

Article

Predicting Acoustic Transmission Loss Uncertainty in Ocean Environments with Neural Networks

Brandon M. Lee ^{1,*} , Jay R. Johnson ^{1,†} and David R. Dowling ^{1,2} 

¹ Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI 48109, USA

² Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, MI 48109, USA

* Correspondence: leebm@umich.edu

† Current address: Cirrus Logic, Inc., Salt Lake City, UT 84111, USA.

Abstract: Computational predictions of acoustic transmission loss (TL) in ocean environments depend on the relevant environmental characteristics, such as the sound speed field, bathymetry, and seabed properties. When databases are used to obtain estimates of these properties, the resulting predictions of TL are uncertain, and this uncertainty can be quantified via the probability density function (PDF) of TL. A machine learning technique for quickly estimating the PDF of TL using only a single, baseline TL calculation is presented here. The technique shifts the computational burden from present-time Monte-Carlo (MC) TL simulations in the environment of interest to ahead-of-time training of a neural network using equivalent MC TL simulations in hundreds of ocean environments. An environmental uncertainty approach which draws information from global databases is also described and is used to create hundreds of thousands of TL-field examples across 300 unique ocean environments at ranges up to 100 km for source frequencies between 50 and 600 Hz. A subset of the total dataset is used to train and compare neural networks with various architectures and TL-PDF-generation methods. Finally, the remaining dataset examples are used to compare the machine-learning technique's accuracy and computational effort to that of prior TL-uncertainty-estimation techniques.

Keywords: transmission loss; environmental uncertainty; underwater acoustics; machine learning; neural networks; supervised learning



Citation: Lee, B.M.; Johnson, J.R.; Dowling, D.R. Predicting Acoustic Transmission Loss Uncertainty in Ocean Environments with Neural Networks. *J. Mar. Sci. Eng.* **2022**, *10*, 1548. <https://doi.org/10.3390/jmse10101548>

Academic Editors: Tracianne B Neilsen and Haiqiang Niu

Received: 16 September 2022

Accepted: 13 October 2022

Published: 20 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many computational tools exist to predict how sound will propagate in an ocean environment given the environment's properties. These tools commonly provide answers by fully or approximately solving the Helmholtz equation with the relevant environmental properties being included via material parameters and boundary conditions. Many applications use computational solvers to predict how sound did propagate, could propagate, or will propagate from a known acoustic source to another location in an ocean environment—often where a receiver such as a hydrophone is located. Uncertainty in the ocean property values used by the solvers may arise from imperfect or incomplete measurements, limited accuracy or resolution in available databases, or the uncertainty accompanying an estimate obtained via inversion. In this last case, there is a rich collection of work on Bayesian techniques for ocean environmental inversion which infer a joint probability distribution function (PDF) for a set of ocean environment properties using recorded acoustic signals [1–5]. In applications where robustness against environmental uncertainty is desired, knowledge of the range of possible values for acoustic amplitude or phase is more beneficial than a single acoustic-field prediction with limited or unknown reliability. Acoustic-field amplitude expressed as transmission loss (TL) [6] is often of interest in ocean acoustic applications and is the primary focus here.

The Monte Carlo (MC) method provides a reliable means for quantifying TL uncertainty in uncertain ocean environments [7]. The MC method proceeds by sampling many

uncertain environmental realizations, computing the TL field in each one, and combining all of the computed results to statistically describe the possible variations in TL. Although the MC method can be used whenever or however the environmental uncertainty is specified, its main drawback is its computational expense. A TL computation must be performed for each of the thousands of uncertain environmental realizations that may be needed to adequately quantify TL uncertainty.

There are several other methods for transferring environmental uncertainty to TL uncertainty which offer varying trade-offs between computational effort, overall accuracy, and adaptability. In particular, real-time applications which frequently require TL-uncertainty knowledge must give preference to prediction speed over accuracy. The uncertainty band, or uBand [8,9], method estimates the uncertainty bounds on frequency-averaged TL predictions by re-scaling the range-averaging window applied on a single center-frequency TL solution. The re-scaling factor relates the uncertain property's variations to an equivalent increase or decrease in the number of propagating modes. The field shifting method [10,11] estimates the PDF of TL produced by N independent, uncertain environmental parameters using only $N + 1$ TL solutions by finding an optimum equivalent spatial shift from the baseline TL field solution for each new TL field solution. Another method uses truncated polynomial chaos expansions [12,13] to produce an approximate compact representation of the stochastic acoustic field. The number of terms necessary for the approximation depends on the complexities of the environmental model and environmental uncertainties, and its coefficients can be estimated with a set of MC acoustic field solutions.

While each of these methods may require less computational effort than MC at the expense of accuracy, there is a limit to their adaptability. The computational effort and implementation effort of these methods increase with the number of uncertain environmental parameters considered. Thus, the use of modern ocean environmental descriptions presents a growing challenge for these methods as descriptions become more detailed with greater measurement density, higher resolution oceanic databases, better modeling of ocean acoustic phenomena, and improved precision of environmental inversions. Although increases in the available resolution of estimated ocean properties such as bathymetry, sound speed, and seabed properties within an environment should allow for more precise predictions of TL, the corresponding effect on the reliability of these new TL predictions is less clear. Unfortunately, this suggests that the task of quantifying TL prediction uncertainty in increasingly detailed ocean environments remains important, even as it is rendered more difficult by the increased detail and complexity in the statistical specification of these environments.

Like the MC method, the ad hoc Area Statistics (AS) method [14] is adaptable enough to quantify TL uncertainty in environments parameterized at the modern-database level of detail. The AS method estimates the PDF of TL at a given receiver location using only a single baseline TL solution, regardless of the number of uncertain environmental parameters. Similar to the uBand and field shifting methods, AS uses information from a baseline TL prediction in a spatial region surrounding the receiver location. The variations in the predicted TL values surrounding the point of interest (POI) are assumed to represent the variations in the TL value that would be seen at the POI due to environmental uncertainty. Thus, the AS method gathers predicted TL values inside a local range-depth box of some size surrounding the POI into a histogram to estimate of the PDF of TL at the POI—a procedure which is very fast. Results from this method showed good agreement with MC PDFs of TL for full-calendar-year variations in ocean environments with reflective bottoms. However, this method struggles for source-receiver ranges less than about 10 km and in ocean environments with an absorbing seabed. Unfortunately, the AS method can only be adjusted by changing the box size, shape, or sample weighting to account for varying sources or degrees of environmental uncertainty, and such adjustments have not yet been successful in doing so.

Machine learning has become increasingly popular for acoustic and underwater acoustic applications [15,16] by providing alternative methods which primarily benefit from

either: (1) the ability to include more complex model dependencies when their inclusion in conventional methods would otherwise not be practical or tractable; or (2) much faster computational speeds than conventional methods. The research effort reported here follows along these lines by using machine learning to obtain an alternative TL PDF estimation method that can: (1) be implemented regardless of the level of detail describing the environment; and (2) make predictions much faster than MC, enabling possible real-time use.

The machine learning method described here uses a trained neural network (NN) to quickly predict the PDF of TL using only a single baseline TL solution. The adaptability of this method is underpinned by supervised machine learning; many example PDFs of TL are produced with the MC method in many different ocean environments at many different source frequencies and depths to create a compound dataset that is partitioned to train and validate the NN. Therefore, the predictions of a successfully trained NN are reliable for new environments if similar environments and the same sources and degrees of environmental uncertainty were represented in the training dataset. A new dataset can simply be created to train another NN for deployment whenever the scope of an application changes to include different types of environments, different descriptions of the environmental properties or uncertainties, or different acoustic source property ranges. The prediction speed advantage of the NN method comes from splitting its computational burden into two uneven steps. First, the overwhelming majority of the necessary total computational effort is completed ahead of time in the creation of a dataset of examples and the training of the NN. After this preparation, the effort in using the trained NN only consists of gathering inputs and computing the TL PDF predictions—a computation orders of magnitude faster than equivalently accurate Monte Carlo computations. A diagram which shows the prediction process for such a trained NN is provided in Figure 1.

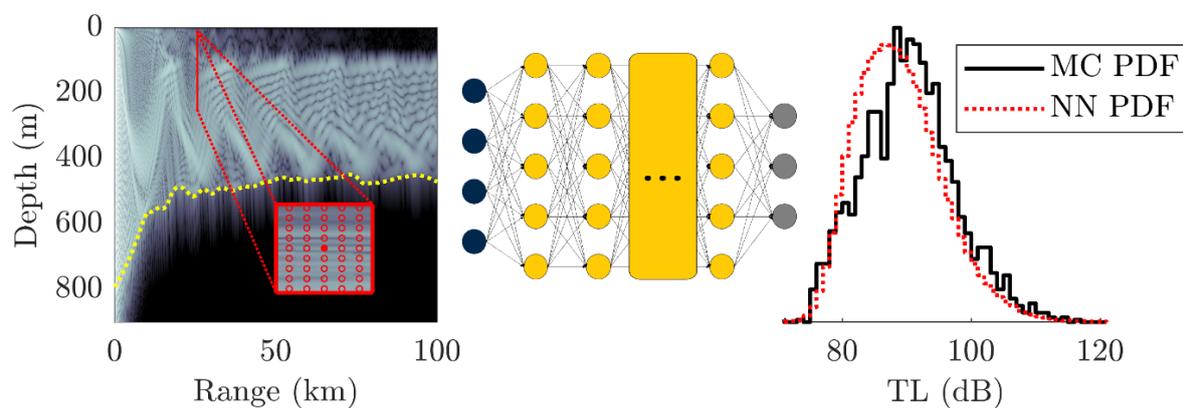


Figure 1. Elements of the proposed machine learning technique for estimating the probability density function (PDF) of transmission loss (TL). A simple, feedforward neural network (NN) diagram is shown between the model’s inputs on the left and outputs on the right. The inputs are values from a single baseline TL solution surrounding the point of interest (POI), and the output is a predicted estimate of the Monte Carlo (MC) PDF of TL at that POI resulting from environmental uncertainty.

Herein, the NN method is implemented and compared to the AS and MC methods which are also capable of quantifying TL uncertainty due to environmental uncertainty in detailed ocean environments described by many uncertain parameters. To create the necessary dataset of realistic examples in detailed environments, an ocean environmental uncertainty approach which leverages open-source databases was developed and is presented in Section 2.1.1. The NN approach to predicting PDFs of TL is developed and presented in Section 2.2. The predictive performance of the NN method is assessed and compared to other methods in Section 3. Finally, the findings of the study are discussed in Section 4.

2. Materials and Methods

2.1. Materials

2.1.1. Ocean Environmental Uncertainty Approach

In order to train and validate the NN method, it was first necessary to develop an approach for generating realistic cases of uncertain ocean acoustic propagation. This approach begins by defining a case's environment by its *when*—the relevant date and time—and its *where*—the acoustic source latitude and longitude, the source-to-receiver-bearing angle, and the maximum source-receiver range. The baseline environmental properties are the best estimate of the true environmental properties at that time and location, as determined from available databases. The case's *when* and *where* are also combined with choices of parameters controlling the degree of uncertainty to produce, with these same databases, an ensemble of uncertain environmental realizations with varying environmental properties. Herein, these properties consist of the bathymetry, range-dependent sound speed field, and seabed properties. The details concerning the implementation of this approach and the databases used in this study can be found in Appendix A.

The impact of the case's environmental uncertainty is then assessed using this ensemble for MC simulations of the acoustic field amplitude. Here, an ensemble of 2000 uncertain environmental realizations was generated, and the RAMGEO parabolic equation solver [17–19] was used to compute the TL field in each one given the same source depth and source frequency. The computational resolution used in range and depth was 1 and 1/5 of the reference wavelength based on past recommendations in [20] (Ch. 6) and [21], and 8 Padé terms were used for the rational approximation of the square-root operator. Each of the environment's sound speed profiles were cubically interpolated in depth, extended to the sea bottom, and linearly interpolated in range to create a new sound speed grid which was input into the RAMGEO solver. For some environments with very absorptive sea bottoms, the sediment properties input into the RAMGEO solver were altered to produce a more computationally efficient half space approximation because no sound reached the basement layer. One TL field solution corresponding to an example baseline environment is shown in Figure 2.

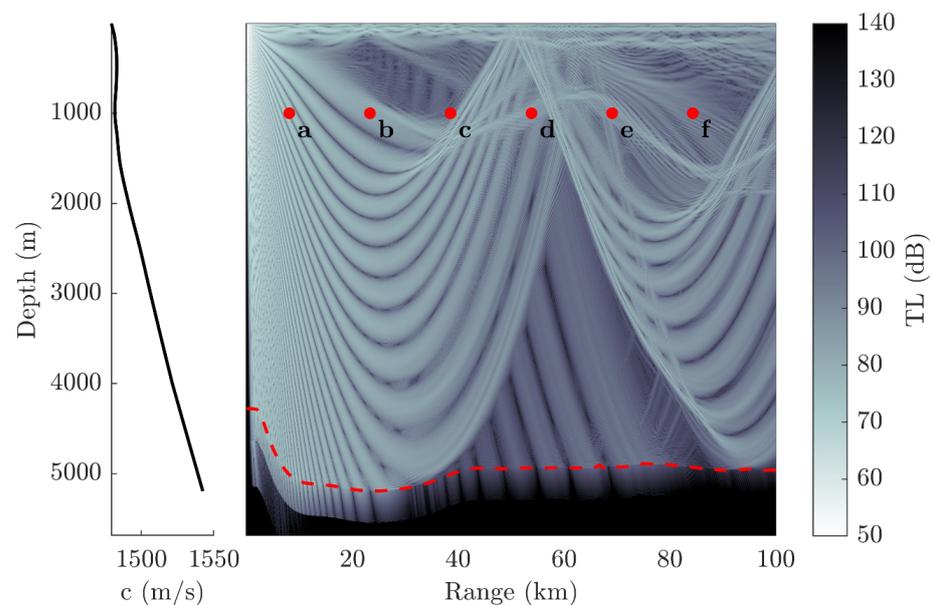


Figure 2. Sample TL-field. In this environment, the acoustic source is located at a range of 0 km and a depth of 72.3 m. The acoustic source frequency is 528.5 Hz. The sound speed profile corresponding to the water column properties at the source location is shown in the left panel. The environment's bathymetry is shown as the dashed red line. The sediment type for this environment is clay, and the sediment thickness for this environment is 350.2 m. The six red points shown in the environment near 1000 m in depth are the example locations for the PDFs of TL shown in Figure 3.

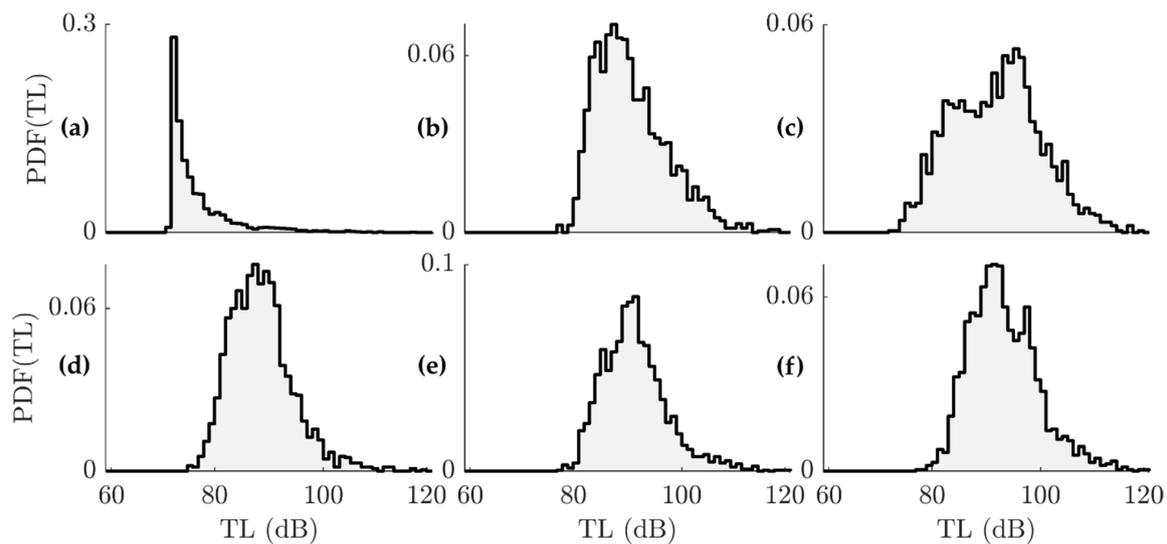


Figure 3. MC PDFs of TL for example receiver locations indicated in Figure 2. The example receiver locations corresponding to the MC PDFs of TL labeled (a–f) are located at ranges of 8.0, 23.3, 38.6, 53.8, 69.0, and 84.3 km and at depths near 1 km.

At any location within the environment, the 2000 TL values at that location were combined into a 100-bin histogram with bin edges spanning 40 to 140 dB at 1 dB spacing. The first and last histogram bins also contained counts for any TL value less than 41 dB or greater than 139 dB respectively. The bin counts were normalized to create a discrete PDF with 100 bins, which is referred to herein as the MC PDF of TL. For any random environmental realization, if a receiver location was within the realization’s seabed, the corresponding value of TL for this location for this random realization was ignored in the construction of the histogram. The MC PDFs of TL at various receiver locations across the example environment of Figure 2 are shown in Figure 3.

2.1.2. Training and Testing Dataset Setup

A dataset of many and varied examples was desired to train and validate the NN method’s predictive performance. Each example corresponds to a computed MC PDF of TL (the example’s target output) and all of the environmental and geometric information needed for that computation (used to generate the example’s collection of potential input values—as described below in Section 2.2.1). In order to greatly reduce the computational cost of producing the total example dataset, multiple examples were considered for each case by selecting many locations (ranges and depths) within that case’s environment.

A case consists of three components:

1. The baseline environment with properties (bathymetry, water column sound speed, seabed properties, etc.) estimated for its geographic location at a nominal date and time;
2. An ensemble of 2000 randomly sampled uncertain environmental realizations with variations in these estimated properties that are consistent with their uncertainties;
3. The acoustic source depth and frequency.

The examples in the example dataset should be representative of those likely to be encountered in the desired application. For this analysis, no specific operational task was assumed. Therefore, cases were randomly chosen to represent a variety of environmental circumstances, and example receiver locations were uniformly spread throughout each case environment’s range and water column.

Each potential case is produced by randomly selecting: a source location on the globe between 50° S and 65° N, limited by the shared coverage of the databases used, a random azimuth, and a random baseline time between 1 January 2019 and 31 December 2019. Next,

the case’s baseline environmental properties were obtained from the relevant databases from the source location along the chosen azimuth. The nominal maximum range for each case is 100 km, but a maximum range as short as 30 km was used instead if it prevented a case’s environment from having a water column depth of 10 m or less at any range. If a potential case was too shallow or the desired properties were not available in the relevant databases, a new potential case was produced; randomly sampled uncertain environmental realizations were handled similarly. The case’s source frequency and source depth were uniform-randomly sampled from 50 to 600 Hz and from 50 to 200 m until the source depth was at least 100 m shallower than the baseline environment’s water column depth at the source. All cases were produced with the environmental uncertainty approach described in Section 2.1.1.

Ten-thousand potential cases were produced randomly with this scheme. From these, 300 were chosen to form a set which is representative of observed sediment types and bathymetries, while still covering the globe. The distributions of mean water column depth, sediment type, and geographic location for all 10,000 potential cases and for the 300 selected cases are shown in Figure 4. Additionally, plots of the TL field solutions obtained with the baseline environmental properties and source properties of each case are provided in the Supplementary Materials.

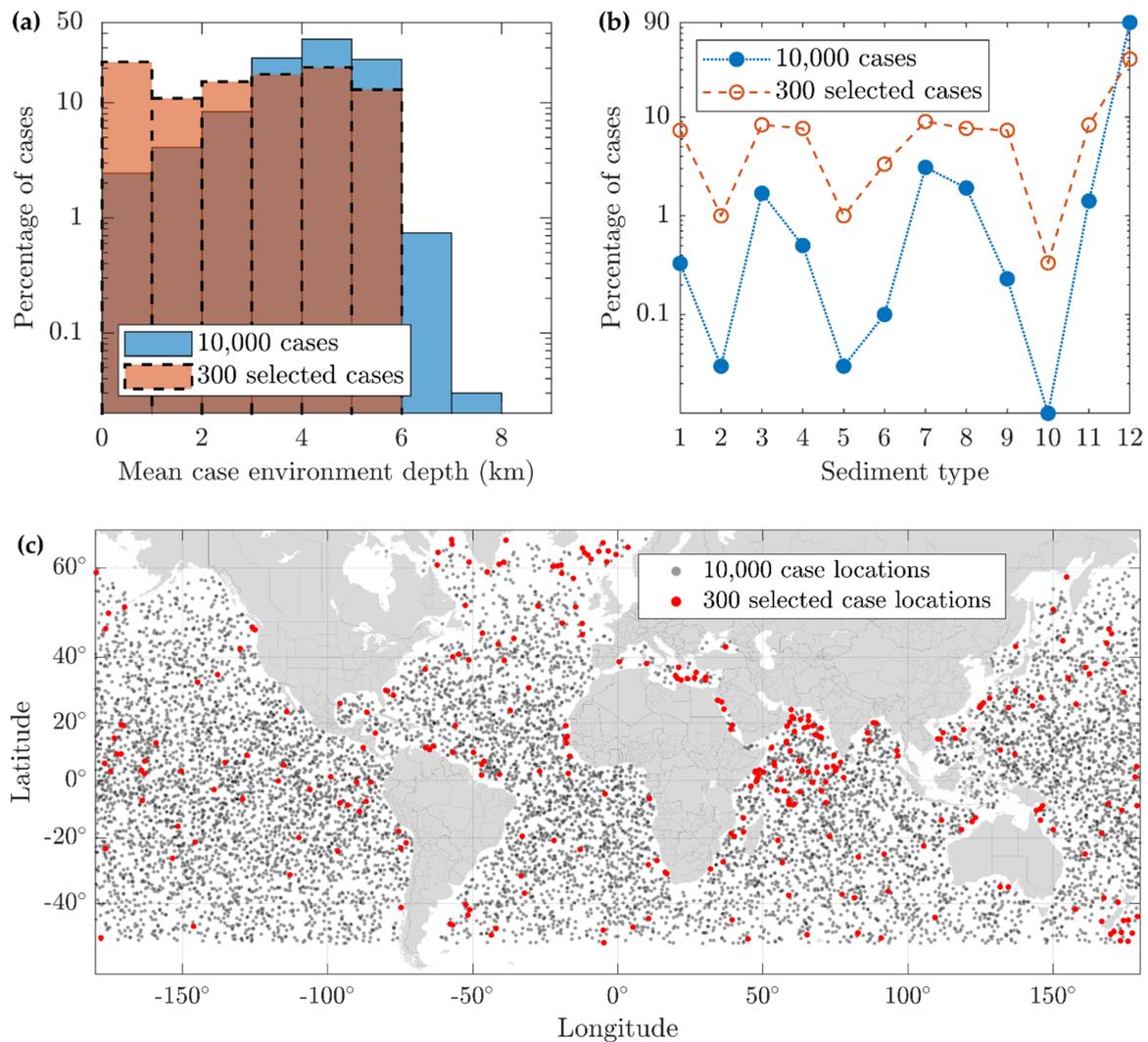


Figure 4. Summary of cases used to create the training and testing datasets. (a) The distributions of mean baseline environment depth for all 10,000 and the 300 selected cases. (b) The distributions of environment sediment type for all 10,000 and the 300 selected cases. The sediment types are—1: cobble

or gravel or pebble, 2: muddy sandy gravel, 3: medium sand or sand, 4: fine sand or silty sand, 5: muddy sand, 6: gravelly mud or sandy silt, 7: medium silt or sand-silt-clay, 8: sandy mud or silt, 9: fine silt or clayey silt, 10: sandy clay, 11: silty clay, and 12: clay. Without careful selection of the final 300 cases, around 90% of the cases would have the clay sediment type. (c) Source locations for all 10,000 and the 300 selected cases.

The 300 selected cases were randomly shuffled and split into three groups. The examples produced from each case were combined in each group to form three datasets: (1) a training dataset (cases 1 to 100); (2) testing dataset 1 (cases 101 to 200); and (3) testing dataset 2 (cases 201 to 300). For each case within the training dataset, up to 25 example receiver locations were chosen every 500 m (from a random starting range) with even vertical spacing within the water column and on the computational output grid. For each case within a testing dataset, this example resolution was increased to up to 50 example receiver locations every 250 m. Some near-source examples with locations corresponding to very wide propagation angles were ignored because of the limitation of the TL-field computational solver. These sets of examples were used to analyze and compare the performance of different NN configurations and other methods of TL PDF prediction.

2.2. Methods

Neural networks are a popular machine learning approach which can explicitly map known inputs to desired outputs through a series of linear and non-linear parametric operations. The parameters, known as NN weights, are optimized to accurately predict example outputs given example inputs in a procedure known as training. A trained NN can make predictions on new examples for the purpose of some application.

With the approach outlined here, NNs are constructed to use as inputs only information from: the baseline environmental and source properties, the baseline TL field solution, and the desired receiver location. The NN output is a prediction of the PDF of TL at that receiver location. These NNs are trained using examples where the relevant inputs are known and the example outputs, MC PDFs of TL, have been computed. Because the example outputs are known and are continuous-valued, this is a supervised learning and regression task in the context of machine learning. In the next sections, further details of these NNs' inputs, outputs, and training are presented. A diagram displaying the connections between these key aspects of the NNs is provided in Figure 5.

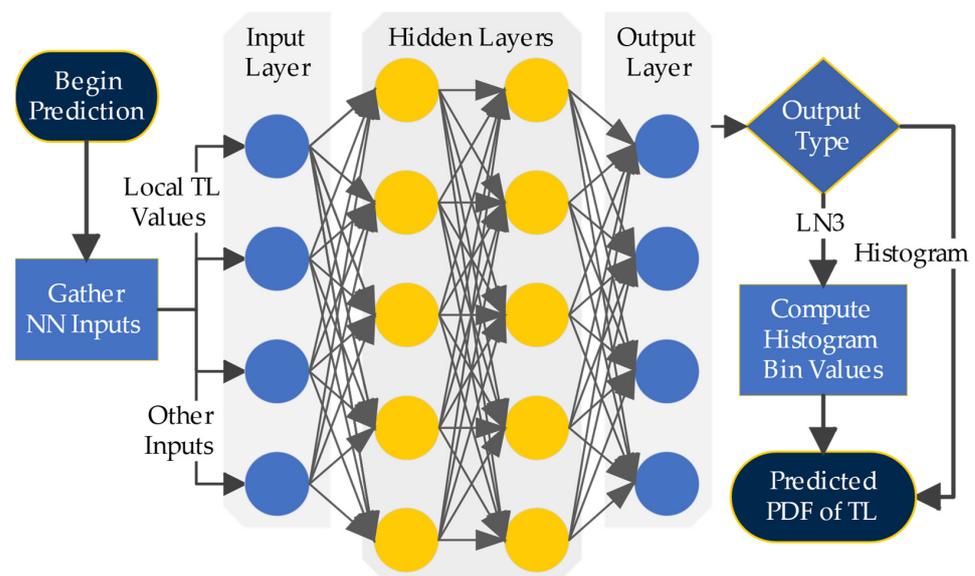


Figure 5. Algorithm block diagram for a NN prediction of a PDF of TL. When presented with a new

example, the NN first gathers its inputs (as described in Section 2.2.1). Those inputs are passed to the input layer of the neural network. The NN shown in this figure is a fully connected feedforward NN. The final values of the NN computation appear in the output layer. Depending on the choice of output type (as discussed in Section 2.2.2), a final computation is performed to produce the predicted PDF of TL. The generic architecture (number of hidden layers, nodes, etc.) of the NN pictured here is not representative of the NNs trained in this study (as detailed in Section 2.2.3).

2.2.1. Neural Network Inputs

Given an acoustic source in an uncertain ocean environment, the MC PDF of TL can be obtained at any receiver location using the procedure described in Section 2.1.1. For an NN to make a prediction of such an example's PDF, it must be given information pertaining to the example. In this analysis, the example information fits into two categories: (1) source-receiver-environment information; and (2) local baseline TL information. By so restricting the NN inputs, the total in situ computational cost of estimating the PDF of TL at a location is orders of magnitude smaller than that of the MC method, which requires thousands of additional TL solutions.

The source-receiver-environment (setup) information contains all of the information necessary to obtain the baseline TL solution. In this analysis, the source-receiver range and receiver depth were the only setup inputs used by the NN. The inclusion of additional setup information did not lead to significant NN performance improvement. This is likely due to the fact that the baseline TL field solution, which is provided in part as an input to the NN, is highly dependent on this information. However, future inclusion of more setup inputs should not greatly increase the NN prediction time if the values of these inputs can be obtained quickly for each example.

The most important information given to the NN as an input is the baseline TL-field values near the POI. For each example, these local baseline TL values were interpolated onto a range-depth grid centered on the POI. The grid height, width, vertical spacing, horizontal spacing, and spacing unit—given in terms of meters or acoustic wavelengths—were considered problem-specific hyperparameters. In this analysis, many NNs were trained with various local TL grids, and their predictive performances were considered and compared. Regardless of the grid size, each relative grid location corresponds to a single input node of the NN's input layer, so every TL and non-TL input are scalar valued. Input feature normalization is performed to reduce any differences in scale between the TL and non-TL inputs.

If an example's grid locations extended beyond the computational domain, the TL values inside of the domain were reflected across the corresponding boundary, an arbitrary choice. Of the testing examples and local grids considered in this analysis, less than 0.07% and between roughly 1% and 13% had a local grid which reached beyond the computational domain in range or depth, respectively. Grid points below the water column received no special attention as the RAMGEO TL field solutions extended into the bottom layers.

2.2.2. Neural Network Outputs and Cost Function

The desired output of the NN is an approximate PDF of TL. Two types of NN outputs were considered for producing a PDF from a NN prediction. Additionally, an error metric used to assess performance and as a training cost function is discussed below.

The first NN output type employed a parametric form for the predicted PDF of TL with parameters provided by the NN's output layer. Here, the predicted TL distribution's cumulative density function was analytically evaluated in order to obtain discrete values for each of the histogram bins for comparison to the MC PDF of TL. In this analysis, a three-parameter log normal (LN3) distribution was used as the parametric PDF form. A LN3-distributed random variable X has a PDF $f(x)$:

$$X \sim f(x; \mu, \sigma, c) = \frac{1}{(x - c)\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(x - c) - \mu)^2}{2\sigma^2}\right) \quad (1)$$

where $\ln(x - c)$ is normally distributed with a mean and standard deviation specified by μ and σ .

This parametric form was chosen over other forms to balance goodness of fit, computed as the maximum likelihood of MC generated TL values, with the corresponding burden in implementation and training. Clearly, there are environmental circumstances where there are better distributions than the LN3 to describe TL PDFs, such as TL PDFs arising from a few short-range, slightly varying propagation paths and TL PDFs produced by the random interference of many long-range propagation paths. However, the NN with a parametric output type needs a single, simple form which can approximate the wide array of MC PDFs of TL. Additionally, the LN3 distribution can be specified simply by the choice of its first three moments, which is valuable because: (1) the mean TL fields are likely similar to the baseline TL fields used for NN inputs, (2) the TL-variance fields may have consistent features relating to source-receiver range, and (3) the TL PDFs are unsurprisingly skewed since TL is a decibel quantity. Interestingly, these and other considerations make the LN3 distribution preferable to the Pearson Type IV distribution. Even though the Pearson Type IV distribution did provide better fits on average of the MC PDFs of TL, it has undefined moments for certain parameter choices, so its implementation with the NN requires adding constraints to the NN predictions that slow training down considerably. As the LN3 output type was implemented here, the three output node values corresponded to the values of the first three moments of the predicted parametric PDF, rather than the parameters of (1). The predicted values for the second moments were ensured positive by using their absolute values, and the magnitude of these predicted values were bounded to maintain single precision numerical stability. Finally, a clear disadvantage of the use of this parametric distribution is its inability to fit multimodal distributions, such as the MC PDF of TL shown in Figure 3c. An example MC PDF of TL, its best fit LN3 PDF, and this best fit LN3 PDF's corresponding bin representation are all shown in Figure 6.

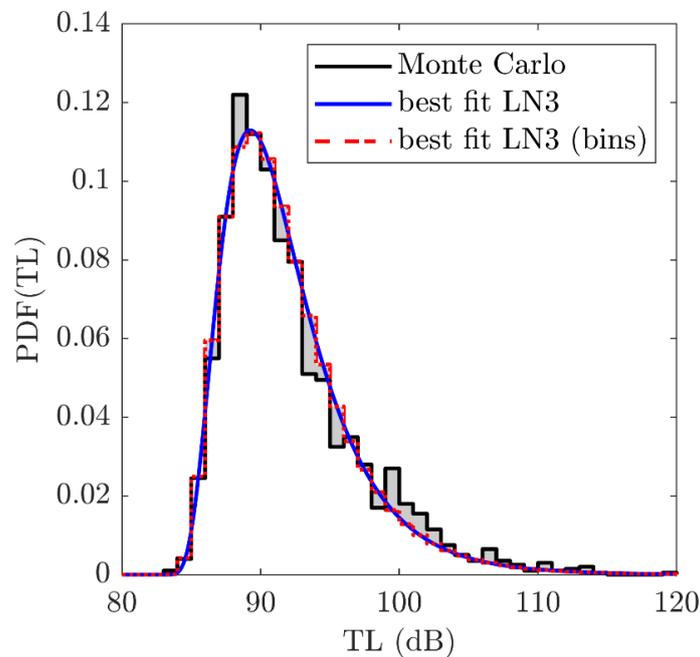


Figure 6. Example MC PDF of TL (solid black line) plotted with its best fit three-parameter log normal (LN3) PDF (solid blue line) and as a histogram (dashed red line). The L_1 error between the MC PDF of TL and its best fit LN3 PDF is visualized as the shaded region and is 0.105 in this example.

With the second NN output type, the NN's output layer directly provided values for each histogram bin. These raw NN predicted bin values were assembled into a discrete PDF by taking their absolute value and rescaling them to have unit bin-sum.

The error metric used to assess the performance for a predicted PDF of TL is the L_1 error, or the integrated absolute difference between the predicted PDF of TL \hat{f} and the MC PDF of TL f .

$$L_1(f, \hat{f}) = \int_{-\infty}^{\infty} |f(x) - \hat{f}(x)| dx \quad (2)$$

This error metric represents the area difference between the two PDFs, and is bounded between 0 and 2, which represent a perfect match and no overlap between the two PDFs, respectively. In the context of the NN outputs described above, the L_1 error can be computed as the sum of the absolute differences of the histogram bin values. An example of the L_1 error between a MC PDF of TL and its best-fit LN3 PDF is shown in Figure 6 and is 0.105. This L_1 error is the lowest a NN with a LN3 output type could achieve when attempting to predict this MC PDF of TL. However, this example's lower-bound L_1 error is much smaller than the average testing L_1 error of the NNs in this study, indicating the approximation introduced by using the LN3 parametric form should not limit the trained NN's predictive accuracy.

2.2.3. Neural Network Architecture and Training

In the method proposed here, training the NN means finding a set of NN weights (the training parameters) which minimize the L_1 errors of predicted PDFs of TL across a set of training examples with inputs and outputs described in the previous subsections. TensorFlow 2.3 [22] was used to train the NNs in this study. It computes the gradient of each prediction's L_1 error with respect to each training parameter via automatic differentiation. These gradients were used to minimize the mean prediction error using the AMSgrad algorithm [23], stochastic mini-batch gradient descent [24], and Weightnorm [25].

There are no restrictions on NN architecture in a generic implementation of the NN method for predicting TL PDFs. In this study, a simple architecture was selected to demonstrate the success of the method while allowing for the possibility of easy reproduction of this work. The NNs implemented here had a number of fully connected, equally sized hidden layers connecting the input layer to the output layer. The hidden layers shared the same choice of non-linear activation function from the following options: (1) the rectified linear unit (ReLU) [26]; (2) the exponential linear unit (ELU) [27]; and (3) the Swish [28] activation function with a constant parameter value of one. A linear activation function was used in the output layer, with the value at each output node taking part in the computation of the predicted PDF and cost function, as discussed above.

Although advanced NN architectures could include convolutional layers [29] to better represent the spatial structure of the local TL grid input feature, the authors did not observe improvements with their inclusion. This is likely due to the inconsistency in absolute grid spacing of the wavelength-unit grids across examples and the sparseness (relative to the wavelength) of the meter-unit grids for the higher-source-frequency examples. Additionally, many techniques have been developed to improve the training of deep NNs by bettering their trainability, such as using skip connections [30] or interlayer normalization [25,31], or by limiting their risk of overfitting through regularization, such as dropout [32]. In this effort, the use of Weightnorm and early stopping (of training) was sufficient. The risk of overfitting and the need for a great number of hidden layers are limited by the noisy quality of the MC PDFs of TL and their incomplete relationship with the local baseline TL values used as NN inputs. Finally, the NNs implemented here received no direct information about the problem's underlying wave propagation physics. Perhaps the limiting of the NN's inputs to almost exclusively include the local baseline TL field accomplishes this to some degree, but one would expect that the inclusion of convolutional layers could more directly represent the spatial derivatives of the field that arise in the equation(s) governing acoustics. Recent work in explicitly informing NNs of the problem's governing physics provide an even more direct approach via Physics-Informed Neural Networks (PINNs) [33], but implementing this technique in the TL PDF prediction problem would present challenges.

Hyperparameter optimization, or NN tuning, was performed to find a suitable configuration of hyperparameters—parameters held constant throughout the training procedure—which minimized the L_1 cross-validation (CV) error averaged over four CV splits. There were seven hyperparameters concerning the architecture of the neural network and its training and five hyperparameters concerning the size and spacing of the local TL grid input to the NN. Additional details concerning the hyperparameter optimization effort can be found in Appendix B.

3. Results

First, many NNs with different output types and configurations were trained and cross-validated on the training dataset, which consists of 491,498 examples across 100 cases. As detailed in Appendix B, 10 h of tuning were performed for the two output types under consideration (histogram and LN3), but only three to four hours were needed to obtain most of the tuning benefits. For each output type, four ‘best’ CV NNs were trained using the ‘best’ hyperparameter configuration—the configuration which produced the lowest mean L_1 CV error. For both output types, these ‘best’ CV NNs were trained for less than 10 min on average. The final histogram NN and the final LN3 NN tested in the following results provide predictions which average the predictions of those four ‘best’ CV NNs, respectively. They were tested and compared to each other on the previously unseen testing dataset 1.

Next, to provide fair comparisons, the performance of the NNs, the AS method, and the MC method were evaluated on testing dataset 2—a dataset on which no method had previously been tested. The computations which produced the following results, including the training and testing of the NNs and the generation of MC PDFs of TL, were performed on the University of Michigan’s Great Lakes HPC cluster (3.0 GHz Intel Zeon Gold 6154).

3.1. Comparing Two Neural Network PDF Output Types

After tuning, the final histogram and LN3 NNs were evaluated on testing dataset 1 which consists of 1,900,603 examples across 100 cases. A visualization of a subset of the L_1 errors for these predictions is provided in Figure 7, which shows the L_1 errors for predictions made by the histogram NN at the example receiver locations in case 101—the first case in testing dataset 1. A similar breakdown of performance for each case within both testing datasets is provided for both final NNs and the AS method in the Supplementary Materials.

It can be seen in Figure 7 that the histogram NN testing errors for this case were greatest at short ranges (<2.5 km). This trend held across all of the testing cases for both the histogram and LN3 NNs. Example A in Figure 6 lends some insight into these inaccurate short-range NN predictions. In this example, the histogram NN overestimates the variance of the MC PDF of TL. This overestimation is not due to an incompatibility between the NN construction and the form of this MC PDF of TL. The histogram NN which produced this example has no constraint on the shape of its predicted PDF. In the case of a NN constrained with the LN3 output type, the NN could still make an accurate prediction of this MC PDF of TL, possibly resulting in an L_1 error as small as 0.046. Instead, it may simply be the case that the NNs perform worse on these near-range, less-variable MC PDFs of TL because smaller errors in the predicted *mean* of these PDFs produce greater L_1 errors when there is a smaller *variance* of the MC PDF of TL. Therefore, the NN training must balance the risky option of correctly predicting a low-variance PDF, which may be very accurate ($L_1 < 0.05$) or very inaccurate ($L_1 \approx 2$), with the safer option of inaccurately predicting a high-variance PDF, which allows for at least some overlap ($L_1 \approx 1.25$ in example A) given the greater margin-of-error for the predicted mean TL value. This explanation was also supported by AS sharing the same systematic underperformance on these near-range examples, as illustrated in the Supplementary Materials.

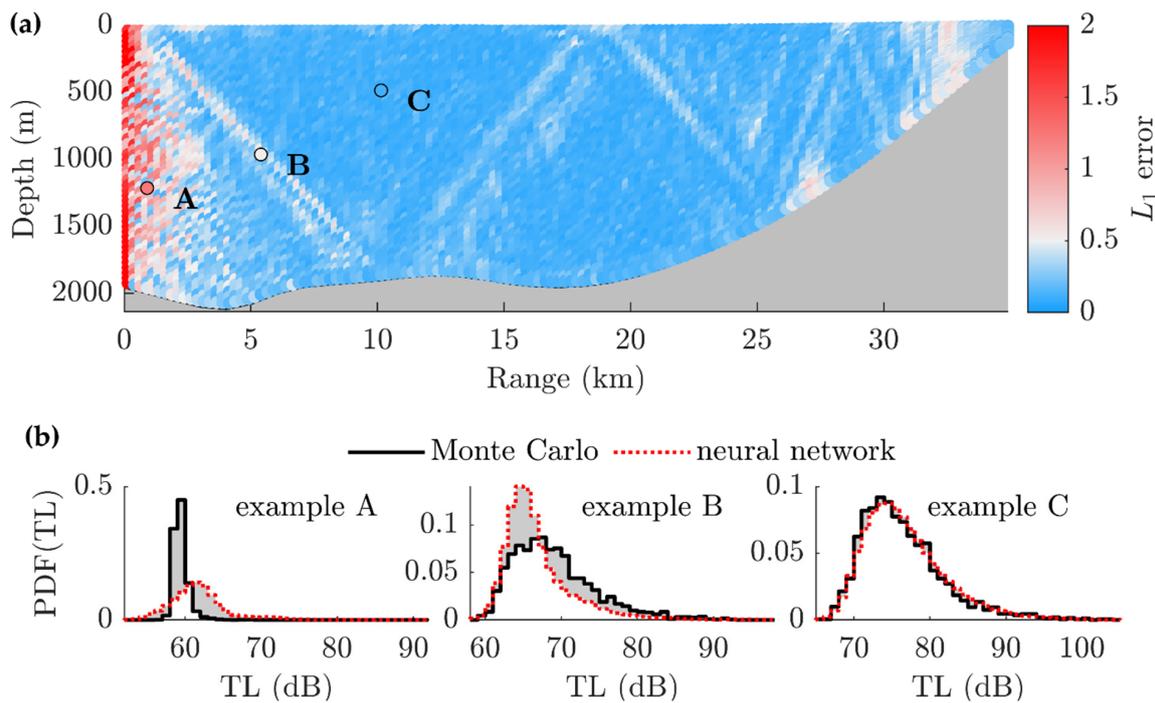


Figure 7. Final histogram NN performance for a sample test case. (a) The L_1 errors of the final histogram NN’s predictions on the examples in case 101 (a testing case) at each example’s receiver location. Results from the three labeled receiver locations are shown in (b). The NN’s predicted PDF of TL (dashed red line) is compared to the MC PDF of TL (solid black line) at three receiver locations and the L_1 error is visualized as the shaded area for each. The L_1 errors for examples A, B, and C are 1.246, 0.507, and 0.117 respectively.

When considering all examples in the testing datasets, there was a fraction of examples with most of their MC TL samples falling in the last histogram bin ($TL \geq 139$ dB) which was conditioned on the choices of the histogram binning definition and the distribution of example receiver ranges. For example, testing dataset 1 contained 55,608 examples (about 3%) with 95% or more MC TL samples contained in the last histogram bin. These high-TL or *quiet* example PDFs were relatively easy to predict for the NN, so they were excluded in the further analysis of the testing results to avoid overstating the performance of any predictive method.

The distributions of the L_1 errors across the examples in testing dataset 1 are provided in Figure 8 for both the histogram and LN3 NNs. The mean testing errors for the histogram and LN3 NNs were 0.3485 and 0.3496 respectively. For reference, the mean L_1 difference: (1) between these same MC PDFs of TL and uniform-randomly generated TL PDFs was estimated to be 1.49; (2) between random pairings of the MC TL PDFs within their cases was estimated to be 1.43; and (3) between the ‘one-sample PDF of TL’ generated from only each example’s baseline TL value was 1.82—i.e., an average of 9% of the probability mass of each MC PDF TL was contained in the same histogram bin as the baseline TL value. Given an error criterion for a given application, a testing success rate can also be computed to determine which might be more successful for that application. For example, if the L_1 error criterion being considered is 0.5 (visualized as the vertical dashed line in Figure 8), then the histogram and LN3 NNs made successful predictions on 80.13% and 80.17% of the testing examples, respectively. In this example, both NNs performed similarly according to either metric, suggesting the handling of the NN output is not the factor limiting NN accuracy. However, the histogram output type is easier to implement and more numerically stable during training.

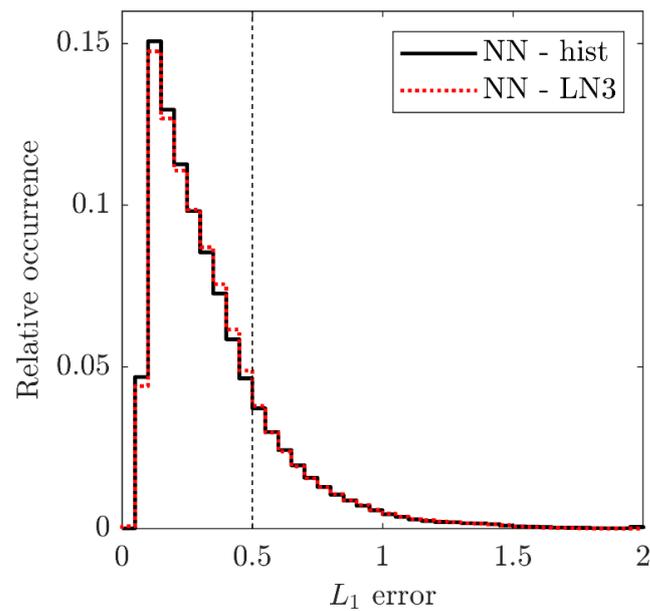


Figure 8. Distributions of testing errors on 1,844,995 examples for the final histogram NN (solid black line) and the LN3 NN (dotted red line) on testing dataset 1. The two distributions are very similar, with most of their differences occurring at L_1 errors less than 0.5 (vertical dashed line).

3.2. Comparing the Neural Network Method with Previous Methods

The final histogram and LN3 NNs were evaluated on testing dataset 2 which consists of 1,884,932 examples across 100 cases. Because the AS method was formulated with example environments generated with a different environmental model and uncertainty approach, the training dataset was used to find an approximate ‘best’ AS box size to use for comparison in this analysis. The best AS box size on the training dataset was 450 m in depth and 5 km in range (centered on the POI), producing a mean L_1 error on the training dataset of 0.3896. This box size is used to evaluate the performance of AS on testing dataset 2.

The distributions of the L_1 errors of the predictions from both final NNs and the AS method on testing dataset 2 are shown in Figure 9. The mean L_1 error on these testing examples is 0.367, 0.372, and 0.405 for the histogram NN, the LN3 NN, and AS method respectively. For an L_1 error criterion of 0.5, the success rate for each method was 77.2%, 76.5%, and 71.8%. In general, both NNs produce more accurate predictions across the examples in this testing dataset than AS. The differences in the preparation and prediction times between the NN and AS methods are discussed below.

To better compare these methods to the MC method, another 2000 uncertain environmental realizations were randomly sampled for each testing case in this dataset and were used to create an alternative MC PDF of TL for each testing example. The difference between an example’s alternative MC PDF of TL produced from all 2000 alternative MC TL samples and the example’s original MC PDFs of TL can be interpreted as a measure of MC convergence. Alternative MC PDFs of TL produced from fewer MC TL samples were considered ‘computationally cheap’ MC PDFs of TL due to their proportional decrease in estimation time and general increase in difference with the original MC PDFs of TL. Alternative MC PDFs of TL were produced at four levels—100 trials, 200 trials, 500 trials, and 2000 trials—for each testing example in testing dataset 2. These alternative MC PDFs of TL were compared to the original MC PDFs of TL via their L_1 difference.

The cumulative distributions of the L_1 errors or L_1 differences for the final NNs, AS method, and the MC method at four levels are shown in Figure 10. Compared to the distributions of L_1 errors of the NNs and AS method, the L_1 differences of the MC method across the testing examples remain nearly constant, especially at large trial numbers. Therefore, a method which produces a distribution of higher L_1 errors has a lower effective speed-up over the full-resolution MC method, considering that a reduction in MC trials

per example that produces a similar set of L_1 differences would also produce a speed-up over the full-resolution MC method. Comparing lower percentile L_1 errors (<50%), the performances of the NNs and the AS method fall between the performances of the MC method with 200 and with 500 trials. Comparing higher percentile L_1 errors (>50%), the performances of the NNs and the AS method become comparable to the MC method with even fewer trials, but these L_1 errors are generally higher for the AS method than the NNs.

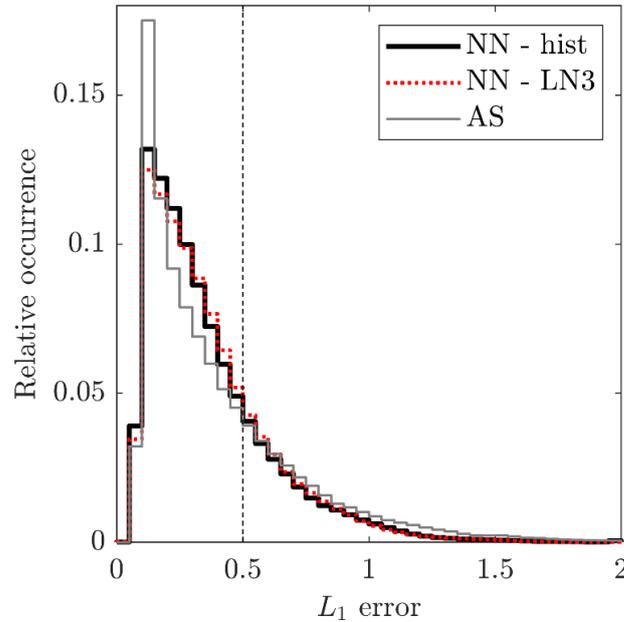


Figure 9. Distributions of testing errors for the final histogram NN (solid black line), the LN3 NN (dotted red line), and AS (thin gray line) on testing dataset 2. A greater proportion of the Area Statistics (AS) method L_1 errors are greater than 0.5 (vertical dashed line) than the NN methods.

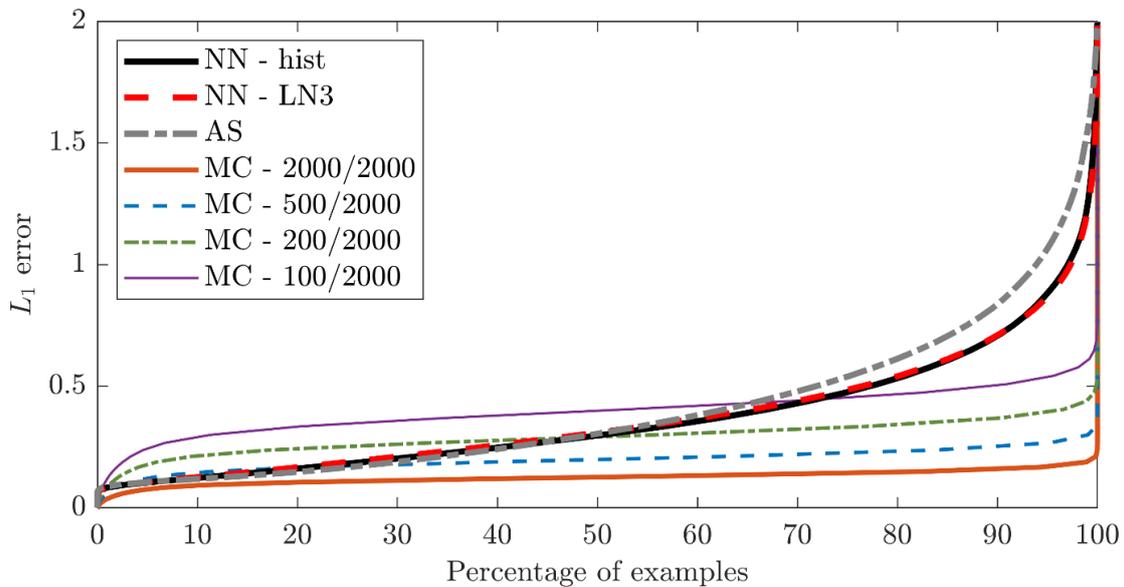


Figure 10. Cumulative distributions of the L_1 errors for the final histogram NN (thick solid black line) and LN3 NN (thick dashed red line) on testing dataset 2 in comparison to the cumulative distribution of the L_1 errors for AS (thick dash-dotted gray line) and the cumulative distributions of L_1 differences for the alternative MC PDFs of TL generated with 2000 trials (thin solid orange line), 500 trials (thin dashed blue line), 200 trials (thin dash-dotted green line), and 100 trials (thin solid purple line).

The average time it took for each method to make the predictions of the examples' TL PDFs for the cases in testing dataset 2 is compared to the mean of the L_1 errors or L_1 differences across all of these predictions in Figure 11. The alternative MC PDFs of TL were the most similar to the original MC PDFs of TL. However, the AS method was roughly 1.5 orders of magnitude faster and the NN method was roughly 2.5 to 3 orders of magnitude faster than the comparable 'cheap' MC PDF of TL predictions. The prediction-speed advantage of these methods became even greater with decreasing numbers of example receiver locations per case environment, since the prediction times: remain nearly constant for the MC method, scale linearly for the AS method, and scale non-linearly for the NN method with vectorized predictions. Reducing the number of examples by 90% and 99% decreased the mean NN prediction times by roughly 71% and 78%, respectively. These prediction times are reported in reference to serial computation. In practice, quicker per-case predictions are possible from parallel TL-field computations for MC trials, parallel AS PDF generation across example locations, and parallel NN evaluations across NNs or example locations.

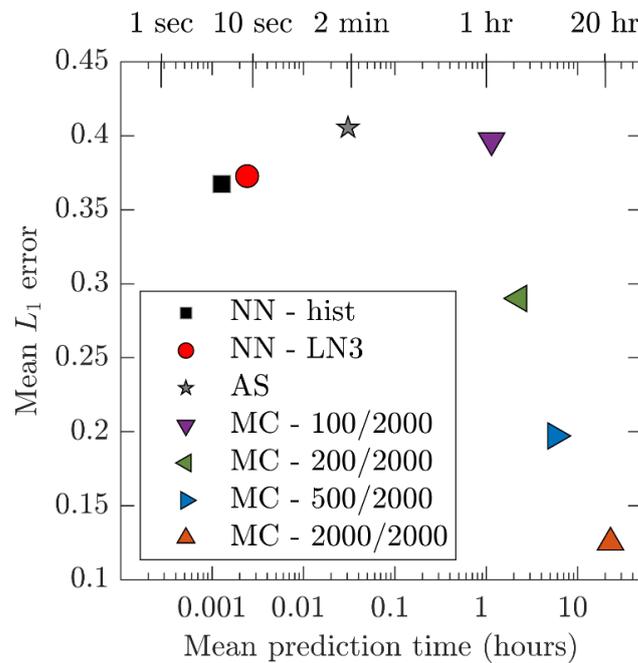


Figure 11. Comparison of the mean testing errors and TL PDF prediction times for each method. The mean prediction time is the average of the 100 testing case prediction times—the time it took to predict the PDF of TL for every example in that case after the baseline TL computation was available.

There is no preparation time for the MC method or AS method, given that the AS box size does not need to be optimized or changed as was needed in this analysis. The preparation time for the NN is approximately equal to the sum of the total prediction time for the MC method on the 100 cases in the training dataset, about 2000 h, and the NN tuning and training time—around 3 to 10 h in this analysis. Here, the NN tuning and training was the relatively fast part of NN preparation, taking only 0.5% of the total preparation time.

4. Discussion

The goal of this research was to produce a fast and flexible method for transferring ocean environmental uncertainty to acoustic transmission loss (TL) uncertainty. The supervised machine learning technique presented here provides a method substantially faster than the 'gold standard' Monte Carlo (MC) method, while maintaining applicability to environments described by many uncertain parameters, at the expense of preparation effort and prediction accuracy. With this technique, a neural network (NN) is trained to predict the probability density function (PDF) of TL from a known acoustic source to a receiver

location in an uncertain ocean environment using only the TL solution computed with the baseline environmental properties.

The speed of this NN method's predictions, which may be suitable for real-time applications, is derived from the inexpensive forward computation of a trained NN and the restriction of the NN's inputs to already available information. The NN method is adaptable because a trained NN can be used to predict TL uncertainty in a previously unseen ocean environment with a very detailed description of its properties and uncertainties as long as its training dataset contains relevant examples. However, the generation of a training dataset and the training of the NN comprise an added, up-front cost of the NN method. Additionally, the accuracy of a trained NN is limited by its incomplete set of inputs and its finite size, training effort, and training example density. In this work, the NN method was implemented, and these trade-offs were evaluated.

First, an environmental uncertainty approach was developed to address the need to generate an ensemble of realizations of possible ocean environments by using available databases and a parametric approach to their uncertainty. From 600,000 such ocean environmental realizations, MC PDFs of TL at nearly 4.3 million locations were assembled into a dataset that was used for NN training and testing. Second, a supervised learning approach was developed to generate and train the NN itself and find suitable values for seven NN hyperparameters, which governed the NN architecture and training, and five input hyperparameters, which defined the relative grid of local TL values given to the NN as inputs. The results provided herein show that a NN trained in ocean environments around the globe, throughout the year, and with various source properties can make predictions in previously unseen environments which agree with the MC PDF of TL within an L_1 error of 0.5 or less with a 76 to 80% success rate.

Details concerning the implementation of the NN method were presented. The hyperparameters considered in this analysis were outlined, and a method for choosing their values was developed, utilized, and analyzed. The hyperparameters values obtained and used here were shared as they might be useful for future implementations. Two types of methods for producing PDF predictions as NN outputs were implemented and compared. With one type, the NN output corresponded to the moments of a three-parameter log normal distribution (LN3). With the other type, the NN outputs corresponded directly to histogram bin values. An equal amount of hyperparameter optimization and training effort was given to produce a NN with each output type. While the performances of these NNs evaluated on the first testing dataset were similar, the histogram output type had the slight edge in prediction accuracy and speed, was easier to implement, and provided better numerical stability during NN training.

Once trained, the NN method was roughly 4 to 5 orders of magnitude faster than the full-resolution, traditional MC method that provided the 'ground truth' PDFs of TL for this analysis. The L_1 error was used to quantify the difference between the NN predictions and the MC PDFs of TL. The distributions of these L_1 errors on the testing datasets were presented. Given an application-specific L_1 error criterion, the predictive success rate of the NN method can be weighed along with its computational speed-up to determine if NN predictions of TL PDFs can support that application.

Another TL PDF prediction method that is faster than the MC method, Area Statistics (AS), was evaluated on the second testing dataset. The NN method was generally more accurate than AS on this testing dataset and made predictions at least as fast as AS. Although AS may not require any ahead-of-time preparation, the NN method does require prior computation in the creation of a training dataset and the training of the NN. However, a trained NN requires no further preparation if its training dataset is representative of the types of ocean environments and acoustic source properties expected to be encountered in operation. Additionally, the AS method did require preparation in this analysis due to the significant difference between the underlying environmental models and uncertainty approaches used here and in its original development.

Both the NN and AS methods generally produced their worst predictions at the shortest ranges (<2.5 km) of each environment when optimized to perform well over examples drawn uniformly in range out to 100 km. On one hand, these short-range examples with low-variance MC PDFs of TL are more difficult to predict accurately given the L_1 error metric's harsh penalization of even correctly shaped, low-variance PDF predictions that are offset just a few dB from perfectly matching the MC PDF of TL. On the other hand, these examples are also more physically tractable than the longer-range examples, since the shorter ranges tend to have fewer, simpler propagation paths less influenced by the uncertain environmental properties. Further investigation could determine if these classes, as well as additional classes which can produce multimodal TL PDFs for example, each require slight alterations to the technique of the NN method. Additionally, the preparation times reported for the NN method display the disparity between the expensive effort of producing the training dataset and the relatively cheap effort needed to train the NN. This imbalance suggests another possible avenue for improving NN predictive accuracy with the training of multiple NNs specialized for particular scenarios—only barely increasing the preparation cost and having almost no effect on the real-time prediction speed of the NN method.

Fundamentally, the NN method provides a quicker alternative to approximate the numerical MC procedure for computing predicted TL uncertainty. Even if a trained NN had a mean prediction L_1 error of zero, its predictions would still only be as accurate as the equivalent MC PDFs of TL; the NN method inherits the limitations of the underlying MC procedure. To fully assess the validity of these methods would require extensive real-world testing and measurements. The difficulty in obtaining a large enough set of measurements that is somehow equivalent to a given amount of environmental uncertainty (such as at nearby locations over a period of time to represent some amount of sound speed uncertainty) likely inhibits the creation of a large-scale real-world dataset of 'cases' that could be used to verify the overall approach or even to train a NN. However, a NN trained and deployed could be verified against individual-point ground truth measurements. For example, a global-generic NN could make 1000 TL PDF predictions all over the globe at specified sample times and a measurement could be made at each location at each sample time. If 50% of the measurements fell within the 25th and 75th percentiles of their respective NN PDF of TL, 30% of the measurements fell below the 30th percentiles, etc., that would be a very successful validation. Likewise, such an effort could be undertaken to refine a system's assumption of its inherent environmental uncertainty if a consistent over- or underestimate of TL uncertainty is observed.

In conclusion, with improved ocean environmental knowledge comes the need for a fast and adaptable means for predicting TL uncertainty arising from ocean environmental uncertainty. As available descriptions of ocean environmental properties become more precise and more accurate, applications might hope to receive these benefits in one form as reduced TL uncertainty. However, the reduced TL uncertainty may provide little to no benefit unless it is actually quantified. Additionally, the increased precision in the descriptions of environmental properties (such as denser estimates of range-, depth-, or time-dependent properties) can make the MC method even more expensive and render other approximate methods unviable. Therefore, real-time applications which rely on TL estimates need methods which are both fast and adjustable in order to benefit from the reduced TL uncertainty provided by improved oceanographic modeling and surveying. The training of NNs to quickly predict the PDFs of TL provides an approach which compares favorably to alternative acoustic-uncertainty prediction methods due to its lower in situ computational cost, better TL PDF prediction accuracy, and/or adaptability to ocean environments specified by modern databases.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jmse10101548/s1>, File S1: Baseline TL for all 300 cases; File S2: Neural Network vs. Area Statistics performance for 200 testing cases.

Author Contributions: Conceptualization, B.M.L., J.R.J. and D.R.D.; methodology, B.M.L., J.R.J. and D.R.D.; software, B.M.L. and J.R.J.; validation, B.M.L. and J.R.J.; formal analysis, B.M.L.; investigation, B.M.L.; resources, D.R.D.; data curation, B.M.L. and J.R.J.; writing—original draft preparation, B.M.L.; writing—review and editing, B.M.L., J.R.J. and D.R.D.; visualization, B.M.L.; supervision, D.R.D.; project administration, D.R.D.; funding acquisition, B.M.L. and D.R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The environmental databases used in this study, the software used to generate the simulated transmission loss dataset, and the software used to train the neural networks are all cited and publicly available.

Acknowledgments: The authors would like to recognize Brandon Lindley and Reynolds Monach of Daniel H. Wagner and Associates for useful conversations and commentary related to this work. This research was supported in part by computational resources and services provided by Advanced Research Computing (ARC), a division of Information and Technology Services (ITS) at the University of Michigan, Ann Arbor.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Appendix A.1. Bathymetry Uncertainty

Bathymetric data for each environment is read from SRTM15+ [34], a global bathymetry and elevation database and model with a spatial sampling of 15 arc seconds in longitude and in latitude. The values on this dense grid are inferred from ship sounding measurements of the bathymetry, satellite measurements of the sea surface, and physical modelling. The baseline bathymetry is obtained by bi-linearly interpolating the water column depth every 1 km along the environment's range.

To model the uncertainty in an environment's bathymetry, each value in the database is assumed to have a normal random error that is correlated with the errors of nearby grid points. The standard deviation of the error for any grid point is assumed to depend on its corresponding water column depth value. The correlation of the errors between any two grid points is modeled as a Gaussian function of the distance between the two grid points with unit-height and a standard deviation, deemed the bathymetry correlation length, which has distance units.

Model parameter values for the standard deviation of the correlated database errors and measurement errors were obtained to maximize the likelihood of 15,283 water column depth observations from 768 areas corresponding to 'Rolling Deck to Repository multi-beam SONAR' surveys from NOAA [35]. The resulting maximum likelihood estimated bathymetry error model is shown in Figure A1, and the database error model was used for the bathymetry uncertainty in all environments. These parameter values were obtained for a piecewise, non-decreasing linear model of the standard deviations of the errors. The estimates for these parameters were stable when neglecting the 0.5% greatest errors between the database and measurements. The inferred bathymetry correlation length is 16.7 km.

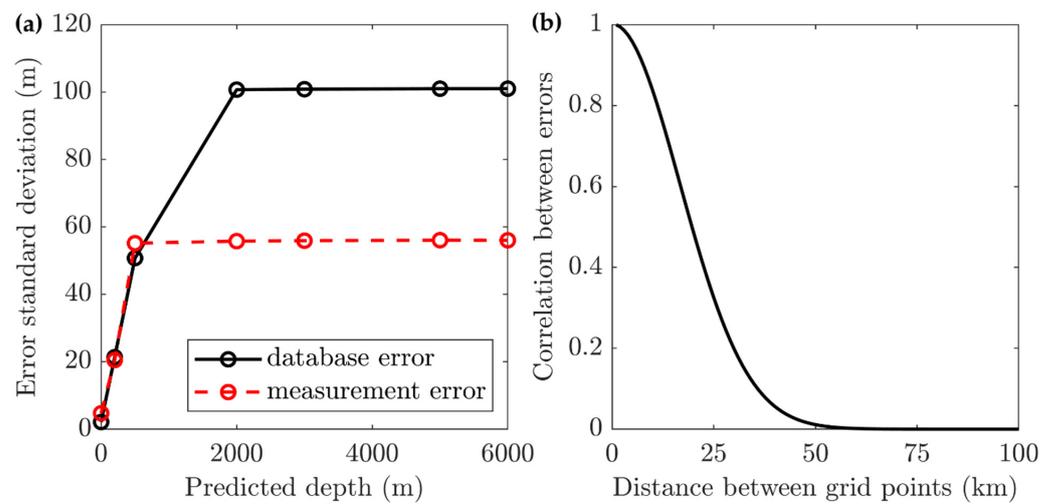


Figure A1. Maximum-likelihood estimated model of bathymetry uncertainty. (a) The model for the standard deviation of the error of the bathymetry database’s predicted value at a grid point given the predicted value (solid line), and the model for the standard deviation of the error of a measured value of water column depth given the measured value (dashed red line). (b) The model for the correlation between the errors in predicted values at two grid points given the distance between the two grid points. The inferred bathymetry correlation length is 16.7 km.

Appendix A.2. Sound Speed Uncertainty

The variability in TL due to sound speed variability can be assessed by considering many possible instances of an environment’s water column properties, as estimated by ocean models [36,37]. Here, sound speed data for each environment was read from the HYCOM Global Ocean Forecasting System (GOFS) 3.1 [38], an ocean model which provides estimated sea water temperature and salinity profiles on a 0.08° longitude × 0.04° latitude grid from 80° South to 90° North. The profiles may contain values at 40 standard depths and are available at three-hour time increments. Given the global location and a corresponding time and date, seawater property profiles are read onto a 2D grid that covers and surrounds the environment of interest. Water column sound speed values are computed using the UNESCO equation [39,40] and are bi-linearly interpolated along the environment’s range to produce an estimated sound speed field for the range-depth extent of the environment.

The uncertainty in sound speed for each environment was assessed by considering an ensemble of different HYCOM estimated sound speed fields, one for each random environmental realization. Each sound speed field corresponds to a different time defined by: time of day, day of year, and year. Therefore, the choice of how sound-speed-field times are randomly sampled for uncertain environmental realizations determines the degree of uncertainty in the sound speed field, with larger time spreads leading to larger sound-speed-field uncertainties.

The process for randomly selecting relevant sound speed fields from the HYCOM database required two steps. First, the intended or baseline time of day and day of year was given a random offset, sampled from a zero-mean normal distribution with a standard deviation having units of time. For the current analysis, this standard deviation (in general a free parameter) was chosen to be $\sigma = 336$ h, or two weeks. Historical HYCOM sound speed fields from the baseline time of the year in previous years were also considered, but with geometrically decreasing likelihood. Thus, the second step required selecting the sound speed field’s year of occurrence as the year of the baseline time with a random non-positive integer offset n , sampled from a finite geometric distribution. The probability $P(n)$ of offsetting n years into the past is given as:

$$P(n) = \frac{p(1-p)^n}{1-(1-p)^{1+N}}, n = 0, 1, 2, \dots, N \tag{A1}$$

where $p = 0.5$, which implies that a sample from $n+1$ years ago is one half as likely a sample from n years ago. Here, the maximum number of past years considered is $N = 10$. These two steps lead to a relevant but randomly selected time for sampling sound-speed profile information. The sampling then occurred by selecting the HYCOM sound speed profile(s) closest to this randomized time for the region of interest. If there were no water temperature or salinity data in the HYCOM database at that time closest to the randomized time, another randomized time was produced and used to extract an alternative sound speed field. For example, there were 63 missing times between January 1, 2019 and December 31, 2019 in the HYCOM data pulled for this study.

With this sample-time-spreading parametric approach to inherent sound speed field uncertainty, different values for the parameters (σ , p , N) may be chosen with specific applications in mind. The NN's predictions of PDFs of TL will reflect the resultant degree of TL uncertainty present in its training dataset.

Appendix A.3. Bottom Property Uncertainty

The seabed for each environment is approximated with a two-layer fluid model comprised of an upper finite-thickness sediment layer, and a lower semi-infinite acoustic basement layer having solid rock properties. The sediment type and thickness for each geographical location were extracted from global databases. For computational purposes, the basement layer extended 900 m below the sediment layer at which point the attenuation was abruptly increased to 10 dB per wavelength to attenuate artificial reflections [19].

The Bottom Sediment Type (BST) Database Version 2.0 [41] was used to estimate the sediment type for each environment of interest. The possible sediment classifications are the 23 High Frequency Environmental Acoustics (HFEVA) model [42] categories. Each HFEVA category corresponds to a nominal bulk grain size (expressed in the logarithmic unit ϕ), sediment density, compression-wave sound speed, and compression-wave attenuation [42] (Sec. IV, Table 2). As a simple means of correcting for the empirically observed attenuation of sandy sediment types (grain size 0.5 to 5.5 ϕ) at low frequencies (<1 kHz), the listed attenuation values were used to compute the attenuation at 1 kHz, which was then decreased to the target frequency using a power law with an exponent of 1.8 [43] (Equation (1)). The nominal grain size for each uncertain environmental realization was given a zero-mean normal random offset and was used to linearly interpolate the realization's sediment acoustic properties. In this analysis, the standard deviation of this grain-size offset was chosen to be 0.2 ϕ .

A range-independent sediment thickness was estimated for each environment using the GlobSed database [44]. Although the sediment thickness of an environment might also be modeled as a random variable, no model was available for the error in the estimated sediment thickness provided in this database, and the quantification of the database error using independent measurements was beyond the scope of this project. Thus, each environment's sediment thickness is held constant between the baseline and uncertain environmental realizations.

For each uncertain environmental realization, the acoustic basement's acoustic properties were computed as a random weighted average between the properties for limestone and basalt bottom types [20] (Table 1.3). For the baseline environment, the exact average of these properties was always used.

Appendix B

This section gives details on the hyperparameter optimization, or NN tuning, that was performed in order to determine a suitable configuration for the twelve hyperparameters of interest. Seven hyperparameters concerned the architecture of the neural network and its training and five hyperparameters concerned the size and spacing of the NN's primary input—the local TL grid. The ranges of values considered and the values corresponding to the best configuration found during the tuning performed for each NN output type in this analysis are summarized in Table A1.

Table A1. List of investigated hyperparameters, the ranges of values explored, and the values corresponding to the best attempted configuration for each output type.

| Hyperparameter (Integer) | Minimum | Maximum | Best Histogram | Best LN3 | |
|------------------------------|---------------------|--------------------|--------------------|--------------------|----------|
| Hidden layers | 1 | 10 | 5 | 4 | |
| Nodes/hidden layer | 2^3 | 2^9 | 2^7 | 2^9 | |
| Mini-batch size | 2^3 | 2^7 | 2^6 | 2^6 | |
| Hyperparameter (Continuous) | Minimum | Maximum | Best Histogram | Best LN3 | |
| Learning rate: hist. | $10^{-4.5}$ | $10^{-2.0}$ | $10^{-3.7}$ | N/A | |
| Learning rate: LN3 | $10^{-5.0}$ | $10^{-3.5}$ | N/A | $10^{-4.5}$ | |
| AMSgrad : β_1 | $1 - (1/10)^{0.95}$ | $1 - (1/10)^{2.5}$ | $1 - (1/10)^{2.3}$ | $1 - (1/10)^{1.0}$ | |
| AMSgrad : β_2 | $1 - (1/10)^{1.0}$ | $1 - (1/10)^{4.0}$ | $1 - (1/10)^{3.1}$ | $1 - (1/10)^{1.1}$ | |
| Hyperparameter (Categorical) | Choices | | | Best Histogram | Best LN3 |
| Activation function | ReLU | ELU | Swish | Swish | Swish |
| Grid points in range | 5 | 11 | 15 | 11 | 11 |
| Grid points in depth | 5 | 11 | 17 | 17 | 17 |
| Spatial units in range | 1 | 2 | | 1 | 1 |
| Spatial units in depth | 0.5 | 1 | | 1 | 1 |
| Spatial unit | 15 m | 1 wavelength | | 15 m | 15 m |

For each hyperparameter configuration considered, four NNs were randomly initialized, trained on random subsets of the training dataset, and cross-validated by monitoring their performance on the remaining training examples. Early stopping of the training occurred if this cross-validation performance stopped improving in order to avoid overfitting. In order to efficiently find a suitable configuration, the probabilistic selection criteria of Bayesian Optimization [45] was implemented via the scikit-optimize library [46] alongside the resource disbursement methodology of the Successive Halving algorithm (SHA) [47]. This tuning scheme permitted quick evaluation of many configurations and thorough investigation of promising ones until a fixed total allowed time had been reached.

The tuning procedure introduced and detailed above was performed twice—once for a histogram output type and once for a LN3 output type. The training dataset contained 491,498 examples across 100 cases. The four cross-validation splits were created by randomly selecting 90 cases to train on and keeping the remaining 10 cases for to evaluate a mean L_1 cross-validation (CV) error on. The tuning aimed to minimize the average of these four mean L_1 CV errors, termed the CV-score for brevity. Every NN was trained for two minutes or until its CV error converged (did not improve after 90 s) or diverged (training became unstable leading to numerical failures). Some well-performing NNs were selected by the optimizer to receive additional training time.

The current best CV-score found within a given amount of tuning time for both output types is shown in Figure A2. The total effort allowed for tuning was about ten hours for each output type, but most of the improvement came within the first three to four hours. The amount of time needed to obtain a suitable NN could be greatly reduced by using a smaller training dataset, by reducing the number of hyperparameters—such as by fixing the local TL grid, or by parallelizing the process to allow for simultaneous training or hyperparameter investigation. The best CV-score for the histogram output type was obtained after 8 h of tuning, and the four CV NNs had an average training time of about 10 min and mean L_1 CV error of 0.3749. The best CV-score for the LN3 output type was obtained in 3 h of tuning, and the four CV NNs had an average training time of about 8 min and mean L_1 CV error of 0.3742. Using the histogram output type provided more stable training than using the LN3 output type, having 0/236 compared to 33/240 attempted NN trainings diverge during the respective tunings. The hyperparameter values corresponding

to the best attempted configuration are available in Table A1. Interestingly, the same local TL grid was indicated for both output types and contained 187 points created by taking 11 points in range taken every 15 m and 17 points in depth taken every 15 m. The preference for a relatively large, source-frequency-independent input grid suggests that the local waveguide-scale TL statistics are more informative of the TL uncertainty than the immediate spatial sensitivity of the TL field given both: (1) the even-distribution of example receiver locations throughout the environments; and (2) the present sources and degrees of uncertainty.

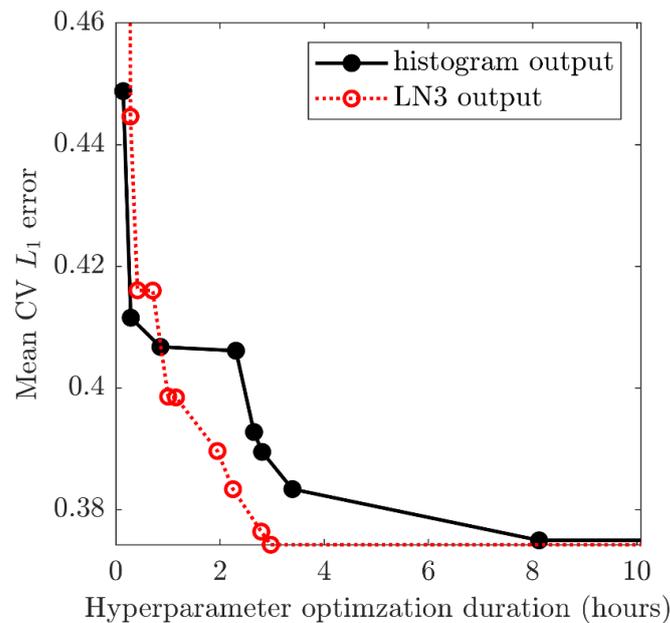


Figure A2. Neural network improvement of cross-validation (CV) performance with tuning effort. The best current CV-scores at any time during the tuning for the histogram output type (solid black line) and the LN3 output type (dotted red line). The first configuration for the LN3 output type NN provided a CV-score of 0.83; this point is omitted from the plot for clarity.

After tuning, four NNs trained with the best configuration for each output type were available. Given an example and its MC PDF of TL, the L_1 error of the average of these four NN predictions will be no worse than the average of the four L_1 errors for each prediction. The mean L_1 errors across the training dataset for the tuned histogram and LN3 NNs decrease from averages of 0.3148 and 0.3330 to 0.3048 and 0.3253 when using the average of the NN predictions. Although making predictions takes four times as long, this improvement in performance may be worth the cost. Therefore, the final histogram NN and final LN3 NN analyzed in Section 3 are the models whose predictions are the average of the four trained CV NNs which correspond to the best attempted configuration for each respective output type.

References

1. Dosso, S.E. Quantifying Uncertainty in Geoacoustic Inversion. I. A Fast Gibbs Sampler Approach. *J. Acoust. Soc. Am.* **2002**, *111*, 129–142. [[CrossRef](#)] [[PubMed](#)]
2. Dosso, S.E.; Nielsen, P.L. Quantifying Uncertainty in Geoacoustic Inversion. II. Application to Broadband, Shallow-Water Data. *J. Acoust. Soc. Am.* **2002**, *111*, 143–159. [[CrossRef](#)] [[PubMed](#)]
3. Dosso, S.E.; Wilmut, M.J. Data Uncertainty Estimation in Matched-Field Geoacoustic Inversion. *IEEE J. Oceanic Eng.* **2006**, *31*, 470–479. [[CrossRef](#)]
4. Jiang, Y.-M.; Chapman, N.R. The Impact of Ocean Sound Speed Variability on the Uncertainty of Geoacoustic Parameter Estimates. *J. Acoust. Soc. Am.* **2009**, *125*, 2881–2895. [[CrossRef](#)] [[PubMed](#)]
5. Chapman, N.R. Perspectives on Geoacoustic Inversion of Ocean Bottom Reflectivity Data. *J. Mar. Sci. Eng.* **2016**, *4*, 61. [[CrossRef](#)]
6. Urlick, R.J. *Principles of Underwater Sound*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1983; Chapter 2, ISBN 978-0-07-066087-8.

7. Huang, C.-F.; Gerstoft, P.; Hodgkiss, W.S. Validation of Statistical Estimation of Transmission Loss in the Presence of Geoacoustic Inversion Uncertainty. *J. Acoust. Soc. Am.* **2006**, *120*, 1932–1941. [[CrossRef](#)]
8. Zingarelli, R.A. A Mode-Based Technique for Estimating Uncertainty in Range-Averaged Transmission Loss Results from Underwater Acoustic Calculations. *J. Acoust. Soc. Am.* **2008**, *124*, EL218–EL222. [[CrossRef](#)]
9. Fabre, J.P.; Wood, W.T. Uncertainty of Transmission Loss Due to Small Scale Fluctuations of Sound Speed in Two Environments. *Proc. Meet. Acoust.* **2013**, *19*, 005008. [[CrossRef](#)]
10. James, K.R.; Dowling, D.R. A Method for Approximating Acoustic-Field-Amplitude Uncertainty Caused by Environmental Uncertainties. *J. Acoust. Soc. Am.* **2008**, *124*, 1465–1476. [[CrossRef](#)]
11. James, K.R.; Dowling, D.R. Pekeris Waveguide Comparisons of Methods for Predicting Acoustic Field Amplitude Uncertainty Caused by a Spatially Uniform Environmental Uncertainty (L). *J. Acoust. Soc. Am.* **2011**, *129*, 589–592. [[CrossRef](#)]
12. Finette, S. A Stochastic Representation of Environmental Uncertainty and Its Coupling to Acoustic Wave Propagation in Ocean Waveguides. *J. Acoust. Soc. Am.* **2006**, *120*, 2567–2579. [[CrossRef](#)]
13. Finette, S. A Stochastic Response Surface Formulation of Acoustic Propagation through an Uncertain Ocean Waveguide Environment. *J. Acoust. Soc. Am.* **2009**, *126*, 2242–2247. [[CrossRef](#)] [[PubMed](#)]
14. Patterson, B.S.; Dowling, D.R. Efficient Estimation of the Probability Density Function of Acoustic Transmission Loss in Uncertain Ocean Environments Using Area Statistics. (*Unclassified*) *JUA(USN)* **2018**, *68*, 194–213.
15. Bianco, M.J.; Gerstoft, P.; Traer, J.; Ozanich, E.; Roch, M.A.; Gannot, S.; Deledalle, C.-A. Machine Learning in Acoustics: Theory and Applications. *J. Acoust. Soc. Am.* **2019**, *146*, 3590–3628. [[CrossRef](#)]
16. Michalopoulou, Z.-H.; Gerstoft, P.; Kostek, B.; Roch, M.A. Introduction to the Special Issue on Machine Learning in Acoustics. *J. Acoust. Soc. Am.* **2021**, *150*, 3204–3210. [[CrossRef](#)]
17. Collins, M.D. A Split-step Padé Solution for the Parabolic Equation Method. *J. Acoust. Soc. Am.* **1993**, *93*, 1736–1742. [[CrossRef](#)]
18. Collins, M.D. Generalization of the Split-step Padé Solution. *J. Acoust. Soc. Am.* **1994**, *96*, 382–385. [[CrossRef](#)]
19. Collins, M.D. *User's Guide for RAM Versions 1.0 and 1.0p*; Naval Research Laboratory: Washington, DC, USA, 1995.
20. Jensen, F.B.; Kuperman, W.A.; Porter, M.B.; Schmidt, H. *Computational Ocean Acoustics*, 2nd ed.; Springer: New York, NY, USA, 2011; ISBN 978-1-4419-8677-1.
21. Smith, K.B. Convergence, Stability, and Variability of Shallow Water Acoustic Predictions Using a Split-Step Fourier Parabolic Equation Model. *J. Comput. Acoust.* **2001**, *9*, 243–285. [[CrossRef](#)]
22. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16), Savannah, GA, USA, 2–4 November 2016.
23. Reddi, S.J.; Kale, S.; Kumar, S. On the Convergence of Adam and Beyond. *arXiv* **2019**, arXiv:1904.09237.
24. Ruder, S. An Overview of Gradient Descent Optimization Algorithms. *arXiv* **2017**, arXiv:1609.04747.
25. Salimans, T.; Kingma, D.P. Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks. In *Advances in Neural Information Processing Systems 29, Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, 5–10 December 2016*; Curran Associates, Inc.: Red Hook, NY, USA, 2016.
26. Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. In Proceedings of the Machine Learning Research 15, Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 11–13 April 2011; Gordon, G., Dunson, D., Dudik, M., Eds.; JMLR: Cambridge, MA, USA, 2011; pp. 315–323.
27. Clevert, D.-A.; Unterthiner, T.; Hochreiter, S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *arXiv* **2016**, arXiv:1511.07289.
28. Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for Activation Functions. *arXiv* **2017**, arXiv:1710.05941.
29. O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *arXiv* **2015**, arXiv:1511.08458.
30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
31. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the Machine Learning Research 37, Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 7–9 July 2015; Bach, F., Blei, D., Eds.; JMLR: Cambridge, MA, USA, 2015; pp. 448–456.
32. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
33. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [[CrossRef](#)]
34. Tozer, B.; Sandwell, D.T.; Smith, W.H.F.; Olson, C.; Beale, J.R.; Wessel, P. Global Bathymetry and Topography at 15 Arc Sec: SRTM15+. *Earth Space Sci.* **2019**, *6*, 1847–1864. [[CrossRef](#)]
35. Rolling Deck to Repository. *Multibeam Bathymetry Database (MBBDB)*; NOAA National Centers for Environmental Information: Boulder, CO, USA, 2004; (accessed on 1 May 2020). [[CrossRef](#)]
36. Duda, T.F.; Lin, Y.-T.; Cornuelle, B.D. Scales of Time and Space Variability of Sound Fields Reflecting Obliquely from Underwater Slopes. *Proc. Meet. Acoust.* **2013**, *19*, 070025. [[CrossRef](#)]
37. Vagle, S.; Burnham, R.E.; O'Neill, C.; Yurk, H. Variability in Anthropogenic Underwater Noise Due to Bathymetry and Sound Speed Characteristics. *J. Mar. Sci. Eng.* **2021**, *9*, 1047. [[CrossRef](#)]

38. Chassignet, E.P.; Hurlburt, H.E.; Smedstad, O.M.; Halliwell, G.R.; Hogan, P.J.; Wallcraft, A.J.; Baraille, R.; Bleck, R. The HYCOM (HYbrid Coordinate Ocean Model) Data Assimilative System. *J. Mar. Syst.* **2007**, *65*, 60–83. [[CrossRef](#)]
39. Fofonoff, N.P.; Millard, R.C., Jr. Algorithms for Computation of Fundamental Properties of Seawater. *UNESCO Tech. Pap. Mar. Sci.* **1983**, *44*, 1–53.
40. Chen, C.; Millero, F.J. Speed of Sound in Seawater at High Pressures. *J. Acoust. Soc. Am.* **1977**, *62*, 1129–1135. [[CrossRef](#)]
41. Naval Oceanographic Office. *Database Description for Bottom Sediment Type (U), Version 2.0*; Naval Oceanographic Office, Acoust. Div., Stennis Space Center: Hancock County, MS, USA, 2003.
42. Washington University Seattle Applied Physics Lab. *APL-UW High-Frequency Ocean Environmental Acoustic Models Handbook*; Applied Physics Laboratory, University of Washington: Seattle, WA, USA, 1994.
43. Holmes, J.D.; Carey, W.M.; Dediu, S.M.; Siegmann, W.L. Nonlinear Frequency-Dependent Attenuation in Sandy Sediments. *J. Acoust. Soc. Am.* **2007**, *121*, EL218–EL222. [[CrossRef](#)] [[PubMed](#)]
44. Straume, E.O.; Gaina, C.; Medvedev, S.; Hochmuth, K.; Gohl, K.; Whittaker, J.M.; Abdul Fattah, R.; Doornenbal, J.C.; Hopper, J.R. GlobSed: Updated Total Sediment Thickness in the World's Oceans. *Geochem. Geophys. Geosyst.* **2019**, *20*, 1756–1772. [[CrossRef](#)]
45. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems 25, Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS 2012), Lake Tahoe, NV, USA, 3–6 December 2012*; Pereira, F., Burges, C.J., Bottou, L., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2012.
46. Scikit-Optimize. Available online: <https://scikit-optimize.github.io/stable/> (accessed on 18 November 2019).
47. Jamieson, K.; Talwalkar, A. Non-Stochastic Best Arm Identification and Hyperparameter Optimization. In *Proceedings of the Machine Learning Research 51, Proceedings of the 18th International Conference on Artificial Intelligence and Statistics, Cadiz, Spain, 9–11 May 2016*; Gretton, A., Christian, R.C., Eds.; JMLR: Cambridge, MA, USA, 2016; pp. 240–248.