Prediction of Acoustic Communication Performance in Marine Robots Using Model-Based Kriging

George P. Kontoudis and Daniel J. Stilwell

***Abstract*— In this paper, we present a data-driven iterative algorithm for accurate prediction of underwater acoustic communication performance at unvisited sites. The prediction algorithm consists of two steps: i) estimation of the covariance matrix; and ii) prediction of the communication performance. The importance of the covariance estimation is highlighted with a multi-stage, model-based iterative methodology that produces unbiased and robust results. The efficiency of the framework has been validated with synthetic data.**

I. Introduction

Coordination of multiple autonomous underwater agents requires effective communication for various cooperative missions [1]. For agents that operate underwater, inter-vehicle communication is usually accomplished using underwater acoustic (UWA) signals. In the majority of the literature, the communication links—interconnecting multiple underwater vehicles—are treated as deterministic, range-dependent functions [2], [3]. Indeed, the *communication performance* is strongly tied to the *vehicle range*, yet the performance is also dependent on environmental effects, including multi-path effect and background noise [4]. In addition to exchange of data, acoustic communication can provide vehicle range information to improve underwater navigation [5].

Our goal in this work is on predicting the UWA communication performance at unvisited locations by using previous measurements, even if the locations are beyond the observation area, i.e. extrapolation. We employ a two-step methodology that is comprised of: i) the estimation of covariance function and its parameters; and ii) the prediction of the communication performance and its corresponding variance. Accurate predictions of anticipated communication performance can be exploited to maintain connectivity in a network of mobile agents and plan better utilization of communication resources. Applications with multiple mobile nodes, such as autonomous underwater vehicles (AUVs), include underwater exploration, mine reconnaissance, and tactical surveillance. Our general approach may be applicable to terrestrial networks, such as aerial and ground communication using WiFi. The main idea is to leverage advances in spatial statistics and UWA communication modeling, to provide a realistic statistical prediction of inter-vehicle communication performance for teams of marine robots.

This work was supported by the Office of Naval Research via grants N00014-18-1-2627 and N00014-19-1-2194.

G. P. Kontoudis and D. J. Stilwell are with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061, USA, email: {gpkont,stilwell}@vt.edu.

Related work: In underwater wireless sensor networks *kriging* [6] (equivalent to Gaussian processes [7], [8]) has been used for several applications. Horner *et al.* [9], proposed a methodology, based partially on ordinary kriging to generate local and global acoustic communication performance maps. A distributed ordinary kriging methodology was used in [10] to estimate coverage holes in large-scale wireless sensor networks. In [11], the acoustic communication performance of micro AUVs was assessed with field trials. In [12], a methodology that combines ordinary kriging and compressive sensing methods, was utilized for prediction of acoustic intensity. Prediction of received signal strength has been used in wireless communications for radio propagation channels. In [13], the authors employ a maximum-likelihood technique to estimate the parameters of the covariance function, logarithmic transformation to estimate the underlying mean, and compressive sensing for prediction with sparse data. The authors in [14], proposed an ordinary kriging prediction framework with detrended data to build radio environment maps, and they also considered positional error for the measurements. In [15], we formulated the UWA communication performance problem using multivariate ordinary kriging. However, in these works it is assumed that the latent process is stationary leading to the formulation of *ordinary kriging*, that ignores model-based parameters. In addition, kriging has been used in wireless communications for the statistical modeling of the environment, yet without a selection method for the the covariance function.

Contributions: The contribution of this paper is twofold. First, we formulate the UWA communication performance problem as a non-stationary random field and propose model-based basis functions to recover stationarity. Second, we introduce an iterative technique with statistical selection of the covariance function and robust estimation of the covariance parameters to identify the most suitable model of the underlying latent process.

Structure: Section II formulates the UWA channel problem, Section III discusses the parameter estimation of the covariance and the spatial prediction, Section IV describes the overall framework and the algorithm, Section V provides simulations and results, and Section VI concludes the paper.

II. Underwater Acoustic Channel

A. Foundations

Basic notions of random fields are discussed in [16]. A *random field* is a stochastic process indexed in the Euclidean space. Let $Z(\mathbf{x})$ be a random field with a covariance matrix $\mathrm{Cov}[Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})] \succ 0$ for all $\mathbf{x}, \mathbf{x}+\mathbf{h} \in \mathcal{D} \subset \mathbb{R}^m$,

where $\mathbf{x} \in \mathbb{R}^m$ denote the spatial coordinates, $\mathbf{h}$ is the separation vector, and m is the dimension of the coordinates, e.g. $m = 2$ for planar coordinates. The *variogram* is a statistical measure of spatial autocorrelation, $2\gamma(\mathbf{h}) := \mathrm{E}[(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2]$, where $\gamma(\mathbf{h})$ is the *semivariogram*. The random field is *intrinsically (strongly) stationary* if both $\mathrm{E}[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 0$ and $\mathrm{Var}[Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})] = 2\gamma(\mathbf{h})$ for all $\mathbf{x}$, $\mathbf{x}+\mathbf{h} \in \mathcal{D}$ are satisfied. An intrinsically stationary random field with a constant mean $\mathrm{E}[Z(\mathbf{x})] = \mu$ and $\mathrm{Cov}[Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})] = C(\mathbf{h})$ is called *second-order (weakly) stationary*, where the function $C(\cdot)$ is the *covariance function*. Second-order stationarity implies the Gaussian assumption. When the semivariogram function is bounded, then the covariance function can be constructed by $C(\mathbf{h}) = \gamma(\infty) - \gamma(\mathbf{h})$. Note that $C(\cdot)$ is a *stationary covariance function*, depending only on the separation vector $\mathbf{h}$ and not on local position $\mathbf{x}$. A bounded semivariogram implies $\gamma(\infty) = \sup_{\mathbf{h}} \gamma(\mathbf{h}) = \sigma^2 + \tau^2 < \infty$ is non-negative and represents the *sill* of the semivariogram with σ^2 the *partial sill* and τ^2 the *nugget*. The nugget represents the variance of the measurement error. When the variogram depends only on the displacement vector norm, i.e. $2\gamma(\mathbf{h}) = 2\gamma(\|\mathbf{h}\|)$, then the variogram is *isotropic*.

B. Problem Formulation

We consider the problem of UWA communication of two vehicles. In Fig. 1, we illustrate two cases of UWA communication between two vehicles at range r, with $\mathbf{x}_\mathrm{t}$ the position of the *transmitting* vehicle and $\mathbf{x}_\mathrm{r}$ the position of the *receiving* vehicle. The first case is shown in Fig. 1-(a) where the success of the communication event depends solely on a maximum communication range Q. This means that if the vehicle range exceeds the communication range $r > Q$, then the communication cannot be accomplished. In practice, this binary approach is unrealistic, as multiple factors may affect the communication of two vehicles, such as scattering, motion-induced Doppler effect, and change of environmental conditions. To this end, we propose multi-dimensional communication performance maps for various ranges as illustrated in Fig. 1-(b). For the evaluation of the communication performance we employ signal-to-noise-ratio (SNR) measurements.

Let the SNR measurements be modeled by,

$$Y(\mathbf{x}; v) = \boldsymbol{\mu}(\mathbf{x}; v) + Z(\mathbf{x}; v) + \epsilon(\mathbf{x}), \tag{1}$$

where $Y(\mathbf{x}; v) \in \mathbb{R}^n$ is the measurement vector describing a non-stationary random field at spatial coordinates $\mathbf{x} \in \mathbb{R}^2$, $\boldsymbol{\mu}(\mathbf{x}; v) \in \mathbb{R}^n$ is the deterministic mean (or spatial *trend*), $Z(\mathbf{x}; v) \sim \mathcal{N}(0, \boldsymbol{\Sigma}(\mathbf{x}; v)) \in \mathbb{R}^n$ is a second-order stationary random field with $\boldsymbol{\Sigma}(\mathbf{x}; v)$ its covariance matrix, and $\epsilon \sim \mathcal{N}(0, \tau^2 I_n)$ is an iid zero-mean Gaussian random field. The mean $\boldsymbol{\mu}$ is the spatial trend that represents large-scale variability, the second-order stationary random field Z captures medium-scale variability, and the white noise ϵ is the small-scale variation of the sensor. The surrogate variable is denoted v and is used to represent model dependence, not explicitly accounted for spatial coordinates $\mathbf{x}$. In Section II-C, we identify a surrogate variable from an UWA propagation channel model.

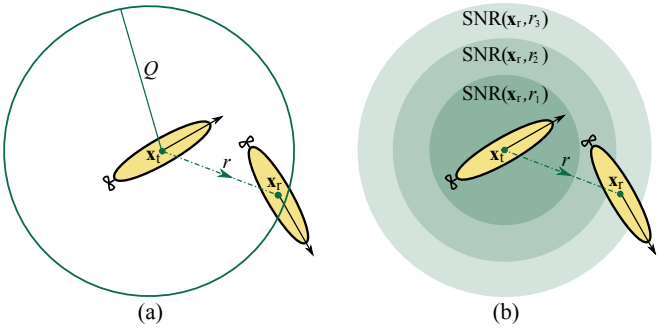

Fig. 1. Communication scenarios of two autonomous underwater vehicles (AUVs) at range r. (a) The communication success relies on a maximum communication range Q. (b) The communication performance using signal-to-noise ratio (SNR) is predicted for specific vehicle ranges.

Assumption 1: The deterministic mean is decomposed by a linear combination of unknown parameters expressed by $\boldsymbol{\mu}(\mathbf{x}; v) = \mathbf{X}(\mathbf{x}; v)\boldsymbol{\beta}$, where $\mathbf{X}(\mathbf{x}; v) \in \mathbb{R}^{n \times p}$ represents the matrix of known basis functions and $\boldsymbol{\beta} \in \mathbb{R}^p$ the vector of the unknown regressor coefficients.

Since the measurements Y are non-stationary, we seek to detrend the measurements, i.e. remove the mean $Y - \boldsymbol{\mu}$, to obtain a stationary random field. Next, with the detrended measurements the covariance matrix $\boldsymbol{\Sigma}$ is estimated with an iterative scheme. After estimating the covariance matrix $\boldsymbol{\Sigma}$, we employ the original measurements Y to perform predictions. A critical component for detrending is the basis functions $\mathbf{X}$, thus we are inspired by the propagation model to design $\mathbf{X}$ and accurately detrend the measurements.

C. Communication Performance

To approximate the communication performance between two agents we use the SNR. In principle, the higher the SNR, the more likely is to detect the signal. Let the power of the transmitted signal to be constant, then the SNR yields, $SNR = P_\mathrm{T} G / P_\mathrm{N}$, where P_T denotes the power of the transmitted signal, G is the channel gain, and P_N is the noise power. To statistically model the channel gain G we employ [17]. The gain G follows a log-normal distribution $\log G \sim \mathcal{N}(\bar{G}, \sigma_\mathrm{G}^2)$, where $\bar{G}$ represents the mean of the log channel gain and σ_G^2 its variance. On the decibel scale, the source level takes the form of $S_\mathrm{l}(f) = 10 \log P_\mathrm{T}$ and the noise level yields $NL(f, \omega) = 10 \log P_\mathrm{N}$. By neglecting variations of water pressure, the gain on the decibel scale $g = 10 \log G$ is Gaussian and expressed as,

$$g(r) = \bar{g}(r) + \nu, \tag{2}$$

where $\nu(t) \sim \mathcal{N}(0, \sigma_\nu^2)$ is a zero-mean Gaussian field, r is the range between vehicles. The mean follows,

$$\bar{g}(r) = g_0 - k_0 10 \log \frac{r}{r_\mathrm{ref}}, \tag{3}$$

where g_0 is a constant gain, r_ref is reference range (e.g., 1 m in our case), and k_0 is the path loss exponent. Note that (1) has identical structure with the model of the UWA

propagation channel model (2). Thus, using (3) we choose v to be the range between transmitting and receiving node, i.e. $v = r$, and the SNR measurements (1) are expressed as,

$$Y(\mathbf{x}; r) = \mathbf{X}(\mathbf{x}; r)\boldsymbol{\beta} + Z(\mathbf{x}; r) + \epsilon(\mathbf{x}). \tag{4}$$

The specific goal of our UWA performance prediction application is summarized in Problem 1.

Problem 1: Predict the communication performance $\hat{Y}$ and the corresponding variance $\mathrm{Var}[\hat{Y}]$ at unvisited locations $\mathbf{x}_0$, provided a set of communication performance measurements Y at locations $\mathbf{x}$ and the vehicle range r.

III. Parameter Estimation and Prediction

In this section, we formulate basis functions $\mathbf{X}$ and use ordinary least squares to estimate the spatial trend $\boldsymbol{\mu}$. Then, we remove the trend by subtracting the mean $\boldsymbol{\mu}$ from the measurements Y. The detrended measurements $Y - \boldsymbol{\mu}$ are employed to estimate the parameters of multiple variogram functions with a maximum likelihood-based method. Next, we use the Bayesian information criterion to select the most suitable variogram function among multiple candidates. With the selected variogram model we construct the covariance matrix $\boldsymbol{\Sigma}$ and use generalized least squares to improve the accuracy of the spatial trend estimator $\boldsymbol{\mu}$. The method iterates until the parameters of the variogram function converge. Finally, we employ the estimated covariance matrix and the original measurements Y to predict the communication performance and its variance with universal kriging.

A. Spatial Trend Modeling

In this section, we seek basis functions $\mathbf{X}$ to model the spatial trend $\boldsymbol{\mu}$ and detrend the measurements $Y - \boldsymbol{\mu}$ and retrieve a stationary random field. The random field in (4) is non-stationary due to the spatial trend. Thus, the original measurements cannot be used to estimate the parameters of the covariance function. A precise model of the trend is of paramount importance for spatial extrapolation. The obvious choice for the elements of the basis function $\mathbf{X}$ is to employ spatial coordinates as covariates. However, surrogate variables—arising from the physical model of the system—are useful covariates to interpret the behavior of the spatial variation [18]. In spatial statistics, polynomial basis functions $\mathbf{X}(\mathbf{x}) = [1, x, y, xy, x^2, y^2]$ are usually used. However, polynomial basis functions do not behave well for extrapolation, due to *unboundedness*, i.e. as $\|\mathbf{x}\| \to \infty$ then $X(\mathbf{x}) \to \infty$. Alternatively, Gaussian radial basis functions (RBF) provide suitable extrapolation results. A Gaussian RBF is given by,

$$X_l(\mathbf{x}; c_l, \sigma^2_{\mathrm{G},l}) = \exp\left(-\frac{(\mathbf{x} - c_l)^2}{2\sigma^2_{\mathrm{G},l}}\right), \tag{5}$$

where c_l is the center of each measurement, e.g., $c_l = 0$ for zero mean measurement error ϵ (4). The corresponding variance is denoted $\sigma^2_{\mathrm{G},l}$ and in practice is a constant value $\sigma^2_{\mathrm{G},l} = \sigma^2_{\mathrm{G}}$ for all l measurements. From (3), it is deduced that the range of the vehicles has a linear-log relationship to the mean. Hence, our proposed hybrid basis function combines Gaussian RBF incorporating spatial coordinates (5) and linear-log range,

$$\mathbf{X}(\mathbf{x}; r) = [1, \exp\left(-\frac{(x - c_x)^2}{2\sigma_x^2}\right), \exp\left(-\frac{(y - c_y)^2}{2\sigma_y^2}\right), r, \log r]. \tag{6}$$

At the initial stage, generalized least squares (GLS) cannot be used for data detrending, as the covariance is unknown. Thus, the mean parameters are initially estimated by the ordinary least squares (OLS),

$$\widehat{\boldsymbol{\beta}}^{(1)}_{\mathrm{OLS}} = \mathbf{X}(\mathbf{x}; r)^{\dagger} Y(\mathbf{x}; r), \tag{7}$$

where $\mathbf{X}^{\dagger} = (\mathbf{X}^{\intercal}\mathbf{X})^{-1}\mathbf{X}^{\intercal}$, $\mathbf{X}^{\dagger} \in \mathbb{R}^{p \times n}$ is the Moore-Penrose pseudoinverse of $\mathbf{X}$. The residual measurements (or detrended data) yield,

$$\tilde{Y}(\mathbf{x}; r) = Y(\mathbf{x}; r) - \mathbf{X}(\mathbf{x}; r)\widehat{\boldsymbol{\beta}}^{(1)}_{\mathrm{OLS}}. \tag{8}$$

Assumption 2: The random field of the underlying latent process is second-order stationary after detrending $\tilde{Y}$.

Assumption 3: The variogram function is isotropic after detrending.

B. Semivariogram

In this section, we present three commonly used semivariogram functions and an optimization method to estimate the initial parameters. A robust estimator of the experimental semivariogram function is proposed in [19], which yields,

$$\hat{\gamma}_{\mathrm{CH}}(\mathbf{h}) = \frac{\left(\frac{\sum_{N(\mathbf{h})} |\tilde{Y}(\mathbf{x}+\mathbf{h}) - \tilde{Y}(\mathbf{h})|^{1/2}}{\mathrm{card}(N(\mathbf{h}))}\right)^4}{0.914 + \frac{0.988}{2\,\mathrm{card}(N(\mathbf{h}))} + \frac{0.090}{\mathrm{card}(N(\mathbf{h}))^2}}, \tag{9}$$

where $N(\mathbf{h}) = \{(o, p) \mid \mathbf{x}_o - \mathbf{x}_p = \mathbf{h}\}$ is the set of measurements at distance $\mathbf{h}$ and $\mathrm{card}(\cdot)$ represents set cardinality. Note that we use the residual measurements $\tilde{Y}$ (8). The robustness relies on a transformation which ensures that the fourth root of the transformed distribution produces small skew.

Next, we provide three isotropic semivariogram functions that will be used as candidate models. The spherical model,

$$\gamma_{\mathrm{s}}(\mathbf{h}; \boldsymbol{\theta}) = \begin{cases} \tau^2 + \sigma^2, & \|\mathbf{h}\| \geq \alpha, \\ \tau^2 + \sigma^2\left(\frac{3\|\mathbf{h}\|}{2\alpha} - \frac{1}{2}\left(\frac{\|\mathbf{h}\|}{\alpha}\right)^3\right), & \|\mathbf{h}\| \leq \alpha, \end{cases} \tag{10}$$

where the semivariogram parameter vector $\boldsymbol{\theta} = [\tau^2\ \sigma^2\ \alpha]^{\intercal} \in \Theta$ contains the nugget, the partial sill, and the semivariogram range with $\Theta = \{\boldsymbol{\theta} \in \mathbb{R}^3 \mid \tau^2 \geq 0, \sigma^2 \geq 0, \alpha \geq 0\}$. Second, the exponential model,

$$\gamma_{\mathrm{e}}(\mathbf{h}; \boldsymbol{\theta}) = \tau^2 + \sigma^2\Big(1 - \exp\Big\{-\frac{\|\mathbf{h}\|}{\alpha}\Big\}\Big). \tag{11}$$

Finally, we consider the Matérn semivariogram function with fixed smoothing parameter at $\kappa = 3/2$ to obtain a mixed polynomial-exponential model,

$$\gamma_{\mathrm{pe}}(\mathbf{h}; \boldsymbol{\theta}) = \tau^2 + \sigma^2\left(1 - \left(1 + \frac{\sqrt{3}\|\mathbf{h}\|}{\alpha}\right)\exp\Big\{-\frac{\sqrt{3}\|\mathbf{h}\|}{\alpha}\Big\}\right), \tag{12}$$

The next step is to formulate an optimization problem to fit the semivariogram models γ (10), (11), (12) and derive the corresponding parameter vector $\boldsymbol{\theta}$. We utilize a weighted least squares (WLS) approach [20] which has the form of,

$$\widehat{\boldsymbol{\theta}}^{(0)}_{\text{CWLS}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{g=1}^{N_g} \text{card}(N(\mathbf{h}_g)) \left(\frac{\widehat{\gamma}_{\text{CH}}(\mathbf{h}_g)}{\gamma(\mathbf{h}_g;\boldsymbol{\theta})} - 1 \right)^2, \quad (13)$$

where N_g is the total number of the separation vectors $\mathbf{h}_g$.

The parameter estimation (13) relies on the residual measurements (8) which incorporate measurement bias. Hence, the estimation is sensitive to the bias of the mean.

C. Restricted Maximum Likelihood

In this section, we seek an unbiased estimator for the parameter vector $\boldsymbol{\theta}$ and a strategy to narrow down the parameter space Θ. An alternative bias-free approach of the maximum likelihood is the restricted maximum likelihood estimation (REML) [21], [22] which makes use of the *error contrasts* to remove the mean dependence from the variance estimates. The main idea is to transform the residual measurements $\tilde{Y}$ from (8) with a matrix $\mathbf{A} \in \mathbb{R}^{n\times(n-p)}$ such that, $\mathbf{A}^\intercal\mathbf{X} = \mathbf{0}$ and $\text{E}[\mathbf{A}^\intercal\tilde{Y}] = 0$, where $\mathbf{X}$ is the basis function (6). In other words, each column vector of matrix $\mathbf{A}$ is orthogonal to all columns of $\mathbf{X}$. Let us define the error contrast, $W \coloneqq \mathbf{A}^\intercal\tilde{Y}$ to obtain $W \sim \mathcal{N}(0, \mathbf{A}^\intercal\boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{A})$. Although $\mathbf{A}$ is not unique, a matrix that satisfies the properties yields, $\mathbf{A} = I_n - \mathbf{X}(\mathbf{X}^\intercal\mathbf{X})^{-1}\mathbf{X}^\intercal$. Apparently, $\mathbf{A}$ does not depend on the estimated mean parameters $\widehat{\boldsymbol{\beta}}_{\text{OLS}}$. The log-restricted likelihood function is defined by,

$$L(\boldsymbol{\theta}|W) = -\frac{1}{2}\Big((n-p)\log(2\pi) + \log|\mathbf{X}^\intercal\mathbf{X}| - \log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \log|\mathbf{X}^\intercal\boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{X}| - \tilde{Y}^\intercal\boldsymbol{\Pi}(\boldsymbol{\theta})\tilde{Y}\Big), \quad (14)$$

$\boldsymbol{\Pi}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} - \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\mathbf{X}(\mathbf{X}^\intercal\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}\mathbf{X}^\intercal\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}$, n is the measurement vector size, and p is the rank of $\mathbf{X}$. Next, the log-restricted likelihood (14) is maximized with respect to $\boldsymbol{\theta} \in \Theta$ to obtain the estimated parameter vector $\widehat{\boldsymbol{\theta}}$. To reduce the search of the parameter space Θ, we propose to use the parameter estimate $\widehat{\boldsymbol{\theta}}^{(0)}_{\text{CWLS}}$ (13) as a center value of the initial set of parameters in the optimization. So far we have computed three covariance parameter vectors $\widehat{\boldsymbol{\theta}}$ corresponding to three candidate models (10), (11), (12). The next step is to find which of these models is the best fit to our data with a statistical model selection technique.

D. Bayesian Information Criterion

The Bayesian information criterion (BIC) [23] is a statistical model selection methodology which is defined as,

$$\text{BIC}(M_k) \coloneqq -2\ln\mathcal{L}(\widehat{\boldsymbol{\theta}}_k \mid \tilde{Y}, M_k) + q\ln n, \quad (15)$$

where $\mathcal{M} = \{M_k = \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}_k) \mid k = 1,\dots,K\}$ is the set of candidate models, $\widehat{\boldsymbol{\theta}}_k$ denotes the REML estimates of $\boldsymbol{\theta}_k$, $q = 3$ is the dimension of the parameter space Θ, $\mathcal{L}(\widehat{\boldsymbol{\theta}}_k \mid \tilde{Y}, M_k)$ represents the likelihood corresponding to the density function $f(\tilde{Y}, M_k \mid \widehat{\boldsymbol{\theta}}_k)$, and n is the measurement size of the vector $\tilde{Y}$. In our case $K = 3$ corresponds to three candidate semivariogram functions (10), (11), (12). In principle, the smaller the BIC the more suitable the candidate model. The advantage of the BIC is *consistency*. That is, if the true model is not listed among the candidate models, the BIC selects the most parsimonious model closest to the true model, by computing the marginal log-likelihood with Laplace approximation.

Since the BIC (15) is computed in the log-scale, its evaluation may be ambiguous. Thus, we employ the posterior density of the BIC [24] which is approximated by,

$$P(M_k \mid \tilde{Y}) \approx \frac{\exp\Big(-\frac{1}{2}\Delta_k\Big)}{\sum_{k=1}^{K}\exp\Big(-\frac{1}{2}\Delta_k\Big)}, \quad (16)$$

where $\Delta_k = \text{BIC}(M_k) - \text{BIC}^\star$ denotes the BIC difference of a candidate model with the minimum BIC candidate model $\text{BIC}^\star = \min_{M_k\in\mathcal{M}}\text{BIC}(M_k)$. Essentially, $P(M_k \mid \tilde{Y})$ is a probability mass function, that provides a probability of suitability of each model to the real model.

E. Iterative Parameter Estimation

For the iterative parameter estimation we utilize the estimated covariance matrix $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(1)})$ as selected by the BIC. The covariance matrix allows the implementation of the generalized least squares (GLS) to improve the estimation of the mean. The parameters of the GLS estimator follow,

$$\widehat{\boldsymbol{\beta}}^{(2)}_{\text{GLS}} = \left(\mathbf{X}^\intercal\boldsymbol{\Sigma}\Big(\widehat{\boldsymbol{\theta}}^{(1)}\Big)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\intercal\boldsymbol{\Sigma}\Big(\widehat{\boldsymbol{\theta}}^{(1)}\Big)^{-1}Y. \quad (17)$$

Subsequently, the residual measurements (8) yield,

$$\tilde{Y}(\mathbf{x};r) = Y(\mathbf{x};r) - \mathbf{X}(\mathbf{x};r)\widehat{\boldsymbol{\beta}}^{(2)}_{\text{GLS}}. \quad (18)$$

In addition, the GLS mean estimation facilitates a more accurate determination of the covariance function. To this end, we employ the detrended measurements (18) and iterate the method. The method terminates when,

$$\|\widehat{\boldsymbol{\theta}}^{(s)} - \widehat{\boldsymbol{\theta}}^{(s-1)}\| \leq \eta \quad (19)$$

where $\eta \in \mathbb{R}^+$ is a small threshold. At every iteration we expect lower BIC values (15). Practically, after the second iteration the algorithm terminates [25, pp. 196–200], [26].

F. Prediction Technique

Let us now describe the universal kriging (UK) technique [6], [27]. UK is the generalized form of kriging and considers spatially varying mean. More specifically, provided measurements $Y \in \mathbb{R}^n$ the prediction follows,

$$\widehat{Y}(\mathbf{x}_0;r) = \sum_{i=1}^{n}\omega_i Y(\mathbf{x}_i;r) = \boldsymbol{\omega}^\intercal Y(\mathbf{x};r), \quad (20)$$

where $\mathbf{x}_0$ is the location of interest, $\boldsymbol{\omega} = [\omega_1 \dots \omega_n]^\intercal \in \mathbb{R}^n$ are the weights we seek to obtain, and Y are the raw measurements (1), i.e. not the residuals $\tilde{Y}$. Next, we formulate the unconstrained minimization problem with multiple Lagrange multipliers $\boldsymbol{\lambda}_{\text{UK}} \in \mathbb{R}^p$ to include the universality conditions

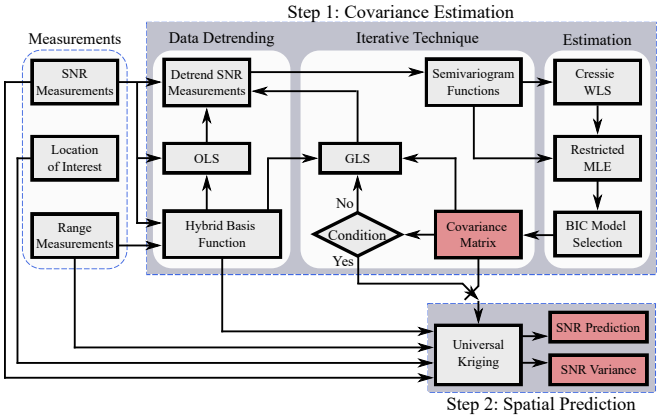

Fig. 2. The structure of the communication performance predictor is a two-step method. The first step is to estimate the covariance matrix and the second to predict the communication performance.

$\boldsymbol{\omega}^\intercal \mathbf{X} = X_0^\intercal$. The solution is provided by,

$$\boldsymbol{\omega}_{\text{UK}} = \boldsymbol{\Gamma}_{\text{UK}}^{-1}\boldsymbol{\gamma}_{\text{UK}}, \tag{21}$$

where $\boldsymbol{\omega}_{\text{UK}} = [\boldsymbol{\omega}^\intercal \ \boldsymbol{\lambda}_{\text{UK}}^\intercal]^\intercal \in \mathbb{R}^{n+p}$ is a stacked vector that contains the weights $\boldsymbol{\omega}$ and the Lagrange multipliers $\boldsymbol{\lambda}_{\text{UK}}$ to minimize the mean square prediction error. The non-singular matrix $\boldsymbol{\Gamma}_{\text{UK}} \in \mathbb{R}^{(n+p)\times(n+p)}$ captures the *redundancy* of measurements and is given by,

$$\boldsymbol{\Gamma}_{\text{UK}} \coloneqq \begin{bmatrix} \boldsymbol{\Gamma} & \mathbf{X} \\ \mathbf{X}^\intercal & \mathbf{0}_{p\times p} \end{bmatrix},$$

The semivariogram vector $\boldsymbol{\gamma}_{\text{UK}} \in \mathbb{R}^{n+p}$ considers the *closeness* of the measurements to the location of interest $\mathbf{x}_0$,

$$\begin{aligned}\boldsymbol{\gamma}_{\text{UK}} =& [\gamma(\mathbf{x}_0,\mathbf{x}_1) \ \dots \ \gamma(\mathbf{x}_0,\mathbf{x}_n) \quad 1 \quad X_2(\mathbf{x}_0) \ \dots \ X_p(\mathbf{x}_0)]^\intercal \\ \coloneqq& \begin{bmatrix} \boldsymbol{\gamma}_0 \\ X_0 \end{bmatrix}.\end{aligned}$$

The unique solution of (21) yields the decoupled coefficients,

$$\boldsymbol{\omega}^\intercal = \left(\boldsymbol{\gamma}_0 + \mathbf{X}(\mathbf{X}^\intercal\boldsymbol{\Gamma}^{-1}\mathbf{X})^{-1}(X_0 - \mathbf{X}^\intercal\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}_0)\right)^\intercal\boldsymbol{\Gamma}^{-1}, \tag{22}$$

and the Lagrange multipliers,

$$\boldsymbol{\lambda}_{\text{UK}} = (\mathbf{X}^\intercal\boldsymbol{\Gamma}^{-1}\mathbf{X})^{-1}(\mathbf{X}^\intercal\boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}_0 - X_0)^\intercal. \tag{23}$$

With (22), (23) the variance is computed by,

$$\sigma^2_{\text{UK}}(\widehat{Y}(\mathbf{x}_0)) = X_0^\intercal\boldsymbol{\lambda}_{\text{UK}} + \boldsymbol{\omega}^\intercal\boldsymbol{\gamma}_0. \tag{24}$$

Note that the solution can be also expressed with respect to the covariance matrix $\boldsymbol{\Sigma}$.

IV. Spatial Prediction Framework

A. Structure

The two-step methodology is shown in Fig. 2. We start by collecting measurements of communication performance (SNR) and vehicle range. The first step is to estimate the covariance matrix and the second to predict the communication performance at unvisited locations.

The covariance estimation comprises of three modules: i) the data detrending; ii) the parameter estimation; and iii) the iterative technique. The data detrending includes the hybrid basis function formulation (6), and OLS computation (7). Next, the detrended measurements are used to compute the candidate semivariogram models (10), (11), (12). The estimation module computes the covariance parameters to be used as initial conditions, by employing the robust experimental semivariogram (9) and a WLS optimization (13). The next stage is the REML estimation (14) that results in three parameter vectors. The last stage of the estimation module considers the selection of the most suitable covariance model among the three candidates with the posterior BIC (16). The last module describes an iterative technique for the selection of the covariance matrix. Since we obtained a covariance matrix, the mean estimates can be improved by computing the GLS (17). Subsequently, we recompute the residuals and run the estimation module to obtain a new covariance matrix. The method iterates until the parameters of the covariance converge (19). The second step uses the measurements, the location of interest, the model-based basis functions, and the covariance matrix from the previous step to solve the universal kriging (21). Finally, we predict the SNR at the location of interest (20) and its corresponding variance (24).

Algorithm 1 Communication Performance Prediction

Input: Y, $\mathbf{x}$, r, $\mathbf{x}_0$, n, p, q, γ, η
Output: $\widehat{Y}$, $\text{Var}[\widehat{Y}]$

$\widehat{\boldsymbol{\theta}}^{(0)} \leftarrow \texttt{initialConditions}(Y)$
$\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(0)}) \leftarrow I_n$; $k \leftarrow 0$;
$\mathbf{X} \leftarrow \texttt{basis}(\mathbf{x}; r)$;
for $s = 1$ to S **do**
  $\widehat{\boldsymbol{\beta}}^{(s)} \leftarrow \texttt{GLS}(Y, \mathbf{X}, \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(s-1)}))$;
  $\tilde{Y}^{(s)} \leftarrow \texttt{detrend}(Y, \mathbf{X}, \widehat{\boldsymbol{\beta}}^{(s)})$;
  for each $\gamma \in \mathcal{C}$ **do**
    $\widehat{\boldsymbol{\theta}}_k^{(s-1)} \leftarrow \texttt{CWLS}(\tilde{Y}, \gamma, \widehat{\boldsymbol{\theta}}^{(s-1)})$;
    $\widehat{\boldsymbol{\theta}}_k^{(s)} \leftarrow \texttt{REML}(\tilde{Y}, \mathbf{X}, n, p, \gamma, \widehat{\boldsymbol{\theta}}_k^{(s-1)})$;
    $M_k \leftarrow \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}_k^{(s)})$;
    $\text{BIC}_k \leftarrow \texttt{BIC}(\tilde{Y}, n, q, \widehat{\boldsymbol{\theta}}_k^{(s)}, M_k)$;
    $k \leftarrow k + 1$;
  end for
  $\text{BIC}^\star \leftarrow \min_{M_k \in \mathcal{M}}\{\text{BIC}(M_k)\}$;
  for $k = 1$ to K **do**
    $\Delta_k \leftarrow \texttt{diffBIC}(\text{BIC}_k, \text{BIC}^\star)$;
  end for
  for $k = 1$ to K **do**
    $P(M_k \mid \tilde{Y}) \leftarrow \texttt{postBIC}(\Delta_k)$;
  end for
  $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(s)}) \leftarrow \max_{M_k \in \mathcal{M}}\{P(M_k \mid \tilde{Y})\}$;
  if $\|\widehat{\boldsymbol{\theta}}^{(s)} - \widehat{\boldsymbol{\theta}}^{(s-1)}\| \leq \eta$ **then**
    break;
  end if
end for
$\widehat{Y}, \text{Var}[\widehat{Y}] \leftarrow \texttt{UK}(Y, r, \mathbf{X}, \mathbf{x}_0, \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(s)}))$;

B. Algorithm

The main routine of the communication performance predictor is presented in Algorithm 1. The `initialConditions` module assigns initial values

to the semivariogram parameter vector $\widehat{\boldsymbol{\theta}}^{(0)}$. The initial covariance matrix estimate $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(0)})$ is set equal to the identity matrix. Next, the algorithm proceeds to the iterative parameter estimation. We consider 3 semivariogram models $\mathcal{C} = \{\gamma_{\text{s}}, \gamma_{\text{e}}, \gamma_{\text{pe}}\}$ (10), (11), (12) at each iteration. In our case, the `basis` is designed for spatial extrapolation (6). Then, the `GLS` function implements (17) to estimate the mean regressor parameters $\widehat{\boldsymbol{\beta}}^{(s)}$. Note, that in the first iteration, the initial covariance matrix is the identity matrix, and thus, the algorithm implements the OLS (7). The function `detrend` computes the residual measurements (or detrended data) $\tilde{Y}^{(s)}$ (8). With the detrended data, the function `CWLS` computes initial values for the estimation of the semivariogram parameter vector $\widehat{\boldsymbol{\theta}}_k^{(s-1)}$ by solving a WLS optimization problem (13). Next, the `REML` module implements the REML (14) to estimate the semivariogram parameter vector $\widehat{\boldsymbol{\theta}}_k^{(s)}$. The `BIC` function calculates the BIC (15) towards a model selection procedure. The `diffBIC` computes the difference of each candidate with the lowest BIC* and then the `postBIC` calculates the posterior BIC (16). The iterative procedure is terminated when the covariance parameters converge to an η-neighborhood (19) (e.g., after two iterations in our case). Finally, we utilize the estimated covariance matrix $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}}^{(s)})$ and the original data Y to solve the universal kriging (20), (24) and obtain SNR prediction $\widehat{Y}$ and its variance $\text{Var}[\widehat{Y}]$.

V. Simulations and Results

A. Simulation Setup

The simulation environment is developed with a well-established, statistical UWA channel model that incorporates 34 parameters and includes contribution from multipath, motion-induced Doppler, surface scattering, and large-scale variability of the channel geometry [17]. This channel model has been exhaustively compared to experimental data from multiple underwater missions, which varied in location, season, time duration, and weather conditions. Moreover, the experimental data reported in [17] demonstrate very close match of the measurement histogram to the Gaussian assumption of the statistical model. Similarly to our case, one of their reported missions in [17, PS experiment] considers mobile agents. Note that experimental results from our field trials verify the fidelity of the simulation [28].

In addition to the channel model, we inject local ambient noise with non-zero Gaussian distributions [4]. The channel gain (2) is computed for signal frequency $f = 25\,\text{kHz}$, bandwidth $B = 5\,\text{kHz}$, surface height $100\,\text{m}$, and vehicle depths $z_1 = 80$ and $z_2 = 50\,\text{m}$. We set the source level $S_{\text{l}} = 180$ dB which is a realistic value for UWA acoustic modems in such frequency f. The large-scale parameters, i.e. path gain and propagation delay, are computed by the Bellhop model [29]. We consider two vehicles whose trajectories are shown in Fig. 3. The black solid line and the dotted black line represent the paths of agent 1 and 2 respectively. Two cases were considered: i) the long-distant prediction; and ii) the short-distant prediction. In the long-distant case, each agent collects 75 SNR measurements and range r. We seek to predict the SNR at 260 and 259 unvisited location for agent 1 and 2. In the short-distant case, each agent collects 250 SNR measurements and range r. We seek to predict the SNR at 85 and 84 unvisited location for agent 1 and 2. In the range-SNR plots (middle and bottom row in Fig. 3), the red shaded areas correspond to the unknown SNR regions, which we pretend to not know *a priori*. After predicting the SNR, we compare to the true values.

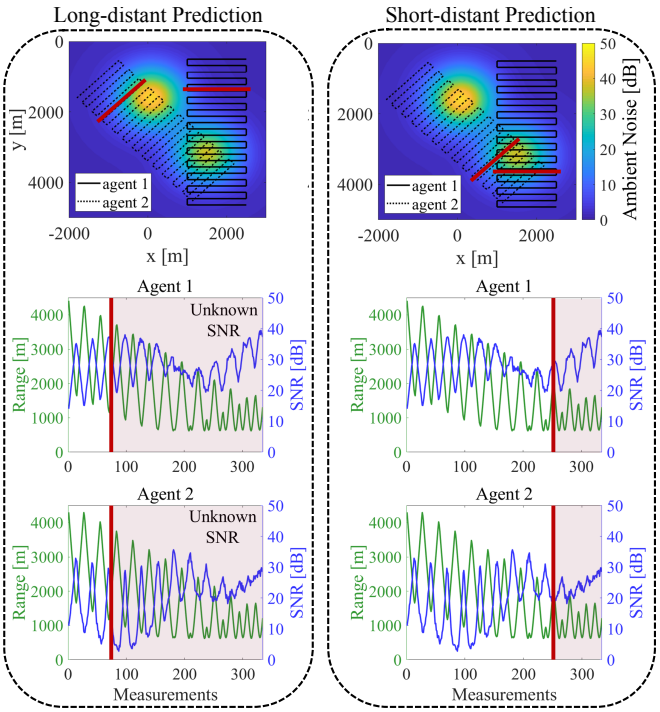

Fig. 3. The color map depicts the ambient noise distribution. The solid red lines separate the areas of collected signal-to-noise ratio (SNR) measurements from the unvisited areas. Each agent collects 75 measurements for long-distant prediction and 250 measurements for short-distant prediction.

The evaluation of the prediction is accomplished with two metrics, the mean average prediction error, $\text{MAPE} = 1/n_u \sum_{u=1}^{n_u} |\widehat{Y}(\mathbf{x}_{0,u}) - Y(\mathbf{x}_{0,u})|$, and the root mean square prediction error, $\text{RMSPE} = \{1/n_u \sum_{u=1}^{n_u} (\widehat{Y}(\mathbf{x}_{0,u}) - Y(\mathbf{x}_{0,u})^2\}^{1/2}$, where n_u is the number of unvisited locations.

B. Simulation Results

We perform 24 simulations including two scenarios, three biases, and four prediction cases to evaluate the accuracy and robustness of our method. The scenarios are: i) 150 data and 519 unvisited locations (long-distant); and ii) 500 data and 169 unvisited locations (short-distant). At each scenario we investigate the effect of the bias to the covariance estimation, i.e. robustness, by adding a systematic error (bias) to the measurements. The added biases are: i) 10; ii) -10; and iii) 0. We perform prediction with four techniques: i) OLS with a polynomial basis; ii) OLS with hybrid basis; iii) GLS with hybrid basis; and iv) UK with hybrid basis. Note that both OLS methods are covariance-free techniques (7), while

TABLE I

COVARIANCE ESTIMATION & PREDICTION ASSESSMENT

Cases			Model Selection, & Covariance Parameters	Prediction Error Values of the Non-Stationary Process							
				Unknown Covariance				Estimated Covariance			
Known Data	Unknown	Bias		OLS (Polynomial)		OLS (Hybrid)		GLS (Hybrid)		UK (Hybrid)	
			BIC-Posterior-Model	MAPE	RMSPE	MAPE	RMSPE	MAPE	RMSPE	MAPE	RMSPE
150	519	−10	313 - 99% - Matérn	9.13	11.61	6.41	**7.89**	**6.35**	7.97	7.22	8.10
		0	$\sigma^2 = 32.93$	15.34	18.43	12.52	14.54	7.72	9.01	**6.94**	**7.95**
Long-distant		10	$\alpha = 1124$	25.24	27.29	22.46	23.68	15.87	17.28	**13.35**	**14.13**
500	169	−10	1048 - 100% - Matérn	10.16	11.24	5.51	7.10	**4.70**	**6.24**	12.67	13.06
		0	$\sigma^2 = 81.02$	4.04	4.81	5.77	6.58	7.10	8.13	**2.90**	**4.15**
Short-distant		10	$\alpha = 1495$	9.89	10.95	14.65	15.37	16.36	17.12	**7.34**	**7.98**

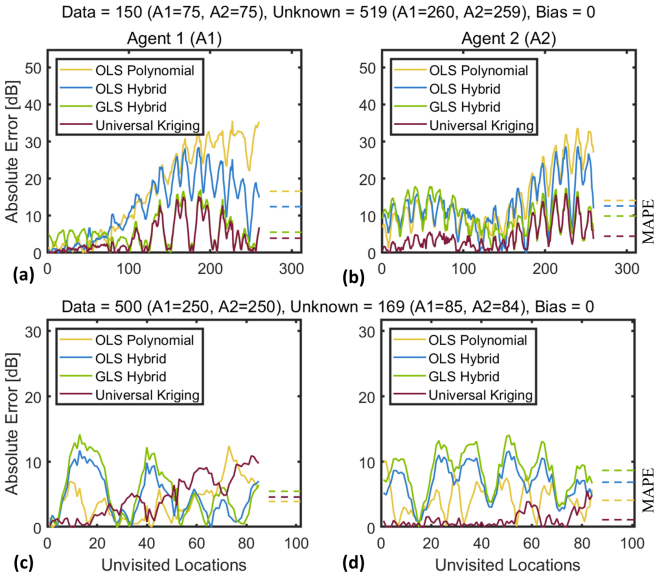

Fig. 4. Absolute local error of zero bias for each agent with four predictors. (a) Long-distant case of agent 1. (b) Long-distant case of agent 2. (c) Short-distant case of agent 1. (d) Short-distant case of agent 2.

the GLS and UK make use of the estimated covariance. In other words, we compare: i) standard polynomial basis to the proposed hybrid basis for the case of OLS; ii) OLS to GLS using the proposed hybrid basis; and iii) the proposed technique to the other three methods. The results are listed in Table I. Next, we illustrate the propagation of absolute error (Fig. 4) in eight simulations on each agent including both scenarios with zero bias for all prediction techniques.

We observe in Table I that the added bias does not affect the covariance matrix estimation, resulting in the same BIC value, semivariogram model, and covariance parameters. Thus, the proposed method is robust, in the sense that it successfully removes the bias from the covariance estimation. Although in all simulation cases the Matérn model (12) is selected according to the posterior distribution of the BIC (16), it is recommended to follow the statistical model selection technique, as for different applications, other semivariogram models may represent the latent process more accurately.

Next, we compare the basis functions of the OLS techniques. Clearly, in the long-distant case, the proposed OLS hybrid basis function outperforms the OLS polynomial in terms of lower MAPE and RMSPE values. Furthermore, in Fig. 4-(a), (b) we observe bounded behavior for long-distant predictions with the OLS hybrid, while the OLS polynomial produces significantly higher absolute errors at distant locations from the acquired measurements. Although for short-distant predictions the OLS polynomial technique appears lower error values, the unbounded behavior advocates to an unreliable methodology for extrapolation.

The covariance estimation allows the implementation of GLS and UK. The covariance-based methods (GLS and UK) outperform the covariance-free methods (OLS) and demonstrate bounded behavior in all cases. Particular, GLS and UK result in significantly lower MAPE and RMSPE from OLS methods (Table I), emphasizing the impact in prediction accuracy of the proposed covariance estimation method.

Next, we compare the covariance-based methods. For zero bias the UK improves the prediction in terms of RMSPE comparing to the GLS over 12% for the long-distant case and 49% for the short-distant case. Similar prediction improvements are reported for positive bias, but not for the negative bias cases. Naturally, one direction of bias (-10 dB) favors the parsimonious techniques of OLS and GLS. Interestingly, the UK errors remain *consistent* with low fluctuations even at areas of high local ambient noise, as illustrated in Fig. 4-(a), (b). In Fig. 4-(c), agent 1 demonstrates local deviation, due to high local ambient noise. Yet, agent 2 yields almost zero absolute error values, as shown in Fig. 4-(d). The proposed model-based UK technique does not exceed 12 dB of local absolute error in any case—even for long-distant prediction—and produces much lower MAPE and RMSPE values, leading to an accurate and reliable methodology for communication performance prediction.

Another powerful tool of UK is the variance of the prediction. Essentially, every predicted value is characterized by a Gaussian probability distribution, interpreting the uncertainty of predicted value. In Fig. 5, the mean prediction values along with the corresponding standard deviation are presented. In the long-distant case, the first error off the standard deviation appears after 130 and 180 unvisited locations for agents 1 and 2 respectively. The longer the distance of the unvisited location from the observation area, the higher

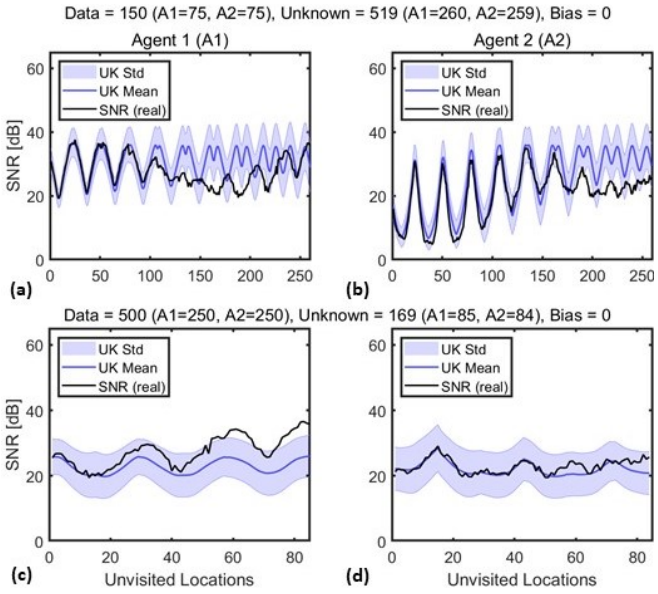

Fig. 5. Mean prediction values for universal kriging with corresponding standard deviation. (a) Long-distant case of agent 1. (b) Long-distant case of agent 2. (c) Short-distant case of agent 1. (d) Short-distant case of agent 2.

the uncertainty of the prediction. In the short-distant case, agent 1 reports the first error off the standard deviation after 60 predictions beyond the observation area, while agent 2 demonstrates no errors from the standard deviation. Interestingly, the short-distant case produces higher variance values from the long-distant case. This reveals that when the vehicles collect measurements from high ambient noise regions the uncertainty of the prediction increases.

VI. Conclusion

This paper proposes a model-based, data-driven framework for prediction of UWA communication performance in AUVs beyond the observation area. We show that the proposed basis function yields bounded error values, even with simple prediction methods for long-distant cases. For subsea environments with high ambient noise, the Matérn model is the most suitable to describe the latent process of communication performance. In the majority of the simulated cases the proposed methodology with model-based universal kriging outperforms the compared techniques. In high ambient noise environments, we observe high prediction uncertainty of communication performance for even short-distant prediction cases. Ongoing work focuses on the analysis of field data [28] and on decentralized methods [30] to implement kriging in multi-robot systems.

References

[1] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, 2005.

[2] Y. Kantaros and M. M. Zavlanos, "Distributed intermittent connectivity control of mobile robot networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3109–3121, 2016.

[3] Y. Sung, A. K. Budhiraja, R. K. Williams, and P. Tokekar, "Distributed assignment with limited communication for multi-robot multi-target tracking," *Autonomous Robots*, vol. 44, no. 1, pp. 57–73, 2020.

[4] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[5] H.-P. Tan, R. Diamant, W. K. Seah, and M. Waldmeyer, "A survey of techniques and challenges in underwater localization," *Ocean Engineering*, vol. 38, no. 14-15, pp. 1663–1676, 2011.

[6] N. Cressie, *Statistics for spatial data*, 2nd ed. Wiley, 1993.

[7] C. E. Rasmussen and C. K. Williams, *Gaussian Processes for Machine Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2006.

[8] R. B. Gramacy, *Surrogates: Gaussian Process Modeling, Design and Optimization for the Applied Sciences*. Chapman Hall/CRC, 2020.

[9] D. Horner and G. Xie, "Data-driven acoustic communication modeling for undersea collaborative navigation," in *IEEE OCEANS*, 2013.

[10] M. Umer, L. Kulik, and E. Tanin, "Spatial interpolation in wireless sensor networks: Localized algorithms for variogram modeling and kriging," *Geoinformatica*, vol. 14, no. 1, p. 101, 2010.

[11] Q. Tao, Y. Zhou, F. Tong, A. Song, and F. Zhang, "Evaluating acoustic communication performance of micro autonomous underwater vehicles in confined spaces," *Frontiers of Information Technology and Electronic Engineering*, vol. 19, no. 8, pp. 1013–1023, 2018.

[12] J. Sun, S. Liu, F. Zhang, A. Song, J. Yu, and A. Zhang, "A kriged compressive sensing approach to reconstruct acoustic fields from measurements collected by underwater vehicles," *IEEE Journal of Oceanic Engineering*, 2020.

[13] M. Malmirchegini and Y. Mostofi, "On the spatial predictability of communication channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 964–978, 2012.

[14] R. Augusto and C. Panazio, "On geostatistical methods for radio environment maps generation under location uncertainty," *Journal of Communication and Information Systems*, vol. 33, no. 1, 2018.

[15] G. P. Kontoudis and D. J. Stilwell, "A comparison of kriging and cokriging for estimation of underwater acoustic communication performance," in *Int. Conf. on Underwater Netw. & Sys.*, 2019, pp. 1–8.

[16] R. J. Adler and J. E. Taylor, *Random fields and geometry*, 1st ed. New York, NY, USA: Springer, 2007.

[17] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 701–717, 2013.

[18] P. J. Diggle and P. J. Ribeiro, *Model-based geostatistics*, 1st ed. New York, NY, USA: Springer, 2007.

[19] N. Cressie and D. M. Hawkins, "Robust estimation of the variogram: I," *Journal of the International Association for Mathematical Geology*, vol. 12, no. 2, pp. 115–125, 1980.

[20] N. Cressie, "Fitting variogram models by weighted least squares," *Journal of the International Association for Mathematical Geology*, vol. 17, no. 5, pp. 563–586, 1985.

[21] P. K. Kitanidis, "Parametric estimation of covariances of regionalized variables," *Journal of the American Water Resources Association*, vol. 23, no. 4, pp. 557–567, 1987.

[22] R. Lark, B. Cullis, and S. Welham, "On spatial prediction of soil properties in the presence of a spatial trend: The empirical best linear unbiased predictor (E-BLUP) with REML," *European Journal of Soil Science*, vol. 57, no. 6, pp. 787–799, 2006.

[23] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[24] A. A. Neath and J. E. Cavanaugh, "The Bayesian information criterion: Background, derivation, and applications," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 199–203, 2012.

[25] R. Webster and M. A. Oliver, *Geostatistics for environmental scientists*, 1st ed. New York, NY, USA: Wiley, 2007.

[26] T. Hengl, G. B. Heuvelink, and A. Stein, "A generic framework for spatial prediction of soil variables based on regression-kriging," *Geoderma*, vol. 120, no. 1-2, pp. 75–93, 2004.

[27] H. Wackernagel, *Multivariate geostatistics: An introduction with applications*, 3rd ed. New York, NY, USA: Springer-Verlag, 2003.

[28] G. P. Kontoudis, S. Krauss, and D. J. Stilwell, "Model-based learning of underwater acoustic communication performance for marine robots," *Robotics and Autonomous Systems*, 2021.

[29] M. B. Porter, "The BELLHOP manual and user's guide: Preliminary draft," 2011.

[30] G. P. Kontoudis and D. J. Stilwell, "Decentralized nested Gaussian processes for multi-robot systems," in *IEEE International Conference on Robotics and Automation*, 2021.