

# POSTERIOR ESTIMATION OF TRANSMISSION LOSS FROM OCEAN ACOUSTIC DATA

Peter Gerstoft<sup>1</sup>, Chen-Fen Huang<sup>1</sup>, W.S. Hodgkiss<sup>1</sup>, and Peter Nielsen<sup>2</sup>

<sup>1</sup> Marine Physical Laboratory/SIO, UCSD, La Jolla, CA 92093-0701, USA

<sup>2</sup> NATO Undersea Research Center, 19138 La Spezia, Italy

*A weakness in sonar performance prediction has been the lack of a means for quantifying the impact of uncertainty in estimates of the ocean environment. This paper describes the estimation of transmission loss and its statistical properties based on observations of ocean acoustic data. This problem is solved by first finding an ensemble of relevant environmental model parameters and then using all of these models to map into the transmission loss domain. In this mapping each environmental model is weighted according to its likelihood function.*

*This algorithm is illustrated on vertical array data acquired during the ASIAEX 2001 East China Sea experiment. The environmental parameters are first estimated. Based on the likelihood that each of these environmental models fits the ocean acoustic data, each model is mapped into transmission loss. This enables us to compute a full probability distribution for the transmission loss in both range and depth.*

## 1. INTRODUCTION

A weakness in sonar performance prediction has been the lack of a means for quantifying the impact of uncertainty in estimates of the ocean environment. Figure 1 offers an overview of posterior estimation of transmission loss from ocean acoustic data. Ocean acoustic data  $\mathbf{d}$  is observed on a vertical or horizontal array. The inverse problem is first solved using Bayes rule to compute the posterior distribution  $p(\mathbf{m}|\mathbf{d})$  of environmental parameters  $\mathbf{m}$ . We are not interested in the environment itself but rather better estimates of the transmission field, the usage domain  $\mathbf{u}(\mathbf{m})$ . Based on the posterior distribution  $p(\mathbf{m}|\mathbf{d})$ , the distribution of  $p(\mathbf{u}(\mathbf{m}))$  is obtained via Monte Carlo simulation. From this transmission-loss distribution, all relevant statistics of the transmission loss can be obtained.

The principle of the inversion is indicated in Fig. 1. Based on the ocean acoustic data  $\mathbf{d}$  we statistically characterize transmission loss (the usage domain  $\mathbf{u}$ ). The vector  $\mathbf{d}$  represents the acoustic data observed at  $N$  hydrophones and the vector  $\mathbf{u}$  represents transmission loss at range and depth. As shown in Fig. 1, this is mapped via a set of

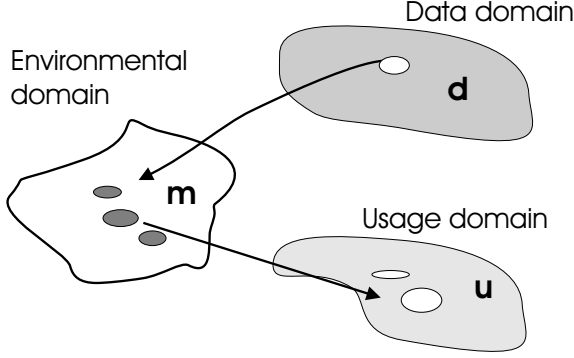


Figure 1: An observation  $\mathbf{d}$  is mapped into a distribution environmental parameters  $\mathbf{m}$  that potentially could have generated it. The environmental parameters are then mapped into the usage domain.

$M$  environmental parameters  $\mathbf{m}$ . The approach involves a number of steps as outlined below:

1. Determine a model for the ocean acoustic environment and select an appropriate propagation model. These two form the mapping  $\mathbf{d}(\mathbf{m})$  from the parameter space  $\mathbf{m}$  to data space  $\mathbf{d}$ .
2. Determine the mapping  $\mathbf{u}(\mathbf{m})$  from the parameter space to usage space. Except for a change in geometry (source depth), here this is similar to  $\mathbf{d}(\mathbf{m})$ , but could be any other model (e.g. some other metric of sonar performance).
3. Find acceptable models  $\mathbf{m}$  from the data. As indicated in Fig. 1, a region around the data can map into several acceptable solutions in the model domain.
4. Map the acceptable models into the usage domain. Several models can map into the same usage region.

As indicated in Fig. 1, the mapping from data to usage domain is non-unique. There are many environmental models that give about the same goodness-of-fit. The maximum likelihood (ML) estimate of the environmental model gives the best fit. Instead of using just one estimated environment, it is proposed to describe the environmental solution probabilistically. This probability is then mapped into the usage domain. Knowing the posterior probability distribution in the usage domain is preferable to having a single point estimate such as the usage domain result corresponding to the ML-solution.

## 2. INVERSE PROBLEM FRAMEWORK

In the Bayesian paradigm, the solution to determining parameters of interest  $\mathbf{m}$  given an observation  $\mathbf{d}$  is characterized by the posterior probability  $p(\mathbf{m}|\mathbf{d})$ . First, the prior information of the model parameter vector is the probability density  $p(\mathbf{m})$ . Then, this information is combined with the likelihood function  $p(\mathbf{d}|\mathbf{m})$  provided by the combination of data and the physical model to give the posterior probability density  $p(\mathbf{m}|\mathbf{d})$  of the model parameters. A clear discussion of inverse theory from a probabilistic point of view is given by [1]. Additional details of Monte Carlo sampling of posterior probabilities can be found in [2, 3, 4]. The solution to the inverse problem is then

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \quad (1)$$

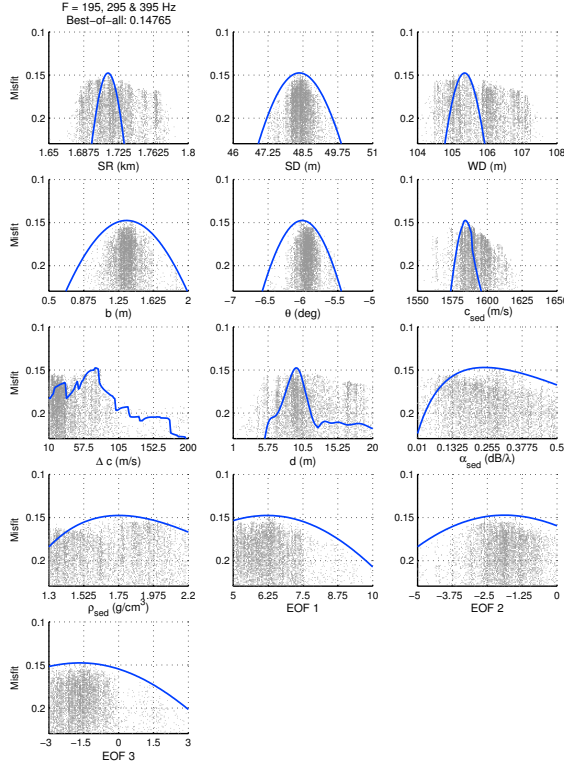


Figure 2: Marginal scatter diagrams of the SAGA search for the model parameters. The vertical axis represents the attained misfit on a linear scale. The thick line is the sensitivity curve of the multi-frequency misfit function using the best-fit model as a baseline.

where the probability of the data  $p(\mathbf{d})$  is independent of the environmental model and here can be neglected. The posterior distribution  $p(\mathbf{m}|\mathbf{d})$  contains all relevant information and from this distribution all of the relevant features of the environment can be found such as the maximum *a posteriori* (MAP) estimator.

We are not interested in the environment itself but rather better estimates in the information usage domain  $\mathbf{u}(\mathbf{m})$ . Based on the posterior distribution  $p(\mathbf{m}|\mathbf{d})$ , the distribution of  $p(\mathbf{u}(\mathbf{m}))$  is obtained and from this distribution all relevant statistics of the usage space can be obtained. In the present application the usage domain is transmission loss.

The posterior distribution of  $p(\mathbf{m})$  is first obtained and from this the probability distribution of  $\mathbf{u}$  is obtained

$$p(\mathbf{u}) = \int_{\mathcal{M}} \delta[\mathbf{u}(\mathbf{m}) - \mathbf{u}] p(\mathbf{m}) d\mathbf{m} \quad (2)$$

where  $\mathcal{M}$  represents the model space. This integral is implemented numerically by using samples from the model space  $\mathbf{m}$  based on the posterior distribution  $p(\mathbf{m}|\mathbf{d})$  and then binning  $\mathbf{u}(\mathbf{m})$ . From the distribution  $p(\mathbf{u})$  all relevant statistical features of the usage domain, e.g., transmission loss, can be obtained, as for example the mean and variation of  $\mathbf{u}$ .

Assuming a simple linear data model as detailed in Ref [2], we obtain the Bartlett objective function  $\phi(\mathbf{m})$  and likelihood function  $\mathcal{L}(\mathbf{m})$ ,

$$\phi(\mathbf{m}) = \text{tr} \mathbf{C} - \frac{\mathbf{d}^\dagger(\mathbf{m}) \mathbf{C} \mathbf{d}(\mathbf{m})}{\|\mathbf{d}(\mathbf{m})\|}, \quad (3)$$

$$\mathcal{L}(\mathbf{m}) = \left[ \frac{N}{\pi \phi(\mathbf{m}^{\text{ML}})} \right]^N \exp \left[ -N \frac{\phi(\mathbf{m})}{\phi(\mathbf{m}^{\text{ML}})} \right]. \quad (4)$$

where  $\mathbf{C}$  is the data covariance matrix and ML refers to the maximum likelihood estimate.

### 3. ASIAEX INVERSION

Data from the 2001 East China sea experiment (see Ref [6]) are used to illustrate the method. A 16-element vertical array was deployed in 105-m deep water. For the inversion, a source towed at 48.5 m depth is used. Matched-field geoacoustic inversion using the selected frequencies 195, 295, and 395 Hz was carried out at  $T = 29$  min over a parameter space of 13 parameters including geometrical, geoacoustic, and ocean sound speed EOF coefficients. Based upon the GPS position of R/V Melville, the source was approximately 1.7 km away from the VLA. Figure 2 shows the marginal dot diagrams for the model parameters. The vertical axis is the achieved misfit using a Bartlett objective function with respect to the parameter sampled during the SAGA optimization [5]. The thick line superimposed on each scatter plot was obtained by using the best-fit model corresponding to the optimal value of the objective function as a baseline and computing the sensitivity for the optimized parameter. We see that the sampled values for the array bow and tilt parameters ( $b$  and  $\theta$ ) are spread mainly inside the sensitivity curve and align mostly with the best-fit values. A similar behavior is observed for the ocean sound speed EOF coefficients but with a wider span. The consistency between the local (line) and global (dots) searches shows that this set of parameters is weakly correlated with the other parameters. For the geoacoustic parameters, most sampled values wander outside the curve. This reveals the more complicated structure of the multi-dimensional search space. Note that the sampled values for the source range (SR) and the water depth (WD) are spread uniformly throughout the range of the parameter interval. This is due to the strong coupling between these two parameters.

For simplicity we assume that only two of the model parameters, water depth and sediment sound speed have any associated uncertainty. Using the optimal values found in the inversion detailed above and varying only the above two parameters gives posterior probability indicated in Fig. 3 based on the likelihood formulation and using the same data as in the inversion.

### 4. TRANSMISSION LOSS PREDICTION

We then use the posterior probability (Fig. 3) to compute the posterior probability  $\mathbf{u}(\mathbf{m})$  using a frequency of 500 Hz and a source depth of 20 m. Except for water depth (bounds 100 to 120 m) and bottom sound speed (bounds 1550 to 1750 m/s), we keep the environment fixed at the values found in the inversion. In the present application we evaluate this using grid integration. First the probability for transmission loss at midwater depth (50 m) is evaluated in Fig. 4. The prior probability assumes even weighting of all the explored environmental models with bounds as indicated in Fig. 3. The prior distribution (top) is spread out over a wide range but the posterior distribution (bottom) is more narrow. For the first 200 m, the propagation is only little influenced by the waveguide parameters and thus there is little difference between posterior and prior distribution. It also is observed that the peaks in the transmission loss curve are much better determined than the nulls.

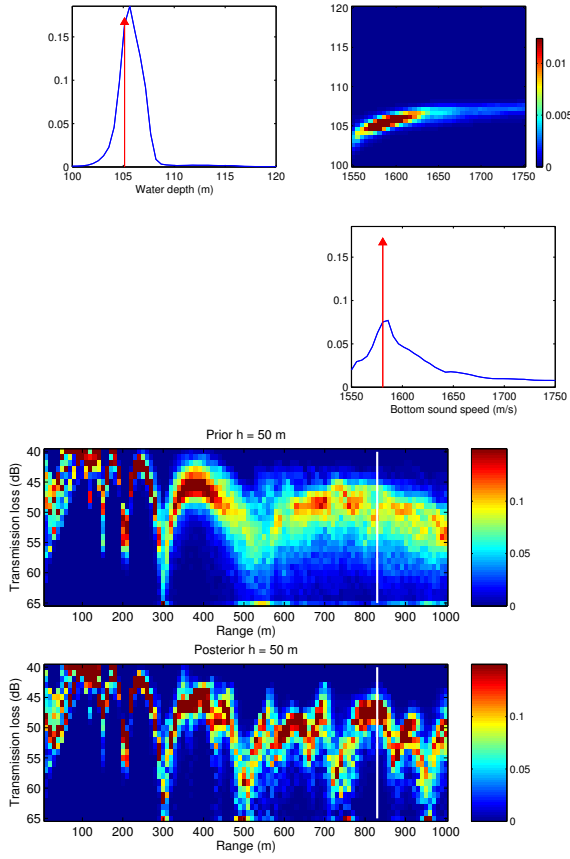


Figure 4: Prior (top) and posterior (bottom) probability distributions for transmission loss versus range at 50-m depth.

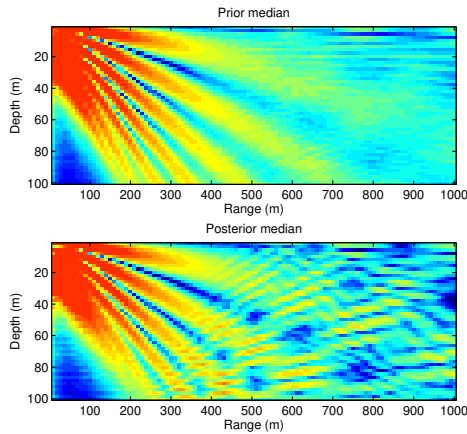


Figure 6: Median transmission loss (dB) based on prior (top) and posterior (bottom) distributions of environmental parameters (water depth and sediment sound speed).

Figure 3: Marginal probabilities for water depth and bottom sound speed. The contour plot shows the 2D distribution. The red vertical arrow indicates the ML solution (saga best fit). Note, that the maximum in the marginal distribution might not correspond to the ML maximum.

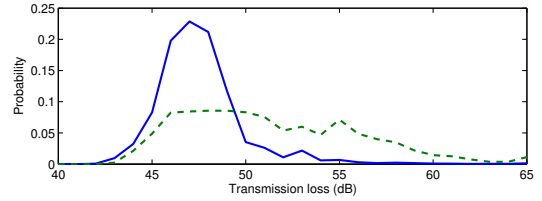


Figure 5: Posterior (solid) and prior (dashed) probabilities at 830-m range and 50-m depth. These correspond to a cut (white lines) though the contours in Fig. 4.

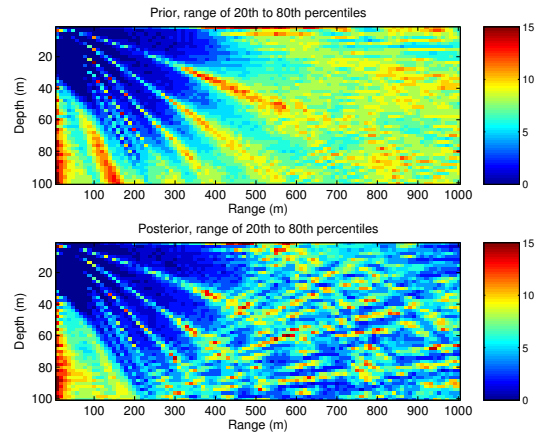


Figure 7: Range (dB) between 20th and 80th percentiles of the transmission loss for prior (top) and posterior (bottom) probabilities.

The nulls are very sensitive to small variations in the waveguide parameters. This easily is seen by taking a cut through the contour plots in Fig. 4 at 830-m range (indicated by a white line, corresponding to a peak in transmission loss curve), as shown in Fig. 5. The posterior (solid) is much more concentrated than the prior (dashed).

As the full probability distributions are available and cannot be assumed to be Gaussian, it is preferable to characterize the distributions with medians and percentiles instead of means and standard deviations. Since the distributions are not symmetric around the medians neither are the percentiles. Contours of the median transmission loss then are computed for the prior and posterior fields, Fig. 6. It is seen that the nulls are more spread out in the prior field than in the posterior field. A good way to understand the uncertainty is to plot the 20th to 80th percentile widths of the prior and posterior fields, Fig. 7. Close to the source, there is little uncertainty for both prior and posterior fields as the sound field is not influenced by the waveguide parameters. Further away from the source, the prior uncertainty increases faster than the posterior, as the waveguide parameters are less well determined. It also is seen that around the nulls of the median fields, Fig. 6, the variations in the fields are the largest, Fig. 7.

## 5. CONCLUSIONS

An algorithm for estimating the statistical properties of transmission loss based on the output from a geoacoustic inversion has been described using a likelihood formulation. The approach taken here then maps the environmental parameters and their uncertainties into transmission loss. The likelihood function is developed assuming the error in the data is Gaussian.

**Acknowledgments:** This work was supported by the Office of Naval Research under grant No. N00014-03-1-0393.

## References

- [1] A. Tarantola, *Inverse Problem Theory: Methods for Data Fitting and Parameter Estimation*, Elsevier, New York, 1987.
- [2] P. Gerstoft and C.F. Mecklenbräuker, "Ocean acoustic inversion with estimation of *a posteriori* probability distributions," *J. Acoust. Soc. Am.*, **104**, pp. 808–817, 1998.
- [3] S.E. Dosso, "Quantifying uncertainties in geoacoustic inversion. I: a fast Gibbs sampler approach," *J. Acoust. Soc. Am.*, **111**, pp. 129–142, 2002.
- [4] K. Mosegaard and A. Tarantola, "Probabilistic approach to inverse problems," Chapter in *International Handbook of Earthquake and Engineering Seismology*, (<http://web.ccr.jussieu.fr/tarantola/>), 2002.
- [5] P. Gerstoft, "Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability distributions," *J. Acoust. Soc. Am.*, **95**, pp. 770–782, 1994.
- [6] C-F. Huang and W.S. Hodgkiss, "Matched Field Geoacoustic Inversion of Low Frequency Source Tow Data From the ASIAEX East China Sea Experiment," *IEEE Oceanic Eng.*, to appear 2004.