ISSN 8755-6839

SCIENCE OF TSUNAMI HAZARDS

Journal of Tsunami Society International

Volume 42 Number 3 2023

TSUNAMI HAZARD: IMPACT OF DATA QUALITY ON A MODELLING AND MAPPING FRAMEWORK

Rudy VanDrie[*1], Gede Pringgana[1], Ni Nyoman Pujianiki,

[1] *Universitas Udayana Bali 60231, INDONESIA*
Email: rudyvandrie@gmail.com

ABSTRACT

The need to review the impact of tsunamis on the island of Bali and Indonesia as a whole is real and warranted. There are in excess 17,500 island, 275million people and multiple Tsunami sources. The last comprehensive assessment of tsunami risk and hazard for parts of Bali and Padang was completed in 2010 with data since identified as unsuitable in accuracy. Several problems are identified to improve tsunami hazard mapping in Indonesia, such as the role of input data accuracy consistency and density of data, a better description of hazard, and the sensitivity of hazard to input data quality. The research objectives include reviewing multiple data sets, undertaking tsunami impact modeling, and developing a data quality metric. Further there is consideration of a future framework approach to enable a nationwide roll-out of modelling on the basis the metric identifying the availability of improved quality of data. Strongly related to this is the ongoing development and review of the Indonesian BATNAS and DEMNAS data. It is recommended that version metadata be developed for the evolving data sets in time. Noting that ongoing improvement in this data is a strong candidate to trigger the need to update Tsunami Modelling and Mapping in Indonesia, either as a whole, or in specific areas as it becomes available.

Keywords: *Tsunami, hazard, models, data, quality, accuracy, DEMNAS, BATNAS, ANUGA*

1. INTRODUCTION

Indonesia is the worlds largest archipelago with over 17,500 islands (see in Figure 1 which is the 14th largest country in the world covering 1,904,569 square kilometres with 275 million people. Making it the 4th most populous country. It is one of the most tectonic seismically active countries in the world where the Indo-Australian Plate and the Pacific Plate and many sub-plates are interacting with the Eurasian plate. The 2004 Indian Ocean earthquake and 2006 Yogyakarta earthquake were two events triggering tsunami. In 2019, 15.4 million tourist visited Indonesian contributing around US$9.8 billion to GDP in 2020 making an interruption due to tsunami potentially very expensive to the entire community.

From this statement it is clear that Indonesia has significant complexity related to Tsunami impact not shared by any other nation. It has a huge number of islands (largest archipeligo in the world), a very large population, strong contribution to its economy through tourism, and sits on one of the most seismically active zones in the world. The length of coastline is estimated to be around 108,000km (second longest in the world) (Pandjaitan, 2020). The 2004 Tsunami and the 2011 Fukashima Tsunami have led to a focused, global research effort on Tsunami's and their impact, this thesis will add to this in a number of ways.

The last comprehensive assessment of Tsunami Risk and hazard undertaken at least for parts of Bali and Padang was conducted in 2009/2010 through GITEWS (GITEWS{DLR / GTZ}, 2010). "Tsunami Hazard Maps for Bali" incorporating;

- 'Multi-scenario Tsunami Hazard Maps for Bali, 1:100,000', and
- 'Multi-scenario Tsunami Hazard Maps for Southern Bali, 1:25,000',

with zoning based on wave height at coast (in line with the InaTEWS warning levels). This work was completed using Mike21 and coarse STRM (30m cell) data. This data is known to be relatively poor (Griffin etal., 2015). The earlier work (Kjell Karlsrud 2009), relied on models using a 100m grid cell using STRM and ASTER terrain data. Of some concern is the ongoing reliance and use by third parties on this mapping product (Sagala etal., 2016) to drive Disaster preparedness. Other attempts to model portions of Indonesia exist but are not comprehensive.

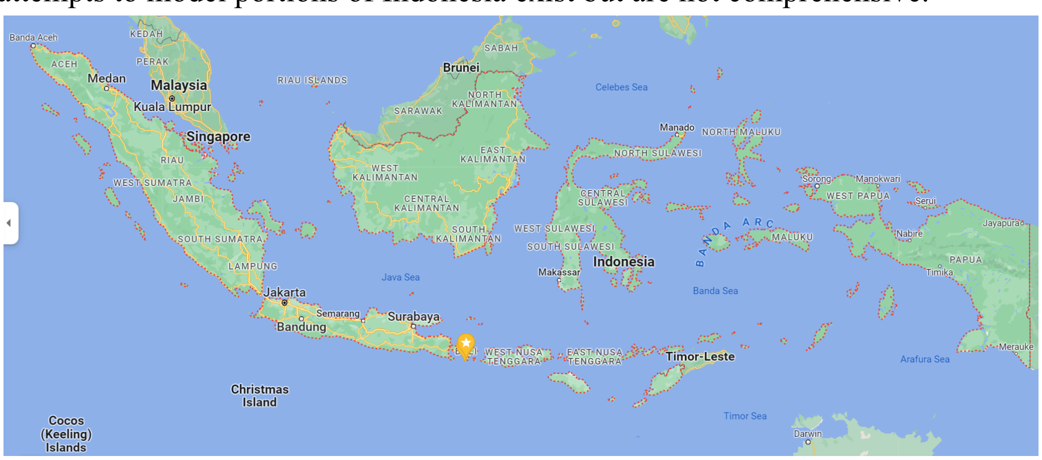


Figure 1. Indonesia Archipelago and Location of Bali

Other attempts to model portions of Indonesia exist but are not comprehensive. (Prerna etal, 2014) undertook tsunami analysis of the Andaman Islands using the STRM data set.

(Fatmawati etal., 2019) undertook run up assessments on a portion of the southern coast of Bali, but it is unknown what terrain data was used for this exercise. (Valentra Etal, 2022) undertook a tsunami impact analysis using higher resolution data in the Lombok Strait. In regard to the inclusion of Buildings research has shown a very real impact on flow behavior and resulting hazard (Murray etal., 2021).

In relation to input data accuracy, (Griffin etal., 2015) provide an assessment for the Tsunami Impact on Padang. However, the notion of a framework to control and assess the accuracy of the input is not directly discussed. It is however significant and relevant to note their findings and conclusions paraphrasing here:

"The results presented in this paper clearly demonstrate that the present generation of freely available global DEMs (i.e., ASTER and SRTM90) are not sufficiently accurate to simulate tsunami inundation with confidence.

Tsunami inundation models developed using DEMs that are currently freely available at a global scale (i.e., ASTER and SRTM) have the potential to dangerously underestimate the inundation extent. These datasets should not be used to assess tsunami inundation zones using hydrodynamic models."

Finally, globally the methods to describe a hazard from Tsunami or indeed any flowing water has not evolved from initial concepts developed in the 1970's. Whilst technology to provide analysis has improved such that there are now numerous tools that provide highly accurate simulation capacity, the core method to define the resulting hazard has remained unchanged and not open to new inputs from the new analysis methods. Hazard has and is defined primarily on momentum or the Velocity times Depth product. Although several researchers have tried to show methods of plausibly improving this (Trieste, 1988) , (VanDrie, 2008) to date no alternate methods are widely adopted beyond the original 1970's formulation of hazard. This work will touch on approaches to hazard, and the notion of using a framework to enable a systematic and consistent approach to model input assessment, model setup and production of output for mapping.

2. METHODS

The framework of thinking in this study is based on the notion that the accuracy of results from models is directly impacted by the accuracy and quality of input data. The old saying; "Rubbish-In, Rubbish-Out" is very true. However, the objective is to try to gain a metric, a notion of measurement of data accuracy and the influence on resulting output. This requires assessment of input data sets, running of multiple models over the same area utilizing each of the data sets and finally reviewing results and the variation in results as a result of only input data differences. Refer to Figure 2 for an outline of the method and approach.

The research concept in this work is to acquire multiple data sets for two project sites. One set of multiple data sources for Ocean Floor bathymetry, one set of above Ocean terrain data. The data sets will be ranked in accuracy based on how the data has been collected. A physically measured (surveyed) data set will be deemed most accurate. For bathymetry single beam sonar is deemed most accurate on the basis that it has been manually acquired. The survey and sonar will become the point of truth data for comparisons. For the additional data sets, a comparison will be undertaken to identify how much error is in the data compared to the point of truth data. The approach here is to utilize two sites with multiple data sets that describe the terrain and bathymetry at various levels of accuracy. Based on the different data sets a tsunami model will be run for each combination of terrain+bathymetry.

The resulting inundation and hazard characteristics on the shore attacked by the tsunami will be compared for each of the models. The comparison of model results from the bathymetric data sets will provide an indication of the sensitivity of hazard based on the drowned terrain changes. The comparison of the dry terrain results for each of the bathymetric data sets, will indicate that a change in offshore data influence the on land hazard. Similarly the comparison of the multiple dry terrain data sets, keeping the bathymetry static will indicate the sensitivity of hazard definition of only the terrain change. The initial steps involve gathering the data sets. There are two distinct different types of data, point data, and surface data (grid). As such a comparison must be based on values located precisely at the points as well as over the 2 dimensional extent of the surfaces created by the points, when compared to the grid data sets.

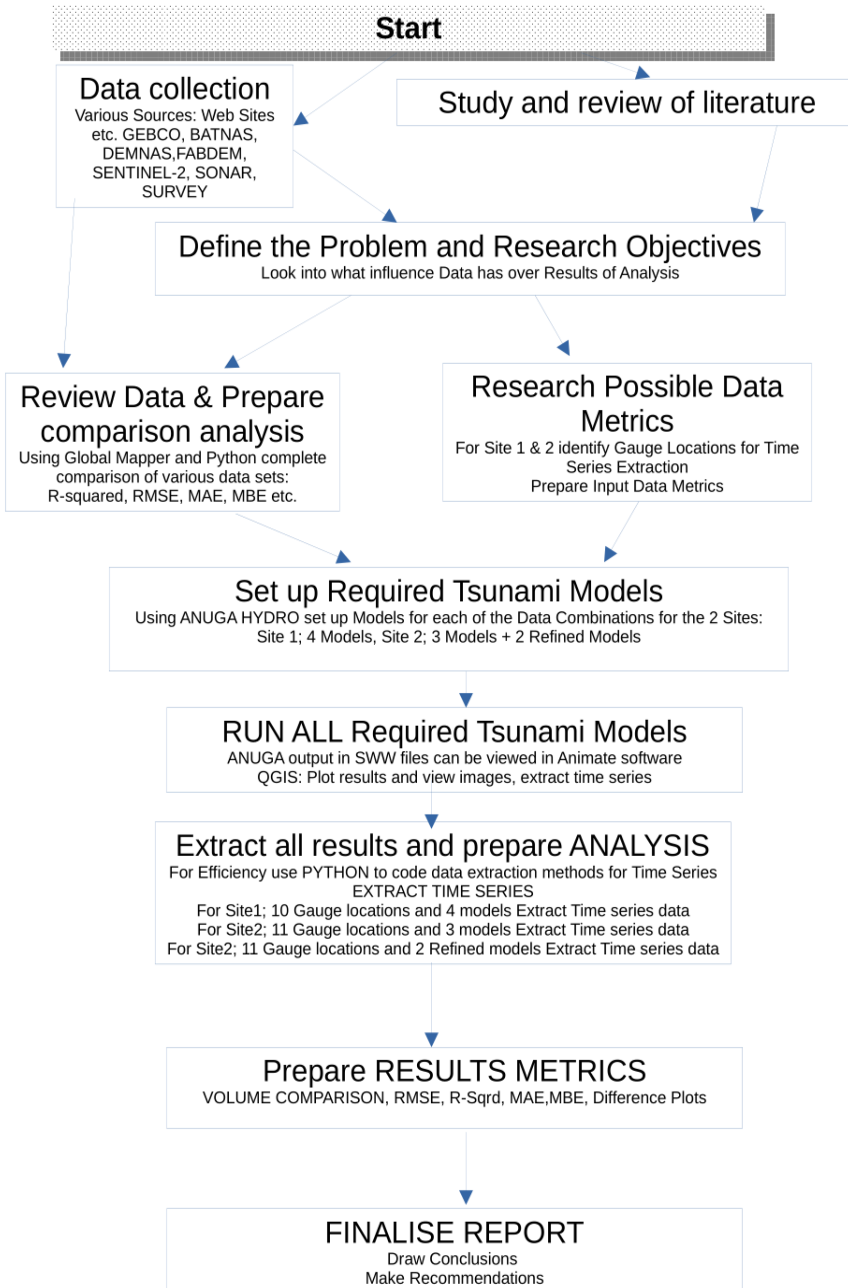


Figure 2. Process Steps

The completion of the input data analysis will attempt to identify a "Quality Metric". From each of the Terrain and Bathymetric data set combinations, a Tsunami Model will be set up apply the exact boundary conditions, such that the only variance in the model is the terrain component of either Bathymetric data or Terrain data. Once the models have all been run results will be extracted at a nominated number of GAUGE locations. The results will include the full time series of the analysis.

ANUGA HYDRO (Zoppou, Roberts, 1999), (Nielsen, Roberts, 2005) is used to model the inundation resulting from a large tsunami wave as it has been shown to replicate extreme flows very well (Mungkasi, van Drie, Roberts, 2013), (Wuppukondur, Baldock, 2020). The focus is not on the details of the type of wave, or replicating a certain event. Instead a simple 10m increase in water height is applied at the boundary which will drive the model to resolve flow characteristics. The focus again is to account for input data difference in the resulting description of hazard. ANUGA has been selected for this task for a number of reasons:

- It uses a flexibly sized triangular mesh (making it easy to produce high levels of detail only where required) (Schlurmann, Kongko, Goseberg, Natawidjaja, Sieh, 2010), which may be automated (Wright, Passalacqua, Simard, Jones, 2022)
- It is extremely stable in the most extreme flow conditions
- It is proven to be a very good model to replicate tsunami inundation
- It is not highly (or overly) sensitive to changes in surface roughness (Cárdenas, Catalán, 2022), (Van Drie, Milevski, Simon,2011).
- It has a number of very useful built-in functions to make extraction of results very easy
- Can run in Parallel (Roberts, Stals, Nielsen, 2007)

The measure of difference can be achieved in a number of ways:

- The elevation at each point directly compared
- The differences measured based on the surfaces created or available
- Using statistical functions such as RMSE, MAE, MBE etc.

The analysis to be undertaken is on the inundation data sets from the ANUGA HYDRO Models described. The data sets include water level (Stage), momentum of the moving water, bed Shear, Depth, Velocity, and Froude number. Note that, Tsunami really are or become debris flows (Synolakis, Bernard; 2006) and this may be a future trigger to re-model when models can be adequately be adapted to account for debris flows. Models may often contain errors (VanDrie, Ghetti, Milevski, 2018) which may also trigger the need for re-modelling.

The focus of this work is in understanding the influence of data quality (Schlurmann etal., 2010) or of error in input data, on the outcome of error in the results of a model to predict hazard. It is usual to take several ERROR METRICS or Key Performance Indicators (KPI's) into account in order to assess findings. Measuring forecast accuracy (or error) is not an easy task as there is no one-size-fits-all indicator. Only experimentation will show you what Key Performance Indicator (KPI) is best. The first distinction required is the difference between the precision of a forecast and its bias. Bias represents the historical average error. That is, will forecasts be, on average, too high or too low. This will give you the overall direction of the error. Precision measures how much spread is between the forecast and the actual value. The precision of a forecast gives an idea of the magnitude of the errors but not their overall direction. Ideally a model

Forecast KPI or Error. Error simply put is the difference in a forecast and known target. Note that if the forecast overshoots the target with this definition, the error will be. positive. If the forecast undershoots the demand, then the error will be negative. The bias has both, is precise and is unbiased is defined as the average error: - As a positive error on one item can offset a negative error on another item, a forecast model can achieve very low bias and not be precise at the same time. Obviously, the bias alone won't be enough to evaluate the forecast precision. But a highly biased forecast is already an indication that something is wrong in the model. The Mean Absolute Percentage Error (MAPE) is one of the most commonly used KPIs to measure forecast accuracy. MAPE is the sum of the individual absolute errors divided by the target value(s). It is the average of the percentage errors. MAPE is not a good forecast KPI as it is a poor-accuracy indicator. MAPE divides each error individually by the target, so it is skewed: high errors during low-target (numbers) will significantly impact MAPE.

Due to this, optimizing MAPE will result in a strange forecast that will most likely undershoot the target. The Mean Absolute Error (**MAE**) is a very good KPI to measure forecast accuracy. As the name implies, it is the mean of the absolute error. One of the first issues of this KPI is that it is not scaled to the average target value. If MAE is 10 for a particular item, you cannot know if this is good or bad. If your average target is 1000, it is, of course, astonishing. Still, if the average demand is 1, this is a very poor accuracy. To solve this, it is common to divide MAE by the average target to get a %. The Root Mean Squared Error (**RMSE)** is a strange KPI but a very helpful one, as we will discuss later. It is defined as the square root of the average squared error.

Actually, many algorithms (especially for machine learning) are based on the Mean Squared Error (MSE), which is directly related to RMSE. Many algorithms use MSE as it is faster to compute and easier to manipulate than RMSE. But it is not scaled to the original error (as the error is squared), resulting in a KPI that cannot be related to the original target scale. Therefore, it should not be used it to evaluate statistical forecast models.

On the question of error weighting:

Compared to MAE, RMSE does not treat each error the same. It gives more importance to the most significant errors. That means that one big error is enough to get a very bad RMSE. RMSE emphasizes the most significant errors, whereas MAE gives the same importance to each error. Generally a forecast of the median will get a good MAE and a forecast of the mean a good RMSE. MAPE promotes a very low forecast as it allocates a high weight to forecast errors when the target (numbers) is low. Optimization of RMSE will seek to be correct on average. In contrast, MAE's optimization will try to be as often overshooting the demand as undershooting the target, which means focusing on the target median. Understanding that a significant difference lies in the mathematical roots of MAE & RMSE is key. One aims at the median, the second aims at the average. The Root Mean Squared Error (RMSE) is one of the two main performance indicators for a regression model. It measures the average difference between values predicted by a model and the actual values. It provides an estimation of how well the model is able to predict the target value (accuracy). Mean Absolute Error (MAE) is one of the most commonly used loss functions for regression problems, MAE helps users to formulate learning problems into optimization problems. It also serves as an easy-to-understand quantifiable measurement of errors for regression problems. In MAE, different errors are not weighted more or less, but the scores increase linearly with the increase in errors. The MAE score is measured as the average of the absolute error values. The Absolute is a mathematical function that makes a number positive. Mean Bias Error (**MBE**) is primarily used to estimate the average bias in the model and to decide if any steps need to be taken to correct the model bias.

Mean Bias Error (MBE) captures the average bias (+ve or -ve) in the prediction. $\mathbf{R^2}$ is the coefficient of determination, and is a measure that provides information about the goodness of fit of a model. In the context of regression it is a statistical measure of how well the regression line approximates the actual data. This can be used to compare data sets given we know that if the elevations were identical in two models the results would also be. In this work the focus in understanding error utilizing: **{ $\mathbf{R^2}$, RMSE, MAE and MBE }**

3. RESULTS AND DISCUSSION

3A. Site Locations and Input Data

The locations (see in Figure 3) were selected purely on the basis of multiple data sets being available for either topography (terrain) or bathymetry as follows:

Site 1, is at Keramas, where there is Sonar data available, GEBCO, BATNAS and Sentinel-2 data. Four data sets in total. Site 2, is at Nyanyi, where there is ground survey points are available and FABDEM and DEMNAS data. Three data sets in total.

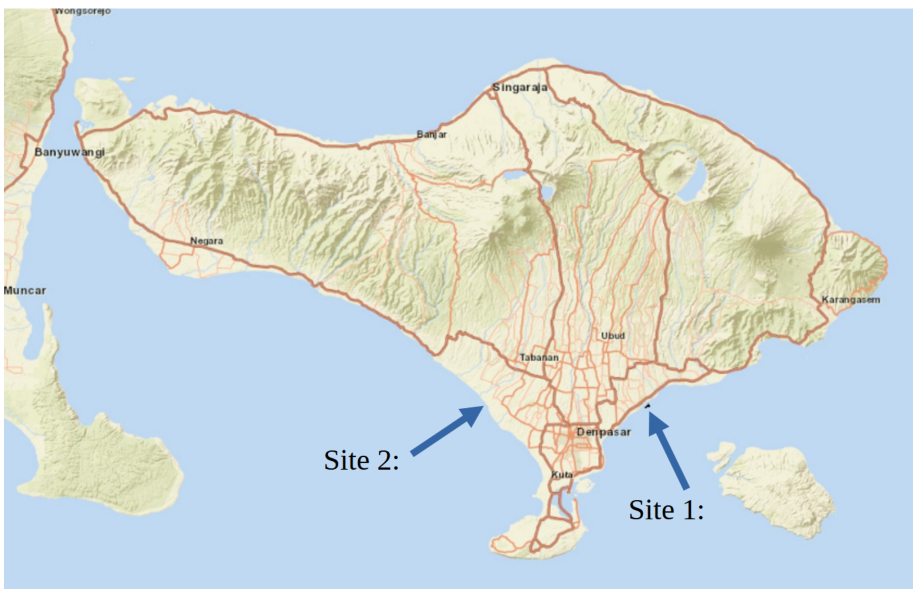


Figure 13. Location of Research Sites

For site 1 the point of truth data is assumed to be the measured SONAR data. This will be compared to the other available data sets being; GEBCO, BATNAS and Sentinel-2. The SONAR data has been filtered to below zero (Figure 4) only and contained within a manageable polygon. There are 2,211 points over an area of 715,525m^2, providing roughly a data density of 1 point per 323.62m^2. Or on average, a data point every 18m in x and y. This data can be used to create a DEM as a surface.

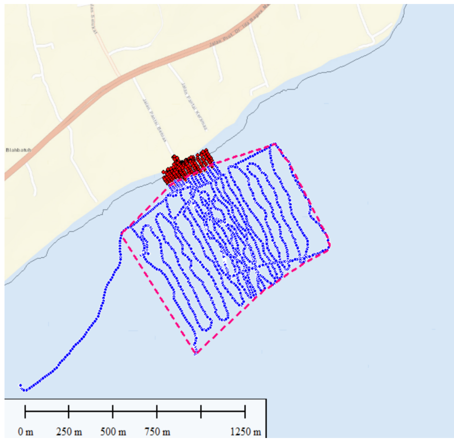


Figure 14. Site 1 SONAR data points

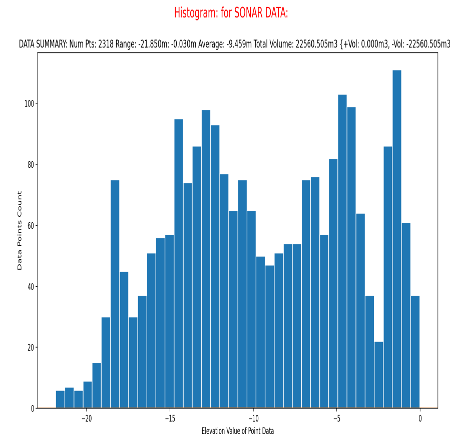


Figure 15. Site 1 SONAR (S) pts

Now that the SONAR data is both a DEM (Figure 6) and a histogram (Figure 5) of point values a comparison can be made of the data differences. This can be achieved both by reviewing the DEMS of the data surfaces and by reviewing the histograms of point elevation values. For the GEBCO data the elevation from the DEM surface is extracted at the same SONAR data points with now the ability to plot the histogram and then compare the differences. The same process can be completed for any other data set.

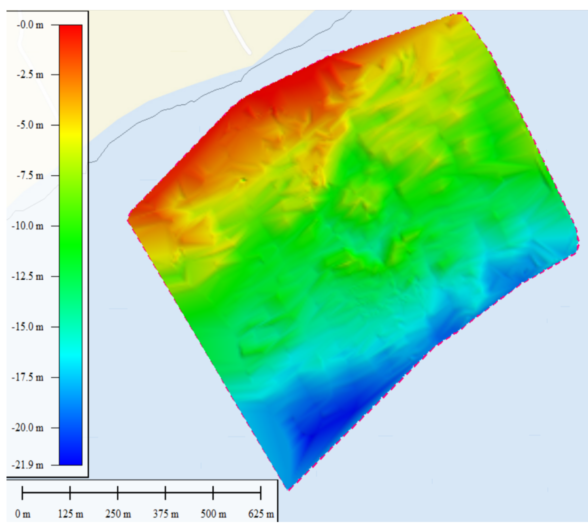
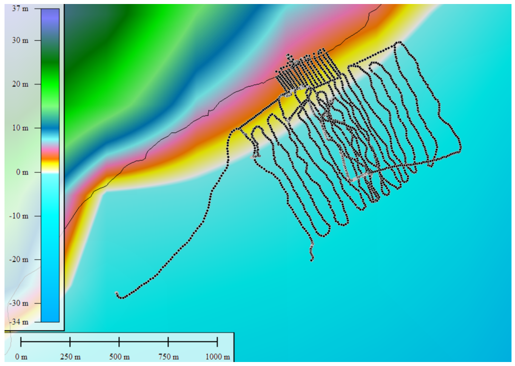


Figure 16. Site 1 SONAR (S) as DEM surface

Extracting the data from GEBCO (Figure 7 & 8) at the same data points provides a method of comparison as does the difference in the DEM as surfaces. Subtracting one surface from another surface provides the difference over the full extent of data change.



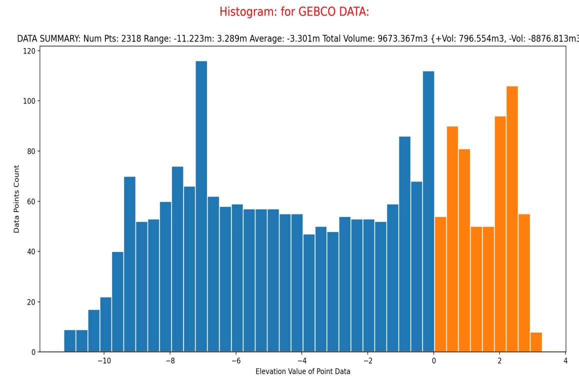
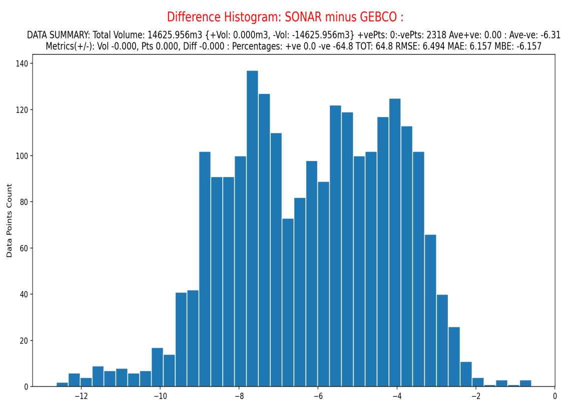


Figure 17. Site 1 GEBCO DATA (G-data)

Figure 18. Site 1 GEBCO Histogram

Subtracting the point elevation values (SONAR minus GEBCO) provides a comparison histogram (Figure 9). Similar subtracting the surface DEMS (Figure 10) provides a spatial view of difference.



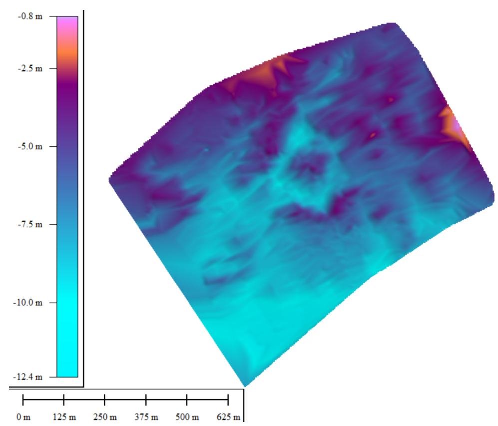
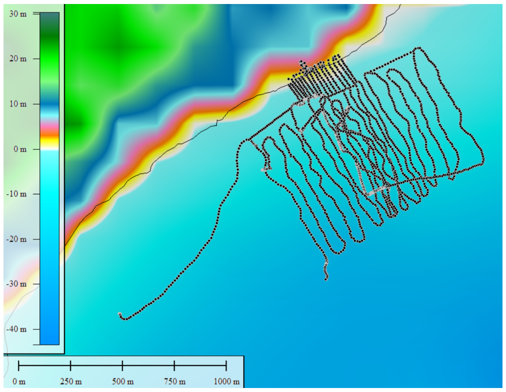


Figure 19. Site 1 Difference Histogram (S minus G)

Figure 10. Site 1 DEM Difference (S minus G)

Similarly with the BATNAS data the data can be extracted at the sonar points (Figure 11 & 12) and compared to the DEM surface. The comparison of surfaces provides a very visually rich comparison. The comparison of the points as a histogram provides a more specific approach to comparison (Figure 13 & 14).



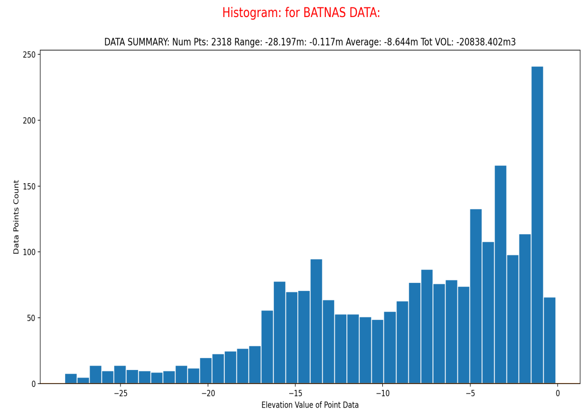
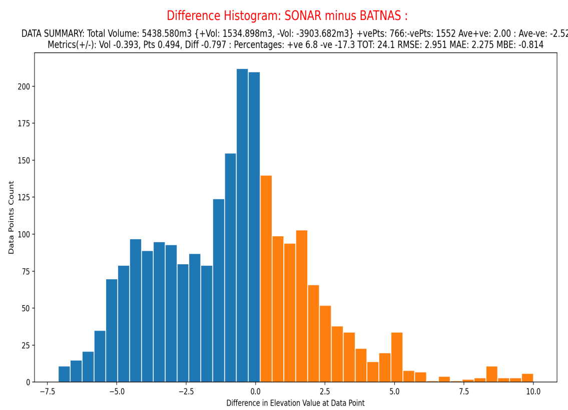


Figure 11. Site 1 BATNAS DATA (B-data)

Figure 22. Site 1 BATNAS Histogram



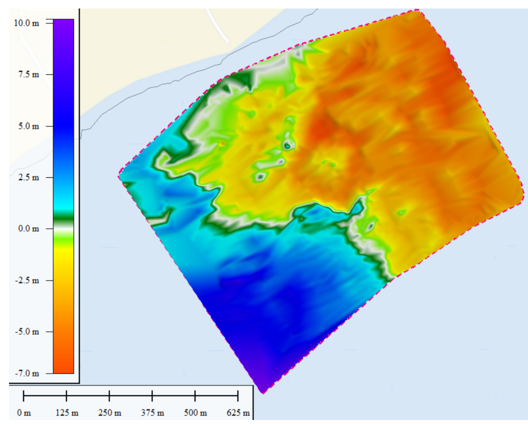
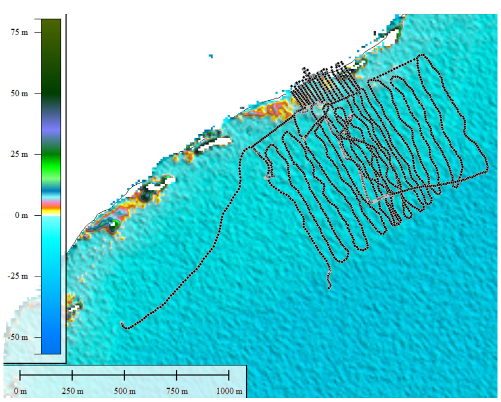


Figure 2312. Site 1 Difference Histogram (S minus B)

Figure 13. Site 1 DEM Difference (S minus B)

The final data set available is from Sentinel-2 as follows (Figure 15 & 16);



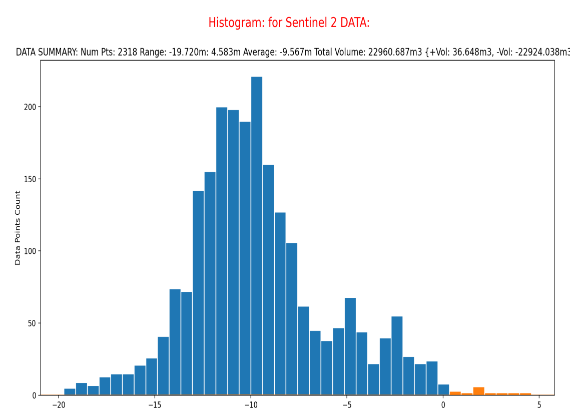


Figure 14. Site 1Sentinel-2 DATA (S2-data)

Figure 26. Site 1 S2 Histogram

Once again comparing to SONAR as points and as a DEM surface (Figure 17 & 18).

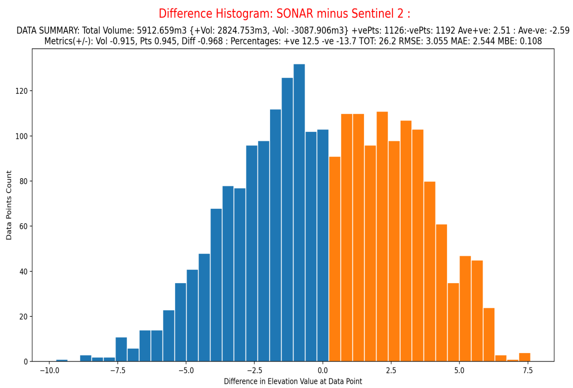


Figure 15. Site 1 Difference Histogram (S minus S2)

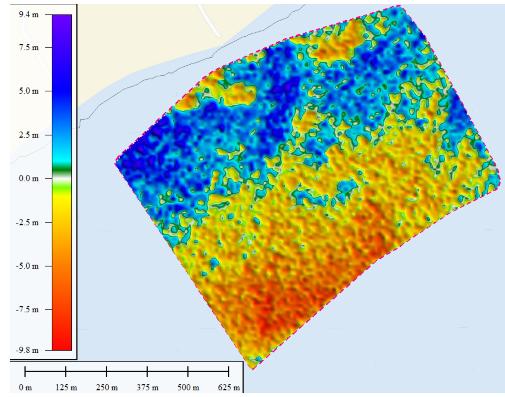


Figure 16. Site 1 DEM Difference (S minus S2)

Another approach to determining an overall difference is to compare how the surface volumes compare. For example, what is the volume contained above each of the surfaces to a specified elevation? This is relatively easy to determine and provides yet another visually rich approach to comparing differences (Figure 19).

The same approach has been applied to site 2 and its available data sets.

3B. Input Data Metrics

Looking at the input data comparing to the point of truth data for Site 1, overall the BATNAS DATA best replicates it based on 2318 points analysis. The **Points Volume Metric** of 24.1% and RMSE of 2.951. However, based on **MBE** of -0.108 the SENTINEL-2 DATA is preferred. On the basis of the **3D surface volume analysis** its suggests BATNAS the best overall candidate with SENTINEL-2 preferred for very shallow water (< 5m depth).

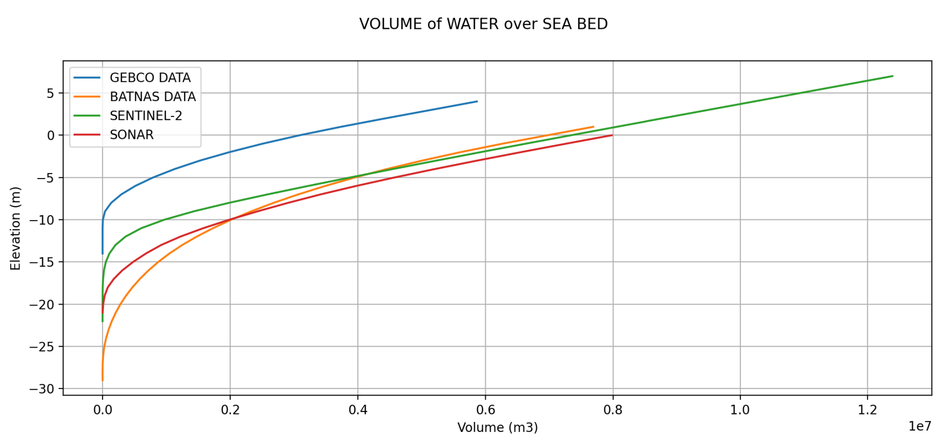


Figure 17. Comparison of water volume over sea bed surface Site 1

For Site 2, the **Points Volume Metric** suggest DEMNAS at 15.9%, which also coincides with the **MBE** of -1.104. The RMSE of 3.026 and MAE of 2.173 prefers the FABDEM data set. Reviewing the terrain based on **3D surface volume** identifies the DEMNAS DATA as the best candidate to replicate the SURVEY.

3C. Tsunami Models

The ANUGA HYDRO MODEL has been set up and run for all of the scenarios described (Figure 20 & 21). The ANUGA HYDRO model produces a single output file (*.SWW) which contains the full time history of conserved quantities (Elevation, Stage, Momentum X and Momentum Y). As it is a finite volume model velocity is not a conserved quantity in the model.

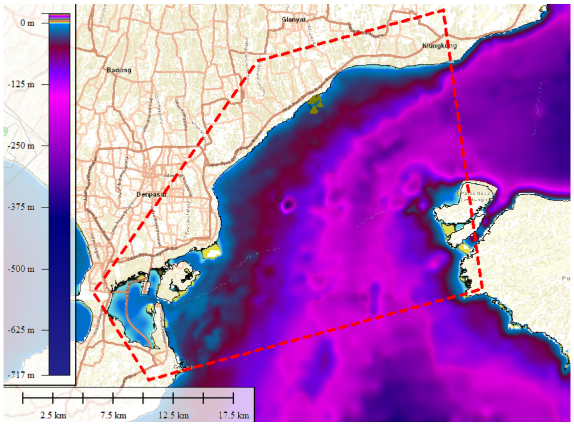
Velocity is derived from Momentum as {V = M/Depth}, where;
M = Sqrt (MomX x MomX + MomY x MomY)
Depth is Stage(Water Elevation) minus Bed Elevation.
This is important as in some model applications such as erosion the elevation can change. For all models since the focus is on the influence of terrain data change, the adopted values of surface roughness are not considered particularly important. For all model runs Manning's N is set at 0.035.

Site 1 has four (4) models with only Bathymetric data being changed between models. The mesh for all models was set at (10x10m) {100m2} cells.

Site 2 has three (3) models with only the terrain data being changed between models.
In addition a further model was set up to test the sensitivity of results to the mesh refinement. The mesh for all models was set at (10x10m) {100m2} cells, except for the refined models which used (5x5m) {25m2}. The Models have been run for 1 hour (3600 seconds in the models).

In addition any surface can be extracted and viewed in QGIS, or exported to any other GIS platform. Further for any location within the models it is possible to extract the conserved quantities and from those derive many other quantities such as:

- Depth
- Velocity
- bed Shear (VxVxD)
- Froude
- Specific Energy (VxV/2g)



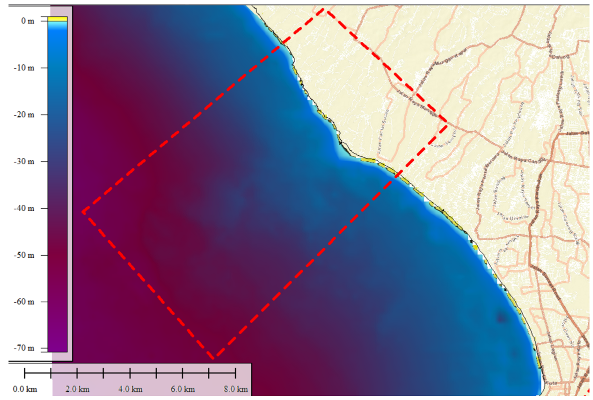
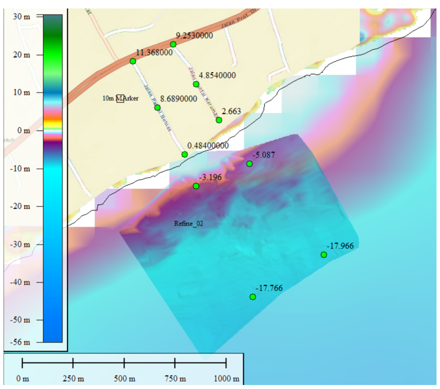


Figure 18. Site 1 Tsunami Model Extent

Figure 19. Site 2 Tsunami Model Extent

Data is to be extracted at Gauge locations for each site as indicated in (Figure 22 & 23).



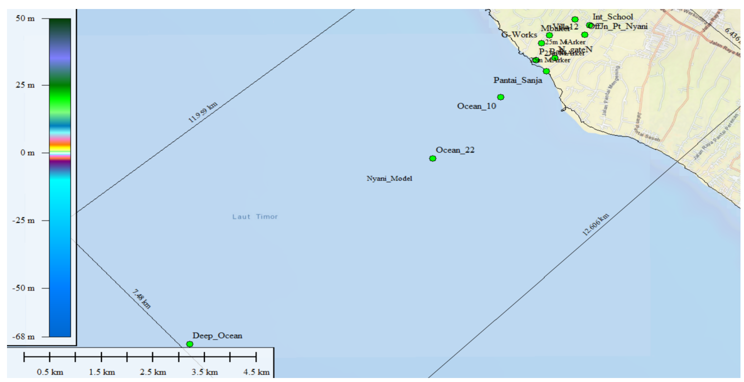


Figure 20. Site 1 Gauge Locations

Figure 21. Site 2 Gauge Locations

The models were run as described and data extracted as both time series and as a surface of maximum momentum (Hazard). For Site 1 the four models (Figure 24-27) produce the maximum momentum plots as shown and these can be used to look into the difference of hazard resulting from the difference in input data for each of the models.

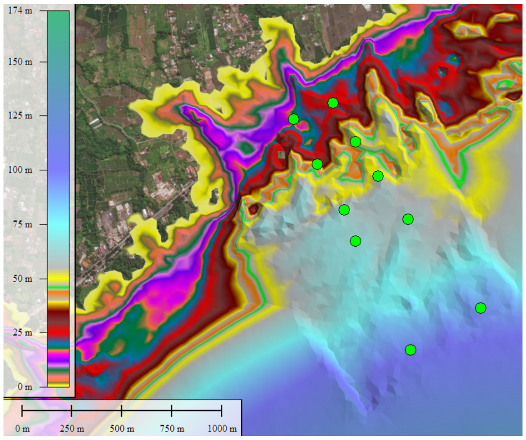


Figure 22. Max Hazard Sonar

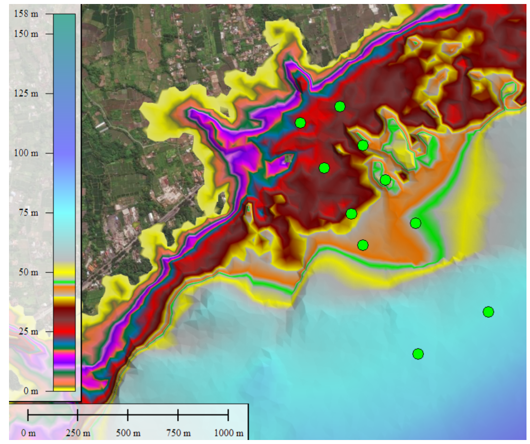


Figure 23. Max Hazard GEBCO

Visually the Gebco model is immediately different. The BATNAS and Sentinel based models are far more similar to the SONAR model, but clearly not identical.

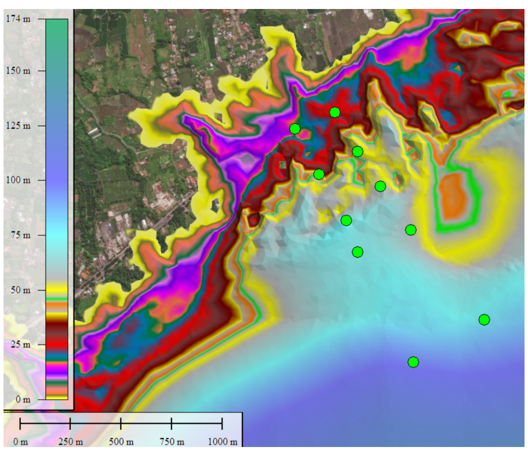


Figure 24. Max Hazard BATNAS

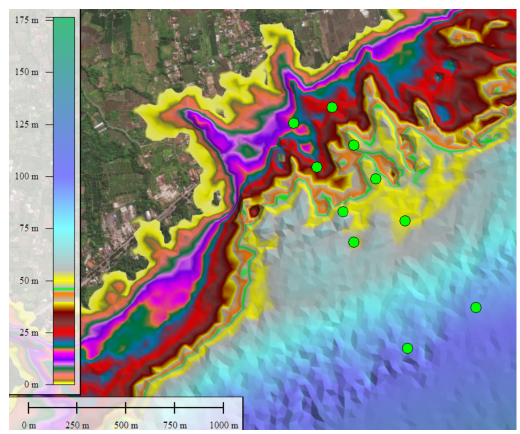
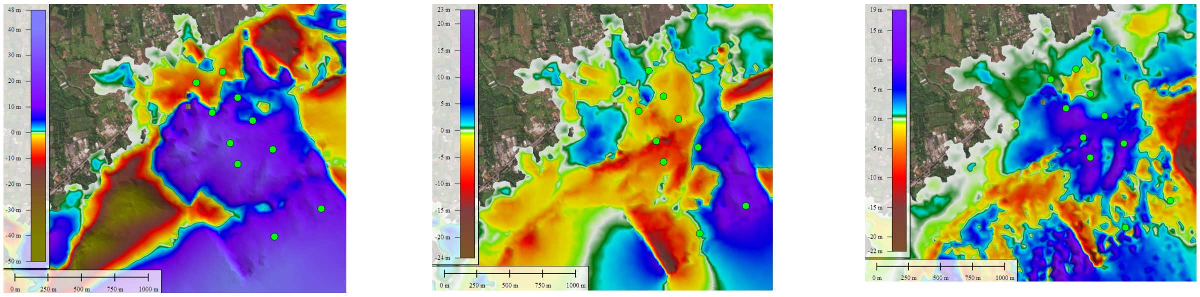


Figure 25. Max Hazard Sentinel

The differences (Figure 28) in the model results can be visually indicated by creating a difference plot of results of each of the models compared to the SONAR model results.



Difference Sonar-Gebco Difference Sonar-Batnas Difference Sonar-Sent

Figure 26. Comparison of Max Hazard Differences for Site 1 models

To drill into more detail time series at the gauge points will be explored further.
The Hazard is defined from Velocity times Depth (Momentum).
The maximum hazard can therefore be plotted spatially for each of the models as for site 1.

For Site 2 the following plots show the maximum momentum (Hazard) for the 3 models (Figure 29-31) run and the difference plots again indicate the resulting variations only as a result of changing the input terrain data.

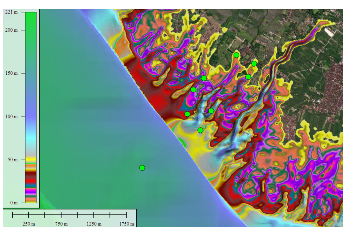


Figure 27. Max Hazard Survey

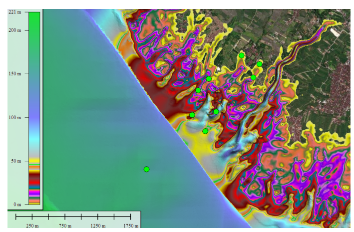


Figure 28. Max Hazard Demnas

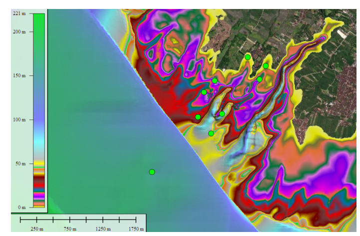
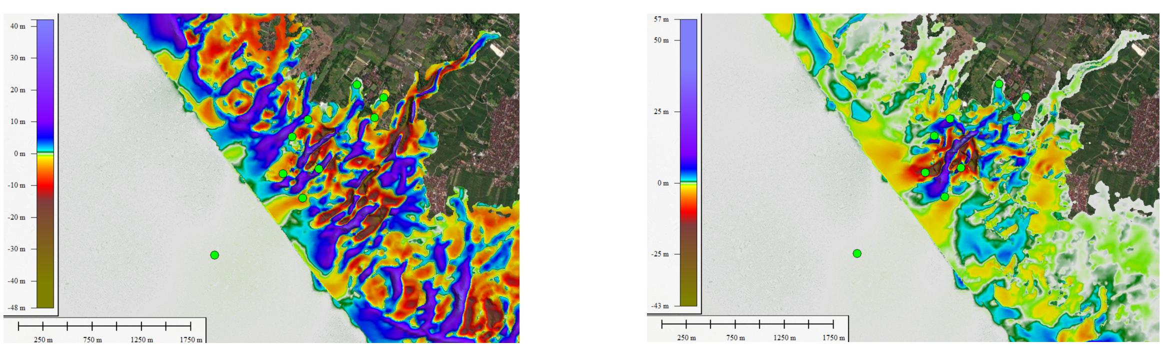


Figure 41. Max Hazard Fabdem

Once again the very obvious difference (Figure 32) here is the Fabdem data set results. Looking at the difference between model results also identifies the extent of differences.



Difference Survey-Demnas Difference Survey-Fabdem

Figure 29. Comparison of Max Hazard Differences for Site 2 models

Once again further details can be extracted from time series behaviour over the entire simulation.

3D. Extracting and comparing Time Series Data

For site 1 there are 10 locations where gauge data has been extracted the following time series plots are an example of those extracted (Figure 33). Site 2 has similar plots for all its gauge points.

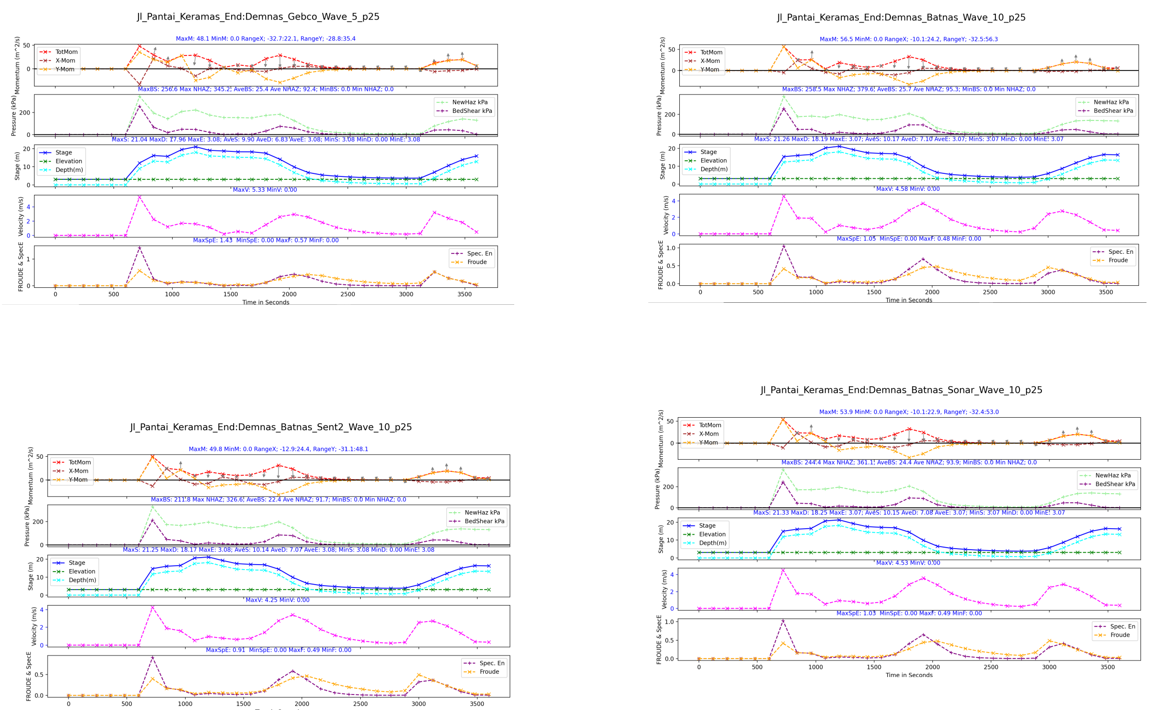


Figure 30. Time Series 4 models at Site 1: Location Jl Pantai Keramas End

The difference plots show the variation between the point of truth model (SONAR) compared to the other data models (GEBCO, BATNAS, SENTINEL). Hence, for Site 1 at each location there are three (3) difference plots (Figure 33-36) . The difference plots statistics have also been accumulated and will be presented here as a form of summary and to draw conclusions. A critically important aspect to be aware of in the nature of tsunami analysis is that of the reflective wave, or the retreating wave on land. The difference at times can be more pronounced in the retreating phase than in the initial wave attack. This is clear in the example shown in the following figures.

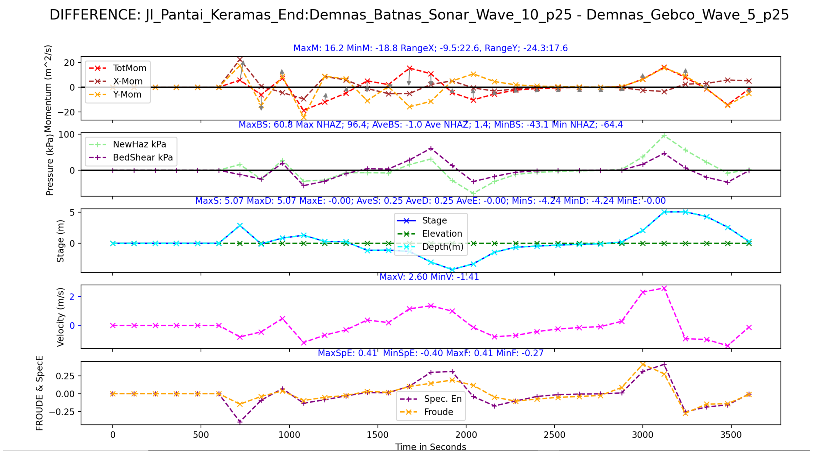


Figure 31. Site 1 Location Jl. Pantai Keramas End Difference SONAR minus GEBCO

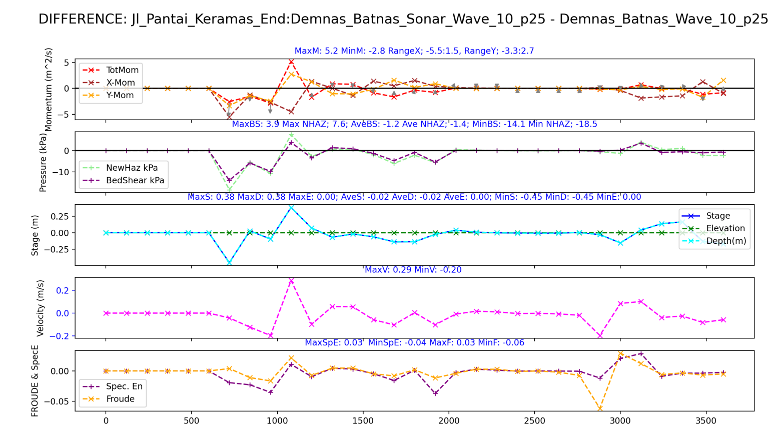


Figure 32. Site 1 Location Jl. Pantai Keramas End Difference SONAR minus BATNAS

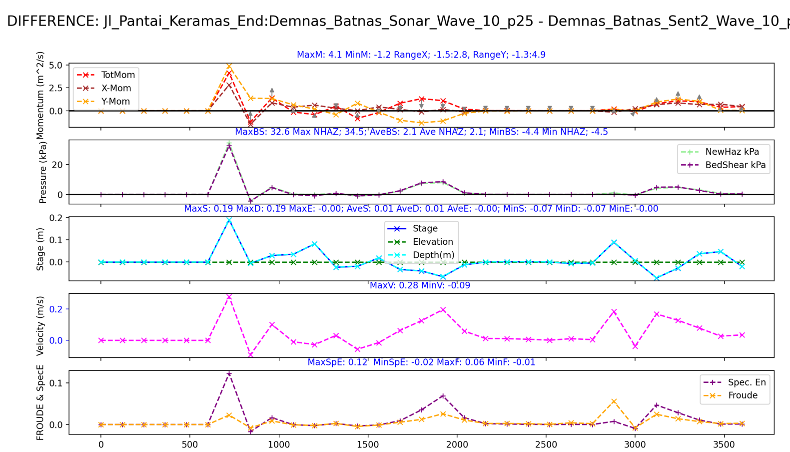


Figure 33. Site 1 Location Jl. Pantai Keramas End Difference SONAR minus SENTINEL

For the 10 Gauge Locations in 3 comparison models at Site 1 the R-squared terms show that for the Bathymetric data changes (only) the impact the surface water elevation (Stage) the least in the SENTINEL model compared to the point of truth SONAR model. Similarly Momentum is best replicated in the SENTINEL model when compared to the point of truth SONAR model. For Bed Shear and a newly considered hazard measure based on pressure, the BATNAS model best replicates the point of truth SONAR model.

For the 11 Gauge Locations in 2 comparison models at Site 2 the R-squared terms show that for the terrain changes (only) the impact on all terms, (Stage, Total Momentum, Bed Shear and the Pressure Hazard term considered), the best model to replicate the point of truth SURVEY model is the DEMNAS based model.

SITE 1:	**SITE 2: Mesh 100**
SONAR - GEBCO	**SURVEY - FABDEM**
Stage 0.9084336	Stage 0.9999042
TotM 0.7677961	TotM 0.9988720
BedShear 0.7769386	BedShear 0.9992434
NewHaz 0.8781592	NewHaz 0.9992782
SONAR - BATNAS	**SURVEY - DEMNAS**
Stage 0.9987725	**Stage 0.9999943**
TotM 0.9922413	**TotM 0.9999616**
BedShear 0.9727753	**BedShear 0.9999540**
NewHaz 0.9908819	**NewHaz 0.9999559**
SONAR - SENTINEL-2	**SITE 2: Mesh 25 (Refined)**
Stage 0.9996891	**SURVEY – DEMNAS Run1**
TotM 0.9926519	**SURVEY – DEMNAS Run2**
BedShear 0.9726684	Stage 0.9999854
NewHaz 0.9818269	TotM 0.9999705
	BedShear 0.9999627
	NewHaz 0.9999642

Table 1 R-Squared Terms for Site 1 and Site 2

The other METRICS considered for the OUTPUT results included **RMSE, MAE and MBE** as follows:

SITE 1:
Demnas_Batnas_Sonar_Wave_10_p25 - **Demnas_Gebco_Wave_5_p25**
Stage RMSE: 2.197 MAE: 1.629 MBE: 0.368
TotM RMSE: 12.728 MAE: 8.935 MBE: 4.213
BedShear RMSE: 35.566 MAE: 20.897 **MBE: -0.060**
NewHaz RMSE: 91.825 MAE: 83.827 MBE: 83.343

Demnas_Batnas_Sonar_Wave_10_p25 - **Demnas_Batnas_Wave_10_p25**
Stage RMSE: 0.252 MAE: 0.132 MBE: 0.053
TotM RMSE: 3.257 MAE: 2.171 MBE: 1.920
BedShear RMSE: 12.381 MAE: 4.442 MBE: -1.237
NewHaz RMSE: 44.880 MAE: 43.628 MBE: 43.628

Demnas_Batnas_Sonar_Wave_10_p25 -
Demnas_Batnas_Sent2_Wave_10_p25
Stage RMSE: 0.127 MAE: 0.059 MBE: 0.017
TotM RMSE: 2.616 MAE: 1.385 MBE: 1.218
BedShear RMSE: 18.781 MAE: 7.006 MBE: -4.363
NewHaz RMSE: 41.685 MAE: 40.994 MBE: 38.013
Table 2 OUTPUT METRICS for Site 1

COMMENTS:
Site 1. Bathymetric Data Change
In running multiple models and only adjusting the Bathymetric data it was found that the Closest estimate to the SONAR data was as a result of the Sentinel-2 DATA with RMSE result for Stage, Total Momentum (VxD) and a New Hazard Term. However, the best RMSE result for BedShear was the BATNAS data.

The same result was clear in the R-squared term also.

SITE 2:
BatNas_DemNas_SURV_Wave_10_p25 -
BatNas_FabDem_Wave_10_p25
Stage RMSE: 0.018 MAE: 0.006 MBE: 0.002
TotM RMSE: 2.571 MAE: 1.941 MBE: -1.160
BedShear RMSE: 8.941 MAE: 6.103 MBE: -5.356
NewHaz RMSE: 8.888 MAE: 6.082 MBE: -5.332

BatNas_DemNas_SURV_Wave_10_p25 -
Batnas_Demnas_Wave_10_p25
Stage RMSE: 0.004 MAE: 0.002 MBE: 0.001
TotM RMSE: 0.426 MAE: 0.313 MBE: -0.097
BedShear RMSE: 1.816 MAE: 1.177 MBE: -0.491
NewHaz RMSE: 1.805 MAE: 1.164 MBE: -0.481

SITE 2: REFINED:
BatNas_DemNas_SURV_Wave_10_M25 -
Batnas_Demnas_Wave_10_M25
Stage RMSE: 0.007 MAE: 0.003 MBE: -0.000
TotM RMSE: 0.364 MAE: 0.187 MBE: 0.001
BedShear RMSE: 1.743 MAE: 0.843 MBE: -0.077
NewHaz RMSE: 1.741 MAE: 0.841 MBE: -0.077
Table 3 OUTPUT METRICS for Site 2

COMMENTS:

Site 2: Terrain Data Change

In running multiple models and only adjusting the terrain it was found that the DEMNAS data provided the closest estimate to GROUND SURVEY for all terms likely to be related to hazard.
The same result was shown in the R-squared term

These results have been plotted in simple graphs showing the input metric change against the output metric change, providing a measure of sensitivity (dy/dx). For site 1 input MBE the Output MBE has a sensitivity is 0.5. For site 2 the sensitivity is 1.85 as shown (Figure 37-38).

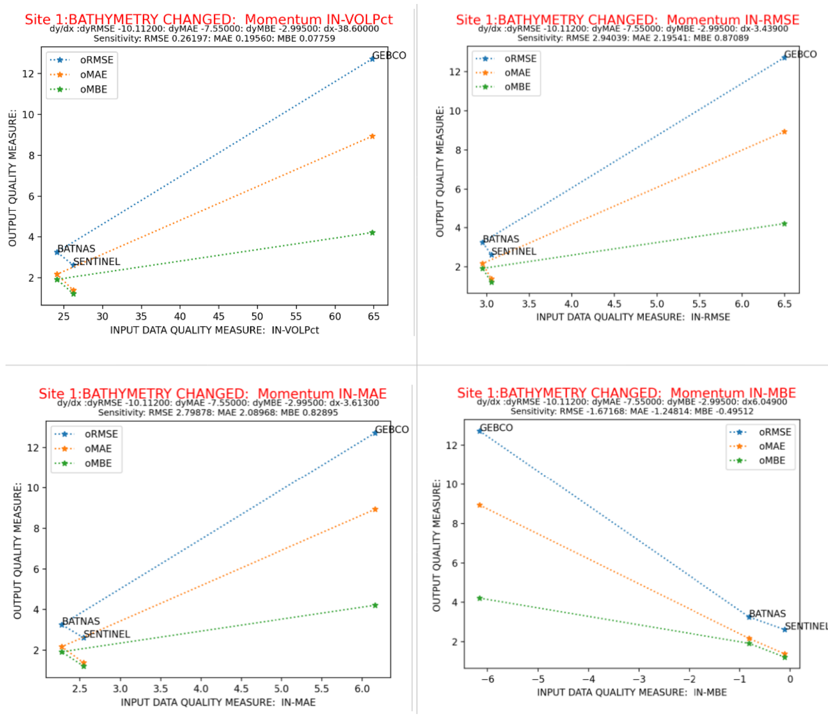


Figure 34. Site 1 RESULTS

Plotting input data and output data variability, indicates sensitivity to change.

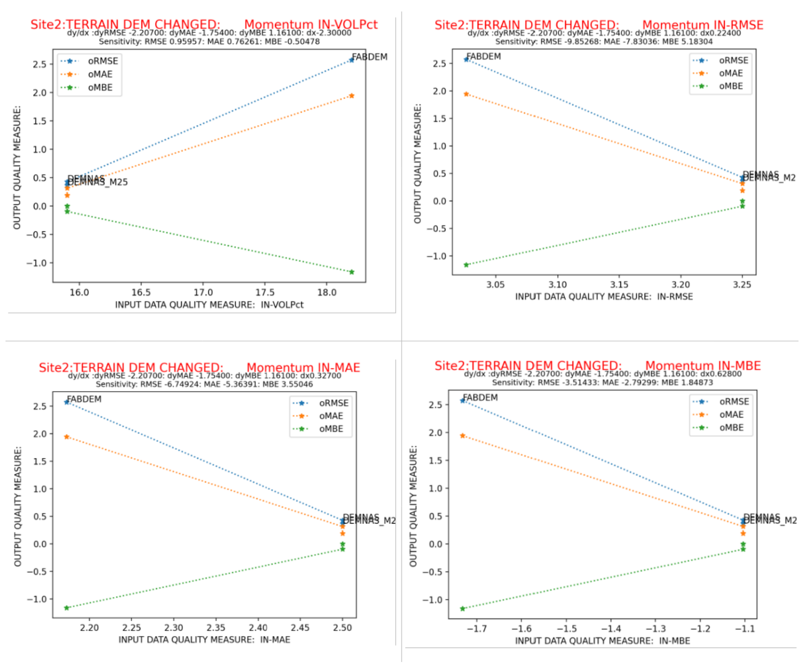


Figure 35. Site 2 RESULTS

4. CONCLUSIONS

For this body of work the objective was to identify the influence of data quality on the resulting determination of hazard for tsunami analysis. The conclusion clearly identifies better candidate metrics for inputs and results that link input data quality to output data quality. The broader objectives were to at least discuss the use of a framework, and to discuss how hazard is defined.

The following can be concluded:

1. The model topographic input data has a great influence on model results.
2. Input data quality has been shown to influence resulting outputs that define tsunami hazard.
3. Input data can be systematically reviewed by the inclusion of data metrics to compare data sets. The input metrics 3D surface volume , Points Volume Metric and Mean Bias Error (MBE) appear to perform better than Root Mean Square Error in identifying candidates for best input model DEMS.
    a. For Site 1 Identifying BATNAS as the best general candidate and SENTINEL-2 in the shallower water (<5m depth) to replicate the SONAR model. For Bathymtric DATA, the model results compared from each of the models for the sites suggests R-Squared performs well to identify a split between BATNAS and SENTINEL-2 whilst RMSE, MAE and MBE select SENTINEL-2 for Stage, Total Momentum, and the suggested new Pressure hazard term whilst identify BATNAS to perform best for Bed Shear in replicating the point of truth SONAR model.
    b. For Site 2 Identifying DEMNAS is the best general candidate to replicate the SURVEY model as land based terrain. For Terrain DATA, the model results compared from each of the models for the sites suggests R-Squared, RMSE, MAE and MBE, each identified the DEMNAS model to perform best for Stage, Total Momentum, Bed Shear and the suggested new Pressure hazard term in replicating the point of truth SURVEY model.
4. It was shown that momentum (current approach to hazard) is sensitive to changes in input data. Even though the extent of measure (range) was limited, the best consistent metric was MBE as compare to RMSE and MAE.
5. The better spread of values from a Pressure based term may be a better candidate for describing hazard compared to the current globally adopted Momentum based term. This made measuring sensitivity more pronounced.
6. Through the extent of work undertaken in relation to the findings presented, it is likely that a framework based approach would stream line many aspects of similar future ventures in tsunami analysis.

In terms of real life adoption and application, the following can be suggested. The need to update time consuming, expensive tsunami modelling requires a specific trigger. One trigger being the availability of updated and improved data. The identification of the level of improvement of new data sets can be estimated through the processes utilised in this thesis, whereby an improvement metric in output can be estimated from the improvement metric in input.

ACKNOWLEDGEMENTS

This body of work could not have been completed without the assistance and guidance of a number of people. Their assistance is greatly appreciated and acknowledged.
Thesis Primary Supervisor: Pak Gede Pringgana; Thesis Secondary Supervisor: Ibu Ni Nyoman Pujianiki; Provision of Data, and support , Pak Komang Gede Putra Airlangga
Staff in the Udayana International Office Administration

REFERENCES

Airlangga KGP (2023) “SONAR Single Beam Sounding Data Thesis: Analysis Of Sentinel-2 Image Utilization For Bathymetric Depth Extraction, (Case Study: Bali Province)".

Bonham, A.J., Hattersley, R.T.(1967),; "Low Level Causeway Water Research Laboratory Report 100".

Cárdenas, Catalán, (2022);"Accelerating Tsunami Modeling for Evacuation Studies through Modification of the Manning Roughness Values", GeoHazards 2022, 3, 492–507. https://doi.org/10.3390/ geohazards3040025

Fatmawati, Yuliara, Riandhita, Kelo, Vellicia, Reswari: (2019) "Tsunami Level Disaster Based on Simulation Scenario of Earthquake Modeling and Seismicity in South Bali 2010-2018"

Filippucci, Brocca, Bonafoni, Saltalippi, Wagner, Tarpanelli, (2022); "Sentinel-2 high-resolution data for river discharge monitoring ", Remote Sensing of Environment 281 (2022) 113255

Flouri, Kaligeris, Alexandrakis, Kampanis, and Synolakis, (2010) ;"Numerical modelling of tsunami waves: Application to the simulation of an earthquake generated tsunami"

GITEWS, DLR / GTZ, Balinese Working Group for Tsunami Hazard Mapping; (2010); "Technical Documentation: Tsunami Hazard Maps for Bali:-Multi-scenario Tsunami Hazard Maps for Bali, 1:100,000 :- Multi-scenario Tsunami Hazard Maps for Southern Bali, 1:25,000; with zoning based on wave height at coast (in line with the InaTEWS warning levels) as well as probability of areas being affected by a tsunami."

Griffin, Latief, Kongko, Harig, Horspool, Hanung, Rojali, Maher, Fuchs, Hossen, Upi, Dewanto, Rakowsky, Cummins; (2015) "An evaluation of onshore digital elevation models for modeling tsunami inundation zones"

Kjell Karlsrud [NGI]; (2009) "Tsunami Risk Reduction Measures phase 2, Tsunami Risk evaluations for Indonesia, Ref: 20061179-00-227-R"

Kuncoro, Setiawan, Kusuma, Manessa, Susetyo; (2021) “Rapid Bathymetry Detection With Sentinel Application Platform (Snap) Using Sentinel Imagery 2a”

Laksono, Widagdo, Aditama, Fauzan, Kovács; (2022) "Tsunami Hazard Zone and Multiple Scenarios of Tsunami Evacuation Route at Jetis Beach, Cilacap Regency, Indonesia"

Marras, Mandli; (2021):"Modeling and Simulation of Tsunami Impact: A Short Review of Recent Advances and Future Challenges"

Mungkasi , VanDrie, Roberts (2013),“Predictions on arrival times of water of the St. Francis dam break flood using ANUGA”, 20th International Congress on Modelling and Simulation MODSIM 2013

Murray, Shand, Higuera; (2021) "Hydrodynamic Modelling of Tsunami Inundation Behaviour in Urban Environments":Australasian Coasts & Ports 2021 Conference – Christchurch, 30 November – 3 December 2021

Nielsen, O., Roberts, S., Gray, d., McPherson, A.,Hitchman, A.(2005); "Hydrodynamic modelling of coastal inundation" :: MODSIM2005

Roberts, Nielsen, Jakeman, (2006):"Simulation of Tsunami and Flash Floods"

Roberts, Stals, Nielsen; (2007) "Parallelisation of a finite volume method for hydrodynamic inundation modelling", ISSN 1446-8735 http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/153

Saberi, Kabolizadeh, Rangzan, Abrehdary (2022); Accuracy assessment and improvement of SRTM, ASTER, FABDEM, and MERIT DEMs by polynomial and optimization algorithm: A case study (Khuzestan Province, Iran) https://doi.org/10.1515/geo-2022-0455

Schlurmann, Kongko, Goseberg, Natawidjaja, Sieh; (2010) "Near-Field Tsunami Hazard Map Padang, West Sumatra: Utilizing High Resolution Geospatial Data And Reasonable Source Scenarios"

Sukmono, Aji 1, Amarrohman, Bashit, Saputra; (2022) "The Extraction of Near- Shore Bathymetry using Sentinel-2A Satellite Imagery: Algorithms and Their Modifications" :: TEM Journal. Volume 11, Issue 1, pages 150-158, ISSN 2217‑8309, DOI: 10.18421/TEM111-17, February 2022.

Sulistian, Aditya, Mugiarto, Istighfarini, Harrys, Sofian, (2021);"Pemutakhiran Digital Elevation Model Nasional Wilayah Laut" (Updating of National Ocean Digital Elevation Model, Case Study: Aceh Waters Area)

Synolakis, Bernard (2006);"Tsunami science before and beyond Boxing Day 2004", DOI: 10.1098/rsta.2006.1824

Tennakoon,(2004); "Parameterisation of 2D Hydrodynamic Models and Flood Hazard Mapping for Naga City, Philippines" ::Thesis, International Institute for Geo-Information Science and Earth Observation Enschede, The Netherlands.

Titov, V.V. (1997): Numerical modeling of long wave runup. Ph.D. thesis, University of Southern California, Los Angeles, State California,

Valentra, Rachmayani, Diastomo, Sapiie; (2022) "Potential Tsunami Due To Submarine Landslide In The South Of Bali Island: A Preliminary Study"

VanDrie, Ghetti, Milevski(2018), “Who is Responsible for Checking Flood Models?“, ASFPM Annual Conference, Phoenix Arizona, USA, 17-21 June

VanDrie R. (2008), “Is There A Better Definition Of Hazard?”,48th Annual NSW Floodplain Management Authorities Conference

Van Drie, Simon, Schymitzek (2009);"ANUGA:- the FREE Ocean Impact model", 12th NSW Local Government IT Conference, Coffs Harbour NSW

Van Drie, Milevski, Simon;(2011) "Validation of a 2-D Hydraulic Model - ANUGA, to undertake Hydrologic Analysis"; 34th IAHR (International Association for Hydro-Environment Engineering and Research),BIENNIAL CONGRESS, 33rd Hydrology and Water Resources Symposium, 10th Conference on Hydraulics in Water Engineering

Varnuska M. (2005); "Surface reconstruction of Geometrical Objects from Scattered Points", Doctoral Thesis, University Of West Bohemia Faculty of Applied Sciences

Wuppukondur, A., & Baldock, T.E.; (2020) "Modelling of tsunami wave overtopping in a converging channel" ::22nd Australasian Fluid Mechanics Conference AFMC2020, Brisbane, Australia,7-10 December 2020. Brisbane, Australia, The University of Queensland. https://doi.org/10.14264/042c06b.

Zoppou, C., and S. Roberts;(1999), "Catastrophic collapse of water supply reservoirs in urban areas" ::Journal of Hydraulic Engineering, American Society of Civil Engineering, 125(7), 686–695.