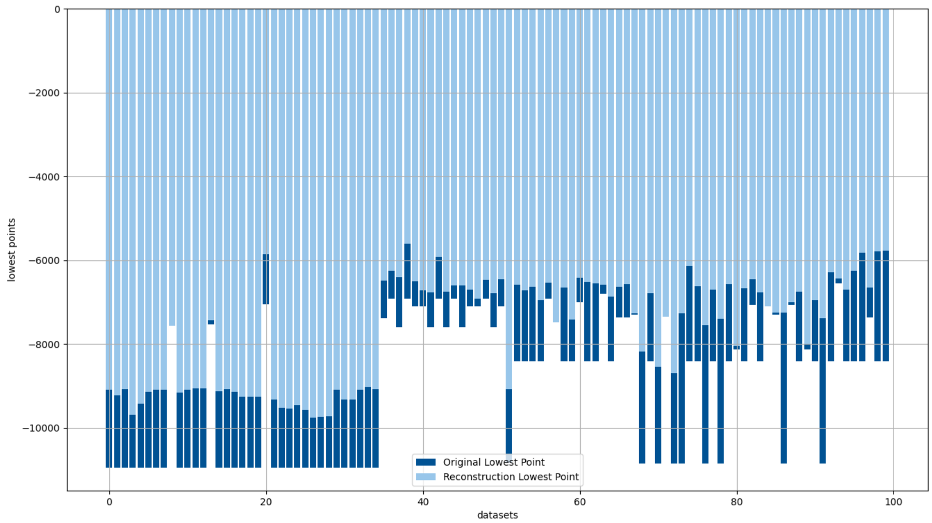

Figure 3.1: Comparison of lowest points of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset

in both directions. It is our assumption that the comparatively larger divergence occurs because the average depth of the dataset is much higher than the depth at its lowest point, while smaller divergences correspond to the fact that the bathymetry is generally more unifomr. We seek to confirm this hypothesis by looking at both the mean and standard deviation of our datasets.

3.1.2 Mean and Standard Deviation

In our next step we compare the lowest points of our datasets to their means as shown in figure 3.2. We see that the lowest points of original and reconstruction of a given dataset are closer when the mean and lowest point of the original are closer together. The fact that the mean and the lowest point often diverge strongly implies that the lowest points of our datasets have a very low influence on their mean. This agrees with the conclusion of Walter's thesis [Wal20] that the implementation struggles with rapid changes in depth over smaller areas.
Interestingly, we get a comparatively strong divergence for the means even in cutouts that contain roughly the same area. The datasets 0-19 mostly correspond to the area

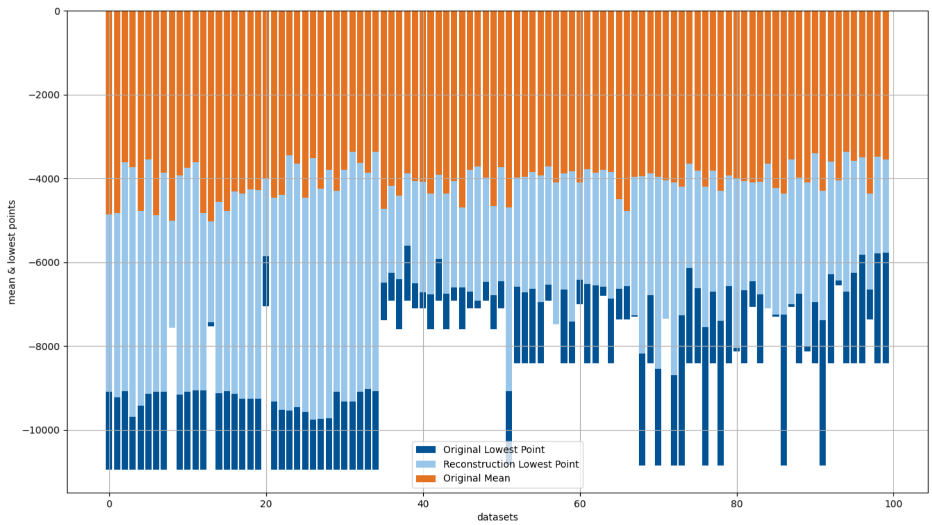

Figure 3.2: Comparison of mean and lowest point of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset

contained in the 2011 Tohoku dataset (figure 2.4). To visualize the overall accuracy of the reconstruction, we subtract the elevation data of the reconstruction from the original data and compute the mean and standard deviation of the resulting dataset. The closer their relation, the closer the mean should be to 0. The results can be seen in figures 3.3 and 3.4. It is clear that the difference in mean between original and reconstruction is rather small on all counts. We therefore conclude that the major attributes of the original dataset are preserved in its reconstruction.

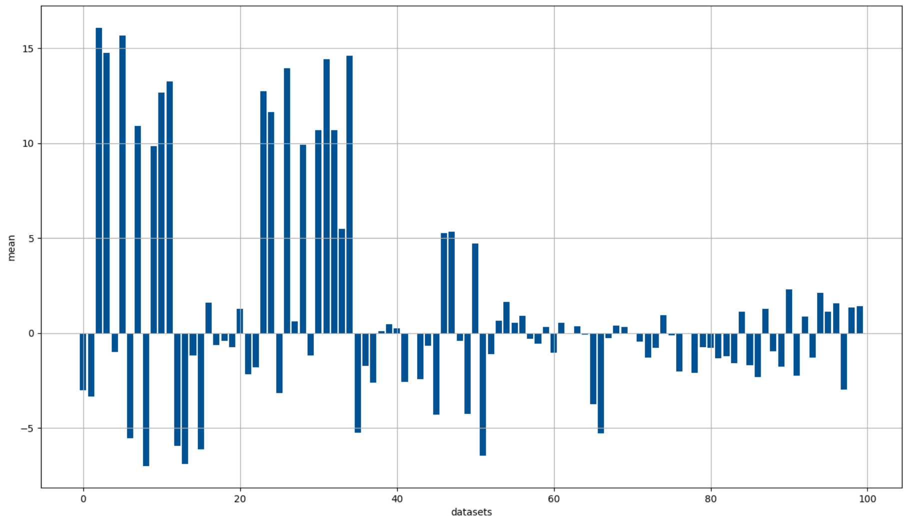

Figure 3.3: Distribution of mean of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset

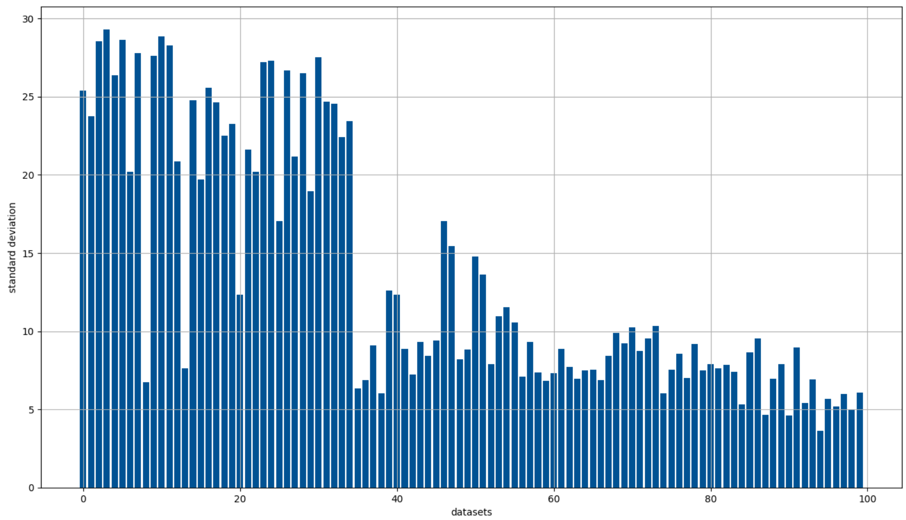

Figure 3.4: Distribution of standard deviation of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset

3.1.3 Error Metrics and Gradient

Through our rudimentary evaluation we have come to the conclusion that the reconstructions of our datasets are broadly accurate when compared to the original data. However, it is clear that there are at least some areas where (major) distortions are occurring as demonstrated by the lowest points. To further test the reasons for discrepancies we use several error metrics on our datasets and examine the relation between our results and the gradient of our bathymetric data.

Root Mean Squared Error

The mean squared error (MSE) is often used as a measure of accuracy between an dataset and an estimation of said dataset. It is computed as the mean squared difference between original and estimate [Tay97]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)} \tag{3.1}$$

where n describes the total number of points and $x_i, \hat{x}_i$ the point of original and estimate respectively.
Though it is more commonly found in statistical or predictive applications, we use it to gauge the accuracy of our reconstructions. It is generally preferable to have an RMSE close to 0.

As figure 3.5 shows, this condition is fulfilled for our datasets. For datasets of $81x10^6$ points that on average encompass a range between 5000 and -10000 a RMSE between 20 and 50 is highly acceptable.

Maximum Error

To contextualize our findings in regards to the RMSE we also compute the maximum error each for our datasets. In this case the maximum error is simply the maximum difference between two corresponding points of the original and reconstructed sets

$$MaxError(x, \hat{x}) = max(|x_i, \hat{x}_i|) \tag{3.2}$$

where $x, \hat{x}$ represent the original and reconstructed datasets.
As we can see in figure 3.6 the maximum error is significantly higher than the RMSE. This difference in results confirms to us that, while significant errors occur in reconstruction for all datasets, said errors do not have a strong impact on overall reconstruction accuracy.

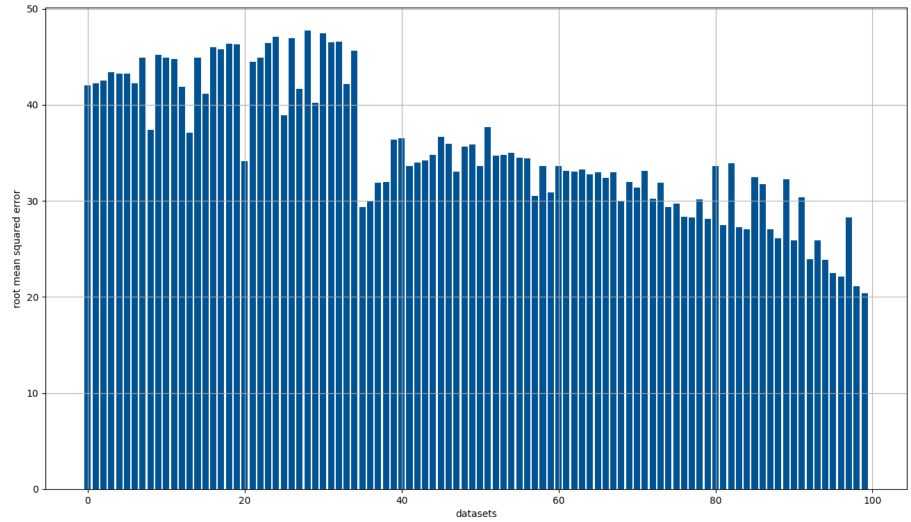

Figure 3.5: Depiction of the overall root mean squared error for all 100 datasets

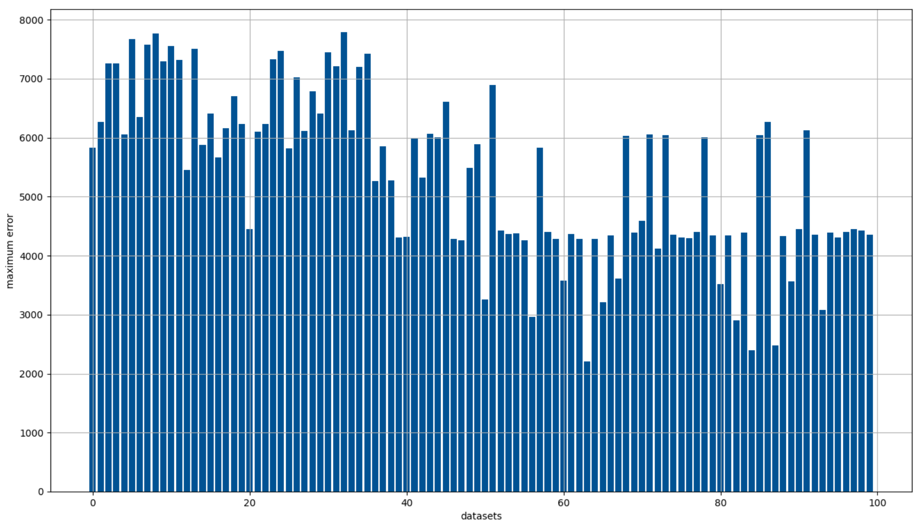

Figure 3.6: Depiction of the maximum error for all 100 datasets

Relation between Error and Gradient

Earlier we mentioned our assumptions that significant errors in the reconstruction process happen mainly in areas of rapid bathymetric change, meaning areas with a steep bathymetric gradient. To confirm this suspicion, we computed the gradient for our datasets as well as the individual errors for all points of said datasets. We then visualized the results. Figure 3.7 lends credence to our hypothesis. The most significant errors clearly occur around sudden drops in depth.

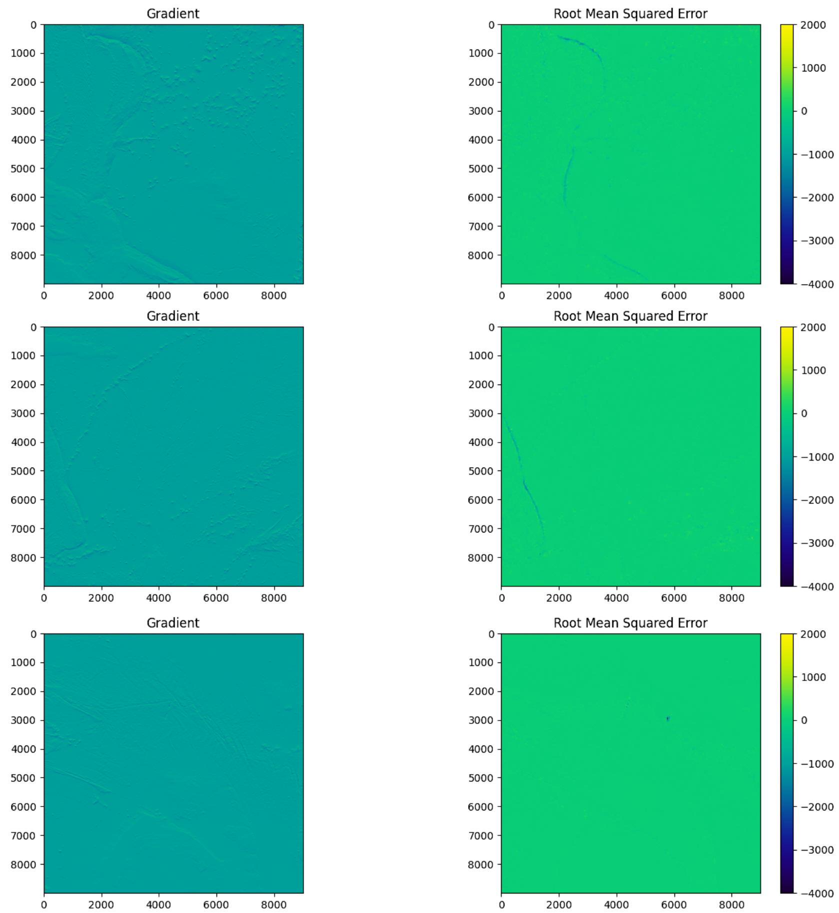

Figure 3.7: Gradient of a dataset as well as error between original and reconstruction for all points

3.2 The Tohoku 2011 Dataset

Figure 2.4 shows that the reconstruction model has issues with the 2011 dataset. We now want to quantify inaccuracies between original and reconstruction with the same methods we used for our GEBCO 2020 cutouts. For simplicity we compare the entire dataset instead of looking at each cutout separately. For reference we also include the values for a cutout of the 2020 Tohoku dataset.

	Lowest Point	Mean	Standard Deviation	RMSE
Original	-9843,52	-4837,17	1260.54	
Reconstruction	-8520,61	-4790,85	1233.29	
Difference	1322,91	46,32	27,25	779,44

Table 3.1: Table with different evaluation metrics for the Tohoku 2011 dataset

	Lowest Point	Mean	Standard Deviation	RMSE
Original	-10952	-5006.20	1211,31	
Reconstruction	-8936	-5002.21	1190.16	
Difference	2016	1,99	21,15	54,58

Table 3.2: Table with different evaluation metrics for the Tohoku 2020 dataset

There are a few notable points in regards to these values. For one, the lowest point differs not only between original and reconstruction, as we expected, but also between the 2011 and 2020 datasets. It is most likely that this happens due to improved accuracy of lower-depth scans for the 2020 dataset but it serves as a reminder that the version of the dataset can have an impact on accuracy. Due to the nature of the reconstruction as seen in figure 2.4 we weren't able to produce a meaningful depiction of the relation between error and gradient.

Notably, both the mean and error of the 2011 dataset are significantly worse than the results for the 2020 dataset while the loss of depth at the lowest point is slightly better. Since there is a significant loss here for both datasets we assume that the comparatively bigger error for the 2020 dataset is related to the correspondingly steeper slope at that point. All in all we conclude that the reconstruction of our GEBCO 2020 datasets is overall very accurate, while the reconstruction of the 2011 dataset suffers in places. In our final step we examine whether these errors have an impact on simulations run on our data.

4 Application of Sampling Methods to Bathymetric Data

This chapter contains the brunt of the thesis, namely the data gathered from simulations on different datasets. First we reevaluate our model on the 2011 Tohoku dataset, then we define a simple wave and use it to test simulations on our randomly chosen datasets from chapter 3.

4.1 The Tohoku 2011 Tsunami Simulation

In our next step we examine several cutouts of the Tohoku dataset to reconfirm the results of Walter's thesis [Wal20]. ExaHyPE allows us to define a buoy plotter for each of these cutouts. Said plotter writes the contents of the Q vector at the specified location to a probe file during runtime. After running the tsunami simulation on both the original 2011 dataset and its reconstructed version, we compare the results measured at the buoy via a simple plot and the time frequency misfit method.
As a first step in our examination we compute the mean of our vectors at the buoy location for all cutouts and compare it to the mean of the reconstructed datasets:

	h	hu	hv	b
Original	5.71809	-3.18648e-06	-4.76697e-06	-5.71809
Reconstruction	5.63703	1.31155e-06	7.30823e-07	-5.63701
Difference	0,08106	4,49803-06	5.49779e-06	0.08107

Table 4.1: Table for mean of the Q vector components for all cutouts

We measure a significant difference for all vectors, but that is not in and out of itself problematic. We already addressed possible causes for errors in the previous chapter. What concerns us now is how impactful these errors are.
Figure 4.1 depicts one noteworthy error: The altitudes are clearly different but the amplitude seems to largely coincide (as seen in figure 4.2). This occurs for all cutouts of the Tohoku dataset and is easily explained by the fact that there was some loss of bathymetric data at the buoy point during reconstruction (figure 4.3). In that case we

consider the plots seen in figure 4.1 and 4.2 a resounding success since they show that the error at the buoy point has no impact on the simulation as a whole. Apart from this depth difference we only note a small time lag that gets more pronounced towards the end of the timeline and increasingly large differences in the lower amplitudes. Since we look at this data in the context of tsunami simulation we are chiefly concerned with accuracy for the larger waves. Differences for smaller water movements are therefore acceptable, provided that these differences do not adversely affect the - to us - more important aspects of the simulation.
We will note that one can see a small shift at the beginning of all plots that show the bathymetry at a buoy point. This is unexpected since we only measure the bathymetric depth at a single point, which means that it should be entirely static. This occurence is easily explained though: Since we use a limiter (as described in section 2.1) there is some refinement of our grid at the beginning of the simulation, which causes our bathymetric value to be corrected to be more accurate.

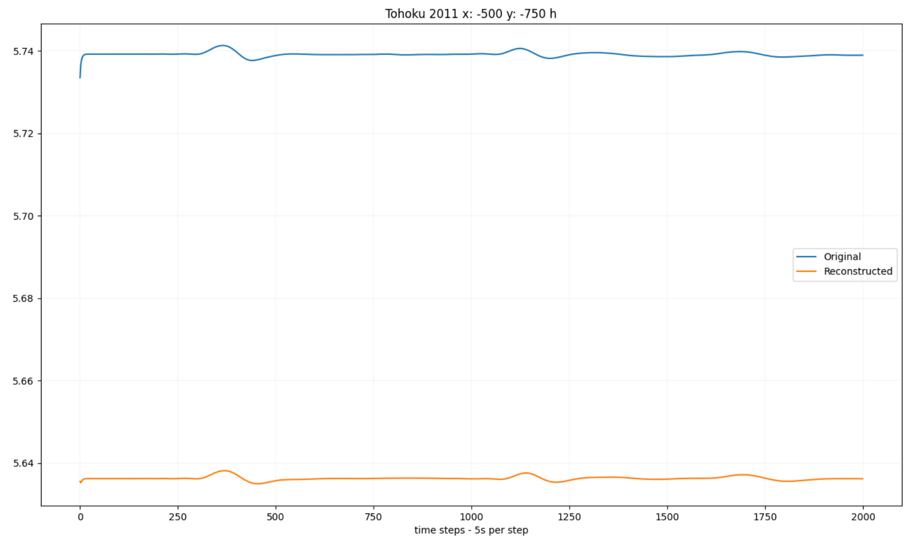

Figure 4.1: Comparison of water height of a cutout of the Tohoku 2011 dataset

Figure 4.2: Comparison of wave height + bathymetry of a cutout of the Tohoku 2011 dataset

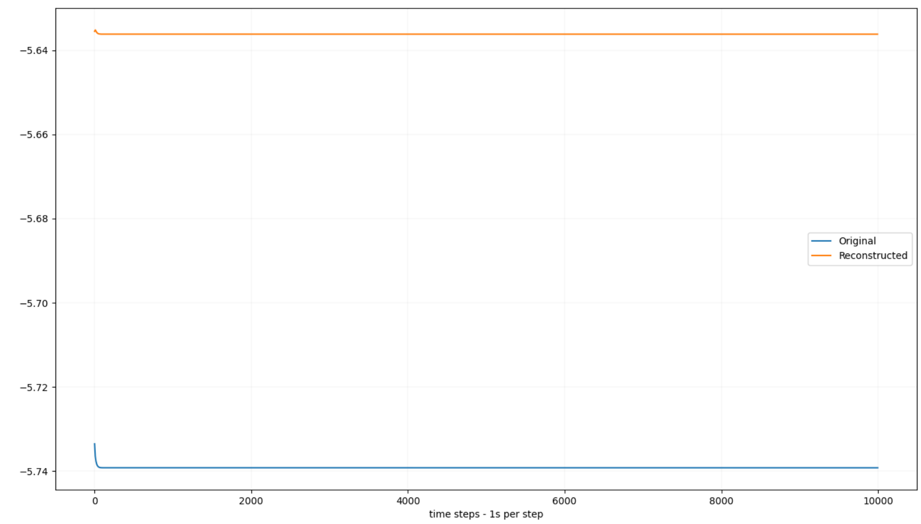

Figure 4.3: Comparison of bathymetry of a cutout of the Tohoku 2011 dataset at the buoy point

Since a visual evaluation alone is not sufficiently accurate, we examine our probe data with the TF Misfit method (see section 2.3).
To this end we use the goodness-of-fit criteria described by Kristekova et al (figure 2.3). Based on these criteria we can conclude that the accuracy of our reconstructed datasets in comparison to the original sets is indeed excellent. With an envelope misfit (EM) and phase misfit (PM) that lie comfortably in the lower part of the interval [0.0 - 0.2], both misfits are negligible. The results depicted in figure 4.2 allow us to draw two conclusions: First that the difference in water level at the buoy has no impact on the course of the tsunami simulation and second that the errors in the lower amplitudes only distort the overall simulation by an insignificant amount.
Since our simulation results are reasonable we concur with [Wal20]'s conclusion that the VAE model produces acceptable results.

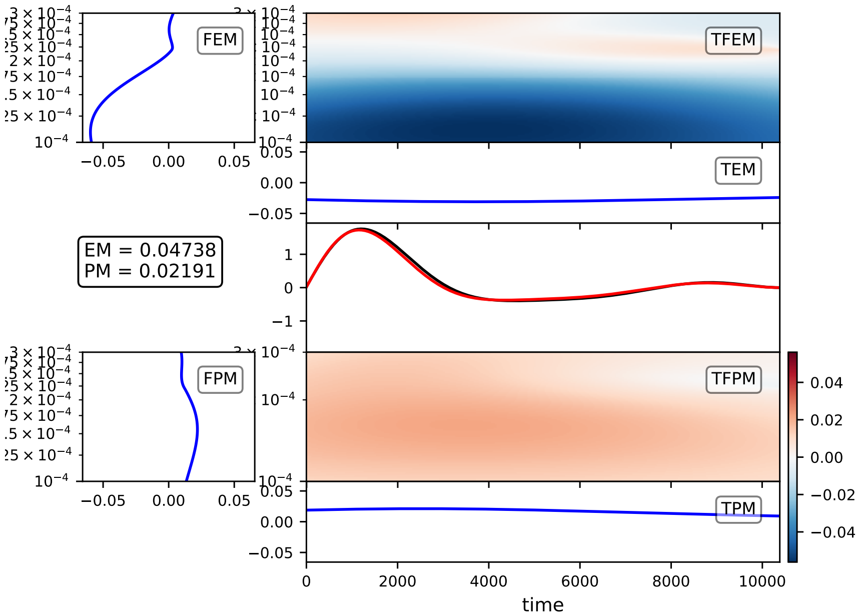

Figure 4.4: Time frequency representation of buoy data for the cutout of the Tohoku 2011 dataset with origin (-500,-750)

	x:0 y:0	x:0 y:-750	x:-500 y:0	x:-500 y:-750
EM	0,03220	0.03523	0,03839	0,04738
PM	0,01276	0,02251	0,01942	0,02191

Table 4.2: Table for EM and PM of water height of the Tohoku 2011 cutouts

4.2 Solitary Gaussian Wave Simulation

For this last step we select 20 of our randomly chosen datasets based on criteria like bathymetric gradient, reconstruction error and distribution of above-water landmasses. On these datasets we again run a simulation and plot the Q vector at a buoy point of our choice. Since we do not have predefined simulation data for any of our randomly chosen cutouts we instead define a simple wave with a Gaussian bell form:

$$f(x) = a \cdot e^{\frac{-(x-b)^2}{c}} \tag{4.1}$$

where a describes the maximum height of the wave, b denotes the location of the wave's maximum on the x-axis and c denotes the width of the wave. A depiction of the wave for different parameters can be seen in figure 4.5.

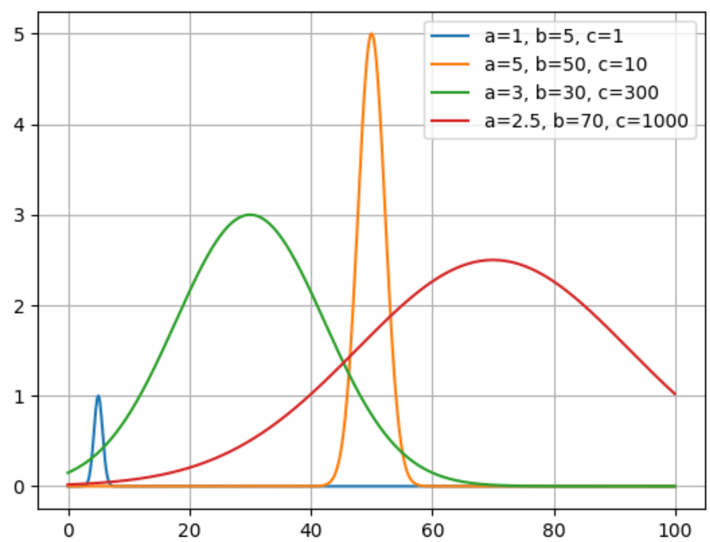

Figure 4.5: Depiction of different Gaussian distributions

Our randomly selected 9000x9000 points large datasets encompass an area of about 1369 km^2. From this area we select a section of 900km^2 and run our simulation for 1000 seconds. During this time a probe placed 5km away from the center of our Gaussian wave reads out our Q vector every second. Of course we use the same area for both the original and reconstructed dataset.
After our simulations have run we compare them with the same methods we used in section 4.1.

	h	hu	hv	b
Original	4.51857	3.84288e-04	-4.83496e-05	-4.48967
Reconstruction	4.30949	3.71144e-04	-2.38280e-05	-4.27915
Difference	0.20908	1.31437e-05	2.45215e-05	0.210515

Table 4.3: Table for mean of the Q vector components for all cutouts

We begin again by determining the mean of the Q vector components over all our datasets. When we compare the results of table 4.1 and table 4.3 the relative error for the GEBCO 2020 datasets is lower for all vectors except for the hv vector. This strengthens the naive hypothesis that a more accurate reconstruction will also lead to a more accurate simulation.
To further determine how our simulations were affected we again look at the data recorded at our buoys. We immediately notice the same error in water level we saw for the Tohoku 2011 dataset in figure 4.1. Again, we first look at plots of the Q vector for our simulations, though this time we can actually compare separate areas with diverging bathymetric makeup. Out of the 20 simulations we have run, figures 4.7, 4.8, 4.9 and 4.10 show relevant plots for the (visually) least and most accurate vectors of our simulations. The plot in figure 4.9 is in fact remarkably accurate, the reconstruction almost perfectly following the original. The results of 4.7 stand in sharp contrast to this result: While the plot of the reconstruction still follows the same shape as the original there is a clear difference in the lower and higher amplitude. To find reasons for this distinction we look at the comparisons we conducted in chapter 3.

	Lowest Point Difference	Mean Difference	Std Dev Difference	RMSE
Cutout 9	-1298	0.09932	6.04262	31.96875
Cutout 10	-991	-4.30066	9.42235	36.65288

Table 4.4: Table with different evaluation metrics for the Tohoku 2020 dataset

The mean is most clearly indicative of a difference in reconstruction quality. Especially

the RMSE does not indicate an error in our reconstruction in such a ways as seen in figure 4.7. To make absolutely sure we are not missing a major divergence we plot the bathymetry of dataset 10 in Paraview (figure 4.10). As we can see the major features are well preserved and (barring minor errors) even the smaller details seem to be retained with a high degree of accuracy. Finally we visualize the datasets themselves and find a clearer indication for our problem. Dataset 10 contains a much more topographically varied area than dataset 9 especially when we look at landmasses above sea level.
We notice a clear correlation between the gradient of a dataset and the accuracy of a simulation run on its reconstruction. This is reasonable since a stronger gradient implies more drastic changes in the bathymetric landscape, which in turn has proven to impact the reconstruction. It should be noted however, that this does not apply to our datasets as a whole: Multiple simulations where coastal runup happened frequently and where a lot of smaller islands disturbed the flow of our wave were no less accurate than several simulations where the underlying bathymetry was much more homogeneous. For comparison one need only look at figures 4.11 and 4.12, the dataset of which corresponds to the same area as the Tohoku 2011 dataset. Our simulation there has to deal with the same difficulties as dataset 10 but the results are much more solid. Based on our findings the simulation error shown in figures 4.7 and 4.8 is in fact an outlier. And even there we only notice a distortion at points instead a completely different wave propagation.

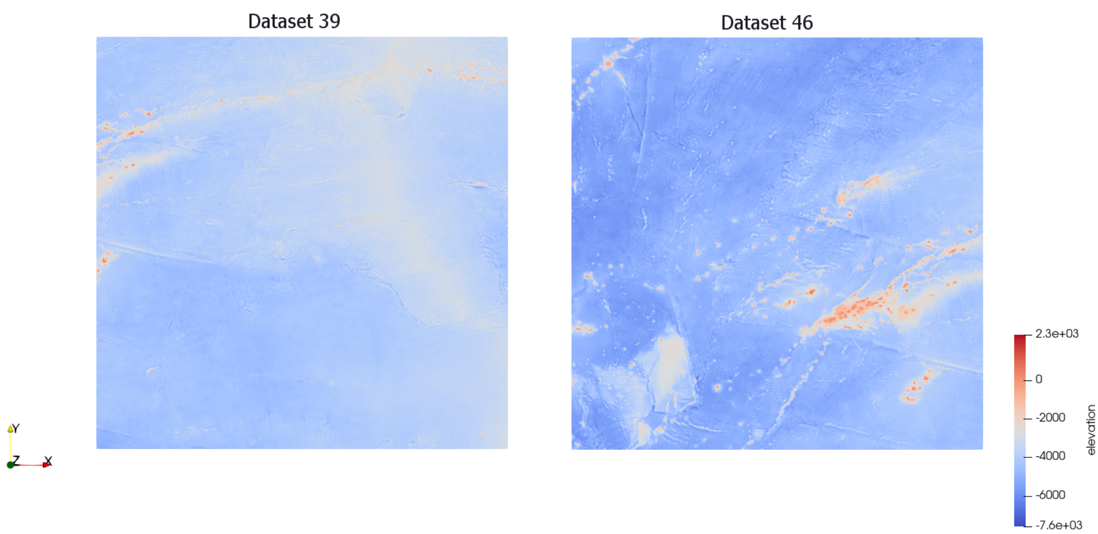

Figure 4.6: Depiction of bathymetry for datasets 9 and 10

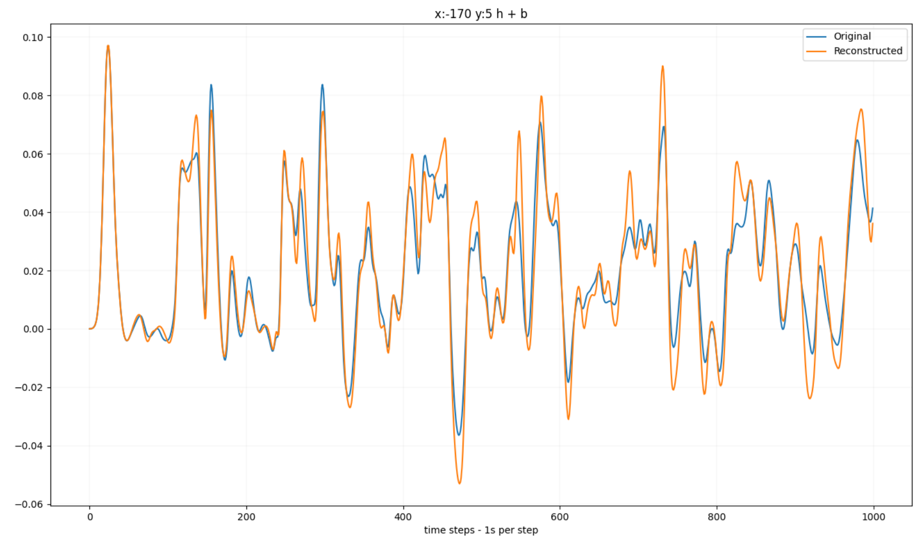

Figure 4.7: Comparison of wave height + bathymetry of cutout 10 of the GEBCO 2020 dataset

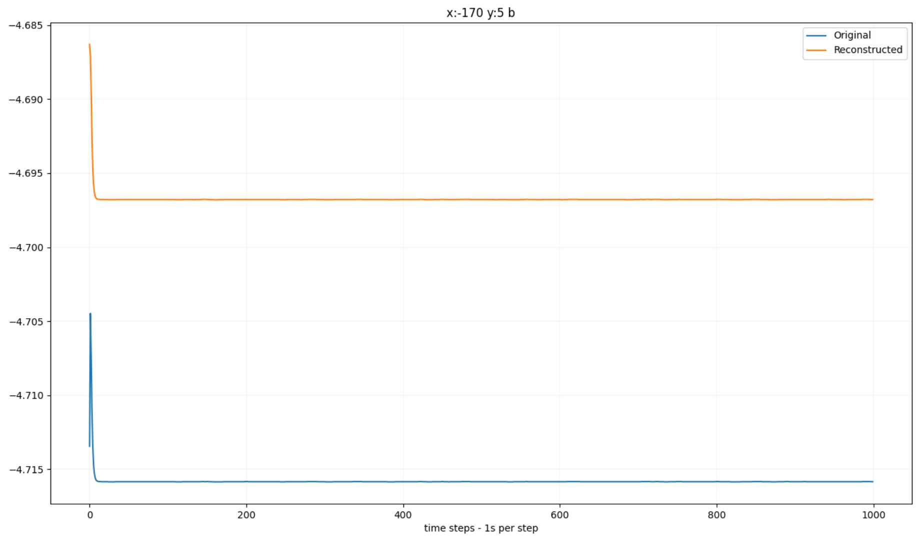

Figure 4.8: Comparison of bathymetry of cutout 10 of the GEBCO 2020 dataset at the buoy point

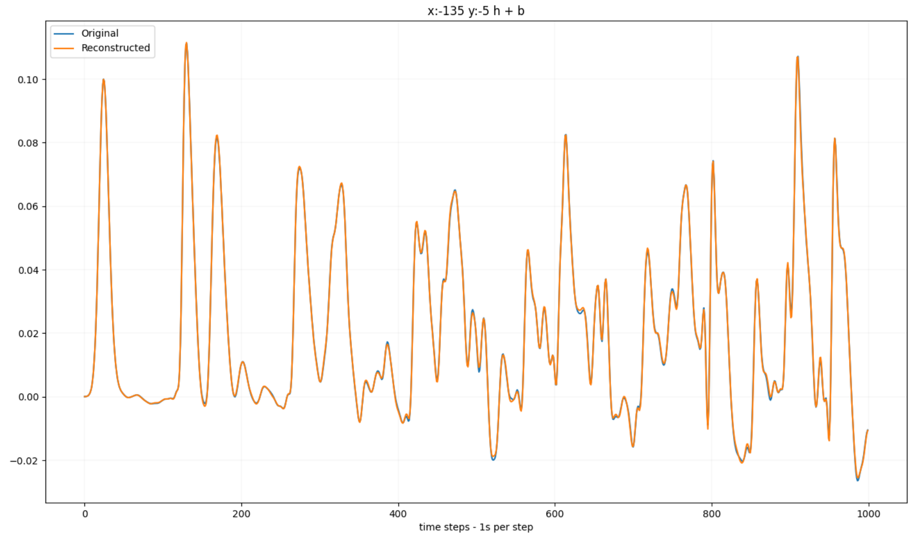

Figure 4.9: Comparison of wave height + bathymetry of cutout 9 of the GEBCO 2020 dataset

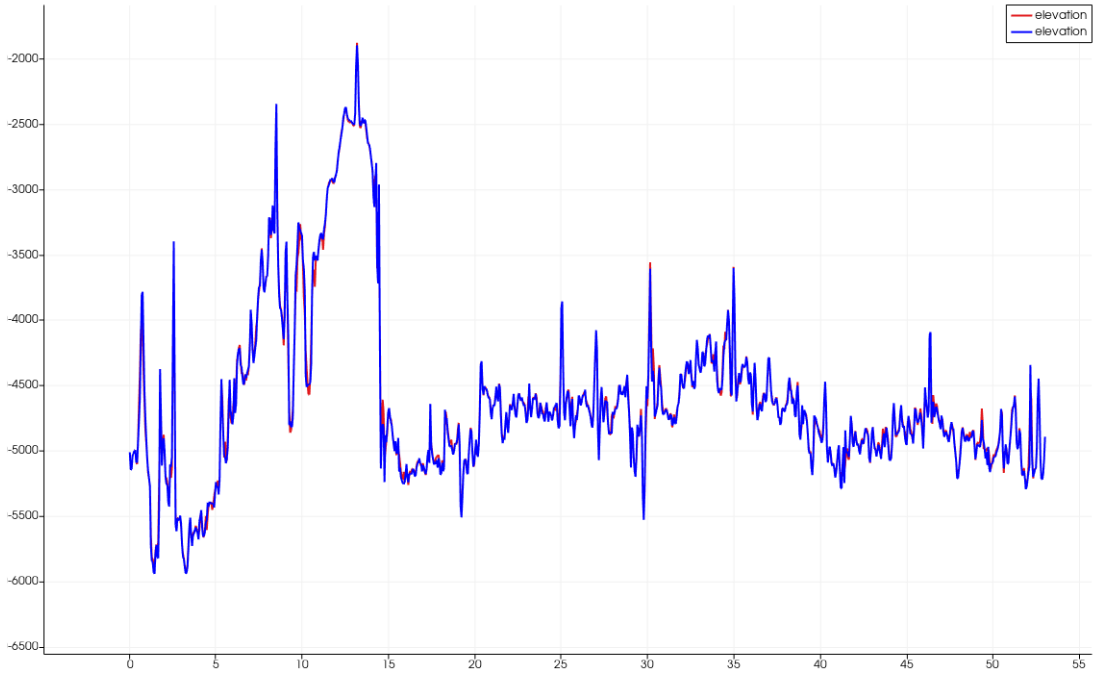

Figure 4.10: Comparison of bathymetry of cutout 10 along the diagonal of the dataset

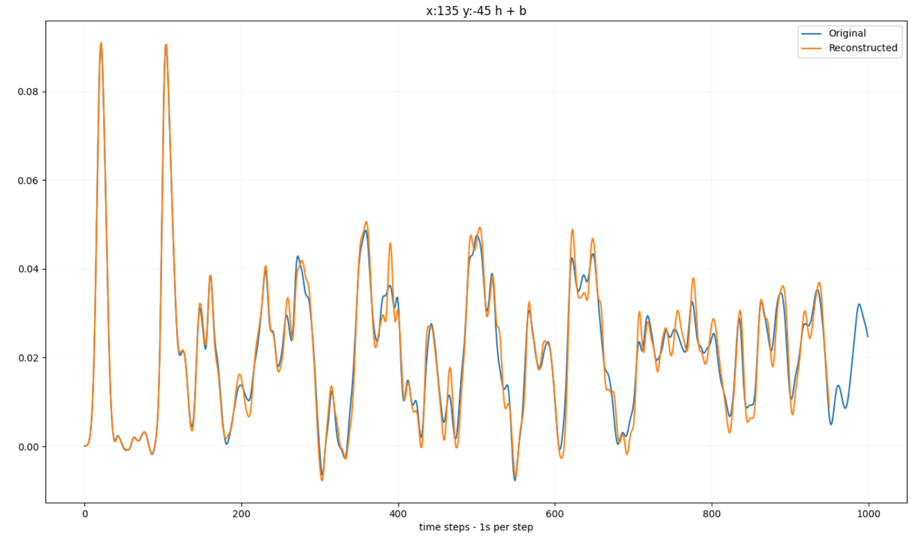

Figure 4.11: Comparison of wave height + bathymetry of cutout 1 of the GEBCO 2020 dataset

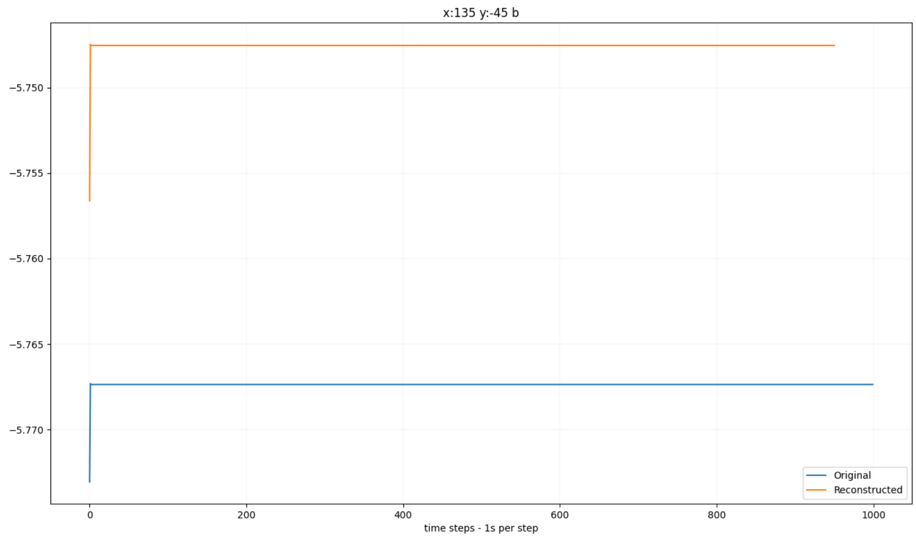

Figure 4.12: Comparison of wave bathymetry of cutout 1 of the GEBCO 2020 dataset at the buoy point

But as before with the Tohoku 2011 dataset we don't want to rely solely on a visual confirmation. Therefore we compute the time frequency misfit for the h vector of our 20 simulations.

Datasets	1	2	3	4	5	6	7
EM	0.04619	0.09672	0.00328	0.11753	0.02549	0.04661	0.07023
PM	0.03805	0.04348	0.00129	0.05779	0.03591	0.03137	0.04622
Datasets	8	9	10	11	12	13	14
EM	0.00909	0.04165	0.99985	0.01994	0.02468	0.05986	0.00980
PM	0.00036	0.00022	0.00974	0.00101	0.00037	0.00132	0.00583
Datasets	15	16	17	18	19	20	-
EM	0.01181	0.00531	0.00531	0.01063	0.00218	0.02853	-
PM	0.00082	0.00359	0.00158	9.80626e-09	0.00101	1.28698e-06	-

Table 4.5: Table with Envelope and Phase Misfit values for buoy data of 20 cutouts of the GEBCO2020 dataset

For all datasets except number 10 our assessment of the envelope and phase misfits are excellent, as per figure 2.3. Dataset 10 however is just on the border of fair, verging into poor. It is also a clear outlier not only during our assessment of plotted data but also in regards to the misfit values.
Looking at table 4.2 and 4.5 side by side we do not see any significant improvement for most of our datasets. Since the misfit for the Tohoku dataset was already excellent this point is not particularly problematic. All in all we can conclude that our simulation results are reasonable across the bank (with one outlier) even though the reconstruction model clearly profits from more homogeneous bathymetries. Looking at the misfit results this distinction is largely cosmetic though. Results for bathymetric datasets with more severe changes in landscape may result in slightly less accurate simulations but these inaccuracies generally don't disturb the overall trajectory of our waves.

5 Conclusion

At the beginning of this thesis we set out to determine whether Walter's model provided accurate enough reconstructions for use in tsunami simulation. For this purpose we examined both the older 2011 Tohoku dataset and the newer, more accurate 2020 GEBCO dataset. We examined first the datasets themselves and then simulations run on said datasets: using real-world displacement data where available and using a simple gaussian wave as stand-in where not.
In conclusion we concur that the model described and implemented by Walter in [Wal20] fulfills its purpose. Even before looking at our simulations the reconstructions are accurate based on our criteria. There is no loss of major landmarks and even fine detail is reasonably well reconstructed on most all datasets.
There are some obvious errors with most reconstructions (most pervasively the loss of depth in areas with strong gradients) but as we saw when running our simulations, this has no drastic effect on their outcome. We did notice a more drastic divergence for one simulation on a more bathymetrically challenging dataset but since we were not able to reduce this divergence on other, similarly varied, datasets we marked it down as an outlier. Still, it makes it clear that the model might not be applicable to all situations.
Nevertheless we conclude that our results satisfy the criteria we set to determine the fitness of Walter's model for further use.

List of Figures

2.1 Deep feature consistent variational autoencoder, image taken from [Hou+16] 5
2.2 Water height at fixed coordinate resulting from tsunami simulations [Wal20] 6
2.3 Discrete goodness of fit values against misfit values, taken from [KKM09] 8
2.4 Comparison between cutouts of the Tohoku area from 2011 (top) and 2020 (bottom) as well as their reconstructions 9

3.1 Comparison of lowest points of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset 11
3.2 Comparison of mean and lowest point of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset 12
3.3 Distribution of mean of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset 13
3.4 Distribution of standard deviation of 100 different 9000x9000 cutouts of the GEBCO 2020 dataset 13
3.5 Depiction of the overall root mean squared error for all 100 datasets . . 15
3.6 Depiction of the maximum error for all 100 datasets 15
3.7 Gradient of a dataset as well as error between original and reconstruction for all points 16

4.1 Comparison of water height of a cutout of the Tohoku 2011 dataset . . . 19
4.2 Comparison of wave height + bathymetry of a cutout of the Tohoku 2011 dataset 20
4.3 Comparison of bathymetry of a cutout of the Tohoku 2011 dataset at the buoy point 20
4.4 Time frequency representation of buoy data for the cutout of the Tohoku 2011 dataset with origin (-500,-750) 21
4.5 Depiction of different Gaussian distributions 22
4.6 Depiction of bathymetry for datasets 9 and 10 24
4.7 Comparison of wave height + bathymetry of cutout 10 of the GEBCO 2020 dataset 25
4.8 Comparison of bathymetry of cutout 10 of the GEBCO 2020 dataset at the buoy point 25

4.9 Comparison of wave height + bathymetry of cutout 9 of the GEBCO 2020 dataset . . . 26
4.10 Comparison of bathymetry of cutout 10 along the diagonal of the dataset 26
4.11 Comparison of wave height + bathymetry of cutout 1 of the GEBCO 2020 dataset . . . 27
4.12 Comparison of wave bathymetry of cutout 1 of the GEBCO 2020 dataset at the buoy point . . . 27

List of Tables

3.1 Table with different evaluation metrics for the Tohoku 2011 dataset . . . 17
3.2 Table with different evaluation metrics for the Tohoku 2020 dataset . . . 17

4.1 Table for mean of the Q vector components for all cutouts 18
4.2 Table for EM and PM of water height of the Tohoku 2011 cutouts 22
4.3 Table for mean of the Q vector components for all cutouts 23
4.4 Table with different evaluation metrics for the Tohoku 2020 dataset . . . 23
4.5 Table with Envelope and Phase Misfit values for buoy data of 20 cutouts of the GEBCO2020 dataset 28

Bibliography

[20220] G. C. G. 2020. Apr. 2020. DOI: 10.5285/a29c5465-b138-234d-e053-6c86abc040b9.

[And04] J. Anderson. "Quantitative measure of the goodness-of-fit of synthetic seismograms." In: *13th World Conference on Earthquake Engineering* (Jan. 2004).

[Hou+16] X. Hou, L. Shen, K. Sun, and G. Qiu. *Deep Feature Consistent Variational Autoencoder*. 2016. arXiv: 1610.00291 [cs.CV].

[KKM09] M. Kristeková, J. Kristek, and P. Moczo. "Time-frequency misfit and goodness-of-fit criteria for quantitative comparison of time signals." In: *Geophysical Journal International* 178 (Aug. 2009), pp. 813–825. DOI: 10.1111/j.1365-246X.2009.04177.x.

[KW14] D. P. Kingma and M. Welling. *Auto-Encoding Variational Bayes*. 2014. arXiv: 1312.6114 [stat.ML].

[Oce] G. B. C. of the Oceans. *GEBCO_2020 Grid*. https://www.gebco.net/data_and_products/gridded_bathymetry_data/gebco_2020/. Accessed: 2020-11-11.

[Rei+19] A. Reinarz, D. E. Charrier, M. Bader, L. Bovard, M. Dumbser, K. Duru, F. Fambri, A. Gabriel, J. Gallard, S. Köppel, L. Krenz, L. Rannabauer, L. Rezzolla, P. Samfass, M. Tavelli, and T. Weinzierl. "ExaHyPE: An Engine for Parallel Dynamically Adaptive Simulations of Wave Problems." In: *CoRR* abs/1905.07987 (2019). arXiv: 1905.07987.

[RHW86] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. "Learning Internal Representations by Error Propagation." In: *Parallel Distributed Processing – Explorations in the Microstructure of Cognition*. MIT Press, 1986. Chap. 8, pp. 318–362.

[Sch+15] A. Schwarz, B. Hazelwood, D. E. Charrier, F. Guera, and J.-M. Gallard. *ExaHyPE Guidebook*. The ExaHyPE consortium, 2015.

[Sul15] T. J. Sullivan. *Introduction to Uncertainty Quantification*. Vol. 63. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-23395-6.

[Tay97] J. R. Taylor. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. University Science Books, 1997.

[TMN01] E. F. Toro, R. C. Millington, and L. A. M. Nejad. "Towards Very High Order Godunov Schemes." In: *Godunov Methods: Theory and Applications*. Ed. by E. F. Toro. Boston, MA: Springer US, 2001, pp. 907–940. ISBN: 978-1-4615-0663-8. DOI: 10.1007/978-1-4615-0663-8_87.

[TS96] Y. Tanioka and K. Satake. "Tsunami generation by horizontal displacement of ocean bottom." In: *Geophysical Research Letters* 23.8 (1996), pp. 861–864. DOI: 10.1029/96GL00736. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/96GL00736.

[Wal20] S. Walter. "Deep Learning for Parameter Reduction on Real-World Topography Data." MA thesis. Technische Universität München, May 2020.