

IDENTIFYING AND REMOVING SLOPPY PARAMETERS IN UNDERWATER
TRANSMISSION LOSS MODELS

by

Hadassah B. Griffin

A senior thesis submitted to the faculty of

Brigham Young University - Idaho

in partial fulfillment of the requirements for the degree of

Bachelor of Science

Department of Physics

Brigham Young University - Idaho

April 2021

Copyright © 2021 Hadassah B. Griffin

All Rights Reserved

BRIGHAM YOUNG UNIVERSITY - IDAHO

DEPARTMENT APPROVAL

of a senior thesis submitted by

Hadassah B. Griffin

This thesis has been reviewed by the research advisor, research coordinator,
and department chair and has been found to be satisfactory.

Date

Mark Transtrum, Advisor

Date

John Paul Johnson, Committee Member

Date

Hatt Richard, Committee Member

Date

R. Todd Lines, Chair

ABSTRACT

IDENTIFYING AND REMOVING SLOPPY PARAMETERS IN UNDERWATER TRANSMISSION LOSS MODELS

Hadassah B. Griffin

Department of Physics and Astronomy

Bachelor of Science

Underwater transmission loss profiles can be used by machine learning systems to identify the type and characteristics of an ocean seabed. When the machine learning system's underlying calculations are based on a physical model, its predictions become more accurate. However, some transmission loss models are sloppy, meaning that some parameters in the model have a large uncertainty when the model is fit to data. These parameters with large parameter uncertainty are “sloppy” parameters, which can be removed from the model without significantly changing the model's predictions. Reducing a transmission loss model to one with fewer parameters yet approximately the same predictions can improve a machine learning system's performance. I was tasked with identifying and removing sloppy parameters from the transmission loss models for five seabeds. By using manifold boundary approximation methods, I successfully reduced three of the five seabed models.

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Mark Transtrum and Dr. Traci Neilsen of Brigham Young University, who mentored me in my undergraduate research process during these past two years. I also want to thank Michael Mortenson, fellow research group member, for his enthusiasm and the work we did together on this topic. Finally, I would like to acknowledge my spouse, who has supported and encouraged me as I work on this research.

Contents

Table of Contents	vi
List of Figures	vii
1 Context and Motivation	1
1.1 Machine Learning and Transmission Loss Models	1
1.2 Sloppy Models	4
2 Methods	9
2.1 Transmission Loss Model Details	10
2.2 Mathematical Representation of Model and Cost	12
2.3 Local Analysis of the Model Manifold: Jacobians and the Fisher Information Matrix	13
2.4 Global Analysis of the Model Manifold: Likelihood Profiles	17
2.5 Analyzing Costs of Reduced Models	21
3 Results and Future Work	22
3.1 Results	22
3.2 Conclusions	28
3.3 Future Work	28
Bibliography	29
A Parameter and Values for Seabed Environments	31
B Local Analysis Identified Sloppy Parameters	33
C Cost Results of Model Reduction	35
D Workflow Diagram and Explanation of the Code	37
E Code Link	39

List of Figures

1.1	A transmission loss plot for an underwater environment with a sandy seabed. The x axis represents the distance range from a sound source, which is located in the top left corner of the plot. The y axis represents the depth below the sound source. The coloring of the contours represents the sound intensity. As a sound gets further from the source in the top left corner, the intensity begins to die away. Image credits given to BYU's underwater acoustics research team.	3
1.2	A simplified diagram which exhibits how a machine learning system processes data. Input data is given to the machine, represented by the red input layer. In the hidden layers, the white circles represent weights. Based on the training a machine learning system has had, these weights will have different values and help the system determine what the final output value or prediction is. The final output layer is represented by the rightmost rectangle. Image credits [1].	3
1.3	A graphical representation of mapping from parameter space to prediction space. Image credits [2].	6
1.4	A generic example of reducing the manifold by approximating it by its boundary. With each iteration, parameters from the model are removed. After multiple approximations, the model is be represented as a hypercorner which still has behaviors of the original model manifold. Image credits [3].	7
2.1	A graphical representation of the model of the sand environment that we used for our numerical simulations. The water layer, single sediment layer, and half-space layer (represented by bottom two blocks) are shown. The values for the 11 parameters used in the model are also shown. The other seabed environments have a similar arrangement of sediments, but contain different sediment values. This plot was created by code from BYU's underwater acoustics research group.	11

2.2	Figure 2.2a shows a cost surface for a 2-D toy model. The surface values are calculated with Equation 2.2, and change based on the values of the parameters. 2.2b shows a contour mapping of this surface. The star in both images represents where the model is optimized for both parameters θ_1 and θ_2 . The denoted widths of the ellipses in 2.2b represent the values of the diagonal elements of the Fisher Information Matrix (FIM). Figures created by Michael Mortenson, graduate student of Brigham Young University.	14
2.3	A Fisher Information Matrix plot created for the transmission loss data for the 100 Hz sand seabed model. The axes represent the different parameters in the model. The color bar shows how correlated these parameters are with respect to each other. Lighter colors represented higher correlated parameters. Darker colors represent very uncorrelated parameters. The uncorrelated ones usually are sloppy parameters. Plot created by Michael Moretenson, BYU graduate student.	16
2.4	Figure 2.4a has contours of the cost surface shown. The brighter colors represent the higher cost interval. The central red dot represents the best fit of parameters one and two. The black X's represent values found when optimizing the model for fixed values of θ_1 . The cost of the model is recorded at these locations and used to create a likelihood profile curve like 2.4b. Figure 2.4a originally created by Yonatan Kurniawan, BYU graduate student.	18
2.5	Contour plot of the cost values for shifting the values of ρ_{hsp} and ρ_2 in the 9 parameter transmission loss model. The dots represent parameter values calculated by the optimization algorithm. Notably, the cost contour and the optimized value agree on a linear relationship for values in the first quadrant. This indicates that these parameters can be removed from the model as a ratio. The diverging likelihood-calculated parameter values are suspected to occur because of some other parameter in the model shifting the optimal values of these parameters as they decrease in value.	20
3.2	All likelihood profiles for the 11 parameter mud model.	24
3.1	All likelihood profiles for the 11 parameter sand model.	24
3.3	Contour plot of the cost values for shifting the values of ρ_{hsp} and c_{hsp} in the 11 parameter sand transmission loss model. The dots represent the values of the parameter calculated by the likelihood profile calculation as it optimized the model. The parameters generally agree with the contour shape. Note that the lowest contour cost value appears for higher values of c_{hsp} . This agrees with our likelihood profiles, which indicate that c_{hsp} must be taken to higher parameter values in order to reduce the model.	27

D.1 Descriptive work flow diagram of the steps that take place to reduce the model.	38
---	----

Chapter 1

Context and Motivation

1.1 Machine Learning and Transmission Loss Models

The acoustic method of analyzing passive sonar has been used by scientists to study ocean characteristics such as ocean depth, sediment type, and the number of sediment layers in a given seabed. Passive sonar refers to analyzing ambient ocean noise, which can come from various sources such as waves crashing on the ocean surface, ship noise, and biological activity [4]. Approaches to interpreting passive sonar information include matched-field processing, image-processing based techniques, and machine learning algorithms [5].

The latter approach for analyzing passive sonar has been the method of choice for the underwater acoustics research group at Brigham Young University (BYU). In a machine learning analysis approach, a computer algorithm is given test data to train with. The algorithm tries to recognize patterns and relationships within the training data in order to make more accurate predictions when given a newer data set. The underwater acoustics group uses machine learning to take passive

sonar data and use that data to identify seabed types and the number of sediment layers in a given ocean environment [6]. This machine learning approach can be useful for cases in which it is difficult to take in situ measurements of a given ocean environment, but we still want to have information about the environment itself. One of the passive sonar data sets analyzed by the machine learning systems are transmission loss profiles. Transmission loss profiles are contour plots that show how sound intensity changes over a range of space at a given frequency. (See, for example, Figure 1.1.) As the machine learning system works on identifying and learning what a particular environment is like based on the transmission loss profile, it also tries to identify characteristics of the environment. For example, the system can try to identify how close a ship passes by, or what type of sediment the seabed below is [5]. If we treat the environment like a model, the characteristics of this environment are like the parameters of a model. Based on the parameters chosen for this modeled environment, weights (parameters in the machine learning process) are made. After the learning process is complete for the machine learning system, these weights help the machine learning identify the environment based on a given acoustic signal (Figure 1.2). Parameters used for transmission loss include the density of the layer that the sound travels through, the sediment layers on the ocean floor, the attenuation (or energy loss) of the sound, and the sound speed.

There are challenges, however, with making the machine learning system efficient at identifying seabeds and what their seabed-specific characteristics are. For example, two distinct “soft” seabed environments can have the same attenuation values for their first sediment layer. So, other values have to be taken into consideration to accurately identify the seabed. A second challenge is that there are limited data sets for the machine learning system to test on. This challenge can be counteracted by creating physics-based models that generate similar artificial data. Some researchers

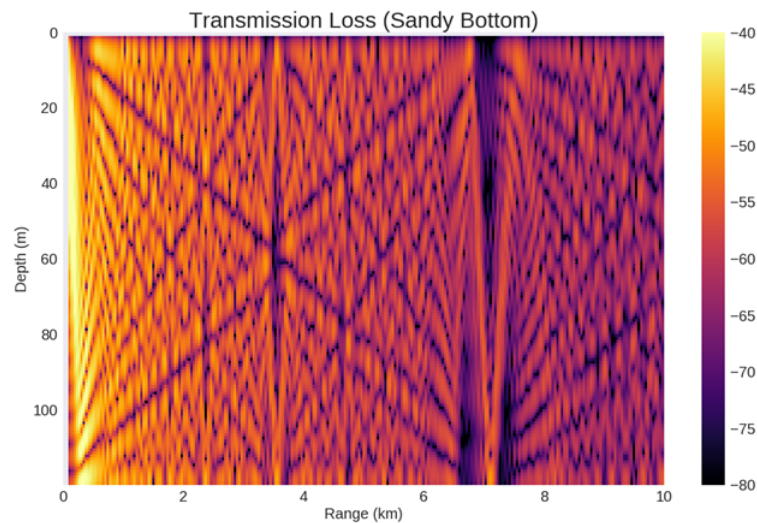


Figure 1.1 A transmission loss plot for an underwater environment with a sandy seabed. The x axis represents the distance range from a sound source, which is located in the top left corner of the plot. The y axis represents the depth below the sound source. The coloring of the contours represents the sound intensity. As a sound gets further from the source in the top left corner, the intensity begins to die away. Image credits given to BYU's underwater acoustics research team.

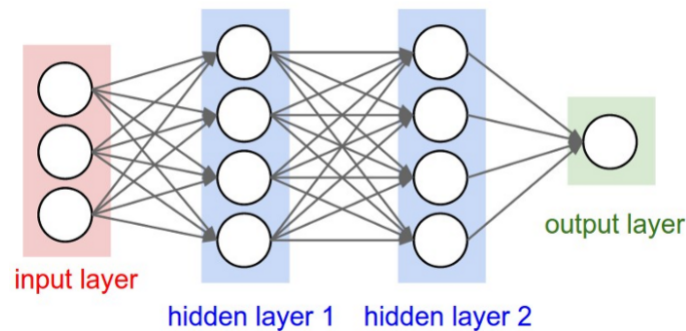


Figure 1.2 A simplified diagram which exhibits how a machine learning system processes data. Input data is given to the machine, represented by the red input layer. In the hidden layers, the white circles represent weights. Based on the training a machine learning system has had, these weights will have different values and help the system determine what the final output value or prediction is. The final output layer is represented by the rightmost rectangle. Image credits [1].

have already applied this method with their own machine learning systems [5].

To create data based on physical models, the BYU acoustics group used ORCA [7], a numerical simulation program that can create transmission loss profiles for various seabed environments and other acoustical factors. ORCA has been successful in creating training data with known parameter values for each environment. However, after training on simulated data, the parameter values that the machine learning system identifies from real data still have uncertainty in them. This inversion process, or the process by which the machine learning system identifies the environment parameter values, inherently has some uncertainty.

To counteract this challenge, work has been done by some researchers to quantify the uncertainty in identifying these sound parameters [4]. This quantification of uncertainty is valuable because it can be used to inform a machine learning system which sound parameters can be estimated from an environment and which ones cannot. This in turn can decrease the run time of the machine learning system's learning process, because it will be informed about values that are identifiable for an environment and that can be reasonably learned. Thus, the machine learning system will take less time trying to identify parameters that are physically unreasonable to detect. This is useful in the case of real-time measurements made by those working in the ocean who need parameter value estimates quicker. Additionally, if the machine learning system is used to identify sound parameter values, the person requesting the information can be informed about the value's uncertainty once it has been quantified.

1.2 Sloppy Models

My research group enters the picture at this point. Dr. Tracianne Neilsen of BYU tasked us with quantifying the uncertainty in sound parameters in transmission loss

profiles for different seabed environments. Like the machine learning group, we used ORCA to generate the data for the transmission loss profiles for different models of seabed environments. We chose to model gravel, sand, silt, clay, and mud seabeds. Each model consisted of 11 parameters. More details about these models and the parameters used will be explained in section 2.1.

Our analysis began by noting that the transmission loss models were sloppy models. Sloppy models are a complex models that have large parameter uncertainty when they are fit to data. [8] These complex models also tend to be nonlinear and have very anisotropic responses to changes in parameters [9]. Anisotropic means that the magnitudes of values in these spaces will vary, depending on the direction one travels.

As a result of the anisotropies, there are some parameters, by themselves or in combination with others, that are “stiff” and others that are “sloppy.” Stiff parameters are ones which cause the predictions of a model to change more as the stiff parameter value changes. Sloppy parameters, on the other hand, can get shifted around by themselves or in combination with others and the model predictions do not change as much. [2].

To analyze the stiffness and sloppiness of parameters, we can use an information geometry approach. Using information geometry allows us to explore anisotropies in the parameter space and to quantify the model’s manifold [8]. To understand what a model manifold is, consider a model to be a mapping from parameter space into a data or prediction space. The model manifold is the geometric surface in prediction space which represents the predictions resulting from all the possible variations of the parameters. (Figure 1.3). Each dimension or degree of freedom in the model manifold represents a parameter in the model.

Usually, this model manifold is bounded [8] [2]. The property of model manifold’s being bounded allows us to measure different distances to the boundaries [2]. These

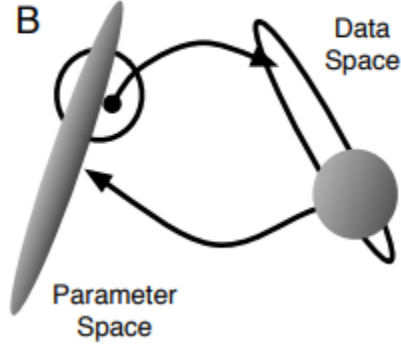


Figure 1.3 A graphical representation of mapping from parameter space to prediction space. Image credits [2].

manifold boundaries represent where parameter values in our model have been taken to extremes such as negative infinity, zero, or positive infinity, depending on the physically allowed values for a parameter. Finding and characterizing these boundaries at extreme values of parameters helps us know how a model’s predictions change based on the value of a parameter. The behavior of the manifold as parameters are taken to its extremes informs us if the parameter is sloppy or stiff. Sloppiness and stiffness can then be quantified to inform us about the uncertainty in a parameter. Additionally, identified sloppy parameters can be removed based on where they become sloppy in the model manifold in order to simplify the model.

The information geometry approach that Dr. Transtrum proposed for our research was manifold boundary approximation method (MBAM). Along with fulfilling the goal of quantifying parameter uncertainties for the transmission loss model, using MBAM would allow us to simplify the models by informing us how to remove sloppy parameters. MBAM has previously been used successfully on models such as loads for power systems [8] and for biological systems [2] [9].

The “boundary approximation method” part of MBAM occurs when parametric

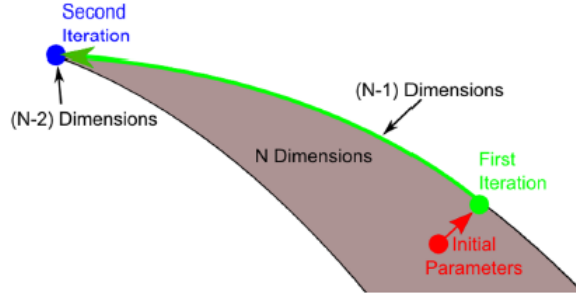


Figure 1.4 A generic example of reducing the manifold by approximating it by its boundary. With each iteration, parameters from the model are removed. After multiple approximations, the model is represented as a hypercorner which still has behaviors of the original model manifold. Image credits [3].

degrees of freedom of the manifold are removed one by one in a systematic way by approximating the entire manifold by its boundary. This works since boundaries of an n -dimensional model manifold are $(n - 1)$ -dimensional manifolds themselves. After a number of approximation steps, the reduced model is represented by a hyper-corner of the original model manifold. If the process is done successfully, this reduced model is a good approximation of the original [8]. An visual example of approximating the manifold for a generic model is given in Figure 1.4.

An advantage to using MBAM on sloppy models is that it is able to identify sloppy parameters on a global and local scale. Work on a global scale is useful for identifying an absolute minimum for nonlinear models, which tend to have multiple minima in the manifold [2]. Global methods are also useful for identifying the manner in which to remove the sloppy parameter, and thereby approximate the model. For our work on the underwater transmission loss model, we planned to first create a computer program that could get the numerical uncertainty in each of these parameter in the model. Second, we planned to identify the sloppy parameters that could be removed from the model. Identifying and removing sloppy parameters successfully can create a reduced model that has numerically similar results. Accomplishing both of these

goals would help the machine learning system that is identifying seabed environments be more informed in its learning process, because the system would know which parameters can be estimated or not. Thus, its run time performance would improve. Additionally, the uncertainty on the parameter predictions would be quantifiable.

I was tasked with identifying and removing the sloppy parameters from these transmission loss models. I used the Jacobians of the model and the Fisher Information Matrix to identify sloppy parameters on the local scale of the model manifold. I then used a likelihood profile method to identify sloppy parameters on the global scale of the manifold. By using both of these methods, I identified sloppy parameters in the model under a certain set conditions for our five seabed models. I also identified the order in which these parameters could be removed from each of the models. For each reduced model, I quantified the cost difference between the original and reduced models. Thus, I was able to show how reliable each of the reduced models would be for their predictions, depending on how many sloppy parameters were removed from the model. Three seabed models were successfully reduced, one had minor success for model reduction, and the final was not reduced successfully.

Chapter 2

Methods

To accomplish the goal of identifying sloppy parameters and removing them from the transmission loss models of different seabeds, I needed to work on the local and global scale of the model manifold in order to obtain reliable results. As explained before, manifold boundary approximation method (MBAM) analysis done at the local and global scales is important in this model reduction process. Local calculations are performed on the region in the model manifold close to the mapped original value of a parameter. Global regions extend beyond this region. The local calculations are computationally faster and can identify potential sloppy parameters that can be compared against the global scale later. The global scale is computationally slower, but the results are more reliable. The possible discrepancy between local and global analysis is due to the fact that local and global features are not guaranteed to be related for nonlinear models. The nonlinearity of a model can lead to it to having multiple local minima on its cost surface. This can frustrate numerical methods searching for the global minima [2]. Therefore, global methods must be used to validate the locally identified manifold features in question.

In the following sections, I give further details on the transmission loss models, the

mathematical representation of the models, the local scale and global scale methods, and how well the original models compare with the reduced ones.

2.1 Transmission Loss Model Details

I was tasked with reducing the models for five different seabed environments. As noted in the first chapter, the ORCA program was used to generate the data.

We selected 11 parameters in our transmission loss model, which had empirically determined values for each seabed. The parameters and their values for each environment are listed in Appendix A. The five seabeds were, in order from hardest to softest: gravel, sand, silt, clay, and mud. We focused on these seabeds since they offered a good spectrum of different model types to try.

The basic arrangement through which the sound would travel was a 75 meter water column, a 35 meter thick layer of the sediment in question, and a half-space, or basement layer, of sediment. The half-space layer is present to ensure that the transmission loss calculation is numerically bounded and is representative of the attenuation of sound deep below the sea floor. An example of the basic environment set up is shown in Figure 2.1.

Other important parameters in our data simulations are as follows. The water above the seabed was modeled as an isotropic water column so that there would be no sound refraction in the water. This was chosen so that we could just focus on the effects of the seabed's sediment on the sound propagation. A frequency of 100 Hertz (Hz) was used for the sound source. Preliminary analysis of the seabed models done by Mackenzie Allen, a student in Dr. Transtrum's research group, identified this frequency as one high enough to distinguish when parameters would begin to exhibit sloppy behavior.

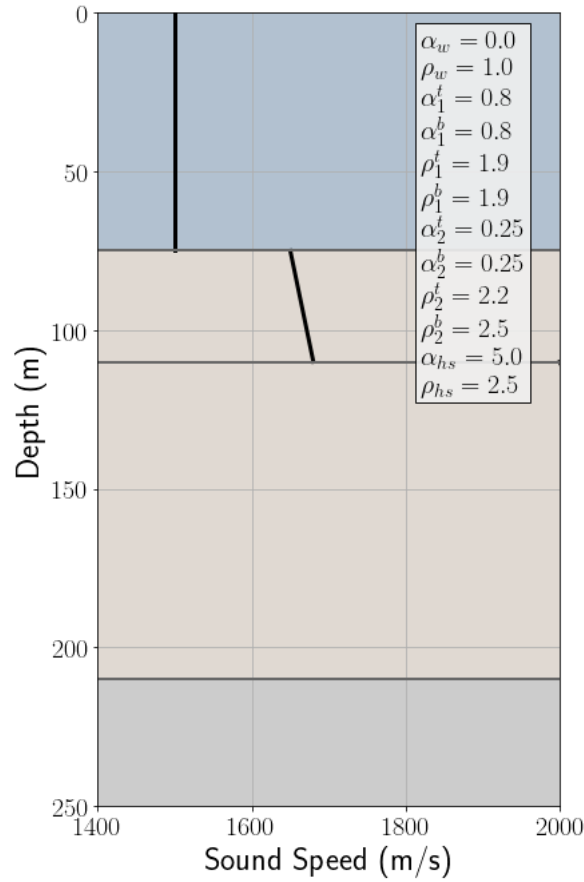


Figure 2.1 A graphical representation of the model of the sand environment that we used for our numerical simulations. The water layer, single sediment layer, and half-space layer (represented by bottom two blocks) are shown. The values for the 11 parameters used in the model are also shown. The other seabed environments have a similar arrangement of sediments, but contain different sediment values. This plot was created by code from BYU's underwater acoustics research group.

The 100 Hz sound source in the model, which simulates the sound that the transmission loss plot is based off of, was positioned 9 kilometers from the array of “receivers.” These receivers are what approximates the sound from the simulated source. 256 receivers were positioned at evenly spaced intervals at depths of 10 - 74 meters below the simulated ocean surface. The large array of receivers was created to capture a continuum of noise from the simulated ocean so that the transmission loss data would be more complete. The code used to build the model for analysis can be found in a code repository link in Appendix E, and is entitled “Environments.py”.

2.2 Mathematical Representation of Model and Cost

The analysis of the transmission loss model requires two things: a mathematical representation of the model, and a method to best fit the model to data.

With regards to the former matter, we mathematically can represent our n -parameter model like so. Let t be our regressor variable. We have sampled data points t_m with observed behaviors y_m and respective uncertainties σ_m . Our mathematical model is $f(t, \theta)$, for parameters θ [2].

With this representation, we can now determine which parameters in the model create the best fit for our data. One approach to assessing this is to assume that our model can reproduce the data, with some error term that accounts for discrepancies:

$$y_m = f(t_m, \theta) + \zeta_m.$$

Here, ζ_m are random variables that are assumed to be independently distributed according to the normal distribution $N(0, \sigma_m^2)$ [2]. We quantify the difference between the predicted and actual data of the fit using residuals:

$$r_m(\theta) = \frac{y_m - f(t_m, \theta)}{\sigma_m}. \quad (2.1)$$

We determine what these residuals are for each θ by finding the parameter values which minimize the cost function:

$$C(\theta) = \frac{1}{2} \sum_m r_m(\theta)^2 \quad (2.2)$$

The method of minimizing this cost function depends on whether we are doing a local or global analysis of our mathematical model. For our local analysis, we evaluate the Fisher Information matrix at the minimized values. More details will be given about this process in Section 2.3. The global method used the Simplex algorithm to minimize the cost along a “likelihood profile”. More information on this will be given in Section 2.4.

2.3 Local Analysis of the Model Manifold: Jacobians and the Fisher Information Matrix

With the model set up, the tools of Manifold Boundary Approximation Method (MBAM) can be applied. We begin with our analysis at the local scale. The following example will illustrate a simplified way that the manifold can be analyzed locally. For visualization purposes, a hypothetical 2-dimensional toy model will be chosen. Refer to Figure 2.2, which shows a cost surface and a contour plot for the toy model. This cost surface is constructed by values of the cost function (Equation 2.2) for different parameter values of the model. The axes in this graphical representation represent values of the two parameters in the model. The center dot represents the best fit of the model for the two parameter values.

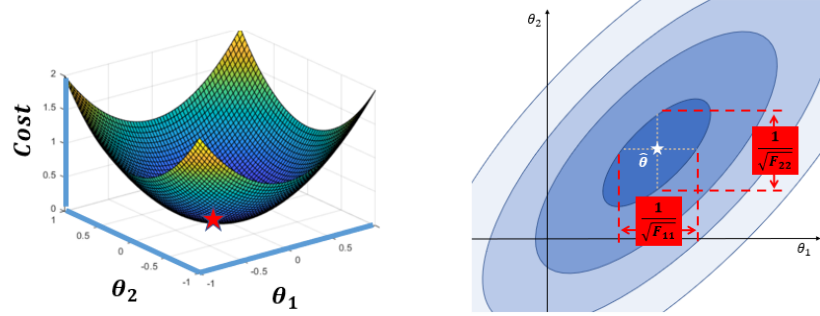


Figure 2.2 Figure 2.2a shows a cost surface for a 2-D toy model. The surface values are calculated with Equation 2.2, and change based on the values of the parameters. 2.2b shows a contour mapping of this surface. The star in both images represents where the model is optimized for both parameters θ_1 and θ_2 . The denoted widths of the ellipses in 2.2b represent the values of the diagonal elements of the Fisher Information Matrix (FIM). Figures created by Michael Mortenson, graduate student of Brigham Young University.

Note that in Figure 2.2b, there are regions where the prediction oval is thicker and some where it is thinner. If a parameter moves in a direction that is thinner for the ellipse, this represents a stiff parameter. It can only be varied a little before moving the prediction of the parameter outside of the model. On the other hand, if a parameter moves in the elongated direction of the ellipse, it is in a confidence region of the contour for longer. This represents a sloppy parameter, which means that it can be changed more without significantly deviating from the model's main prediction.

Since there is a dimension for each parameter in a model on the model manifold, we have an 11-dimensional object to analyze. Envisioning this object is rather difficult since (1) we live in a 3-dimensional space and (2) the plots we can make to represent data in more than 3-dimensions require creative workarounds for each dimension you want to add in. Fortunately, there are tools available to help us identify the information we need in this multi-dimensional prediction space. The tools we used for local analysis were the Jacobians of the model and the corresponding Fisher

Information Matrix (FIM).

The Jacobian matrix is a matrix of the derivatives of the cost function of a model with respect to the parameters in that model [2]:

$$J_{m\mu} = \partial_\mu r_m \quad (2.3)$$

where r_m is defined by Equation 2.1. The first expression on the right hand side is defined as

$$\partial_\mu \equiv \frac{\partial}{\partial \theta_\mu}, \quad (2.4)$$

which denotes the partial derivatives with respect to each parameter in the model. The subscripts in Equation 2.3 indicate the positions within the matrix.

The FIM is a square matrix composed of the Jacobian transpose matrix and Jacobian matrix of a model:

$$F = J^T J. \quad (2.5)$$

FIM and the Jacobians are useful tools because they inform about the uncertainty in each parameter value. In the case of Jacobians, we can do a singular value decomposition on the Jacobian matrix to gain valuable information:

$$J = U \Sigma V^T \quad (2.6)$$

The symbol Σ in 2.6 represents a matrix whose diagonal contains the singular values of the Jacobian. The singular values are the square roots of the eigenvalues of the FIM matrix: $\sigma_i = \sqrt{\lambda_i}$. V is a matrix whose columns represent the the local eigenparameters of the FIM with eigenvalues $\lambda_i = \sigma_i^2$. U matrix columns represent eigenpredictions. The singular values are of note because they inform us how much the

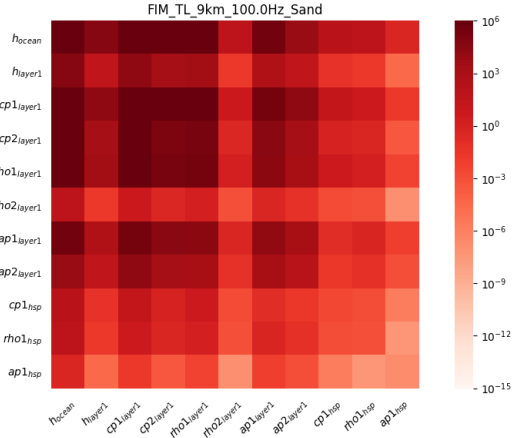


Figure 2.3 A Fisher Information Matrix plot created for the transmission loss data for the 100 Hz sand seabed model. The axes represent the different parameters in the model. The color bar shows how correlated these parameters are with respect to each other. Lighter colors represented higher correlated parameters. Darker colors represent very uncorrelated parameters. The uncorrelated ones usually are sloppy parameters. Plot created by Michael Moretenson, BYU graduate student.

the mapped eigenvectors of U stretch in prediction space. Therefore, if the singular values of the Jacobian have low values, that means that the prediction made by their respective parameters do not change a lot. Therefore, those particular parameters are sloppy and have a high uncertainty [2]. The diagonals of the FIM matrix quantify uncertainty in each parameter. The FIM also informs us how correlated parameters are to each other within the model. See, for example, Figure 2.3. The brighter colors on this plot represent higher parameter correlations. The darker colors represent low parameter correlations.

In our toy 2-dimensional model in Figure 2.2b, there are red brackets spanning the width of the ellipse of the first contour surrounding the optimal value of the parameters in the model. This value is the same as $\frac{1}{\sqrt{FIM_{ii}}}$, where the indices indicate the corresponding diagonals of the FIM for each parameter. Within this bracketed region, the parameter can vary while still yielding the same cost.

Along with using the FIM to visually and numerically identify which parameters were more correlated, I used the Jacobians of each model to identify frequencies at which the highly uncorrelated parameters became sloppy. The Jacobians for different seabeds and a range of frequencies were calculated using code contained in the code repository link in Appendix E entitled “CalculateJacobians.py”. After these calculations were completed, I created a program, also contained in the repository, that performed the singular value decomposition on the Jacobians. If the singular values were below a certain threshold, the parameters were identified as sloppy and the frequency at which this occurred was noted. This local analysis gave preliminary guesses for what parameter could be sloppy the global scale.

2.4 Global Analysis of the Model Manifold: Likelihood Profiles

With local analysis completed on the manifold by using the Jacobians and the Fisher Information Matrix (FIM), the next step was to run calculations on the global scale.

We used the global analysis method of creating likelihood profiles of for each parameter in a given model. A likelihood profile is a cost curve traced out by a parameter that is held at a fixed value while the rest of the parameters in the model are shifted around until the minimum cost is found.

We will return to our 2-dimensional toy model example to help us visualize how this process is done. In Figure 2.4, the yellow ellipse in the center of the plot represents the analysis that we performed before on the local scale in Figure 2.2b. Notably, instead of having just ellipses as before, there is a nonlinear canyon shape. This contour surface is created by error-calculating algorithms that directly calculate the cost of holding one parameter value fixed and moving around the others in the model.

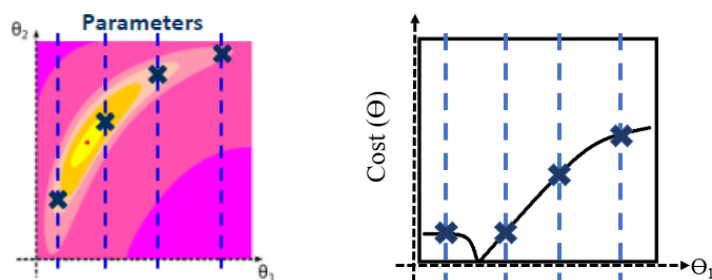


Figure 2.4 Figure 2.4a has contours of the cost surface shown. The brighter colors represent the higher cost interval. The central red dot represents the best fit of parameters one and two. The black X's represent values found when optimizing the model for fixed values of θ_1 . The cost of the model is recorded at these locations and used to create a likelihood profile curve like 2.4b. Figure 2.4a originally created by Yonatan Kurniawan, BYU graduate student.

We are interested in finding how the cost of the function changes as we shift a parameter value. We take a parameter, in the 2-dimensional toy model case, on the x-axis of our graph. We fix its value at a point, denoted by the blue dashed lines on the graph. We then shift the other parameter value on the y-axis around until a value that minimizes the cost of the model is located, which on these plots is represented with a black x. We keep shifting our parameter on the x-axis and repeat this process, recording the minimal cost value as we go. If the optimization of the y-axis's parameter is done correctly, the cost values recorded at each step will follow the shape of the cost canyon.

The cost value at these x's trace out the likelihood curve. The shape of this cost curve is informative, because it tells us how the cost of a model changes when a parameter of interest is held fixed. If we continue to examine Figure 2.4b, we can see that as θ_1 is taken to lower values, the cost flattens out and has a lower magnitude. On the other hand, increasing the value of θ_1 increases the cost of the model. If the cost has a relatively low magnitude, it indicates that if we take θ_1 to an extremely low value in the model, it becomes sloppy. This then informs us of a possible way

to remove θ_1 from the model. On the other hand, we know that for large θ_1 values, the model differs more between the original and predicted value. That means that θ_1 acts like a stiff parameter in that direction. Therefore, we should not take it to high values if we want to remove it from the model.

Thus, by using a simple error and shifting method to generate the contour plot and a global optimization method to trace out the likelihood profile, we have two tools to verify if a parameter in our 2-D toy model is indeed sloppy, as well as identify values the parameter can be fixed at to reduce the model.

Returning to our 11 parameter transmission loss model, its analysis is more difficult than our 2-dimensional toy model. In this case, it is better to first calculate the likelihood values. The contours are calculated for 2-dimensional cross sections of the model, for two parameters of interest. The likelihood profiles with low costs for their parameters show us the ones to investigate with the contour plots. We used our contour plots to verify that our optimization algorithm was working effectively, as well as check possible ways to remove parameters. Though some parameters can be removed from the model by fixing their values at an extreme, other parameters correlate with one another and must be handled differently. An example of this is where parameters act like a ratio with respect to one another. The contour plots can help identify this behavior, as is demonstrated in Figure 2.5.

We calculated the contour costs of shifting our original parameter value using the script in the repository link entitled “CalculateContourCost.py”. We calculated the likelihood profiles using the script “CalculateLikelihoodSimplex.py”. This latter script used a package called “likelihood_simplexMethod.py”, which used the Simplex method as its global optimization method to minimize the cost function [10] [11]. The calculated values were plotted using code from “PlotContourCost.py” and “AnalyzeLikelihood.py”.

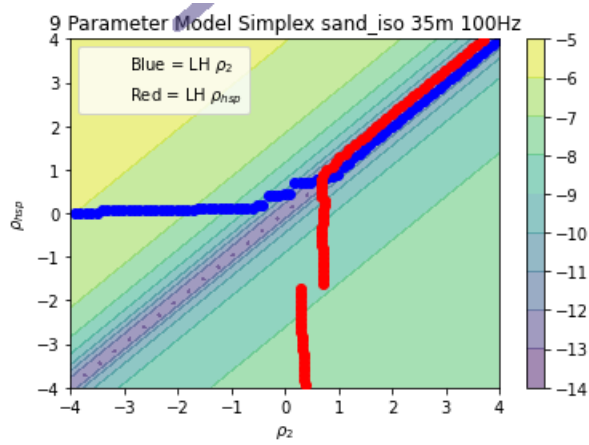


Figure 2.5 Contour plot of the cost values for shifting the values of ρ_{hsp} and ρ_2 in the 9 parameter transmission loss model. The dots represent parameter values calculated by the optimization algorithm. Notably, the cost contour and the optimized value agree on a linear relationship for values in the first quadrant. This indicates that these parameters can be removed from the model as a ratio. The diverging likelihood-calculated parameter values are suspected to occur because of some other parameter in the model shifting the optimal values of these parameters as they decrease in value.

For each full 11 parameter version of the 5 environments, likelihood profiles were created. The parameter identified with the lowest cost plateau would be selected as the first to be removed. A reduced model was created for each fixed sloppy parameter, based on values identified in the likelihood profile. For the reduced model, the process of creating a likelihood profile of parameters, identifying, and removing continued until the parameters identified as sloppy from the local calculations were removed. Contour plots were made at each stage as well to validate the results of the likelihood profiles, as well as check for parameters that correlated with each other.

At each stage of parameter removal, new best fit parameters were calculated and used for the reduced model using the code from “Find_New_Best_Fits.py” in the linked repository in Appendix E. This was done since removing a parameter for the model, even a sloppy one, can shift the original values in the full model.

2.5 Analyzing Costs of Reduced Models

After I finished creating the reduced models for transmission loss for each seabed, I needed to test how well they compared to their original version of the model. I did this by using the Simplex algorithm [11] to find the optimal parameter values θ of the reduced model when it was put into a cost function (Equation 2.2). These values were compared against the data generated for the 11 parameter model.

Data for the comparison was created using the code “DataGenerator.py”. The model cost comparison was done using the code “Fits.py”.

Chapter 3

Results and Future Work

In this chapter, I explain the results, conclusion, and future work that can be done with regards to my research. In summary, I was able to identify sloppy parameters. Using that knowledge, I successfully reduced the transmission loss models for the seabed environments for gravel, sand, and silt. Clay had some success, while mud proved to be a difficult model to reduce.

3.1 Results

For the local analysis of the model manifold of the 11 parameter transmission loss model, we calculated the Jacobian matrices (Equation 2.3). We did a singular value decomposition (Equation 2.6) on the Jacobian to identify which parameters were sloppy, based on the singular values. From the Jacobian calculations in a scanning range of 64 - 256 Hz, for a tolerance of 10^{-6} , we found that the four parameters α_{hsp} , c_{hsp} , ρ_{hsp} , and ρ_2 became sloppy for the three hardest seabed environments: gravel, sand, and silt. The clay seabed only had α_{hsp} as a sloppy parameter, and no other sloppy parameters were detected for the mud seabed. Some details can be

seen in tables 3.1 and 3.2. Tables with the remaining parameters identified as sloppy from the Jacobians can be found in Appendix B.

Table 3.1 Sand Sloppy Parameters listed

Sand Sloppy Parameters	
α_{hsp}	108.9617 Hz
ρ_2	150.3648 Hz
c_{hsp}	156.8160 Hz
ρ_{hsp}	165.8502 Hz

Table 3.2 Clay Sloppy Parameters listed

Clay Sloppy Parameters	
α_{hsp}	222.5941 Hz

What is worth noting from these results are that these four parameters appear consistently in the harder environments. This is useful information because it indicates that there may be a relationship between the model manifolds for the harder seabed environments where these parameters are sloppy. Therefore, we can have a good guess on which parameters are sloppy in other seabed models that combine characteristics of the harder seabeds we modeled.

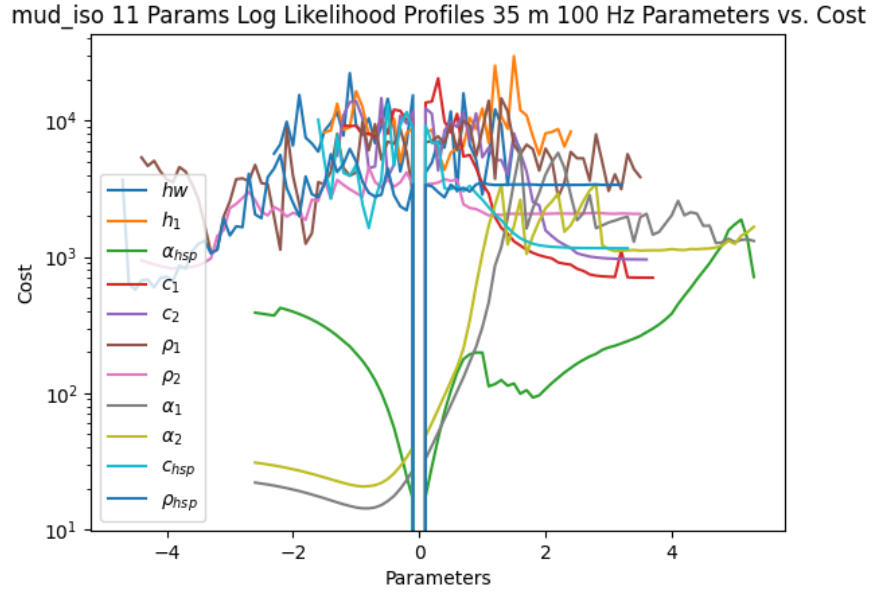


Figure 3.2 All likelihood profiles for the 11 parameter mud model.

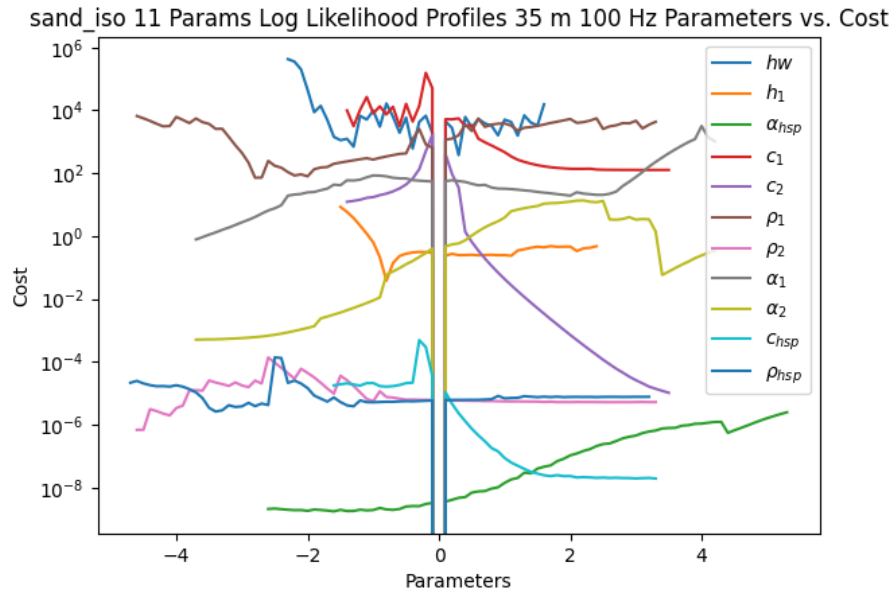


Figure 3.1 All likelihood profiles for the 11 parameter sand model.

With the local-scale sloppy parameters identified, we could go on to validate that they are indeed sloppy on a global scale. For the global analysis, the likelihood profiles

identified sloppy parameters and the recommended order to remove them. Each time we removed a sloppy parameter, we re-calculated the likelihood profiles of the reduced model, in case the sloppy parameters of the model shifted around. This changing of sloppy parameter order for each model reduction occurred in the mud environment.

Table 3.3 shows parameters whose values were fixed to reduce the model.

Table 3.3 Parameters removed in reduced models, as well as value assigned in order to remove. ρ_2/ρ_{hsp} indicates a ratio of the two parameters. “()” indicate that the model was not investigated.

Reductions of 11 Parameter Model					
Model	Gravel	Sand	Silt	Clay	Mud
10	$\alpha_{hsp} = 0$	$\alpha_{hsp} = 0$	$\alpha_{hsp} = 0$	$c_{hsp} = 10^6$	$\alpha_1 = 0$
9	$c_{hsp} = 10^6$	$c_{hsp} = 10^6$	$c_{hsp} = 10^6$	$\alpha_{hsp} = 0$	$\alpha_{hsp} = 0$
8	$\rho_{hsp} = 10^1$; ρ_2/ρ_{hsp}	$\rho_{hsp} = 10^1$; ρ_2/ρ_{hsp}	$\rho_{hsp} = 10^1$; ρ_2/ρ_{hsp}	$\rho_{hsp} = 10^1$; ρ_2/ρ_{hsp}	$c_{hsp} = 10^6$
7	()	()	()	()	$\alpha_2 = 0$
6	()	()	()	()	$\rho_2 = 0$
5	()	()	()	()	$\rho_{hsp} = 0$

A likelihood profile for a harder environment, sand, is shown in Figure 3.1. Interesting results to note are that compared to the local analysis, α_{hsp} generally drops first for global and local analysis, except for the softer environments of clay and mud. The removal of the parameters for gravel, sand, and silt occurred in the same order. Clay switched the order of removing c_{hsp} before α_{hsp} , though examination of the likelihood profile and contour plot show that their cost plateaus are very close to each other. Mud behaved quite differently, on the other hand. The sloppiest parameter, α_1 , did not appear in local analysis or as a sloppy parameter in the other seabed environments. See Figure 3.2. According to Dr. Neilsen, this behavior in mud can

be expected since the sound absorption of mud is smaller.

The reduction of the mud model was the most difficult since the order of plateaus for the likelihoods changed for each parameter removal. Notably, the magnitude of the cost plateaus is much higher than the harder environment counterparts. Therefore, the reductions of the mud model are not as reliable.

With the exception of the mud environment, analysis of the contour plots of parameters ρ_{hsp} and ρ_2 showed that these parameters correlated as a ratio. Therefore, we removed them by fixing ρ_{hsp} at a high trial-and-error found value that did not saturate the numerical approximation. We then set ρ_2 as a ratio of this value in the model calculation. The behavior of the parameters that led us to determine that it was a ratio can be seen in Figure 2.5.

The contour plots in general agreed with the trends of the parameter values. For an example, see Figure 3.3. Cases such as the contour plot of ρ_2 vs ρ_{hsp} (see Figure 2.5), where the optimized parameter values diverge, are suspected to occur since we are working with a 2-dimensional cross section of an 11-dimensional problem. Another parameter could be causing the divergence here, or other minor divergences in other plots. Dr. Transtrum determined that the divergence in the above mentioned case was still an acceptable result which did not interfere with the analysis we conducted. We attempted to find the parameter causing the divergence by plotting the optimized parameter value in a given model one-on-one against the remaining parameter values. However, we were unsuccessful in identifying the parameter responsible for the divergence.

Some parameter values on contour plots appeared to move strangely, but that occurred in cases where the cost plateau was so flat that the optimization algorithm appears to be trying to explore to find the best fit.

When comparing the global and local analysis result, it is clear that although the

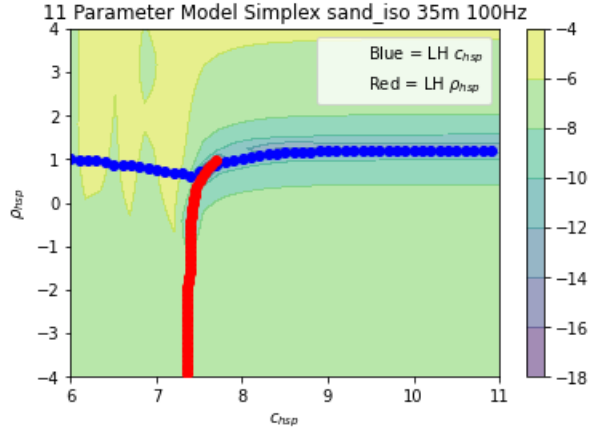


Figure 3.3 Contour plot of the cost values for shifting the values of ρ_{hsp} and c_{hsp} in the 11 parameter sand transmission loss model. The dots represent the values of the parameter calculated by the likelihood profile calculation as it optimized the model. The parameters generally agree with the contour shape. Note that the lowest contour cost value appears for higher values of c_{hsp} . This agrees with our likelihood profiles, which indicate that c_{hsp} must be taken to higher parameter values in order to reduce the model.

local analysis could predict the majority of the sloppy parameters, we still needed to do a global analysis to confirm the order that they get removed in. The discrepancy in the order of removing parameters from both methods is likely due to (1) our local Jacobians sampled the frequency over a range, and therefore, there are some modes and frequencies at which a parameter becomes sloppy before others. (2) Our global analysis sampled a specific frequency, and so if we tested other frequencies, it is possible that the order of removing parameters would change.

The results of reducing our model shown in a table in Appendix C. Dr. Transtrum and Dr. Neilsen determined that a threshold of cost 10^{-3} for the comparisons would determine if the reduced model was acceptable. Gravel, sand, and silt seabed reduced models were well within this threshold. Clay reduced models, however, had a cost magnitude of 10^{-1} . Mud reduced models had a much greater disparity for reduced versions, ranging from cost magnitudes of 10^1 to 10^4 . In con-

sideration of the high costs of the likelihood profiles and the absence of identified dropped parameters from the Jacobians, this is to be expected.

3.2 Conclusions

Based on what we found, we were able to identify four parameters that consistently were sloppy for all of the seabed models, except for mud. Harder seabeds like gravel, sand, and silt had lower costs for removing parameters. Therefore, if a machine learning system is using a reduced model, the approximations will work better for harder seabed environments.

3.3 Future Work

For future work, we would extend our test model for different frequencies, range of noise source, sediment layer thickness, and number of sediment layers. This would be done in order to see how the model manifold changes, as well the model's corresponding sloppy parameters. By extending our test model, we could create a larger catalog of sloppy parameter values. These values can be used to inform a machine learning system about which parameters for transmission loss data are identifiable for a given seabed, frequency, range, etc. The informed machine learning system can then produce results that are more reliable, as well as calculate good approximations relating to transmission loss data in a faster manner than a full-model calculation.

Additionally, the order of dropped parameters can be used to further identify the structure of the model manifold for different seabed models. The different manifolds can be compared and contrasted to discover more about the transmission loss model in general.

Bibliography

- [1] A. Abdelfattah, “Image Classification using Deep Neural Networks — A beginner friendly approach using TensorFlow,”.
- [2] M. K. Transtrum, B. B. MacHta, and J. P. Sethna, “Geometry of nonlinear least squares with applications to sloppy models and optimization,” *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **83** (2011).
- [3] M. K. Transtrum and P. Qiu, “Model Reduction by Manifold Boundaries,” *Phys. Rev. Lett.* **113**, 098701 (2014).
- [4] M. Siderius and J. Gebbie, “Environmental information content of ocean ambient noise,” *Citation: The Journal of the Acoustical Society of America* **146**, 1824 (2019).
- [5] C. Frederick, S. Villar, and Z.-H. Michalopoulou, “Seabed Classification using Physics-based Modeling and Machine Learning,” , 2020.
- [6] D. F. Van Komen, T. B. Neilsen, D. B. Mortenson, M. C. Acree, D. P. Knobles, M. Badiey, and W. S. Hodgkiss, “Seabed type and source parameters predictions using ship spectrograms in convolutional neural networks,” *The Journal of the Acoustical Society of America* **149**, 1198–1210 (2021).

-
- [7] E. K. Westwood, C. T. Tindle, and N. R. Chapman, “A normal mode model for acousto-elastic ocean environments,” *The Journal of the Acoustical Society of America* **100**, 3631–3645 (1996).
- [8] C. C. Youn, A. T. Saric, M. K. Transtrum, and A. M. Stankovic, “Information geometry for model reduction of dynamic loads in power systems,” In , (Institute of Electrical and Electronics Engineers Inc., 2017).
- [9] B. K. Mannakee, A. P. Ragsdale, M. K. Transtrum, and R. N. Gutenkunst, “Sloppiness and the geometry of parameter space,”.
- [10] J. A. Nelder and R. Mead, “A Simplex Method for Function Minimization,” *The Computer Journal* **7**, 308–313 (1965).
- [11] M. H. Wright, “”Direct Search Methods: Once Scorned, Now Respectable,” *Numerical Analysis* pp. 191–208 (1995).

Appendix A

Parameter and Values for Seabed Environments

The parameters values for each seabed environment are listed in the table below. Before we used these values to calculate the Jacobians and optimizations, the values were scaled and made into dimensionless quantities so that they could be compared against each other. The conversion was calculated by Dr. Transtrum and Dr. Neilsen. It occurs in the `Environments.py` package found in the code repository link in Appendix E.

The written out parameter names, in order of appearance on the table, are: height of the water layer, height of the sediment layer, sound speed at the top of the sediment layer, sound speed at the bottom of the sediment layer, density at the top of the sediment layer, density at the bottom of the sediment layer, sound attenuation at the top of the sediment layer, sound attenuation at the bottom of the sediment layer, sound speed in the half-space layer, density of the half-space layer, sound attenuation of the half-space layer.

Table A.1

Parameter Values for Seabed Transmission Loss Models					
Parameter	Gravel	Sand	Silt	Clay	Mud
h_w	75 m	75 m	75 m	75 m	75 m
h_1	35 m	35 m	35 m	35 m	35 m
c_1	1800 m/s	1650 m/s	1575 m/s	1500 m/s	1470 m/s
c_2	1830 m/s	1680 m/s	1611 m/s	1536 m/s	1506 m/s
ρ_1	2.0 g/cm ³	1.9 g/cm ³	1.7 g/cm ³	1.5 g/cm ³	1.6 g/cm ³
ρ_2	2.0 g/cm ³	1.9 g/cm ³	1.7 g/cm ³	1.5 g/cm ³	1.6 g/cm ³
α_1	0.6 dB/m/kHz	0.8 dB/m/kHz	1.0 dB/m/kHz	0.2 dB/m/kHz	0.035 dB/m/kHz
α_2	0.6 dB/m/kHz	0.8 dB/m/kHz	1.0 dB/m/kHz	0.2 dB/m/kHz	0.035 dB/m/kHz
c_{hsp}	2000 m/s	2000 m/s	2000 m/s	2000 m/s	2000 m/s
ρ_{hsp}	2.2 g/cm ³	2.2 g/cm ³	2.2 g/cm ³	2.2 g/cm ³	2.2 g/cm ³
α_{hsp}	0.25 dB/m/kHz	0.25 dB/m/kHz	0.25 dB/m/kHz	0.25 dB/m/kHz	0.25 dB/m/kHz

Appendix B

Local Analysis Identified Sloppy Parameters

Remaining tables of sloppy parameters not reported in results section.

Table B.1 Gravel Sloppy Parameters listed

Gravel Sloppy Parameters	
α_{hsp}	113.6365 Hz
ρ_{hsp}	165.8502 Hz
ρ_2	180.3867 Hz
c_{hsp}	185.5100 Hz

Table B.2 Silt Sloppy Parameters listed

Silt Sloppy Parameters	
α_{hsp}	89.5690 Hz
c_{hsp}	152.4852 Hz
ρ_2	154.6354 Hz
ρ_{hsp}	163.5440 Hz

Table B.3 Mud Sloppy Parameters listed

Mud Sloppy Parameters	
()	()

Appendix C

Cost Results of Model Reduction

Table containing the results of the costs of the models reduced from the original 11 parameter model.

Table C.1 Costs of reduced models

Cost of Model Reduction		
Seabed	Parameters	Cost
Gravel	10	1.52E-06
Gravel	9	7.67E-06
Gravel	8	1.04E-05
Sand	10	1.48E-09
Sand	9	1.21E-07
Sand	8	2.70E-07
Silt	10	5.23E-10
Silt	9	8.79E-09
Silt	8	4.67E-08
Clay	10	1.03E-01
Clay	9	1.07E-01
Clay	8	2.62E-01
Mud	10	2.56E01
Mud	9	4.95E+02
Mud	8	1.63E+04
Mud	7	1.76E+04
Mud	6	1.41E+04
Mud	5	1.83E+04

Appendix D

Workflow Diagram and Explanation of the Code

Figure D.1 contains a descriptive work flow diagram that shows the steps taken to analyze and reduce the models. Each seabed transmission loss model went through this process. The code corresponding to each step is contained in the code repository referenced in Appendix E. The script title and order of appearance is indicative of what step in the analysis the respective code corresponds to.

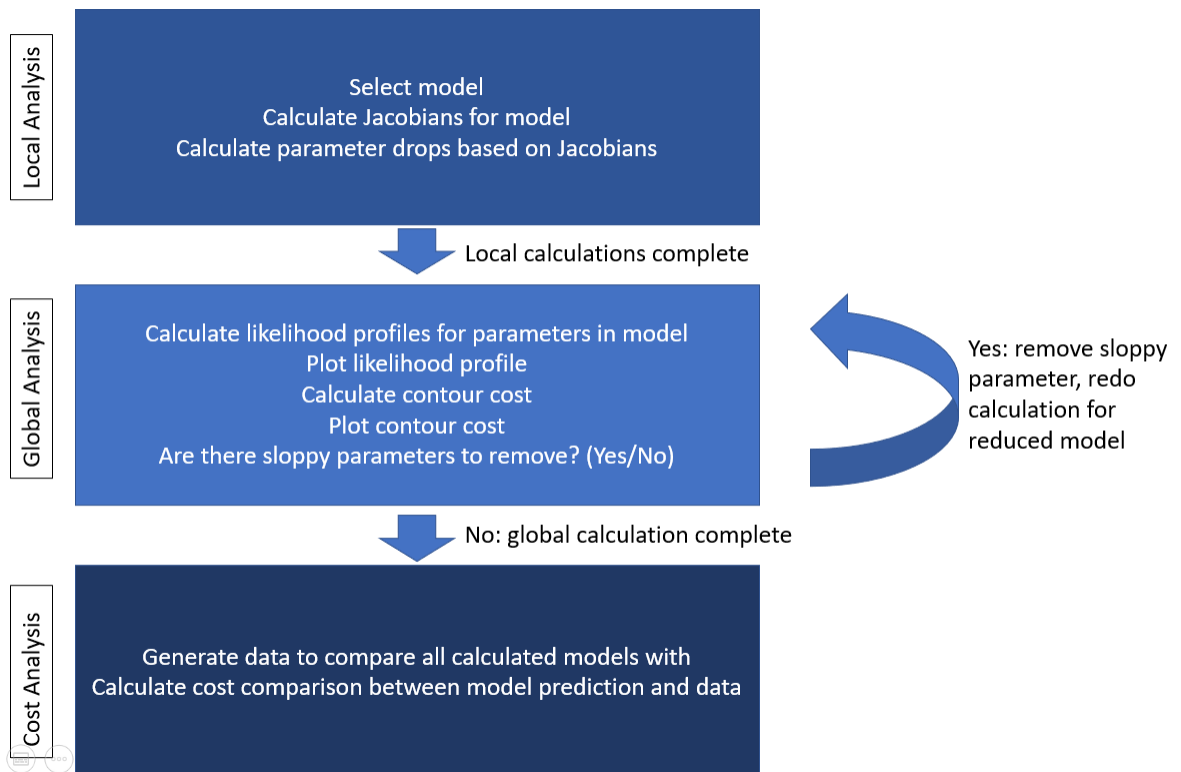


Figure D.1 Descriptive work flow diagram of the steps that take place to reduce the model.

Appendix E

Code Link

The main scripts of code referenced in this paper can be found on the following web page: <https://byui-physics.github.io/main/theses.html>.

Note: representative code scripts are the “AnalyzeLikelihood.py” and “Environments.py”. The former script has different variations based off of it, depending on the plot I needed. “Environments.py” was adapted for each reduced model. Since there are multiple versions of these files, the basic script versions are given here.