Uncertainty Quantification of Underwater Sound Propagation Using Surrogate Models

Mikhail S. Lytaev(✉)

St. Petersburg Federal Research Center of the Russian Academy of Sciences, 14-th Linia, V.I., No. 39, Saint Petersburg 199178, Russia
mikelytaev@gmail.com

Abstract. Accurate prediction of underwater sound propagation is essential for sonar performance assessment, marine ecology studies, and ocean acoustic tomography. However, the inherent uncertainty in environmental parameters, particularly the sound speed profile (SSP), introduces significant variability in transmission loss predictions. In this work, we propose a Kriging-based surrogate modeling framework for uncertainty quantification (UQ) of underwater acoustic propagation. The sound speed variability is parameterized using empirical orthogonal functions (EOFs) derived from Argo float data collected in the Baltic Sea. A wide-angle parabolic equation solver, implemented within the PyWaveProp library, serves as the high-fidelity model, and Kriging surrogate models from the SMT library are employed to construct a computationally efficient approximation that maps uncertain EOF coefficients to transmission loss fields. Latin hypercube sampling is used for experimental design, and the surrogate accuracy is validated against direct Monte Carlo simulations. Numerical experiments demonstrate that the mean transmission loss field under SSP uncertainty can differ significantly from the deterministic prediction, while the qualitative structure of the field remains largely preserved even under large perturbation amplitudes.

Keywords: Underwater acoustics · Uncertainty quantification · Kriging · Gaussian process regression · Surrogate modeling · Parabolic equation method

1 Introduction

Underwater sound propagation modeling plays a central role in a wide range of applications, including sonar performance prediction, underwater communication, marine mammal impact assessment, and ocean acoustic tomography [13]. The acoustic field in the ocean is governed by the wave equation, which depends on environmental parameters such as the sound speed profile (SSP), water depth (bathymetry), and geoacoustic properties of the seabed. In practice, these parameters are known only approximately, introducing uncertainty into the predicted acoustic field.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026, LNCS 16767, pp. 51–66, 2027.
https://doi.org/10.1007/978-3-032-30491-9_4

Uncertainty quantification (UQ) aims to characterize the impact of input uncertainties on model outputs. The reference approach is Monte Carlo (MC) simulation: draw N independent realizations of the uncertain inputs, run the acoustic solver for each, and estimate output statistics from the ensemble. While MC converges at a rate $\mathcal{O}(N^{-1/2})$ regardless of input dimension, it requires $N = 10^3$–10^5 solver evaluations to converge–prohibitive when a single parabolic-equation (PE) solve takes seconds to minutes. This has motivated the development of surrogate-based UQ methods, which replace the expensive simulator with a fast-to-evaluate approximation [24].

Several surrogate modeling strategies have been explored for underwater acoustic UQ. Polynomial chaos expansions (PCE) represent the model output as a series of orthogonal polynomials in the uncertain inputs, enabling analytical computation of statistical moments and sensitivity indices [3,22,25]. Finette [9,17] pioneered the application of Karhunen–Loève decomposition combined with polynomial chaos to the stochastic acoustic waveguide problem, demonstrating rapid convergence for the mean field and intensity. Khazaie et al. [16] compared PCE, Kriging, and hybrid PC-Kriging surrogates for shallow-water acoustic UQ, finding that all three methods drastically reduce the required sample size compared to brute-force MC.

Dynamically orthogonal (DO) methods take an intrusive approach, reformulating the stochastic differential equation in a low-rank representation [1,2]. While achieving orders-of-magnitude speedup over MC, DO methods require modification of the acoustic solver code, limiting their applicability to legacy or black-box models. Moreover, existing DO formulations for underwater acoustics have been developed only for the narrow-angle PE, and their extension to wide-angle approximations remains an open problem.

Deep learning surrogates have also been applied to underwater acoustics. Du et al. [8] developed a physics-informed neural network (PINN) for acoustic field prediction by embedding the Helmholtz equation as a loss constraint. Pham and Sarkar [19] trained a neural network to estimate transmission loss uncertainty from a single baseline calculation. Indu Kant Deo et al. [15] combined neural encoders with a sparse variational Gaussian process output layer for three-dimensional TL prediction with calibrated uncertainty bands.

Kriging, also known as Gaussian process (GP) regression, occupies a favorable position among these approaches: it is non-intrusive, provides both a prediction and an associated predictive variance, and supports adaptive sampling strategies [23]. Despite these advantages, Kriging remains underexplored for underwater acoustic propagation problems, particularly in the context of SSP uncertainty parameterized by empirical orthogonal functions (EOFs).

In this paper, we present a Kriging-based surrogate modeling framework for UQ of underwater sound propagation. The high-fidelity model is a wide-angle PE solver implemented in the PyWaveProp library [20]. We propose a parameterization of the uncertain sound speed profile using EOF coefficients derived from Argo float data [14] in the Baltic Sea, which provides a compact, data-driven input space for the surrogate model. We systematically compare

four Kriging variants from the SMT (Surrogate Modeling Toolbox) library [6] and validate the surrogate predictions against direct adaptive MC simulations.

The remainder of the paper is organized as follows. Section 2 reviews the relevant background on underwater acoustic modeling and Kriging. Section 3 describes the proposed methodology. Section 4 presents numerical experiments and results. Section 5 concludes the paper.

2 Background

2.1 Underwater Acoustic Propagation

The acoustic pressure field $p(\mathbf{r}, t)$ in the ocean satisfies the Helmholtz equation in the frequency domain:

$$\nabla^2 p(\mathbf{r}) + k^2(\mathbf{r})\, p(\mathbf{r}) = -\delta(\mathbf{r} - \mathbf{r}_s),$$

where $k(\mathbf{r}) = \omega / c(\mathbf{r})$ is the wavenumber, ω is the angular frequency, $c(\mathbf{r})$ is the spatially varying sound speed, and $\mathbf{r}_s$ denotes the source position.

The parabolic equation (PE) method provides an efficient numerical solution for range-dependent environments by factoring the Helmholtz equation into a one-way marching equation [7]:

$$\frac{\partial \psi}{\partial r} = \mathrm{i} k_0 \left(\sqrt{1 + X} - 1 \right) \psi,$$

where ψ is the reduced pressure field, r is the range, k_0 is a reference wavenumber, and X is an operator involving depth derivatives and the index of refraction. Transmission loss (TL) is the standard metric for characterizing acoustic propagation:

$$\mathrm{TL}(r, z) = -20 \log_{10} \left| \frac{p(r, z)}{p_0} \right|,$$

where p_0 is a reference pressure.

2.2 Kriging (Gaussian Process Regression)

Kriging models the response $y = f(\mathbf{x})$ as a realization of a Gaussian process:

$$Y(\mathbf{x}) = \boldsymbol{\beta}^T \mathbf{f}(\mathbf{x}) + Z(\mathbf{x}),$$

where $\boldsymbol{\beta}^T \mathbf{f}(\mathbf{x})$ is a trend (regression) term and $Z(\mathbf{x})$ is a zero-mean GP with covariance

$$\mathrm{Cov}[Z(\mathbf{x}), Z(\mathbf{x}')] = \sigma^2\, R(\mathbf{x}, \mathbf{x}'; \boldsymbol{\theta}),$$

defined for any two points $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, where σ^2 is the process variance, R is a correlation function (e.g., Matérn or squared exponential), and $\boldsymbol{\theta}$ are hyperparameters estimated by maximum likelihood.

Given a training set $\{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^N$, let $\mathbf{R} \in \mathbb{R}^{N \times N}$ denote the correlation matrix with entries $R_{ij} = R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}; \boldsymbol{\theta})$, let $\mathbf{F} \in \mathbb{R}^{N \times p}$ be the regression matrix

with rows $\mathbf{f}(\mathbf{x}^{(i)})^T$, and let $\mathbf{r}(\mathbf{x}^*) = [R(\mathbf{x}^{(1)}, \mathbf{x}^*; \boldsymbol{\theta}), \ldots, R(\mathbf{x}^{(N)}, \mathbf{x}^*; \boldsymbol{\theta})]^T$ be the vector of cross-correlations between the training points and a new point $\mathbf{x}^*$. The Kriging predictor is then:

$$\hat{y}(\mathbf{x}^*) = \mathbf{f}(\mathbf{x}^*)^T \hat{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x}^*)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}}),$$

with an associated prediction variance $\hat{\sigma}^2(\mathbf{x}^*)$ that quantifies surrogate uncertainty.

The hyperparameters $\boldsymbol{\theta}$ and the process variance σ^2 are estimated by maximizing the concentrated log-likelihood function [23]:

$$\ln L(\boldsymbol{\theta}) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\hat{\sigma}^2) - \frac{1}{2}\ln|\mathbf{R}|,$$

where $\hat{\sigma}^2 = \frac{1}{N}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}})$ is the maximum likelihood estimate of the process variance and $|\mathbf{R}|$ is the determinant of the correlation matrix. Equivalently, the hyperparameters are found by minimizing the negative log-likelihood $-\ln L(\boldsymbol{\theta})$. This optimization problem is typically non-convex and is solved using gradient-based methods with multiple restarts.

In this work, we compare four Kriging variants from the SMT library [6], which differ primarily in how they handle the correlation matrix and the hyperparameter optimization:

- **KRG** (full Kriging) estimates a separate length-scale parameter θ_j for each input dimension $j = 1, \ldots, d$, so that the correlation function takes the product form $R(\mathbf{x}_i, \mathbf{x}_k; \boldsymbol{\theta}) = \prod_{j=1}^{d} r(|x_{ij} - x_{kj}|; \theta_j)$, where $r(\cdot\,; \theta_j)$ is a one-dimensional kernel (e.g., squared exponential or Matérn). The number of hyperparameters scales linearly with d, making the likelihood optimization increasingly expensive in high dimensions.
- **KPLS** [5] (Kriging with partial least squares) addresses the curse of dimensionality by first applying PLS regression to identify a small number $h \ll d$ of latent components $\mathbf{t}_l = \mathbf{X}\mathbf{w}_l$ $(l = 1, \ldots, h)$ that maximize the covariance between inputs and outputs. The correlation function is then parameterized in the reduced space, requiring only h hyperparameters instead of d. This dramatically reduces the cost of likelihood optimization while preserving accuracy when the response varies primarily along a low-dimensional subspace.
- **KPLSK** [4] is a two-stage hybrid: it first fits a KPLS model to obtain an inexpensive initial estimate of the hyperparameters, then uses these as starting values for a full Kriging optimization in the original d-dimensional space. This combines the fast convergence of KPLS with the flexibility of KRG, often achieving accuracy close to full Kriging at a fraction of the cost.
- **GPX** is a Kriging implementation based on the `egobox` library [18], a Gaussian process toolbox written in Rust. It uses the same mathematical formulation as KPLS but benefits from optimized linear algebra routines, yielding significantly faster training and prediction times.

3 Methodology

3.1 Problem Formulation

We consider a set of uncertain environmental parameters $\mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{X} \subset \mathbb{R}^d$ with a known joint probability distribution $\pi(\mathbf{x})$. The quantity of interest (QoI) is the transmission loss field $\mathrm{TL}(r, z; \mathbf{x})$ computed by the PE solver at specified receiver locations.

The goal of UQ is to estimate statistical moments and probability distributions of the QoI:

$$\mathbb{E}[\mathrm{TL}] = \int_{\mathcal{X}} \mathrm{TL}(\mathbf{x})\, \pi(\mathbf{x})\, d\mathbf{x}, \quad \mathrm{SD}[\mathrm{TL}] = \sqrt{\mathbb{E}[\mathrm{TL}^2] - (\mathbb{E}[\mathrm{TL}])^2}.$$

3.2 EOF Parameterization of Sound Speed Uncertainty

Empirical orthogonal function (EOF) analysis [11,21], also known as principal component analysis (PCA), provides a data-driven decomposition of a spatially distributed random field into orthogonal modes ordered by explained variance. It has been widely used in physical oceanography to characterize the variability of temperature, salinity, and sound speed profiles.

Let $\{c_k(z)\}_{k=1}^K$ denote a set of K observed sound speed profiles, each sampled (or interpolated) onto a common depth grid $\{z_1, \ldots, z_J\}$. Define the ensemble mean profile

$$\bar{c}(z_j) = \frac{1}{K} \sum_{k=1}^{K} c_k(z_j),$$

and the anomaly matrix $\mathbf{A} \in \mathbb{R}^{K \times J}$ with entries $A_{kj} = c_k(z_j) - \bar{c}(z_j)$. The spatial covariance matrix is

$$\mathbf{C} = \frac{1}{K-1} \mathbf{A}^T \mathbf{A} \in \mathbb{R}^{J \times J}.$$

Since $\mathbf{C}$ is symmetric positive semi-definite, it admits an eigendecomposition $\mathbf{C} = \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^T$, where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_J]$ is an orthogonal matrix of eigenvectors and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_J)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_J \geq 0$. The eigenvectors $\mathbf{u}_i$ are the EOF modes (spatial basis functions) and the eigenvalues λ_i represent the variance captured by each mode.

In practice, the EOF modes are computed via the singular value decomposition (SVD) of the anomaly matrix, $\mathbf{A}^T = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T$, which is numerically more stable than forming $\mathbf{C}$ explicitly [10]. The columns of $\mathbf{U}$ are the EOF modes and the singular values satisfy $\sigma_i^2/(K-1) = \lambda_i$. The principal component (PC) coefficients, representing the projection of each profile onto the EOF basis, are given by $\alpha_{ki} = \sum_j A_{kj}\, u_{ij}$. The fraction of total variance explained by the first M modes is

$$\rho_M = \frac{\sum_{i=1}^{M} \lambda_i}{\sum_{i=1}^{J} \lambda_i}.$$

For UQ, we use the truncated EOF expansion to parameterize the sound speed perturbation. The perturbed profile is expressed as:

$$c(z) = \bar{c}(z) + \sum_{i=1}^{M} \xi_i \varphi_i(z),$$

where $\bar{c}(z)$ is the mean profile, $\varphi_i(z)$ are the EOF modes (discretized as $\mathbf{u}_i$), and $\xi_i \sim \mathcal{N}(0, \sigma^2_{\mathrm{EOF}})$ are independent Gaussian random coefficients. The truncation order M is chosen so that ρ_M is close to unity; the rapid spectral decay of the EOF eigenvalues (the first five modes capture 96.1% of the total variance in the Baltic Sea dataset) justifies the use of a low-dimensional parameterization. The EOF basis is optimal in the L^2 sense: no other linear decomposition of the same rank captures more variance [21].

3.3 Surrogate Construction

It is important to note that the surrogate models employed in this work are real-valued: a separate scalar Kriging model is trained independently for each spatial field point (r_j, z_j) of interest, mapping the EOF coefficients $\boldsymbol{\xi}$ to the transmission loss $\mathrm{TL}(r_j, z_j; \boldsymbol{\xi})$ at that location. The surrogate modeling workflow consists of the following steps:

1. **Experimental design.** Generate N sample points $\{\mathbf{x}^{(i)}\}_{i=1}^{N}$ in the input space using Latin hypercube sampling (LHS) [12].
2. **Forward model evaluation.** Run the PE solver at each sample point to obtain the training responses $\{y^{(i)}\}_{i=1}^{N}$.
3. **Kriging fitting.** Train the Kriging model by estimating the trend coefficients $\hat{\boldsymbol{\beta}}$ and the correlation hyperparameters $\hat{\boldsymbol{\theta}}$ via maximum likelihood estimation. We compare four Kriging variants from the SMT library [6]: full Kriging (KRG), Kriging with partial least squares (KPLS), KPLS with additional Kriging refinement (KPLSK), and GPX.
4. **Validation.** Assess surrogate accuracy using RMSE against direct Monte Carlo results.

3.4 Uncertainty Propagation

Once the surrogate is trained and validated, uncertainty propagation is performed by evaluating the Kriging predictor on a large Monte Carlo sample $\{\mathbf{x}^{(j)}\}_{j=1}^{N_{\mathrm{MC}}}$ drawn from $\pi(\mathbf{x})$, where $N_{\mathrm{MC}} \gg N$. This yields empirical estimates of the mean, variance, and full probability distribution of the QoI at negligible computational cost.

4 Numerical Experiments

4.1 Problem Setup

We consider a shallow-water underwater acoustic environment at a frequency of $f = 100$ Hz. The source is located at a depth of $z_s = 20$ m with a beam

width of 2°. The water column depth is $H = 100$ m. The computational domain extends to a range of $r_{\max} = 10$ km and a depth of $z_{\max} = 150$ m (including a 50 m sediment layer). The wide-angle parabolic equation is solved using the PyWaveProp library [20] with a Padé $(2, 3)$ rational approximation, using range and depth steps of $\Delta r = 20$ m and $\Delta z = 1$ m, respectively. The seabed is modeled as a homogeneous half-space with a sound speed of 1700 m/s, density of 1.5 g/cm^3, and attenuation of 1.0 dB/λ.

The sound speed variability is modeled using EOFs derived from Argo float data collected in the Baltic Sea. The mean sound speed profile ranges from approximately 1424 m/s (near the thermocline minimum at ∼57 m depth) to 1468 m/s (near the surface). A total of 82 EOF modes were extracted; however, the variance is strongly concentrated in the leading modes: mode 1 alone explains 70.5% of the total variability, the first two modes capture 88.2%, and the first five modes account for 96.1%. This rapid spectral decay justifies the use of a low-dimensional EOF parameterization for UQ.

Figure 1 illustrates the effect of varying the number of EOF modes M and the perturbation standard deviation (SD)σ on the resulting sound speed profile ensemble. The 3σ spread of the SSP ensemble grows both with M and σ: for $M = 1$, $\sigma = 2.5$, the total spread is only 43.8 m/s (barely exceeding the mean profile range), whereas for $M = 5$, $\sigma = 20$, it reaches 79.1 m/s. The variability is concentrated in the thermocline region (∼9–30 m depth), where the SD peaks at ∼3.5 m/s for $M = 3$, $\sigma = 10$, while below ∼65 m the variability drops below 0.7 m/s. Notably, the EOF modes vanish at the surface and bottom boundaries, so the SSP uncertainty is exactly zero at these extremes regardless of the perturbation amplitude.

In further experiments, the uncertain input parameters are the first $M = 3$ EOF coefficients $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$, each drawn independently from a zero-mean Gaussian distribution $\xi_i \sim \mathcal{N}(0, \sigma_{\text{EOF}}^2)$. The perturbed sound speed profile is then $c(z) = \bar{c}(z) + \sum_{i=1}^{M} \xi_i \varphi_i(z)$, where $\bar{c}(z)$ is the mean profile and $\varphi_i(z)$ are the EOF modes. Four levels of SSP uncertainty are considered: $\sigma_{\text{EOF}} = 2.5$, 5, 10, and 20 m/s, spanning mild to extreme environmental variability.

4.2 Monte Carlo Uncertainty Propagation

To provide a ground-truth reference for the surrogate-based UQ, we performed direct (full-model) adaptive Monte Carlo simulations for four levels of SSP uncertainty ($\sigma_{\text{EOF}} = 2.5$, 5, 10, 20 m/s). The adaptive algorithm draws samples in batches of 100, checks convergence of the running mean and SD against a 1% relative tolerance after a minimum of 500 samples, and terminates early if the criterion is met (maximum: 50,000 samples). All four runs converged, requiring 3,000 samples for $\sigma_{\text{EOF}} = 2.5$ and 4,200 samples for the remaining cases. Table 1 summarizes the results at a receiver depth of $z_r = 20$ m.

Several trends are evident. First, the spatially averaged mean TL is remarkably stable across perturbation levels, shifting by only 0.23 dB from $\sigma_{\text{EOF}} = 2.5$ (−54.46 dB) to $\sigma_{\text{EOF}} = 20$ (−54.23 dB). This indicates that the ensemble-averaged propagation loss is insensitive to SSP perturbation amplitude at the

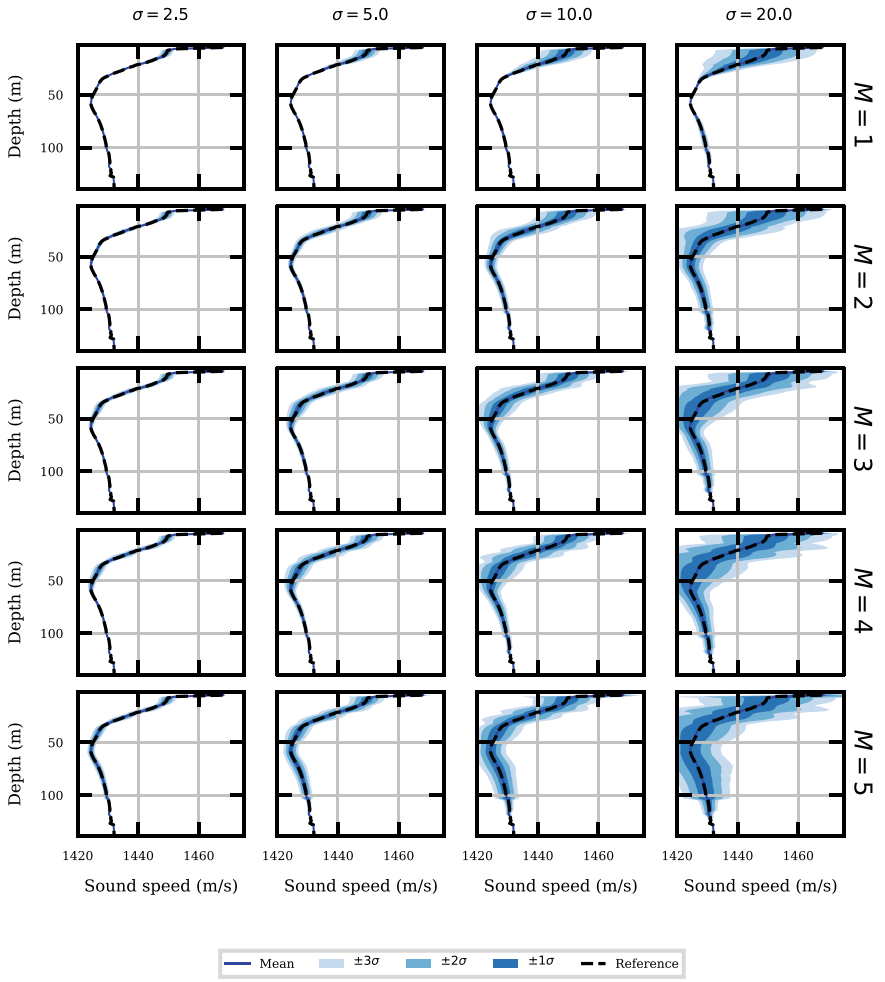

Fig. 1. Sound speed profile ensembles for different numbers of EOF modes M and perturbation SDs σ. Shaded regions indicate $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$ intervals.

global level. However, the pointwise mean TL can differ from the deterministic reference by several decibels (see Fig. 2, bottom panel), which is significant for applications that rely on point-by-point comparisons between measurements and model predictions.

Second, the spatial variability of the mean TL (captured by its spatial SD) decreases monotonically from ± 4.73 dB to ± 2.13 dB as σ_{EOF} increases–the interference pattern is progressively "washed out" by averaging over increasingly diverse SSP realizations.

Third, the pointwise TL uncertainty (SD TL) grows sub-linearly with input uncertainty: an eightfold increase in σ_{EOF} (from 2.5 to 20 m/s) yields only a

Table 1. Direct adaptive Monte Carlo results for different SSP perturbation levels. Mean TL and SD TL are spatial statistics (mean ± spatial SD) of the pointwise mean and SD of the transmission loss field along range at $z_r = 20$ m.

σ_{EOF} (m/s)	Samples	Time (s)	Mean TL (dB)	SD TL (dB)	Max SD TL (dB)
2.5	3 000	137	-54.46 ± 4.73	1.67 ± 1.63	7.90
5	4 200	236	-54.43 ± 4.06	2.76 ± 1.89	8.26
10	4 200	241	-54.36 ± 3.07	3.85 ± 1.70	7.51
20	4 200	235	-54.23 ± 2.13	4.55 ± 1.19	7.83

2.7× increase in mean SD TL (from 1.67 to 4.55 dB), suggesting saturation due to nonlinear mode coupling at large perturbations.

Fourth, the maximum SD TL remains in the range 7.5–8.3 dB regardless of σ_{EOF}, indicating localized sensitivity hotspots, likely near interference nulls and shadow-zone boundaries, where the TL is highly sensitive to SSP changes at any perturbation level.

Figure 2 compares the mean transmission loss, its SD, and the pointwise difference from the deterministic reference for all four perturbation levels. Figure 3 shows the individual $\pm 1\sigma$ uncertainty bands.

4.3 Surrogate Comparison

Four surrogate methods from the SMT library [6] were compared: full Kriging (KRG), Kriging with partial least squares (KPLS), KPLS with additional Kriging refinement (KPLSK), and GPX. Each method was evaluated with $N = 10, 30, 100$, and 300 training samples generated via Latin hypercube sampling [12]. The experiments were repeated for three levels of EOF perturbation SD ($\sigma_{\text{EOF}} = 2.5$, 5, 10), representing increasing environmental uncertainty. Tables 3, 4 and 5 report the training-plus-prediction time per field point (in seconds), the mean RMSE, the SD RMSE (total SD combining aleatory and epistemic uncertainty, compared against direct Monte Carlo), and the epistemic RMSE (uncertainty due to the surrogate model itself). A detailed description of each metric is provided in Table 2.

Several trends emerge from the surrogate comparison. First, full Kriging (KRG) consistently achieves the lowest mean RMSE at every noise level when a sufficient number of training samples is available. This confirms that the full correlation structure is beneficial for this problem. However, KRG is also the most expensive method per field point: at $N = 300$ it requires 6.6–7.6 s/pt, roughly 12–14× slower than GPX at the same sample size.

Second, GPX offers a dramatic speedup, 0.03–0.55 s/pt across all configurations, but its accuracy plateaus or even degrades at $N = 300$. This non-monotonic behavior suggests overfitting or hyperparameter instability in the GPX implementation for this problem class.

Third, KPLSK provides a practical middle ground. Its accuracy at $N = 100$ is close to that of full KRG, while requiring roughly half the time per point. For

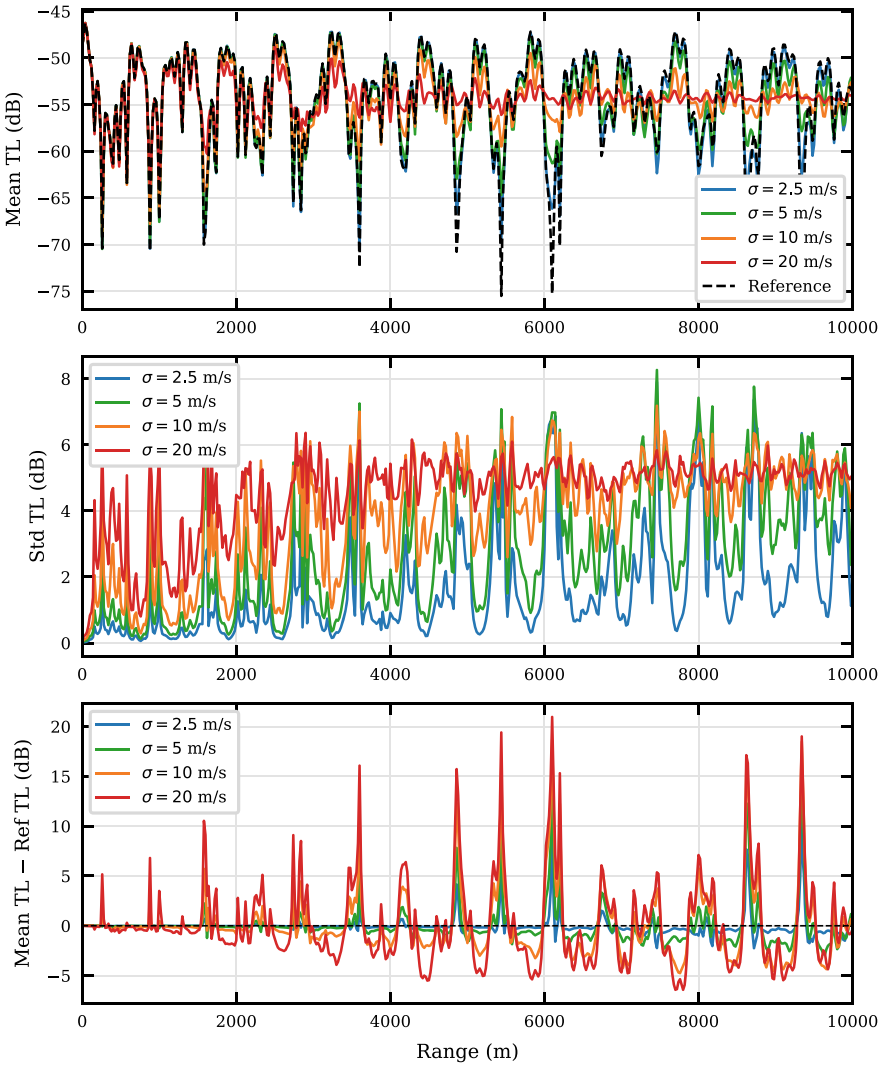

Fig. 2. Direct Monte Carlo results for $\sigma_{\mathrm{EOF}} = 2.5$, 5, 10, and 20 m/s: mean transmission loss (top), SD (middle), and pointwise difference from the deterministic reference (bottom).

moderate uncertainty ($\sigma_{\mathrm{EOF}} \leq 5$) and $N \geq 100$, KPLSK achieves sub-0.32 dB mean RMSE, which is adequate for most practical UQ applications.

The decomposition of prediction uncertainty into total and epistemic components reveals additional insights. The mean epistemic SD (reflecting surrogate model uncertainty) decreases with increasing N for KRG and KPLSK, confirming that these models become more confident as training data grows. In contrast, GPX exhibits elevated epistemic uncertainty at intermediate sample sizes, con-

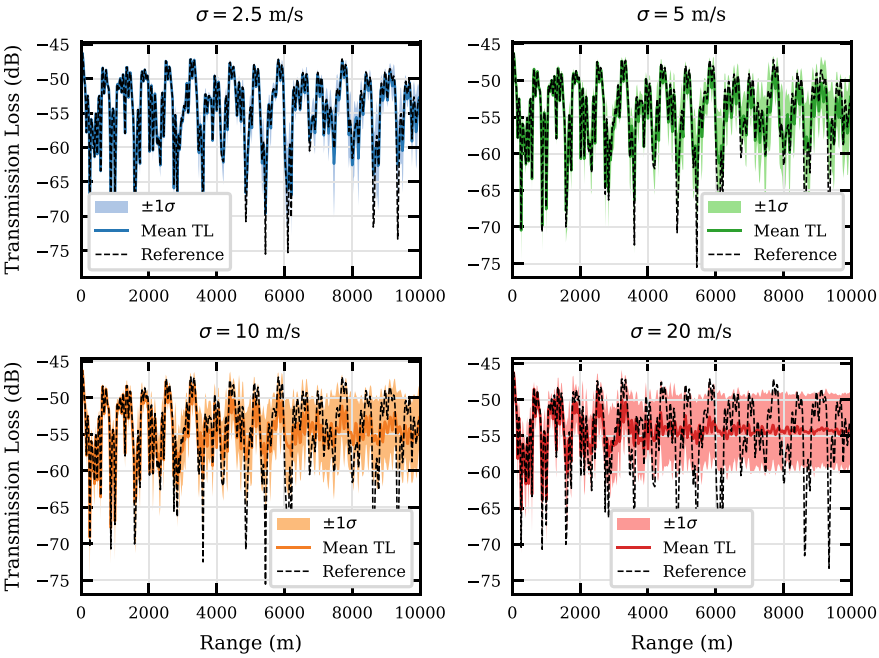

Fig. 3. Transmission loss with $\pm 1\sigma$ uncertainty bands for each EOF perturbation level ($\sigma_{\mathrm{EOF}} = 2.5$, 5, 10, 20 m/s). Dashed line shows the deterministic reference.

Table 2. Description of metrics reported in Tables 3, 4 and 5.

Column	Description
Method	Surrogate model type: KRG (full Kriging), KPLS (Kriging with partial least squares), KPLSK (KPLS with Kriging refinement), or GPX
N	Number of Latin hypercube training samples
Time/pt (s)	Total wall-clock time (training + prediction) per spatial field point, in seconds
Mean RMSE	Root mean square error of the surrogate-predicted mean TL field against the direct Monte Carlo mean, in dB
SD RMSE	Root mean square error of the total (aleatory + epistemic) SD of TL against the direct Monte Carlo SD, in dB. Measures how well the surrogate captures the overall uncertainty, combining both the physical variability due to SSP uncertainty and the surrogate model uncertainty
Mean Epist. SD	Spatial mean of the epistemic (surrogate model) SD of TL, in dB. Quantifies the surrogate's own prediction uncertainty arising from limited training data

Table 3. Surrogate accuracy comparison for $\sigma_{\mathrm{EOF}} = 2.5$.

Method	N	Time/pt (s)	Mean RMSE	SD RMSE	Mean Epist. SD
KRG	10	0.5017	0.7187	0.7641	0.8730
	30	1.0780	0.2488	0.3396	0.7116
	100	1.7031	0.0749	0.1466	0.5104
	300	7.5733	0.1153	0.0764	0.5104
KPLS	10	0.1026	1.2516	0.7761	1.6167
	30	0.2452	0.4188	0.4150	1.2701
	100	0.7021	0.2231	0.3145	1.0868
	300	3.3923	0.2730	0.6042	0.5790
KPLSK	10	0.1506	0.9737	0.7319	1.1323
	30	0.3549	0.2930	0.3341	0.9599
	100	0.8787	0.1140	0.1648	0.5669
	300	3.3725	0.4032	0.5363	0.4989
GPX	10	0.0359	0.7363	0.7511	0.7751
	30	0.0482	0.6519	0.5870	0.8113
	100	0.1187	0.1096	0.2821	0.6822
	300	0.5314	0.4469	0.7200	0.2422

Table 4. Surrogate accuracy comparison for $\sigma_{\mathrm{EOF}} = 5$.

Method	N	Time/pt (s)	Mean RMSE	SD RMSE	Mean Epist. SD
KRG	10	0.4176	1.4256	1.1377	1.4842
	30	0.7153	1.0140	0.7579	1.4319
	100	1.2103	0.2447	0.3208	0.8110
	300	6.8701	0.1341	0.2995	0.5848
KPLS	10	0.0786	1.1715	1.1384	2.4355
	30	0.1437	1.5743	0.8073	2.0606
	100	0.3319	0.4893	0.8173	1.3021
	300	2.2528	0.3508	0.6319	0.9131
KPLSK	10	0.1259	1.1194	1.1541	2.0379
	30	0.2275	1.1781	0.7762	1.4594
	100	0.5135	0.3130	0.6992	0.7605
	300	3.2145	0.2562	0.5584	0.5669
GPX	10	0.0293	1.4840	1.1657	1.3222
	30	0.0421	1.0869	0.6629	2.0065
	100	0.1196	0.2980	0.4665	1.4636
	300	0.5474	0.5254	1.1012	0.6254

Table 5. Surrogate accuracy comparison for $\sigma_{\mathrm{EOF}} = 10$.

Method	N	Time/pt (s)	Mean RMSE	SD RMSE	Mean Epist. SD
KRG	10	0.3310	2.4563	1.9933	1.7990
	30	0.6102	0.8858	0.9423	2.8066
	100	1.1901	0.6842	0.7071	2.1414
	300	6.6497	0.1741	0.3467	1.3658
KPLS	10	0.0645	2.0116	1.7691	3.5809
	30	0.1090	1.1218	1.0079	3.5206
	100	0.2613	0.8275	1.1308	2.7229
	300	1.7783	0.4117	1.6946	1.4364
KPLSK	10	0.1097	2.4120	2.0413	2.0813
	30	0.1908	0.8856	0.9491	2.8347
	100	0.4355	0.6784	0.8009	2.1666
	300	2.5251	0.3153	1.6476	1.0544
GPX	10	0.0285	1.9631	1.9193	2.4935
	30	0.0416	1.1036	1.0867	3.0295
	100	0.1094	0.8129	0.7652	2.6499
	300	0.5407	1.3705	2.4934	0.5575

sistent with its non-monotonic accuracy behavior. The SD RMSE (combining aleatory and epistemic uncertainty, compared against direct Monte Carlo) generally decreases with N, indicating improved overall uncertainty estimation as training data grows. At large N, the epistemic component diminishes and the remaining error is dominated by the aleatory contribution, reflecting the shift from surrogate inadequacy to irreducible input variability.

Finally, increasing σ_{EOF} from 2.5 to 10 roughly triples the mean RMSE across all methods and sample sizes, reflecting the greater complexity of the input–output mapping when the SSP variability spans a wider range. This underscores the importance of choosing an EOF perturbation scale that is consistent with the actual oceanographic uncertainty.

Since each surrogate is a scalar model trained independently per spatial point, the computational advantage of the surrogate approach depends critically on the number of field points at which uncertainty estimates are required. When UQ is needed at a small number of receiver locations–for example, at a fixed sensor array or along a single transect–the one-time cost of training a handful of Kriging models is amortized over a large number of surrogate evaluations, yielding substantial speedups over direct Monte Carlo. However, when the full two-dimensional (or three-dimensional) transmission loss field must be characterized, a separate surrogate must be trained for every grid point, and the aggregate training cost can exceed that of brute-force Monte Carlo, which naturally pro-

duces the entire field in each realization. In such scenarios, direct Monte Carlo may be the more practical choice.

5 Conclusion

We presented a Kriging-based surrogate modeling framework for UQ of underwater sound propagation. The sound speed variability was parameterized using EOFs derived from Argo float data collected in the Baltic Sea; the first five modes capture 96.1% of the observed variance, providing a compact and physically meaningful input space for the surrogate. The wide-angle parabolic equation solver from the PyWaveProp library was used as the high-fidelity forward model, and Kriging implementations from the SMT library served as surrogate approximations.

A systematic comparison of four surrogate methods (KRG, KPLS, KPLSK, GPX) revealed that full Kriging achieves the best accuracy, while KPLSK offers a favorable accuracy–cost trade-off for moderate uncertainty levels. GPX provides orders-of-magnitude speedup per field point but exhibits accuracy limitations for this problem class. The decomposition of prediction uncertainty into aleatory and epistemic components showed that the total (aleatory + epistemic) SD provides a more complete comparison against direct Monte Carlo, and that as the surrogate improves with more training data, the dominant error source shifts from model inadequacy to irreducible input variability. Increasing the EOF perturbation scale from $\sigma_{\mathrm{EOF}} = 2.5$ to 10 roughly triples the surrogate error, highlighting the need for sufficient training data under high environmental uncertainty.

Direct adaptive Monte Carlo simulations with 3,000–4,200 full-model evaluations revealed two key findings. First, the mean transmission loss field under SSP uncertainty can differ significantly from the deterministic prediction on a point-by-point basis (up to several decibels), even though the spatially averaged TL remains stable. This has important practical implications: when comparing simulation results with field measurements, one should account for the bias introduced by environmental uncertainty rather than relying on a single deterministic run. Second, the qualitative structure of the transmission loss field (the overall pattern of convergence and shadow zones) is remarkably robust to SSP uncertainty. The TL uncertainty grows sub-linearly with SSP perturbation amplitude, while localized sensitivity hotspots ($\sim$8 dB) persist regardless of the perturbation level.

Future work will explore several directions. First, vector-valued surrogate models that jointly predict the transmission loss at multiple spatial points could share information across locations, improving accuracy and eliminating the need to train a separate model per field point–thereby making surrogate-based UQ competitive even for full two-dimensional field characterization. Second, adaptive sampling strategies (e.g., expected-improvement or variance-reduction criteria) would allow the training set size to be determined automatically rather than fixed a priori, reducing the number of required forward model evaluations

while ensuring a target surrogate accuracy. Additional extensions include higher-dimensional input spaces incorporating bathymetry and seabed parameters, and validation against measured oceanographic data.

Acknowledgments. This study was supported by the Russian Science Foundation grant No. 25-71-00028.

References

1. Ali, W.H.: Dynamically orthogonal equations for stochastic underwater sound propagation. Ph.D. thesis, Massachusetts Institute of Technology (2019)
2. Ali, W.H., Lermusiaux, P.F.: Dynamically orthogonal narrow-angle parabolic equations for stochastic underwater sound propagation. Part II: applications. J. Acoust. Soc. Am. **155**(1), 656–672 (2024)
3. Blatman, G., Sudret, B.: Adaptive sparse polynomial chaos expansion based on least angle regression. J. Comput. Phys. **230**(6), 2345–2367 (2011)
4. Bouhlel, M.A., Bartoli, N., Otsmane, A., Morlier, J.: An improved approach for estimating the hyperparameters of the kriging model for high-dimensional problems through the partial least squares method. Math. Probl. Eng. **2016**(1), 6723410 (2016)
5. Bouhlel, M.A., Bartoli, N., Otsmane, A., Morlier, J.: Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction. Struct. Multidiscip. Optim. **53**(5), 935–952 (2016)
6. Bouhlel, M.A., Hwang, J.T., Bartoli, N., Lafage, R., Morlier, J., Martins, J.R.: A python surrogate modeling framework with derivatives. Adv. Eng. Softw. **135**, 102662 (2019)
7. Collins, M.D.: A split-step padé solution for the parabolic equation method. J. Acoust. Soc. Am. **93**(4), 1736–1742 (1993)
8. Du, L., Wang, Z., Lv, Z., Wang, L., Han, D.: Research on underwater acoustic field prediction method based on physics-informed neural network. Front. Mar. Sci. **10**, 1302077 (2023)
9. Finette, S.: A stochastic response surface formulation of acoustic propagation through an uncertain ocean waveguide environment. J. Acoust. Soc. Am. **126**(5), 2242–2247 (2009)
10. Golub, G.H., Van Loan, C.F.: Matrix Computations. JHU Press (2013)
11. Hannachi, A., Jolliffe, I.T., Stephenson, D.B., et al.: Empirical orthogonal functions and related techniques in atmospheric science: a review. Int. J. Climatol. **27**(9), 1119–1152 (2007)
12. Helton, J.C., Davis, F.J.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Saf. **81**(1), 23–69 (2003)
13. Jensen, F.B., Kuperman, W.A., Porter, M.B., Schmidt, H., Tolstoy, A.: Computational Ocean Acoustics, vol. 2011. Springer, New York, NY (2011). https://doi.org/10.1007/978-1-4419-8678-8
14. Johnson, G.C., et al.: Argo–two decades: global oceanography, revolutionized. Ann. Rev. Mar. Sci. **14**(1), 379–403 (2022)
15. Kant Deo, I., Venkateshwaran, A., Jaiman, R.K.: A physics-guided probabilistic surrogate modeling framework for digital twins of underwater radiated noise. arXiv e-prints arXiv:2509.25730 (2025)

16. Khazaie, S., Wang, X., Komatitsch, D., Sagaut, P.: Uncertainty quantification for acoustic wave propagation in a shallow water environment. Wave Motion **91**, 102390 (2019)
17. Khine, Y.Y., Creamer, D.B., Finette, S.: Acoustic propagation in an uncertain waveguide environment using stochastic basis expansions. J. Comput. Acoust. **18**(04), 397–441 (2010)
18. Lafage, R.: Egobox, a Rust toolbox for efficient global optimization. J. Open Sour. Softw. **7**(78), 4737 (2022)
19. Lee, B.M., Johnson, J.R., Dowling, D.R.: Predicting acoustic transmission loss uncertainty in ocean environments with neural networks. J. Mar. Sci. Eng. **10**(10), 1548 (2022)
20. Lytaev, M.S.: Rational interpolation of the one-way helmholtz propagator. J. Comput. Sci. **58**, 101536 (2022)
21. Preisendorfer, R.W.: Principal Component Analysis in Meteorology and Oceanography. Elsevier, Amsterdam (1988)
22. Sudret, B.: Global sensitivity analysis using polynomial chaos expansions. Reliab. Eng. Syst. Saf. **93**(7), 964–979 (2008)
23. Williams, C.K., Rasmussen, C.E.: Gaussian Processes for Machine Learning, vol. 2. MIT Press, Cambridge, MA (2006)
24. Xiu, D.: Numerical Methods for Stochastic Computations: A Spectral Method Approach. Princeton University Press (2010)
25. Xiu, D., Karniadakis, G.E.: The wiener-askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput. **24**(2), 619–644 (2002)