IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 13, 2018, accepted November 7, 2018, date of publication November 22, 2018,
date of current version December 27, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2882890

Online Modeling and Prediction of the Large-Scale Temporal Variation in Underwater Acoustic Communication Channels

WENSHENG SUN, (Student Member, IEEE), AND ZHAOHUI WANG, (Member, IEEE)
Electrical and Computer Engineering Department, Michigan Technological University, Houghton, MI 49931-1295, USA

Corresponding author: Zhaohui Wang (zhaohuiw@mtu.edu)

This work was supported by NSF under grants CNS-1551067 and ECCS-1651135 (CAREER).

ABSTRACT Influenced by environmental conditions, underwater acoustic communication channels exhibit dynamics on various time scales. The channel dynamics within a short transmission duration have been extensively studied in existing research. In this paper, we investigate online modeling and prediction of slowly-varying channel parameters in a long term, by exploiting their inherent temporal correlation and correlation with water environmental conditions. Examples of those parameters include the locally-averaged channel properties within a transmission, such as the average channel-gain-to-noise-power ratio, the fast fading statistics, the average delay spread, and the average Doppler spread. Adopting a data-driven perspective, this paper models the temporal evolution of a slowly-varying channel parameter of interest as the summation of a time-invariant component, a time-varying process that can be explicitly represented by available environmental parameters, and a Markov latent process that describes the contribution from unknown or unmeasurable physical mechanisms. An algorithm is developed to recursively estimate the unknown model parameters and predict the channel parameter of interest, based on the sequentially collected channel measurements and environmental parameters in real time. We further extend the above model and the recursive algorithm to channels that exhibit periodic (a.k.a. *seasonal*) dynamics, by introducing a multiplicative seasonal autoregressive process to model the seasonal correlation. The proposed models and algorithms are evaluated via extensive simulations and data sets from two shallow-water experiments. The experimental results reveal that the average channel-gain-to-noise-power ratio, the fast fading statistics, and the average delay spread can be well predicted.

INDEX TERMS Underwater acoustic channels, large-scale channel dynamics, online channel modeling and prediction, recursive estimation, seasonal channels.

I. INTRODUCTION

Underwater acoustic (UWA) channels exhibit large temporal dynamics. Influenced by environmental conditions, such as water-air interface characteristics, the sound speed profile and the distribution of ambient acoustic sources, the impulse response of an UWA channel could fluctuate on various time scales: seasonal, diurnal, tidal cycles, minutes in the presence of internal waves, and seconds with ocean swells [2]–[4]. Extensive research has been devoted to the statistical modeling and countermeasures of fast channel variation within a transmission that consists of one or multiple packets [5]–[7]. The study on the large-scale channel variation, namely, the temporal evolution of slowly-varying channel parameters over a long term, (e.g., hours, days, months, or years), has been very limited. Examples of those slowly-varying parameters include the locally-averaged channel parameters within a transmission, such as the average channel-gain-to-noise-power ratio (also referred to as *channel SNR*), the average delay spread, the average Doppler spread, and the statistics of fast channel variations. Compared to the fast channel variation, the large-scale channel variation can be attributed to the large-scale change of environmental conditions [7], hence holds a great potential of being predictable.

In this work, we develop a *data-driven* approach for online modeling and prediction of slowly-varying channel parameters in the real-time UWA communication system by exploiting their correlation with water environmental conditions. Prediction of those parameters will allow *proactive adaptation* of higher-level transmission strategies to

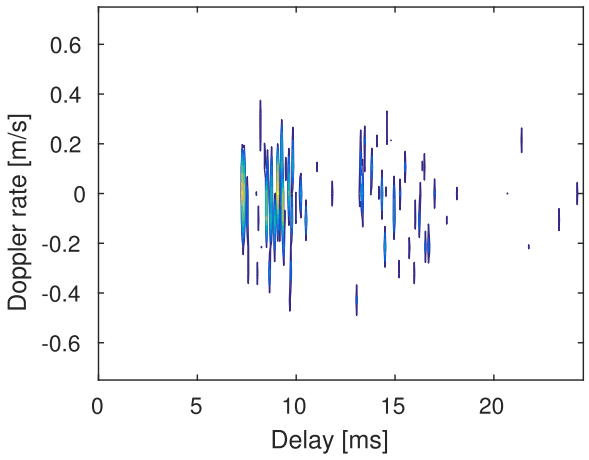


FIGURE 1. An example of the estimated channel impulse response within an OFDM block duration of 129.46 milliseconds in SPACE08.

the channel dynamics. In the sequel, we will first briefly describe our observations on the large-scale channel variation in two field experiments, and then summarize existing approaches to modeling the large-scale channel variation. An overview of this work is presented in the end of this section.

A. OBSERVATION OF THE LARGE-SCALE CHANNEL VARIATION IN FIELD EXPERIMENTS

We introduce the results from two field experiments to illustrate the large-scale channel dynamics. The SPACE08 experiment was conducted in an oceanic environment where a waveform of 1 minute and within the frequency band $[8, 18]$ kHz was transmitted every 2 hours to a receiver located 200 meters away. The waveform consists of 60 short blocks, and each block is modulated by the zero-padded (ZP) orthogonal frequency-division multiplexing (OFDM) technique and has a duration of 129.46 ms. Fig. 1 provides an example of the channel impulse response (CIR) which is estimated based on the received waveform of one OFDM block during a transmission. The KW-AUG14 experiment was conducted in the Keweenaw Waterway near Michigan Tech, August 2014 where a waveform of 8.83 seconds and within the frequency band $[14, 20]$ kHz was transmitted every 15 minutes to a receiver located 312 meters away. The waveform consists of 20 ZP OFDM-modulated blocks, and each of duration 250 ms. Detailed descriptions of the two experiments can be found in Section VII. For each experiment, the CIR can be estimated based on each received OFDM block, and the estimated CIRs within each transmission can be used to calculate the locally-averaged channel parameters of the transmission; rigorous descriptions can be found in Section II-A. In Fig. 2, we plot the evolution of several locally-averaged channel parameters throughout all transmissions in the two field experiments. For both experiments, one can observe that the average channel SNR is negatively correlated with the wind speed and possibly correlated with the temperature. The Nakagami-m fading parameter in KW-AUG14 exhibits negative correlation with the wind speed and the temperature, while the correlation is not obvious in SPACE08. The average root mean square (RMS) delay spread is correlated with the wind speed negatively in SPACE08 while positively in KW-AUG14. Moreover, a diurnal pattern of the slowly-varying channel parameters can be observed in KW-AUG14. Correlations between UWA channel parameters and water environmental conditions have also been revealed in other field experiments; see, e.g., [2], [3], [8]–[10]. In this work, following the terminology in time series analysis, *we refer to the UWA channels with periodic dynamics (e.g., diurnal or monthly) as seasonal channels [11], where the "seasonal cycle" does not necessarily correspond to the seasons in an astronomical year.*

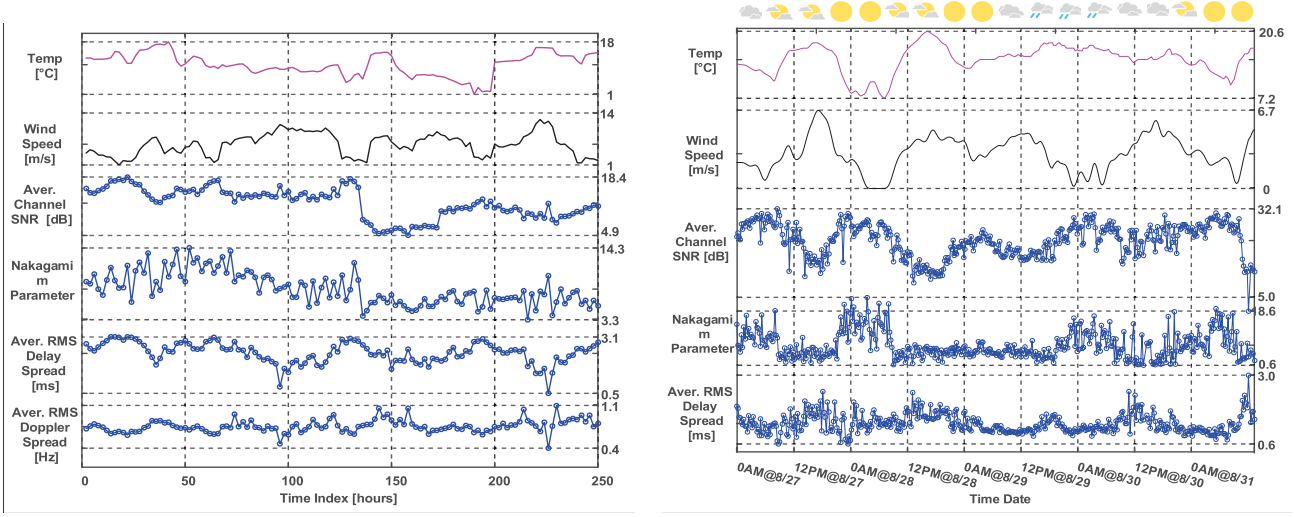


FIGURE 2. Evolution of several slowly-varying channel parameters in two field experiments. The sequences of the average channel SNR are scaled by corresponding transmission power levels. RMS: root mean square. (a) SPACE08. (b) KW-AUG14.

B. EXISTING METHODS FOR MODELING THE LARGE-SCALE UWA CHANNEL VARIATION

Existing methods for UWA channel modeling can be grouped into three categories: wave propagation theory-based modeling, empirical channel modeling and statistical channel modeling. Compared to the latter two approaches, the wave propagation theory-based model [12] yields the highest accuracy. However, it is a deterministic method for a fixed geometry and environmental description, hence cannot accommodate random environmental dynamics.

Using measurements in various water settings, marine engineers have built empirical models that relate the transmission loss and the ambient noise level with water environmental parameters, such as water temperature, salinity, pH, surface wind speed, rainfall rate, and sea state; see, e.g., [13]–[16]. Consider that the acoustic propagation property and the ambient acoustic environment are site-dependent. The empirical model parameters are often computed via curve fitting based on field measurements.

In addition, statistical methods have been widely used to characterize the statistical distribution of the signal transmission loss along each path or an equivalent power loss after combining the signals propagating along multiple paths. Compared to the characterization of channel fast fading [7], [17], [18], studies on the modeling of the large-scale channel variation have been very limited. Based on field measurements, a lognormal distribution of the locally-averaged transmission loss was proposed in [7] and [19], and the possibility of modeling the temporal evolution of the locally-averaged transmission loss as a first-order autoregressive (AR) process was discussed in [20].

It is worth noting that existing channel modeling methods mainly work in an offline manner. They are used either to evaluate the channel conditions before the system deployment, or to characterize the channel behaviors based on field measurements after the system is recovered.

C. OUR WORK

The goal of this work is to develop a method for online modeling and prediction of the large-scale channel variation *during* the system deployment based on sequentially collected channel measurements and water environmental parameters. To this end, a *data-driven* perspective is adopted to exploit the inherent correlation of the large-scale channel variation and its correlation with water environmental conditions.

Specifically, for a slowly-varying channel parameter of interest, we model its temporal evolution as the summation of (i) a time-invariant component, (ii) a time-varying process that can be explicitly represented by available water environmental parameters, and (iii) a hidden Markov latent process which accounts for the contribution from unknown or unmeasurable physical mechanisms. After casting the evolution model into a state-space representation, and following the maximum likelihood (ML) principle and the expectation-maximization (EM) algorithm [21], a low-complexity algorithm is developed to *recursively* estimate the unknown model parameters based on sequentially obtained channel measurements and environmental parameters during the system operation, which then allows to predict the slowly-varying channel parameter in the near future. The proposed modeling method and the recursive algorithm are further extended to seasonal channels, where *a multiplicative seasonal AR process* [11] is introduced to model the seasonal correlation.

The effectiveness of the proposed models and recursive algorithms are evaluated via simulations and data sets from two shallow-water experiments, the SPACE08 and the KW-AUG14. The slowly-varying channel parameters that are examined using the experimental data sets include the average channel SNR, the fast fading statistics, the average RMS delay spread, and the average RMS Doppler spread. The results reveal that superior modeling and prediction performance can be achieved by exploiting the correlation between the large-scale channel variation and water environmental parameters as well as the seasonal correlation in seasonal channels.

Remark 1: The developed algorithms can be applied to real-time operating UWA communication systems. Specifically, the model parameters can be updated recursively in time step-by-step based on newly obtained channel measurements during recent acoustic transmissions as well as newly obtained environmental parameters.[1] The updated model allows the prediction of the large-scale channel variation based on the forecast of environmental conditions. The prediction could guide higher-level proactive adaptation of future transmission strategies, such as the transmission schedule, the transmission power and rate, and the modulation scheme [22]. It has been shown in an early study [23] that even with moderate channel prediction performance, proactive adaptation of the transmission schedule improves energy efficiency more than 20% than a benchmark method that transmits each packet upon its arrival with minimal transmission power that meets a predetermined SNR threshold.

The rest of the paper is organized as follows. The data-driven modeling method is presented in Section II. A recursive algorithm for the model parameter estimation is developed in Section III. Extension of the proposed model and the recursive algorithm to seasonal channels is presented in Section IV. The model order selection for practical UWA channels is discussed in Section V. Simulations and experimental data processing results are presented in Sections VI and VII, respectively. Conclusions are drawn in Section VIII.

Notation: Bold upper case letters and lower case letters are used to denote matrices and column vectors, respectively. $\mathbf{A}^{\mathrm{T}}$ denotes the transpose of matrix $\mathbf{A}$. $\det(\mathbf{A})$ denotes the determinant of matrix $\mathbf{A}$. $[\mathbf{a}]_m$ denotes the mth element of vector $\mathbf{a}$, and $[\mathbf{A}]_{m,k}$ denotes the (m, k)th element of matrix $\mathbf{A}$. $\mathbb{E}[x]$ denotes the expectation of random variable x.

[1] The environmental parameters can be collected by sensors equipped on the communication nodes (e.g., surface buoys and underwater nodes), or sent from a remote control center to surface buoys via radio-frequency links.

II. A DATA-DRIVEN METHOD FOR MODELING LARGE-SCALE CHANNEL VARIATIONS

In this section, we will first provide several examples of slowly-varying parameters in UWA channels, and then develop a data-driven model for the temporal evolution of a slowly-varying channel parameter of interest.

A. SLOWLY-VARYING PARAMETERS IN UWA CHANNELS

The UWA channel features multiple time-varying sound propagation paths. Denote $N_{\rm pa}$ as a generic representation of the number of paths. The CIR at time t is

$$h(t;\tau)=\sum_{p=1}^{N_{\rm pa}} A_p(t)\delta(\tau-\tau_p(t)), \tag{1}$$

where $A_p(t)$ and $\tau_p(t)$ are the time-varying amplitude and delay of the pth path, respectively.

For an UWA transmission with $N_{\rm bl}$ short blocks, the channel is often assumed block-stationary and could change from one block to another. For the ℓth block in the kth transmission, the CIR can be approximated as

$$h_{k,\ell}(t;\tau)=\sum_{p=1}^{N_{{\rm pa},k,\ell}} A_{p,k,\ell}\delta(\tau-(\tau_{p,k,\ell}-a_{p,k,\ell}t)), \tag{2}$$

where $N_{{\rm pa},k,\ell}$ denotes the number of paths, and for each path, e.g., the pth path, the amplitude is approximated as a constant $A_{p,k,\ell}$, and the delay variation is approximated by a first-order polynomial $(\tau_{p,k,\ell}-a_{p,k,\ell}t)$ with $\tau_{p,k,\ell}$ being the initial delay and $a_{p,k,\ell}$ being the Doppler rate, respectively. Estimation of the path parameters is typically performed in each block based on training symbols. An example of the estimated CIR based on the pilot subcarriers in one OFDM block in the SPACE08 experiment is depicted in Fig. 1. In addition, the channel SNR in the ℓth block of the kth transmission can be denoted as

$$\zeta_{k,\ell}:=\frac{1}{N_{0,k,\ell}}\sum_{p=1}^{N_{{\rm pa},k,\ell}}|A_{p,k,\ell}|^2, \tag{3}$$

where $N_{0,k,\ell}$ is the noise power in the ℓth block.

Different from the fast variation of path parameters, the structure of the CIR could change slowly from one transmission to another in accordance with environmental conditions. Corresponding to the multiple ($N_{\rm bl}$) individual CIRs in the kth transmission, several examples of slowly-varying channel parameters are in the following.

- *The average channel SNR* in decibel (dB), defined as

$$\overline{\zeta}_{\rm dB}[k]:=\frac{1}{N_{\rm bl}}\sum_{\ell=1}^{N_{\rm bl}}10\log_{10}(\zeta_{k,\ell}). \tag{4}$$

- *The fast fading statistics.* Despite the fast variation of path parameters within one transmission, the statistics of the fast variation could change slowly from one transmission to another. In this work, we adopt a Nakagami-m channel fading model [24], and examine the temporal evolution of the fading parameter m from one transmission to another. For the kth transmission, the fading parameter m can be estimated based on the block SNRs $\{\zeta_{k,1},\cdots,\zeta_{k,N_{\rm bl}}\}$ that follow a corresponding Gamma distribution.
- *The average RMS delay spread* that quantifies the channel dispersion in delay [25],

$$\tau_{\rm spread}[k] := \frac{1}{N_{\rm bl}}\sum_{\ell=1}^{N_{\rm bl}}\frac{\sqrt{\sum_{p=1}^{N_{{\rm pa},k,\ell}}|A_{p,k,\ell}|^2(\tau_{p,k,\ell}-\bar{\tau}_{k,\ell})^2}}{\sqrt{\sum_{p=1}^{N_{{\rm pa},k,\ell}}|A_{p,k,\ell}|^2}}, \tag{5}$$

with

$$\bar{\tau}_{k,\ell}:=\frac{\sum_{p=1}^{N_{{\rm pa},k,\ell}}|A_{p,k,\ell}|^2\tau_{p,k,\ell}}{\sum_{p=1}^{N_{{\rm pa},k,\ell}}|A_{p,k,\ell}|^2}. \tag{6}$$

- *The average RMS Doppler spread* that quantifies the channel dispersion in the Doppler rate, denoted by $a_{\rm spread}[k]$, which can be similarly defined as $\tau_{\rm spread}[k]$ through replacing $\tau_{p,k,\ell}$ by $a_{p,k,\ell}$ in (5) and (6).

In the next subsection we will develop a data-driven method to model the temporal evolution of a slowly-varying channel parameter of interest. Estimation of the model parameters will be pursued in Sections III and IV.

B. A DATA-DRIVEN MODEL FOR SLOWLY-VARYING CHANNEL PARAMETERS

Consider the temporal evolution of a slowly-varying channel parameter of interest, which is represented by process $\{\alpha[k]\}$, with k being an integer time index. We model the process $\{\alpha[k]\}$ as the summation of a time-invariant component γ_0, a time-varying process $\{g[k]\}$ that can be explicitly represented by available and relevant water environmental parameters, and a latent process $\{x[k]\}$ that describes the contribution from unknown or unmeasurable physical mechanisms, namely,

$$\alpha[k]=\gamma_0+g[k]+x[k],\qquad \forall k. \tag{7}$$

Specifically about the processes $\{g[k]\}$ and $\{x[k]\}$,

- The process $\{g[k]\}$ can be taken as a function of L types of available and relevant environmental parameters $\{\phi_\ell[k];\ell=1,\cdots,L\}$. Consider the potentially nonlinear relationship between the slowly-varying channel parameter and water environmental parameters [13]–[16]. The function can be represented by the Maclaurin series expansion,

$$\begin{aligned} g[k]=&\sum_{\ell=1}^{L}c_\ell\phi_\ell[k]+\sum_{\ell_1=1}^{L}\sum_{\ell_2=1}^{L}c_{\ell_1,\ell_2}\phi_{\ell_1}[k]\phi_{\ell_2}[k]\\ &+\sum_{\ell_1=1}^{L}\sum_{\ell_2=1}^{L}\sum_{\ell_3=1}^{L}c_{\ell_1,\ell_2,\ell_3}\phi_{\ell_1}[k]\phi_{\ell_2}[k]\phi_{\ell_3}[k]+\cdots \end{aligned} \tag{8}$$

where the expansion coefficients are unknown and could be slowly time-varying. Estimation of the expansion coefficients based on channel measurements and environmental parameters is challenged by their infinite dimensionality.

To make the problem tractable, a finite number of important summands on the right side of (8) can be selected to approximate the function. Specifically, we include the infinite elements on the right side of (8) in a set $\mathcal{E}[k] := \{\phi_1[k], \cdots, \phi_L[k], \phi_1^2[k], \phi_1[k]\phi_2[k], \cdots\}$, and denote $\mathcal{I}$ as an index set of N_u important elements within $\mathcal{E}[k], \forall k$. The important elements can form a finite set $\mathcal{U}[k] := \{u_1[k], \cdots, u_{N_u}[k]\}$, which yields the approximation,

$$g[k] \approx \sum_{n=1}^{N_u} b_n u_n[k], \tag{9}$$

where $\{b_n\}$ denote the coefficients of the N_u important elements.

- The latent process $\{x[k]\}$ is modeled as a Markov process with memory length of P,

$$x[k] = \sum_{p=1}^{P} a_p x[k-p] + w[k], \tag{10}$$

where the coefficients $\{a_p\}$ are unknown and could be slowly time-varying, and the process noise $w[k]$ follows a zero-mean Gaussian distribution with variance σ_w^2, namely, $w[k] \sim \mathcal{N}(0, \sigma_w^2)$.

The latent process order P and the index set $\mathcal{I}$ of important elements in $\mathcal{E}[k]$ can be determined via a model-order selection criterion. A detailed discussion is presented in Section V.

Denote $y[k]$ as the measurement of the slowly-varying channel parameter at time k. We have

$$y[k] = \gamma_0 + x[k] + g[k] + v[k], \tag{11}$$

where $v[k]$ is an equivalent noise term which consists of modeling inaccuracy and the measurement noise, and is assumed $v[k] \sim \mathcal{N}(0, \sigma_v^2)$, independent from the process noise $w[k]$ in (10).

Define $\mathbf{a} := [a_1, \cdots, a_P]^{\mathrm{T}}$, $\mathbf{b} := \left[b_1, \cdots, b_{N_u}\right]^{\mathrm{T}}$, $\mathbf{x}[k] := [x[k], \cdots, x[k-P+1]]^{\mathrm{T}}$, and $\mathbf{u}[k] := \left[u_1[k], \cdots, u_{N_u}[k]\right]^{\mathrm{T}}$. The system model can be compactly represented as

$$x[k] = \mathbf{a}^{\mathrm{T}}\mathbf{x}[k-1] + w[k], \tag{12a}$$
$$y[k] = \gamma_0 + x[k] + \mathbf{b}^{\mathrm{T}}\mathbf{u}[k] + v[k]. \tag{12b}$$

Define $\mathbf{w}[k] := [w[k], 0, \cdots, 0]^{\mathrm{T}}$, $\mathbf{h} := [1, 0, \cdots, 0]^{\mathrm{T}}$, and

$$\mathbf{A} := \begin{bmatrix} a_1 & a_2 & \cdots & a_{P-1} & a_P \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

We have the state-space representation of the system model,

$$\mathbf{x}[k] = \mathbf{A}\mathbf{x}[k-1] + \mathbf{w}[k], \tag{13a}$$
$$y[k] = \gamma_0 + \mathbf{h}^{\mathrm{T}}\mathbf{x}[k] + \mathbf{b}^{\mathrm{T}}\mathbf{u}[k] + v[k]. \tag{13b}$$

Should the parameters in the set $\mathbf{\Theta} := \{\gamma_0, \mathbf{a}, \mathbf{b}, \sigma_w^2, \sigma_v^2\}$ be known, the latent process can be tracked via the Kalman filter [26]. In the next section, we will develop a recursive algorithm to estimate the unknown model parameters while tracking the latent process based on the measurements $\{y[k]\}$ and the environmental parameter vectors $\{\mathbf{u}[k]\}$. The estimated model parameters allow multiple-step-ahead prediction of the slowly-varying channel parameter. For notation convenience, in the sequel we use $x[k]$ and x_k, $y[k]$ and y_k, $\mathbf{x}[k]$ and $\mathbf{x}_k$ interchangeably, and denote $\mathbf{x}_{k_1}^{k_2} := \{x_{k_1}, \cdots, x_{k_2}\}$ and $\mathbf{y}_{k_1}^{k_2} := \{y_{k_1}, \cdots, y_{k_2}\}$.

III. A RECURSIVE ALGORITHM FOR CHANNEL MODELING AND PREDICTION

Following the ML principle [26], the unknown parameters in $\mathbf{\Theta}$ could be estimated at each time step (e.g., time k) by maximizing the log-likelihood function of the complete data set, $L_k(\mathbf{\Theta}) := \ln f(\mathbf{y}_0^k, \mathbf{x}_{-1}, \mathbf{x}_0^k|\mathbf{\Theta})$. However, note that the latent process $\{x_{k'}\}$ is not observable. The EM algorithm [21] can be applied to estimate the unknown parameters iteratively through an expectation step and a maximization step. Specifically, at time k,

- *Expectation:* Given a parameter set estimation $\hat{\mathbf{\Theta}}$, the expectation of the log-likelihood function can be approximated as

$$\mathbb{E}[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}] = \int\int \left[\ln f(\mathbf{y}_0^k, \mathbf{x}_{-1}, \mathbf{x}_0^k|\mathbf{\Theta})\right] \times f(\mathbf{x}_{-1}, \mathbf{x}_0^k|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}) d\mathbf{x}_{-1} d\mathbf{x}_0^k. \tag{14}$$

- *Maximization:* The parameter set estimation can be updated as

$$\hat{\mathbf{\Theta}}^{(\text{new})} = \arg \max_{\mathbf{\Theta}} \mathbb{E}[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}]. \tag{15}$$

The iterative operation terminates when the number of iterations reaches a pre-determined value or the change of the parameter set estimation is less than a pre-determined threshold.

Note that (14) can be decomposed as

$$\mathbb{E}[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}] = \int [\ln f(\mathbf{x}_{-1}|\mathbf{\Theta})] f(\mathbf{x}_{-1}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}) d\mathbf{x}_{-1} + \sum_{k'=0}^{k} \int [\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \mathbf{\Theta})] \times f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}) dx_{k'} d\mathbf{x}_{k'-1}. \tag{16}$$

The expectation $\mathbb{E}[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}]$ is computed based on the probability density function (PDF) $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}), \forall k' \le k$. For a given estimation $\hat{\mathbf{\Theta}}$, finding $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}})$ requires to process all the data points. Hence, the original EM algorithm is not amenable to online implementation.

We next propose an approximation to $\mathbb{E}[L_k(\boldsymbol{\Theta})|\hat{\boldsymbol{\Theta}}]$ that enables the development of a low-complexity recursive algorithm for the model parameter estimation and channel tracking.

A. APPROXIMATION FOR RECURSIVE OPERATION

The approximation to $\mathbb{E}[L_k(\boldsymbol{\Theta})|\hat{\boldsymbol{\Theta}}]$ is made in several steps.

First, we approximate the expectation in (16) by

$$\begin{aligned}\mathbb{E}[L_k(\boldsymbol{\Theta})|\hat{\boldsymbol{\Theta}}] \approx{}& \ln f(\mathbf{x}_{-1}|\boldsymbol{\Theta})] \\ &+ \sum_{k'=0}^{k} \int [\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})] \\ &\times f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}}) dx_{k'} d\mathbf{x}_{k'-1}, \quad (17)\end{aligned}$$

where the expectation of $[\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})]$ is performed with respect to $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}})$ instead of $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k}, \hat{\boldsymbol{\Theta}})$. This removes the dependence of $\{x_{k'}, \mathbf{x}_{k'-1}\}$ on future measurements.

Secondly, denote $\hat{\boldsymbol{\Theta}}_{k'}$ as the parameter set estimation at time k'. We make a further approximation to (17) through replacing $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}})$ by $f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}}_{k'})$, $\forall k' < k$, namely,

$$\begin{aligned}\mathbb{E}[L_k(\boldsymbol{\Theta})|\hat{\boldsymbol{\Theta}}] \approx{}& \ln f(\mathbf{x}_{-1}|\boldsymbol{\Theta}) \\ &+ \sum_{k'=0}^{k-1} \int [\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})] \\ &\times f(x_{k'}, \mathbf{x}_{k'-1}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}}_{k'}) dx_{k'} d\mathbf{x}_{k'-1} \\ &+ \int [\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})] \\ &\times f(x_k, \mathbf{x}_{k-1}|\mathbf{y}_0^{k}, \hat{\boldsymbol{\Theta}}) dx_k d\mathbf{x}_{k-1}. \quad (18)\end{aligned}$$

The approximations in (17) and (18) enable recursive computation of the summands on the right side of (18).

Thirdly, note that the joint PDF $f(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}})$ can be decomposed as

$$\begin{aligned}&f(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) \\ &\quad = f(\mathbf{x}_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) \delta(\mathbf{x}_k, \mathbf{x}_{k-1}) \\ &\quad = \frac{1}{c_0} f(\mathbf{x}_k, \mathbf{x}_{k-1}, y_k|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) \delta(\mathbf{x}_k, \mathbf{x}_{k-1}) \\ &\quad = \frac{1}{c_0} f(y_k|\mathbf{x}_k, \hat{\boldsymbol{\Theta}}) f(\mathbf{x}_k|\mathbf{x}_{k-1}, \hat{\boldsymbol{\Theta}}) \\ &\qquad \times f(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) \delta(\mathbf{x}_k, \mathbf{x}_{k-1}), \quad (19)\end{aligned}$$

where c_0 is a normalization constant, and the function $\delta(\mathbf{x}_k, \mathbf{x}_{k-1})$ is introduced to constrain the equity of common elements in $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$. We approximate the joint PDF by

$$\begin{aligned}\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) :={}& \frac{1}{c_0'} f(y_k|\mathbf{x}_k, \hat{\boldsymbol{\Theta}}) \\ &\times f(\mathbf{x}_k|\mathbf{x}_{k-1}, \hat{\boldsymbol{\Theta}}) \tilde{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1}) \delta(\mathbf{x}_k, \mathbf{x}_{k-1}), \quad (20)\end{aligned}$$

where c_0' is a normalization constant, and the approximation is made through replacing $f(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}})$ in (19) by $\tilde{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1})$ in (20), with $\tilde{f}(\mathbf{x}_{k'}|\mathbf{y}_0^{k'}, \hat{\boldsymbol{\Theta}}_{k'})$ defined as the marginalization PDF of $\mathbf{x}_{k'}$ with respect to $\tilde{f}(x_{k'}, \mathbf{x}_{k'-1}|y_{k'}, \mathbf{y}_0^{k'-1}, \hat{\boldsymbol{\Theta}}_{k'})$, $\forall k'$.

Finally, based on (18) and (20), the expectation $\mathbb{E}[L_k(\boldsymbol{\Theta})|\hat{\boldsymbol{\Theta}}]$ is approximated by $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}})$ which is recursively defined as

$$\begin{aligned}Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}) = \lambda Q_{k-1}(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_{k-1}) &+ \int [\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})] \\ &\times \tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}) d\mathbf{x}_{k-1} dx_k, \quad (21)\end{aligned}$$

where $\lambda \in (0, 1]$ is a forgetting factor that accounts for the temporal variation of unknown model parameters.

Based on (20) and (21), a recursive algorithm will be developed for the model parameter estimation and channel tracking, while at each time step, operations similar to the expectation and the maximization in the EM algorithm are iteratively performed.

B. A LOW-COMPLEXITY RECURSIVE ALGORITHM

Denote $\hat{\boldsymbol{\Theta}}_k^{(i)} = \{\hat{\gamma}_{0,k}^{(i)}, \hat{\mathbf{a}}_k^{(i)}, \hat{\mathbf{b}}_k^{(i)}, \hat{\sigma}_{w,k}^{2,(i)}, \hat{\sigma}_{v,k}^{2,(i)}\}$ as the parameter set estimation in the ith iteration at time k. The function $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$ is computed through finding the expectation of $[\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})]$ with respect to the PDF $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_k^{(i)})$ (c.f. (21)). The parameter set estimation can then be updated as $\hat{\boldsymbol{\Theta}}_k^{(i+1)} = \arg\max_{\boldsymbol{\Theta}} Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$.

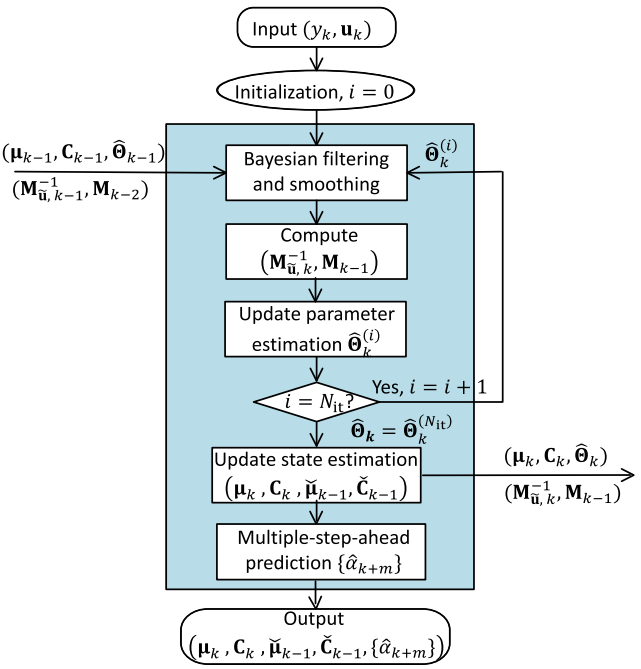


FIGURE 3. The proposed low-complexity recursive algorithm at time k.

At the outset, the proposed low-complexity recursive algorithm is depicted in Fig. 3. Denote $\tilde{f}(\mathbf{x}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k) = \mathcal{N}(\boldsymbol{\mu}_k, \mathbf{C}_k)$ (c.f. (20)). At time k, the algorithm takes the PDF $\tilde{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1}) = \mathcal{N}(\boldsymbol{\mu}_{k-1}, \mathbf{C}_{k-1})$, the parameter set estimation $\hat{\boldsymbol{\Theta}}_{k-1}$, auxiliary quantities $\{\mathbf{M}_{k-2}, \mathbf{M}_{\tilde{\mathbf{u}},k-1}^{-1}\}$ (to be

defined shortly; computed at time $(k-1)$), the measurement y_k and the environmental parameter vector $\mathbf{u}_k$ as input. Set $\hat{\boldsymbol{\Theta}}_k^{(0)} = \hat{\boldsymbol{\Theta}}_{k-1}$. Given the parameter set estimation $\hat{\boldsymbol{\Theta}}_k^{(i)}$, the Kalman filtering and smoothing can be performed to compute the expectations of quantities in $[\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})]$ with respect to the PDF $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_k^{(i)})$, namely, the second summand in (21). The parameter set estimation can then be updated through maximizing $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$. The updated parameter estimation can then be used for the Kalman filtering and smoothing in the next iteration. The iterative operation terminates when the number of iterations reaches a pre-determined threshold N_{it}. We set $\hat{\boldsymbol{\Theta}}_k^{(N_{\text{it}})} = \hat{\boldsymbol{\Theta}}_k$ as the final parameter set estimation at time k. Based on $\hat{\boldsymbol{\Theta}}_k$, the PDF $\tilde{f}(\mathbf{x}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k) = \mathcal{N}(\boldsymbol{\mu}_k, \mathbf{C}_k)$ is computed via the Kalman filtering. The PDF $\tilde{f}(\mathbf{x}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k)$, the parameter set estimation $\hat{\boldsymbol{\Theta}}_k$ and $\{\mathbf{M}_{k-1}, \mathbf{M}_{\tilde{\mathbf{u}},k}^{-1}\}$ that are computed at time k, will be used for the recursive operation at time $(k+1)$. Additionally, based on the parameter set estimation $\hat{\boldsymbol{\Theta}}_k$ and the state estimation $\boldsymbol{\mu}_k$, multiple-step-ahead prediction of the slowly-varying channel parameter can be achieved.

Next, we describe in details the component of the recursive and iterative parameter estimation, the Kalman filtering and smoothing, and the multiple-step-ahead prediction.

1) RECURSIVE AND ITERATIVE PARAMETER ESTIMATION

The parameter estimation can be updated by maximizing $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$. Note that $f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta}) = f(y_k|x_k, \boldsymbol{\Theta})f(x_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})$. Substitute $f(y_k|x_k, \boldsymbol{\Theta}) = \mathcal{N}(\gamma_0 + x_k + \mathbf{b}^{\mathrm{T}}\mathbf{u}_k, \sigma_v^2)$ and $f(x_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{a}^{\mathrm{T}}\mathbf{x}_{k-1}, \sigma_w^2)$ into the log-likelihood function in (21). Denote $\tilde{\mathbf{u}}_k := [1, \mathbf{u}_k^{\mathrm{T}}]^{\mathrm{T}}$. Set the partial derivative of $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$ with respect to each unknown parameter to zero. A set of recursive equations can be obtained; see the detailed derivation in Appendix,

$$\hat{\mathbf{a}}_k^{(i+1)} = \hat{\mathbf{a}}_{k-1} + \mathbf{M}_{k-1}^{-1} \times \left(\mathbb{E}[x_k\mathbf{x}_{k-1}] - \mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\hat{\mathbf{a}}_{k-1}\right), \tag{22a}$$

$$\hat{\sigma}_{w,k}^{2,(i+1)} = \hat{\sigma}_{w,k-1}^2 + \frac{1-\lambda}{1-\lambda^k} \times \left\{\mathbb{E}\left[\left(x_k - \hat{\mathbf{a}}_k^{(i+1),\mathrm{T}}\mathbf{x}_{k-1}\right)^2\right] - \hat{\sigma}_{w,k-1}^2\right\}, \tag{22b}$$

$$\begin{bmatrix}\hat{\gamma}_{0,k}^{(i+1)} \\ \hat{\mathbf{b}}_k^{(i+1)}\end{bmatrix} = \begin{bmatrix}\hat{\gamma}_{0,k-1} \\ \hat{\mathbf{b}}_{k-1}\end{bmatrix} + \frac{\mathbf{M}_{\tilde{\mathbf{u}},k}^{-1}\tilde{\mathbf{u}}_k}{\lambda + \tilde{\mathbf{u}}_k^{\mathrm{T}}\mathbf{M}_{\tilde{\mathbf{u}},k}^{-1}\tilde{\mathbf{u}}_k} \times \left(y_k - \mathbb{E}[x_k] - \hat{\gamma}_{0,k-1} - \hat{\mathbf{b}}_{k-1}^{\mathrm{T}}\mathbf{u}_k\right), \tag{22c}$$

$$\hat{\sigma}_{v,k}^{2,(i+1)} = \hat{\sigma}_{v,k-1}^2 + \frac{1-\lambda}{1-\lambda^{k+1}} \times \left\{\mathbb{E}\left[(y_k - x_k - \hat{\gamma}_{0,k}^{(i+1)} - \hat{\mathbf{b}}_k^{(i+1),\mathrm{T}}\mathbf{u}_k)^2\right] - \hat{\sigma}_{v,k-1}^2\right\}, \tag{22d}$$

with two matrices defined as

$$\mathbf{M}_{k-1} := \lambda\mathbf{M}_{k-2} + \mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}], \tag{23a}$$

$$\mathbf{M}_{\tilde{\mathbf{u}},k} := \lambda\mathbf{M}_{\tilde{\mathbf{u}},k-1} + \tilde{\mathbf{u}}_k\tilde{\mathbf{u}}_k^{\mathrm{T}}. \tag{23b}$$

The expectations in (22) and (23) are performed with respect to $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_k^{(i)})$ (c.f. (21)).

2) KALMAN FILTERING AND SMOOTHING

Computation of the expectations in (22) and (23) requires the marginalization of the joint PDF $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_k^{(i)})$ with respect to $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$, respectively. Denote the marginal PDFs as $\tilde{f}(\mathbf{x}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\boldsymbol{\mu}_k^{(i)}, \mathbf{C}_k^{(i)})$ and $\breve{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\breve{\boldsymbol{\mu}}_{k-1}^{(i)}, \breve{\mathbf{C}}_{k-1}^{(i)})$. Given the expansion of the joint PDF in (20), the marginalization can be performed through the Kalman filtering and smoothing [26], as detailed next.

Define $\hat{\mathbf{A}}_k^{(i)}$ and $\hat{\mathbf{C}}_{\mathbf{w},k}^{(i)}$ as the matrices corresponding to $\hat{\mathbf{a}}_k^{(i)}$ and $\hat{\sigma}_{w,k}^{2,(i)}$, respectively. Based on $\tilde{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1})$ and the system model in (13), the mean and the covariance matrix of $\mathbf{x}_k$ in the marginal PDF $\tilde{f}(\mathbf{x}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ can be formulated as

$$\boldsymbol{\mu}_k^{(i)} = \hat{\mathbf{A}}_k^{(i)}\boldsymbol{\mu}_{k-1} + \mathbf{k}_k^{(i)}(y_k - \hat{\mathbf{a}}_k^{(i),\mathrm{T}}\boldsymbol{\mu}_{k-1} - \hat{\gamma}_{0,k}^{(i)} - \hat{\mathbf{b}}_k^{(i),\mathrm{T}}\mathbf{u}_k), \tag{24a}$$

$$\mathbf{C}_k^{(i)} = (\mathbf{I} - \mathbf{k}_k^{(i)}\mathbf{h}^{\mathrm{T}})\mathbf{P}_k^{(i)}, \tag{24b}$$

where the Kalman gain $\mathbf{k}_k^{(i)} = \mathbf{P}_k^{(i)}\mathbf{h}\left(\hat{\sigma}_{v,k}^{2,(i)} + \mathbf{h}^{\mathrm{T}}\mathbf{P}_k^{(i)}\mathbf{h}\right)^{-1}$ and the prediction mean square error (MSE) matrix $\mathbf{P}_k^{(i)} = \hat{\mathbf{A}}_k^{(i)}\mathbf{C}_{k-1}\hat{\mathbf{A}}_k^{(i),\mathrm{T}} + \hat{\mathbf{C}}_{\mathbf{w}}^{(i)}$. We further have $\mathbb{E}[\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \mathbf{C}_k^{(i)} + \boldsymbol{\mu}_k^{(i)}\boldsymbol{\mu}_k^{(i),\mathrm{T}}$.

The marginal PDF $\breve{f}(\mathbf{x}_{k-1}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ can be obtained via the one-step backward smoothing, with the mean and the covariance matrix formulated as

$$\breve{\boldsymbol{\mu}}_{k-1}^{(i)} = \boldsymbol{\mu}_{k-1} + \mathbf{J}_{k-1}^{(i)}(\boldsymbol{\mu}_k^{(i)} - \hat{\mathbf{A}}_k^{(i)}\boldsymbol{\mu}_{k-1}), \tag{25a}$$

$$\breve{\mathbf{C}}_{k-1}^{(i)} = \mathbf{C}_{k-1} + \mathbf{J}_{k-1}^{(i)}(\mathbf{C}_k^{(i)} - \mathbf{P}_k^{(i)})\mathbf{J}_{k-1}^{(i),\mathrm{T}}, \tag{25b}$$

where the gain matrix $\mathbf{J}_{k-1}^{(i)} = \mathbf{C}_{k-1}\hat{\mathbf{A}}_k^{(i),\mathrm{T}}(\mathbf{P}_k^{(i)})^{-1}$. We further have $\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \breve{\mathbf{C}}_{k-1}^{(i)} + \breve{\boldsymbol{\mu}}_{k-1}^{(i)}\breve{\boldsymbol{\mu}}_{k-1}^{(i),\mathrm{T}}$.

Based on the joint PDF $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_k^{(i)})$, the correlation between $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ can be obtained as

$$\mathbb{E}[\mathbf{x}_k\mathbf{x}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \mathbf{C}_k^{(i)}\mathbf{J}_{k-1}^{(i),\mathrm{T}} + \boldsymbol{\mu}_k^{(i)}\breve{\boldsymbol{\mu}}_{k-1}^{(i),\mathrm{T}}. \tag{26}$$

The expectations $\mathbb{E}[x_k|\hat{\boldsymbol{\Theta}}_k^{(i)}]$, $\mathbb{E}[x_k\mathbf{x}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}]$, and $\mathbb{E}[x_k^2|\hat{\boldsymbol{\Theta}}_k^{(i)}]$ to be used in (22) can be extracted from $\mathbb{E}[\mathbf{x}_k|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \boldsymbol{\mu}_k^{(i)}$, $\mathbb{E}[\mathbf{x}_k\mathbf{x}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}]$, and $\mathbb{E}[\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}]$, respectively.

3) MULTIPLE-STEP-AHEAD PREDICTION

Based on the parameter set estimation $\hat{\boldsymbol{\Theta}}_k$ and the state estimation $\boldsymbol{\mu}_k$ (denoted next also as $\hat{\mathbf{x}}_k$), the m-step-ahead prediction of the slowly-varying channel parameter can be recursively computed based on the system model in (12). Specifically,

$$\hat{x}_{k+m} = \hat{\mathbf{a}}_k^{\mathrm{T}}\hat{\mathbf{x}}_{k+m-1}, \tag{27a}$$

$$\hat{\alpha}_{k+m} = \hat{\gamma}_{0,k} + \hat{x}_{k+m} + \hat{\mathbf{b}}_k^{\mathrm{T}}\mathbf{u}_{k+m}, \tag{27b}$$

for $m = 1, 2, \cdots$, where $\mathbf{u}_{k+m}$ can be obtained from meteorological forecast sources, e.g., [27].

Remark 2: Although this work assumes periodic channel measurements, the proposed model and the recursive algorithm can be applied to the scenario with non-periodic channel measurements through replacing the discrete-time state-space model in (13) by a continuous-time state-space model (c.f. [26, Ch. 9]).

Remark 3: The proposed model and the recursive algorithm subsume a linear regression method that models the temporal evolution of the slowly-varying channel parameter only based on available environmental parameters without introducing the latent process, namely, the model in (11) degrades to $y_k = \gamma_0 + g_k + w_k$. The model parameters γ_0 and $\mathbf{b}$ can be recursively estimated via (22c).

4) COMPUTATIONAL COMPLEXITY

The computational complexity of the proposed algorithm at each time step is analyzed in the following. At the outset, we would like to note that in practical systems, the values of P and N_u are typically very small. In Section VII, $N_{\text{it}} = 20$ and $N_u = 2$ are used for the experimental data processing, and the value of P varies from 1 to 4 for different channel parameters.

- *Kalman filtering and smoothing:* For (24), calculation of the Kalman gain vector $\mathbf{k}_k^{(i)}$ of length P has (P^2+2P) arithmetic multiplications (AMs), (P^2+P+2) arithmetic additions (AAs) and 1 arithmetic division (AD). Calculation of the MSE matrix $\mathbf{P}_k^{(i)}$ of size $(P \times P)$ has $(2P^3)$ AMs and $(2P^3+2P^2)$ AAs. Eq. (24a) has (P^2+2P+N_u) AMs and (P^2+3P+N_u+4) AAs, and Eq. (24b) has (P^3+P^2) AMs and (P^3+2P^2) AAs. The total computations associated with (24) include $(3P^3+3P^2+4P+N_u)$ AMs, $(3P^3+6P^2+4P+4+N_u)$ AAs and 1 AD. For (25), calculation of the gain matrix $\mathbf{J}_{k-1}^{(i)}$ of size $(P \times P)$ has $(2P^3)$ AMs and $(2P^3)$ AAs for the matrix multiplication and a complexity of $\mathcal{O}(P^3)$ for the inversion of matrix $\mathbf{P}_k^{(i)}$. Eq. (25a) has $(2P^2)$ AMs and $(2P^2+4P)$ AAs, and Eq. (25b) has $(2P^3)$ AMs and $(2P^3+4P^2)$ AAs. The total computations associated with (25) include $(4P^3+2P^2)$ AMs, $(4P^3+6P^2+4P)$ AAs and a $(P \times P)$ matrix inversion with complexity $\mathcal{O}(P^3)$. In addition, calculation of $\mathbb{E}[\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}}|\hat{\mathbf{\Theta}}_k^{(i)}]$ and of $\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}|\hat{\mathbf{\Theta}}_k^{(i)}]$ each has (P^2) AMs and $(2P^2)$ AAs. Calculation of the correlation matrix in (26) has (P^3+P^2) AMs and (P^3+2P^2) AAs. Therefore, the total computations for Kalman filtering and smoothing include $(8P^3+8P^2+4P+N_u)$ AMs, $(8P^3+18P^2+8P+6+N_u)$ AAs, 1 AD and the inversion of a $(P \times P)$ matrix with complexity $\mathcal{O}(P^3)$.
- *Parameter estimation:* For (22a), calculation of the matrix $\mathbf{M}_{k-1}$ of size $(P \times P)$ in (23a) has (P^2) AMs and $(2P^2)$ AAs. Eq. (22a) has $(2P^2)$ AMs, $(2P^2+4P)$ AAs and the inversion of $\mathbf{M}_{k-1}$ with complexity $\mathcal{O}(P^3)$. Eq. (22b) has (P^2+3P+1) AMs and (P^2+2P+6) AAs. Eq. (22c) involves the calculation and the inversion of matrix $\mathbf{M}_{\tilde{\mathbf{u}},k}$ of size $(N_u+1) \times (N_u+1)$. Note that the Woodbury matrix identity [26] can be applied for the recursive computation of $\mathbf{M}_{\tilde{\mathbf{u}},k}^{-1}$ based on $\mathbf{M}_{\tilde{\mathbf{u}},k-1}^{-1}$. Therefore, the total computations associated with (22c) include $(5N_u^2+13N_u+9)$ AMs, $(4N_u^2+12N_u+16)$ AAs and 2 ADs. Lastly, Eq. (22d) has (N_u+4) AMs and (N_u+9) AAs. Therefore, the total computations to update the parameter estimations include $(3P^2+3P+5N_u^2+14N_u+14)$ AMs, $(3P^2+6P+4N_u^2+13N_u+31)$ AAs, 2 ADs and the inversion of a $(P \times P)$ matrix with complexity $\mathcal{O}(P^3)$.
- *Iterative operations:* The proposed algorithm performs N_{it} iterations of computations in (22) - (26). Therefore, the computations at each time step include $N_{\text{it}}(8P^3+11P^2+7P+5N_u^2+15N_u+14)$ AMs, $N_{\text{it}}(8P^3+21P^2+14P+4N_u^2+14N_u+37)$ AAs, $(3N_{\text{it}})$ ADs and a complexity of $\mathcal{O}\big(2N_{\text{it}}(P^3)\big)$ for matrix inversion.

IV. MODELING AND PREDICTION IN SEASONAL CHANNELS

The UWA channel could exhibit periodic variations, such as the diurnal pattern as depicted in Fig. 2(b). In this type of channels, the slowly-varying channel parameter in one cycle could be highly correlated with those in previous cycles. Following the terminology in time series analysis [11], we refer to such type of channels as *seasonal channels*.

The data-driven model in (11) applies to seasonal channels. However, different from non-seasonal channels, the latent process in seasonal channels will be represented by a multiplicative seasonal AR process $(\text{AR}(P) \times (P_{\text{se}})_S)$ [11], whose polynomial representation in the lag operator D is a multiplication of the polynomial of an AR(P) process, $(1-\sum_{p=1}^{P} a_p D^p)$, and the polynomial of a seasonal AR(P_{se}) process, $(1-\sum_{q=1}^{P_{\text{se}}} \xi_q D^{qS})$, where $S \gg P$ denotes the seasonal cycle. The latent process in the time domain can be represented as

$$x[k] = \sum_{p=1}^{P} a_p x[k-p] + \sum_{q=1}^{P_{\text{se}}} \xi_q x[k-qS] - \sum_{p=1}^{P}\sum_{q=1}^{P_{\text{se}}} a_p \xi_q x[k-qS-p] + w[k]. \quad (28)$$

The proposed recursive algorithm for non-seasonal channels could be applied to seasonal channels by defining a long state vector $[x_k, x_{k-1}, \cdots, x_{k-P_{\text{se}}S-P+1}]^{\mathrm{T}}$ of length $(P+P_{\text{se}}S)$. This, however, will incur very large computational and storage cost. In this section, we will exploit the structure of (28), and develop a low-cost recursive algorithm for seasonal channels. To make the exposition easier, we focus on a simple scenario with $P_{\text{se}} = 1$, namely,

$$x[k] = \sum_{p=1}^{P} a_p x[k-p] + \xi x[k-S] - \xi \sum_{p=1}^{P} a_p x[k-S-p] + w[k], \quad (29)$$

while the developed algorithm can be extended to the scenario $P_{\rm se} > 1$ with slight modification.

Based on (29), we introduce an auxiliary random variable,

$$z_k := x_k - \xi x_{k-S}, \tag{30}$$

which according to (29), forms an AR process,

$$z_k = \sum_{p=1}^{P} a_p z_{k-p} + w_k. \tag{31}$$

Define $\mathbf{z}_k := [z_k, \cdots, z_{k-P+1}]^{\rm T}$. We have the state-space representation of (31),

$$\mathbf{z}_k = \mathbf{A}\mathbf{z}_{k-1} + \mathbf{w}_k, \tag{32}$$

where $\mathbf{A}$ and $\mathbf{w}_k$ are defined as in (13). The latent process can be reformulated as

$$x_k = \mathbf{a}^{\rm T}\mathbf{z}_{k-1} + \xi x_{k-S} + w_k. \tag{33}$$

Note that according to the principle of orthogonality [26], x_{k-S} is independent of z_k (c.f. (30)) and correspondingly $(\mathbf{a}^{\rm T}\mathbf{z}_{k-1})$ (c.f. (31)).

A. APPROXIMATION FOR RECURSIVE OPERATION

We redefine the unknown parameter set as $\mathbf{\Theta} := \{\gamma_0, \mathbf{a}, \xi, \mathbf{b}, \sigma_w^2, \sigma_v^2\}$. Based on (33), the log-likelihood function $[\ln f(\mathbf{y}_0^k, \mathbf{x}_{-1}, \mathbf{x}_0^k|\mathbf{\Theta})]$ can be decomposed as

$$L_k(\mathbf{\Theta}) = \sum_{k'=0}^{k} \ln f(x_{k'}, y_{k'}|\mathbf{z}_{k'-1}, x_{k'-S}, \mathbf{\Theta}) + \ln f(x_{-1}, \cdots, x_{-S}|\mathbf{\Theta}). \tag{34}$$

Similar to non-seasonal channels, for the development of a recursive algorithm, an approximation to $\mathbb{E}\big[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}\big]$ can be made through several steps. Particularly about the joint PDF $f(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}})$, it can be decomposed and approximated as

$$\begin{aligned}
&f(x_k, \mathbf{z}_{k-1}, x_{k-S}|y_k, \mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) \\
&= \frac{1}{c_1} f(x_k, \mathbf{z}_{k-1}, x_{k-S}, y_k|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) && \text{(35a)} \\
&= \frac{1}{c_1} f(y_k|x_k, \hat{\mathbf{\Theta}}) f(x_k|\mathbf{z}_{k-1}, x_{k-S}, \hat{\mathbf{\Theta}}) \\
&\quad \times f(\mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) && \text{(35b)} \\
&\approx \frac{1}{c_1'} f(y_k|x_k, \hat{\mathbf{\Theta}}) f(x_k|\mathbf{z}_{k-1}, x_{k-S}, \hat{\mathbf{\Theta}}) \\
&\quad \times f(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) f(x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) && \text{(35c)} \\
&= \frac{1}{c_1'} f(y_k|x_k, \hat{\mathbf{\Theta}}) f(z_k|\mathbf{z}_{k-1}, \hat{\mathbf{\Theta}}) f(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) \\
&\quad \times f(x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) \delta(z_k, x_k - \xi x_{k-S}) && \text{(35d)} \\
&= \frac{1}{c_1'} f(y_k|x_k, \hat{\mathbf{\Theta}}) f(\mathbf{z}_k|\mathbf{z}_{k-1}, \hat{\mathbf{\Theta}}) \\
&\quad \times f(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) f(x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) \\
&\quad \times \delta(z_k, x_k - \xi x_{k-S}) \delta(\mathbf{z}_k, \mathbf{z}_{k-1}), && \text{(35e)}
\end{aligned}$$

where c_1 and c_1' are normalization constants, $\delta(z_k, x_k - \xi x_{k-S})$ is introduced to ensure the equity in (30), and the approximation from (35b) to (35c) is made by assuming that $f(\mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) \approx f(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}) f(x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}})$. We further approximate the above PDF by

$$\begin{aligned}
&\tilde{f}(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}) \\
&:= \frac{1}{c_1''} f(y_k|x_k, \hat{\mathbf{\Theta}}) \\
&\quad \times f(\mathbf{z}_k|\mathbf{z}_{k-1}, \hat{\mathbf{\Theta}}) \tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}_{k-1}) \tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\mathbf{\Theta}}_{k-S}) \\
&\quad \times \delta(z_k, x_k - \xi x_{k-S}) \delta(\mathbf{z}_k, \mathbf{z}_{k-1}),
\end{aligned} \tag{36}$$

where c_1'' is a normalization constant, and the approximation is made through replacing $f(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}})$ and $f(x_{k-S}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}})$ in (35e) by $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}_{k-1})$ and $\tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\mathbf{\Theta}}_{k-S})$, respectively, with $\tilde{f}(\mathbf{z}_{k'}|\mathbf{y}_0^{k'}, \hat{\mathbf{\Theta}}_{k'})$ and $\tilde{f}(x_{k'}|\mathbf{y}_0^{k'}, \hat{\mathbf{\Theta}}_{k'})$ defined as the marginalization of $\tilde{f}(x_{k'}, \mathbf{z}_{k'-1}, x_{k'-S}|\mathbf{y}_0^{k'}, \hat{\mathbf{\Theta}})$ with respect to $\mathbf{z}_{k'}$ and $x_{k'}$, $\forall k'$.

Similar to the non-seasonal channel, the expectation $\mathbb{E}[L_k(\mathbf{\Theta})|\hat{\mathbf{\Theta}}]$ can be approximated by $Q_{{\rm se},k}(\mathbf{\Theta}|\hat{\mathbf{\Theta}})$ which is recursively defined as

$$\begin{aligned}
Q_{{\rm se},k}(\mathbf{\Theta}|\hat{\mathbf{\Theta}}) &= \lambda Q_{{\rm se},k-1}(\mathbf{\Theta}|\hat{\mathbf{\Theta}}_{k-1}) \\
&\quad + \int [\ln f(x_k, y_k|\mathbf{z}_{k-1}, x_{k-S}, \mathbf{\Theta})] \\
&\quad \times \tilde{f}(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}) dx_k d\mathbf{z}_{k-1} dx_{k-S}.
\end{aligned} \tag{37}$$

B. A LOW-COMPLEXITY RECURSIVE ALGORITHM

The proposed algorithm for seasonal channels operates recursively in a similar fashion as that for non-seasonal channels. Denote $\tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\mathbf{\Theta}}_{k-S}) = \mathcal{N}(\mu_{k-S}, \sigma_{k-S}^2)$ and $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}_{k-1}) = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{z},k-1}, \mathbf{C}_{\mathbf{z},k-1})$. At time k, the algorithm takes $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}_{k-1}), \tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\mathbf{\Theta}}_{k-S})$, the parameter set estimation $\hat{\mathbf{\Theta}}_{k-1}$, auxiliary quantities $(\tilde{\mathbf{M}}_{\mathbf{a},k-2}, \tilde{m}_{\xi,k-2}, \mathbf{M}_{\tilde{\mathbf{u}},k-1}^{-1})$ (to be defined shortly; computed at time $(k-1)$), the measurement y_k and the environmental parameter vector $\mathbf{u}_k$ as input, and sets $\hat{\mathbf{\Theta}}_k^{(0)} = \hat{\mathbf{\Theta}}_{k-1}$. The parameter set estimation and the Bayesian filtering and smoothing can be performed iteratively, until the number of iterations reaches a pre-determined threshold $N_{\rm it}$. The final parameter set estimation at time k is set as $\hat{\mathbf{\Theta}}_k = \hat{\mathbf{\Theta}}_k^{(N_{\rm it})}$. Based on $\hat{\mathbf{\Theta}}_k$, the PDFs $\tilde{f}(\mathbf{z}_k|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}_k) = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{z},k}, \mathbf{C}_{\mathbf{z},k})$ and $\tilde{f}(x_k|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}_k) = \mathcal{N}(\mu_k, \sigma_k^2)$ are computed via the Bayesian filtering. The PDF $\tilde{f}(\mathbf{z}_k|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}_k)$, $\hat{\mathbf{\Theta}}_k$ and $(\tilde{\mathbf{M}}_{\mathbf{a},k-1}, \tilde{m}_{\xi,k-1}, \mathbf{M}_{\tilde{\mathbf{u}},k}^{-1})$ that are computed at time k, will be used for the recursive operation at time $(k+1)$. The PDF $\tilde{f}(x_k|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}_k)$ will be used for the recursive operation at time $(k+S)$. Additionally, based on the parameter set estimation $\hat{\mathbf{\Theta}}_k$ and the state estimation $\boldsymbol{\mu}_{\mathbf{z},k}$ and $\{\mu_{k'}; k' \le k\}$, multiple-step-ahead prediction of the slowly-varying channel parameter can be achieved.

We next briefly describe the recursive and iterative parameter estimation and the multiple-step-ahead prediction.

A detailed description of the Bayesian filtering and smoothing is presented in Section IV-C.

1) RECURSIVE AND ITERATIVE PARAMETER ESTIMATION

At time k, given the parameter set estimation in the ith iteration, $\hat{\boldsymbol{\Theta}}_k^{(i)}$, and following the same procedure as in Section III-B.1, the parameter set estimation can be updated through maximizing $Q_{\mathrm{se},k}(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$. Specifically, the estimations of $\{\gamma_0, \mathbf{b}_k, \sigma_v^2\}$ can be updated according to the same equations as in (22). The estimations of $\{\mathbf{a}, \xi, \sigma_w^2\}$ can be updated as

$$\hat{\mathbf{a}}_k^{(i+1)} = \hat{\mathbf{a}}_{k-1} + \tilde{\mathbf{M}}_{\mathbf{a},k-1}^{-1}\left\{\mathbb{E}\left[z_k\mathbf{z}_{k-1}\right] - \mathbb{E}\left[\mathbf{z}_{k-1}\mathbf{z}_{k-1}^{\mathrm{T}}\right]\hat{\mathbf{a}}_{k-1}\right\}, \tag{38a}$$

$$\hat{\xi}_k^{(i+1)} = \hat{\xi}_{k-1} + \tilde{m}_{\xi,k-1}^{-1} \times\left\{\mathbb{E}\left[(x_k - \hat{\mathbf{a}}_k^{(i+1),\mathrm{T}}\mathbf{z}_{k-1})x_{k-S}\right] - \mathbb{E}\left[x_{k-S}^2\right]\hat{\xi}_{k-1}\right\}, \tag{38b}$$

$$\hat{\sigma}_{w,k}^{2,(i+1)} = \hat{\sigma}_{w,k-1}^2 + \frac{1-\lambda}{1-\lambda^k} \times\left\{\mathbb{E}\left[\left(x_k - \hat{\xi}_k^{(i+1)}x_{k-S} - \hat{\mathbf{a}}_k^{(i+1),\mathrm{T}}\mathbf{z}_{k-1}\right)^2\right] - \hat{\sigma}_{w,k-1}^2\right\}, \tag{38c}$$

where $\tilde{\mathbf{M}}_{\mathbf{a},k-1}$ and $\tilde{m}_{\xi,k-1}$ are defined as

$$\tilde{\mathbf{M}}_{\mathbf{a},k-1} := \lambda\tilde{\mathbf{M}}_{\mathbf{a},k-2} + \mathbb{E}[\mathbf{z}_{k-1}\mathbf{z}_{k-1}^{\mathrm{T}}], \tag{39a}$$

$$\tilde{m}_{\xi,k-1} := \lambda\tilde{m}_{\xi,k-2} + \mathbb{E}[x_{k-S}^2]. \tag{39b}$$

The expectations are performed with respect to $\tilde{f}(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$.

2) MULTIPLE-STEP-AHEAD PREDICTION

Based on the parameter set estimation $\hat{\boldsymbol{\Theta}}_k$ and the state estimation $\boldsymbol{\mu}_{\mathbf{z},k}$ and $\{\mu_{k'}; k' \le k\}$ (denoted next also as $\hat{\mathbf{z}}_k$ and $\{\hat{x}_{k'}; k' \le k\}$, respectively), the m-step-ahead prediction of the latent process and the slowly-varying channel parameter can be obtained recursively as

$$\hat{z}_{k+m} = \hat{\mathbf{a}}_k^{\mathrm{T}}\hat{\mathbf{z}}_{k+m-1}, \tag{40a}$$

$$\hat{x}_{k+m} = \hat{z}_{k+m} + \hat{\xi}_k\hat{x}_{k+m-S}, \tag{40b}$$

$$\hat{\alpha}_{k+m} = \hat{\gamma}_{0,k} + \hat{x}_{k+m} + \hat{\mathbf{b}}_k^{\mathrm{T}}\mathbf{u}_{k+m}, \tag{40c}$$

for $m = 1, 2, \cdots$, where $\mathbf{u}_{k+m}$ can be obtained from meteorological forecast sources, e.g., [27].

C. BAYESIAN FILTERING AND SMOOTHING

To find the expectations in (38) and (39), we compute the marginalization of $\tilde{f}(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ with respect to $x_k, \mathbf{z}_k, \mathbf{z}_{k-1}$, and x_{k-S}, respectively. Denote the marginal PDFs as

$$\tilde{f}(x_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\mu_k^{(i)}, \sigma_k^{2,(i)}),$$
$$\tilde{f}(\mathbf{z}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\boldsymbol{\mu}_{\mathbf{z},k}^{(i)}, \mathbf{C}_{\mathbf{z},k}^{(i)}),$$
$$\breve{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i)}, \breve{\mathbf{C}}_{\mathbf{z},k-1}^{(i)}),$$
$$\breve{f}(x_{k-S}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)}) = \mathcal{N}(\breve{\mu}_{k-S}^{(i)}, \breve{\sigma}_{k-S}^{2,(i)}).$$

Based on $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1})$, $\tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\boldsymbol{\Theta}}_{k-S})$, and the model in (33), x_k can be predicted and with the prediction MSE,

$$\mu_{k|k-1}^{(i)} = \hat{\mathbf{a}}_k^{(i),\mathrm{T}}\boldsymbol{\mu}_{\mathbf{z},k-1} + \hat{\xi}_k^{(i)}\mu_{k-S},$$
$$\sigma_{k|k-1}^{2,(i)} = \hat{\mathbf{a}}_k^{(i),\mathrm{T}}\mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{a}}_k^{(i)} + \hat{\xi}_k^{(i),2}\sigma_{k-S}^2 + \sigma_{\mathrm{w},k}^{2,(i)}.$$

Based on the measurement y_k and the measurement model in (12b), the mean and the variance of x_k in the marginal PDF $\tilde{f}(x_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ can be formulated as

$$\mu_k^{(i)} = \mu_{k|k-1}^{(i)} + \frac{\sigma_{k|k-1}^{2,(i)}}{\sigma_{\mathrm{v},k}^{2,(i)} + \sigma_{k|k-1}^{2,(i)}} \times(y_k - \mu_{k|k-1}^{(i)} - \hat{\gamma}_{0,k}^{(i)} - \hat{\mathbf{b}}_k^{(i),\mathrm{T}}\mathbf{u}_k), \tag{41a}$$

$$\sigma_k^{2,(i)} = \frac{\sigma_{\mathrm{v},k}^{2,(i)}\sigma_{k|k-1}^{2,(i)}}{\sigma_{\mathrm{v},k}^{2,(i)} + \sigma_{k|k-1}^{2,(i)}}. \tag{41b}$$

We further have $\mathbb{E}[x_k^2|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \sigma_k^{2,(i)} + (\mu_k^{(i)})^2$.

Similar operation can be applied to $\mathbf{z}_k$. Define $\hat{\mathbf{A}}_k^{(i)}$ and $\hat{\mathbf{C}}_{\mathbf{w},k}^{(i)}$ as the matrix corresponding to $\hat{\mathbf{a}}_k^{(i)}$ and $\hat{\sigma}_{w,k}^{2,(i)}$, respectively. Based on $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}}_{k-1})$, $\mathbf{z}_k$ can be predicted as $\hat{\mathbf{A}}_k^{(i)}\boldsymbol{\mu}_{\mathbf{z},k-1}$, with the prediction MSE matrix $\mathbf{P}_{\mathbf{z},k}^{(i)} = \hat{\mathbf{A}}_k^{(i)}\mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{A}}_k^{(i),\mathrm{T}} + \hat{\mathbf{C}}_{\mathbf{w},k}^{(i)}$. Note that the measurement y_k can be represented as

$$y_k = \gamma_0 + \mathbf{h}^{\mathrm{T}}\mathbf{z}_k + \xi x_{k-S} + \mathbf{b}^{\mathrm{T}}\mathbf{u}_k + v_k.$$

Define the gain vector $\mathbf{k}_{\mathbf{z},k}^{(i)} = \mathbf{P}_{\mathbf{z},k}^{(i)}\mathbf{h}\big(\hat{\xi}_k^{(i),2}\sigma_{k-S}^2 + \hat{\sigma}_{k,v}^{(i)} + \mathbf{h}^{\mathrm{T}}\mathbf{P}_{\mathbf{z},k}^{(i)}\mathbf{h}\big)^{-1}$. The mean and the covariance matrix of $\mathbf{z}_k$ in the marginal PDF $\tilde{f}(\mathbf{z}_k|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ can be formulated as

$$\boldsymbol{\mu}_{\mathbf{z},k}^{(i)} = \hat{\mathbf{A}}_k^{(i)}\boldsymbol{\mu}_{\mathbf{z},k-1} + \mathbf{k}_{\mathbf{z},k}^{(i)} \times\big(y_k - \hat{\mathbf{a}}_k^{(i),\mathrm{T}}\boldsymbol{\mu}_{\mathbf{z},k-1} - \hat{\xi}_k^{(i)}\mu_{k-S} - \hat{\gamma}_{0,k}^{(i)} - \hat{\mathbf{b}}_k^{(i),\mathrm{T}}\mathbf{u}_k\big), \tag{42a}$$

$$\mathbf{C}_{\mathbf{z},k}^{(i)} = (\mathbf{I} - \mathbf{k}_{\mathbf{z},k}^{(i)}\mathbf{h}^{\mathrm{T}})\mathbf{P}_{\mathbf{z},k}^{(i)}. \tag{42b}$$

We further have $\mathbb{E}[\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \mathbf{C}_{\mathbf{z},k}^{(i)} + \boldsymbol{\mu}_{\mathbf{z},k}^{(i)}\boldsymbol{\mu}_{\mathbf{z},k}^{(i),\mathrm{T}}$.

Furthermore, the marginal PDF $\breve{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$ can be obtained via the one-step backward smoothing. Denote the gain matrix $\mathbf{J}_{\mathbf{z},k-1}^{(i)} = \mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{A}}_k^{(i),\mathrm{T}}(\mathbf{P}_{\mathbf{z},k}^{(i)})^{-1}$. The mean and the covariance matrix of $\mathbf{z}_{k-1}$ in the marginal PDF can be formulated as

$$\breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i)} = \boldsymbol{\mu}_{\mathbf{z},k-1} + \mathbf{J}_{\mathbf{z},k-1}^{(i)}(\boldsymbol{\mu}_{\mathbf{z},k}^{(i)} - \hat{\mathbf{A}}_k^{(i)}\boldsymbol{\mu}_{\mathbf{z},k-1}), \tag{43a}$$

$$\breve{\mathbf{C}}_{\mathbf{z},k-1}^{(i)} = \mathbf{C}_{\mathbf{z},k-1} + \mathbf{J}_{\mathbf{z},k-1}^{(i)}(\mathbf{C}_{\mathbf{z},k}^{(i)} - \mathbf{P}_{\mathbf{z},k}^{(i)})\mathbf{J}_{\mathbf{z},k-1}^{(i),\mathrm{T}}. \tag{43b}$$

We further have $\mathbb{E}[\mathbf{z}_{k-1}\mathbf{z}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \breve{\mathbf{C}}_{\mathbf{z},k-1}^{(i)} + \breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i)}\breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i),\mathrm{T}}$. Based on the joint PDF $\tilde{f}(x_k, \mathbf{z}_{k-1}, x_{k-S}|\mathbf{y}_0^k, \hat{\boldsymbol{\Theta}}_k^{(i)})$, we also have $\mathbb{E}[\mathbf{z}_k\mathbf{z}_{k-1}^{\mathrm{T}}|\hat{\boldsymbol{\Theta}}_k^{(i)}] = \mathbf{C}_{\mathbf{z},k}^{(i)}\mathbf{J}_{\mathbf{z},k-1}^{(i),\mathrm{T}} + \boldsymbol{\mu}_{\mathbf{z},k}^{(i)}\breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i),\mathrm{T}}$.

Given $\tilde{f}(x_{k-S}|\mathbf{y}_0^{k-S}, \hat{\mathbf{\Theta}}_{k-S})$, $\tilde{f}(\mathbf{z}_{k-1}|\mathbf{y}_0^{k-1}, \hat{\mathbf{\Theta}}_{k-1})$, and the measurement representation,

$$y_k = \gamma_0 + \mathbf{a}^{\mathrm{T}}\mathbf{z}_{k-1} + \xi x_{k-S} + \mathbf{b}^{\mathrm{T}}\mathbf{u}_k + w_k + v_k,$$

the mean and the variance in the marginalized PDF $\breve{f}(x_{k-S}|\mathbf{y}_0^k, \hat{\mathbf{\Theta}}_k^{(i)})$, can be formulated as

$$\breve{\mu}_{k-S}^{(i)} = \mu_{k-S} + \frac{\hat{\xi}_k^{(i)}\sigma_{k-S}^2\big(y_k - \hat{\mathbf{a}}_k^{(i),\mathrm{T}}\boldsymbol{\mu}_{\mathbf{z},k-1} - \hat{\xi}_k^{(i)}\mu_{k-S} - \hat{\gamma}_{0,k}^{(i)} - \hat{\mathbf{b}}_k^{(i),\mathrm{T}}\mathbf{u}_k\big)}{\hat{\mathbf{a}}_k^{(i),\mathrm{T}}\mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{a}}_k^{(i)} + \sigma_{\mathrm{w},k}^{2,(i)} + \sigma_{\mathrm{v},k}^{2,(i)} + \hat{\xi}_k^{2,(i)}\sigma_{k-S}^2}, \tag{44a}$$

$$\breve{\sigma}_{k-S}^{2,(i)} = \frac{(\hat{\mathbf{a}}_k^{(i),\mathrm{T}}\mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{a}}_k^{(i)} + \sigma_{\mathrm{w},k}^{2,(i)} + \sigma_{\mathrm{v},k}^{2,(i)})\sigma_{k-S}^2}{\hat{\mathbf{a}}_k^{(i),\mathrm{T}}\mathbf{C}_{\mathbf{z},k-1}\hat{\mathbf{a}}_k^{(i)} + \sigma_{\mathrm{w},k}^{2,(i)} + \sigma_{\mathrm{v},k}^{2,(i)} + \hat{\xi}_k^{2,(i)}\sigma_{k-S}^2}. \tag{44b}$$

We further have $\mathbb{E}[x_{k-S}^2|\hat{\mathbf{\Theta}}_k^{(i)}] = \breve{\sigma}_{k-S}^{2,(i)} + (\breve{\mu}_{k-S}^{(i)})^2$.

The expectations to be used in (38) and (39) can be directly extracted from the above results. In particular, given (30), the expectation $\mathbb{E}[x_k x_{k-S}|\hat{\mathbf{\Theta}}_k^{(i)}]$ can be computed based on $\mathbb{E}[z_k^2|\hat{\mathbf{\Theta}}_k^{(i)}]$, $\mathbb{E}[x_k^2|\hat{\mathbf{\Theta}}_k^{(i)}]$ and $\mathbb{E}[x_{k-S}^2|\hat{\mathbf{\Theta}}_k^{(i)}]$. Note that x_{k-S} and $(\mathbf{a}^{\mathrm{T}}\mathbf{z}_{k-1})$ are independent. We have $\mathbb{E}[x_{k-S}(\mathbf{a}_k^{(i),\mathrm{T}}\mathbf{z}_{k-1})|\hat{\mathbf{\Theta}}_k^{(i)}] = \breve{\mu}_{k-S}^{(i)}(\mathbf{a}_k^{(i),\mathrm{T}}\breve{\boldsymbol{\mu}}_{\mathbf{z},k-1}^{(i)})$.

V. MODEL ORDER SELECTION

The non-seasonal latent process in (10) can be regarded as a degraded seasonal latent process in (28) with a seasonal order of zero. The orders (P, P_{se}) and the index set $\mathcal{I}$ of N_u important elements within $\mathcal{E}[k]$, $\forall k$ for the process $\{g_k\}$, can be determined via the minimum description length (MDL) criterion [28], as described in the following.

We stack the channel measurements $\{y[k]\}$ into a long vector $\mathbf{y}$ of length K. Stack the coefficients of the seasonal AR(P_{se}) process into a vector $\boldsymbol{\xi} := [\xi_1, \cdots, \xi_{P_{\mathrm{se}}}]^{\mathrm{T}}$ (c.f. (28)). Define a long vector $\boldsymbol{\theta} := \big([1\ \mathbf{a}^{\mathrm{T}}] \otimes [1\ \boldsymbol{\xi}^{\mathrm{T}}]\big)^{\mathrm{T}}$ of length $(P+1)\times(P_{\mathrm{se}}+1)$ and with $\otimes$ denoting the Kronecker product.

Based on (11) and (28), we have

$$\mathbf{y} = \mathbf{H}(\gamma_0, \mathbf{b})\boldsymbol{\theta} + \mathbf{n}, \tag{45}$$

where $\mathbf{H}(\gamma_0, \mathbf{b})$ is a matrix containing unknown parameters, and its kth row is formed by γ_0, $\{y[k']; k' < k\}$ according to (P, P_{se}), and by the elements in $\{\mathcal{E}[k']; k' \le k\}$ that are indexed by $\mathcal{I}$ and weighed by $\mathbf{b}$, and $\mathbf{n}$ is a noise vector, with $n[k] \sim \mathcal{N}(0, \sigma_n^2)$.

The optimal values of (P, P_{se}) and the index set $\mathcal{I}$ can be determined according to the MDL criterion [28],

$$\min_{(P,P_{\mathrm{se}},\mathcal{I})} \ \frac{K}{2}\ln\hat{\sigma}_n^2 + \frac{1}{2}(P + P_{\mathrm{se}} + N_u)\ln K, \tag{46}$$

where $\hat{\sigma}_n^2 = \frac{1}{K}\mathbf{y}^{\mathrm{T}}\mathbf{P}^{\perp}(\hat{\gamma}_0, \hat{\mathbf{b}})\mathbf{y}$ is the ML estimation of the noise variance, with $\mathbf{P}^{\perp}(\hat{\gamma}_0, \hat{\mathbf{b}}) := \mathbf{I} - \mathbf{H}(\hat{\gamma}_0, \hat{\mathbf{b}})(\mathbf{H}^{\mathrm{T}}(\hat{\gamma}_0, \hat{\mathbf{b}})\mathbf{H}(\hat{\gamma}_0, \hat{\mathbf{b}}))^{-1}\mathbf{H}^{\mathrm{T}}(\hat{\gamma}_0, \hat{\mathbf{b}})$, and $(P+P_{\mathrm{se}}+N_u)$ is the number of model parameters. The ML estimation $\hat{\gamma}_0$ and $\hat{\mathbf{b}}$ can be found based on (45) by iterative computational methods. In real applications, consider that the large-scale phenomena of water environments change very slowly. The model order selection can be carried out once in a while by a central processing station, after it collects the measured slowly-varying channel parameters from underwater nodes. Given small values of (P, P_{se}) and limited types of environmental parameters, the optimization problem in (46) can be solved via exhaustive search.

VI. SIMULATION RESULTS

The proposed recursive algorithms are evaluated via Monte Carlo simulations. In each simulation setting, we consider 400 Monte Carlo runs, and each run contains a time series of a slowly-varying channel parameter of 3000 samples. The time series is generated according to the model specified in (12). In each Monte Carlo run, the time-invariant component γ_0 is randomly selected uniformly from [3, 30]. The latent processes in non-seasonal channels are generated as AR(P) processes according to (10), while the latent processes in seasonal channels are generated as multiplicative seasonal AR processes (AR(P)$\times$(1)$_{96}$) according to (29), with the seasonal coefficient ξ randomly selected uniformly from $[-1, 1]$. Two types of environmental parameters are considered. The time sequences of environmental parameters are generated independently as AR(P) processes. The AR coefficients of each process are obtained based on a minimum-phase polynomial whose roots are randomly chosen within the unit circle in the complex plane. The process $\{g_k\}$ is generated as a linear combination of the time sequences of the two types of environmental parameters $\phi_i[k]$,

$$g[k] = \sum_{l=1}^{L=2} b_l\phi_l[k]. \tag{47}$$

The combinational coefficients in $\mathbf{b} = [b_1, b_2]^{\mathrm{T}}$ are randomly selected according to a uniform distribution over $[0.2, 1]\times\zeta$, where ζ is a scalar for controlling the energy ratio between the process $\{g_k\}$ and the latent process $\{x_k\}$. Specifically, we define the energy ratio

$$\eta := \frac{\sum_{k=1}^{K} x_k^2}{\sum_{k=1}^{K}(x_k^2 + g_k^2)}, \tag{48}$$

to control the contribution of the latent process $\{x_k\}$ and the contribution of the process $\{g_k\}$ in the generated time series $\{\alpha_k\}$, with $K = 3000$. When $\eta = 1$, the sequence $\{\alpha_k\}$ only consists of γ_0 and the latent process. When $\eta = 0$, the sequence $\{\alpha_k\}$ only consists of γ_0 and the process $\{g_k\}$. The value of ζ can be computed based on a pre-selected value of η. The energy ratio between the summed process $\{x_k + g_k\}$ and the measurement noise is set to a moderate value of 8 dB.

The normalized mean square error (NMSE) is taken as the performance metric, which is computed after the convergence of the model parameter estimation. Specifically, for vector $\mathbf{a}$,

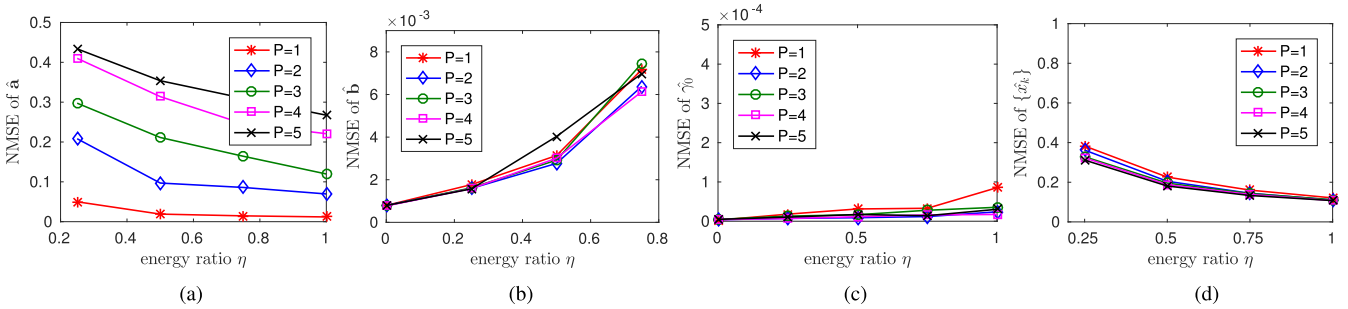


FIGURE 4. Non-seasonal channels: Normalized mean square estimation error of the model parameters and the latent process. (a) latent process coeff.: $\hat{\mathbf{a}}$. (b) env process coeff.: $\hat{\mathbf{b}}$. (c) time-invariant component: $\hat{\gamma}_0$. (d) latent process: $\{\hat{x_k}\}$

the estimation NMSE is computed as

$$\frac{1}{N}\sum_{k=k_0}^{K}\frac{\|\mathbf{a}-\hat{\mathbf{a}}_k\|_2^2}{\|\mathbf{a}\|_2^2}, \tag{49}$$

where $\hat{\mathbf{a}}_k$ is the estimation at time k, k_0 is the time index when the estimation converges, $N := (K - k_0 + 1)$, and $\|\cdot\|_2$ denotes the ℓ_2 norm. The estimation NMSE of other model parameters can be similarly computed. The estimation NMSE of the latent process is computed as

$$\frac{\frac{1}{N}\sum_{k=k_0}^{K}(x_k-\hat{x}_k)^2}{\frac{1}{N}\sum_{k=k_0}^{K}x_k^2}. \tag{50}$$

The NMSE of the m-step-ahead prediction of the slowly-varying channel parameter is computed as

$$\frac{\frac{1}{N-m}\sum_{k=k_0}^{K-m}(\alpha_{k+m}-\hat{\alpha}_{k+m})^2}{\frac{1}{N}\sum_{k=k_0}^{K}(\alpha_k-\bar{\alpha})^2}, \tag{51}$$

with $\bar{\alpha}$ being the average of the sequence $\{\alpha_k\}$. In the proposed algorithm for non-seasonal channels and for seasonal channels, the forgetting factor is set as $\lambda = 0.99$. The proposed algorithms in all the simulation settings converge within about $k_0 = 800$ time steps.

A. NON-SEASONAL CHANNELS

The recursive algorithm for non-seasonal channels will be evaluated in two scenarios. The first scenario assumes perfect prior knowledge of the latent process order P and has access to both types of environmental parameters, while the second scenario does not assume the prior knowledge of the latent process order and may not have access to all the environmental parameters. The second scenario is closer to real world applications.

1) MODELING AND PREDICTION WITH CHANNEL GENERATION KNOWLEDGE

The proposed recursive algorithm is evaluated using the sequences of $\{\alpha_k\}$ that are generated according to different latent process orders and different values of the energy ratio η. The estimation NMSEs of $\mathbf{a}$, $\mathbf{b}$, γ_0 and the latent process are depicted in Fig. 4. One can see that as the energy ratio η increases, the estimation NMSE of $\mathbf{a}$ and of the latent process decreases, while the estimation NMSE of $\mathbf{b}$ increases.

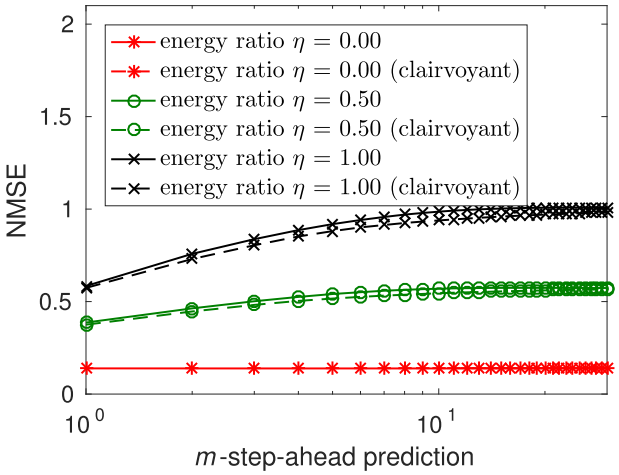


FIGURE 5. Non-seasonal channels: Prediction performance with the channel generation knowledge. Clairvoyant: the Kalman filter performance with perfect knowledge of model parameters.

In addition, as the latent process order P increases, the estimation NMSE of $\mathbf{a}$ increases drastically, whereas the estimation NMSEs of $\mathbf{b}$, γ_0 and the latent process are less sensitive to the order change. The vector $\mathbf{b}$ and the time-invariant component γ_0 can be accurately estimated with the NMSE less than 10^{-2} and 10^{-4}, respectively.

Corresponding to the latent process order $P = 2$ and different values of the energy ratio η, Fig. 5 depicts the m-step-ahead prediction performance of the proposed algorithm. As a performance upper bound, the m-step-ahead prediction NMSE of the Kalman filter with perfect knowledge of the model parameters is also plotted. One can observe that the proposed algorithm achieves a performance very close to the performance upper bound. Additionally, the prediction accuracy improves as the contribution of the latent process decreases (i.e., as η decreases). In other words, the channel can be more accurately modeled and predicted when it has less contribution from unknown physical mechanisms or unavailable environmental parameters.

2) MODELING AND PREDICTION WITHOUT CHANNEL GENERATION KNOWLEDGE

We generate the sequences of $\{\alpha_k\}$ according to the latent process order $P = 2$ and different values of the energy ratio η.

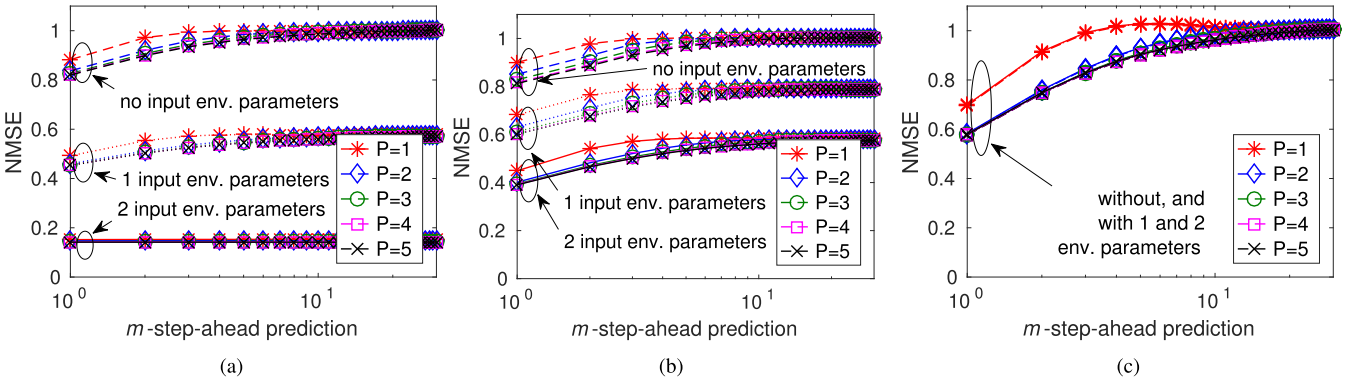


FIGURE 6. Non-seasonal channels: Prediction performance without the channel generation knowledge. $P = 2$. (a) energy ratio $\eta = 0$. (b) energy ratio $\eta = 0.5$. (c) energy ratio $\eta = 1$.

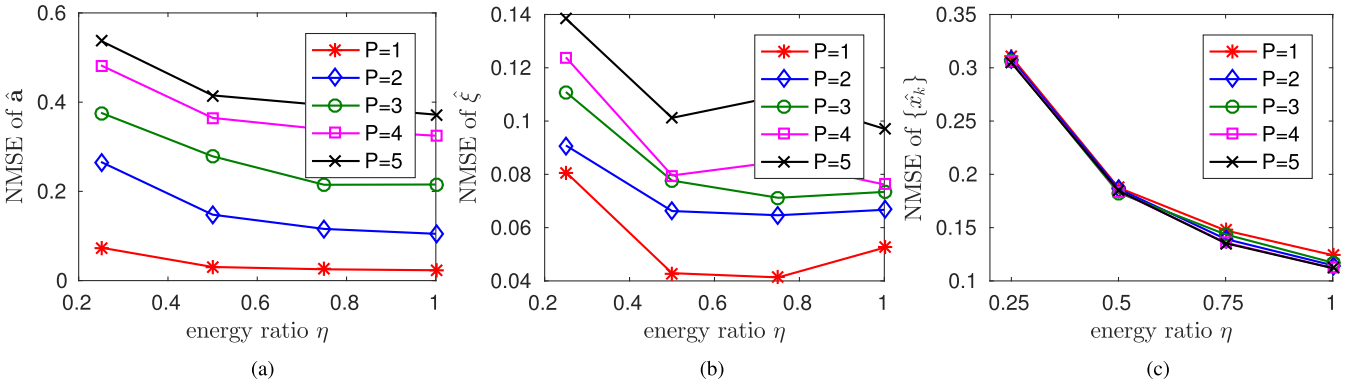


FIGURE 7. Seasonal channels: Normalized mean square estimation error of the model parameters and the latent process. (a) $\hat{\mathbf{a}}$. (b) $\hat{\xi}$. (c) $\{\hat{x_k}\}$.

Without the knowledge of $P = 2$ and potentially in lack of one or both types of environment parameters, the m-step-ahead prediction performance of the proposed algorithm is shown in Fig. 6, where different orders of the latent process are examined for channel modeling and prediction. One can see that the prediction performance improves when more environmental parameters are incorporated and when the contribution of the latent process decreases. Furthermore, for each energy ratio, performance improvement can be observed when the order of the latent process increases from 1 to the true value of 2, while the improvement is less obvious for further increase. Additionally, when the energy ratio equals to one, namely, the sequence $\{\alpha_k\}$ only consists of the time-invariant component γ_0 and the latent process, incorporation of the environmental parameters into the channel modeling does not lead to obvious performance degradation.

B. SEASONAL CHANNELS

Following the seasonality in the KW-AUG14 experiment, we consider a seasonal cycle of $S = 96$. We next evaluate the proposed algorithm for seasonal channels with and without the channel generation knowledge.

1) MODELING AND PREDICTION WITH CHANNEL GENERATION KNOWLEDGE

For the sequences of $\{\alpha_k\}$ with different values of P and different values of η, the estimation NMSEs of $\mathbf{a}$, ξ and the latent process are depicted in Fig. 7. The estimation NMSEs of $\mathbf{b}$, γ_0 are almost identical to those in Fig. 4 for non-seasonal channels. Comparing the NMSEs in Figs. 4 and 7, one can see that the estimation NMSE of $\mathbf{a}$ in seasonal channels is larger than that in non-seasonal channels, primarily because of the nonlinear relationship between $\mathbf{a}$ and ξ. Furthermore, the estimation NMSE of the latent process in seasonal channels is less than that in non-seasonal channels, thanks to the seasonal correlation of the latent process.

Corresponding to $P = 2$ and different values of the energy ratio η, Fig. 8 shows the m-step-ahead prediction performance of the proposed algorithm. As a performance upper bound, the m-step-ahead prediction NMSE of the Kalman filter with perfect knowledge of the model parameters is also plotted. One can obtain similar observations as those in non-seasonal channels. However, compared to the simulation results in Fig. 5, less NMSE can be achieved in the seasonal channel, benefiting from the seasonal correlation of the latent process.

2) MODELING AND PREDICTION WITHOUT CHANNEL GENERATION KNOWLEDGE

We generate the sequences of $\{\alpha_k\}$ in seasonal channels with $P = 2$ and different values of the energy ratio η. Without the knowledge of $P = 2$ and potentially in the lack of one or both types of environmental parameters, the m-step-ahead

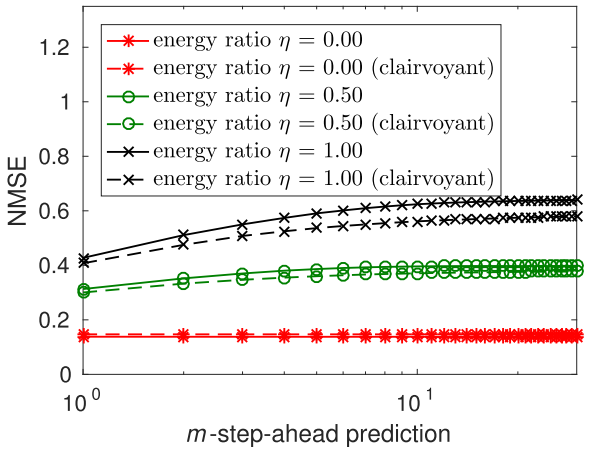


FIGURE 8. Seasonal channels: Prediction performance with the channel generation knowledge. Clairvoyant: the Kalman filter performance with perfect knowledge of model parameters.

prediction performance of the proposed algorithm is shown in Fig. 9, where different values of P are examined. Compared to Fig. 6, similar observations can be obtained, while as the energy ratio η increases, higher prediction accuracy can be achieved in the seasonal channel, benefiting from the seasonal correlation of the latent process.

VII. EXPERIMENTAL DATA PROCESSING

The proposed models and algorithms are evaluated using measurements from two shallow-water field experiments: one is the Surface Processes and Acoustic Communication Experiment (SPACE08) conducted from Oct. 14 to Nov. 1, 2008 near the coast of Martha's Vineyard, MA, and the other was conducted in the Keweenaw Waterway, MI in Aug. 2014, abbreviated as KW-AUG14. In SPACE08, a waveform of 1 minute and within the frequency band [8, 18] kHz was transmitted every 2 hours at a fixed power level. The waveform consists of 60 ZP OFDM-modulated blocks with parameters specified in Table 1. In KW-AUG14, a waveform of 8.83 seconds and within the frequency band [14, 20] kHz was transmitted every 15 minutes at a fixed power level. The waveform consists of 20 ZP OFDM-modulated blocks with parameters specified in Table 1. The CIR is estimated

TABLE 1. OFDM parameters in SPACE08 and KW-AUG14.

OFDM parameters		SPACE08	KW-AUG14
center frequency	f_c	13 kHz	17 kHz
bandwidth	B	9.77 kHz	6 kHz
total # of subcarriers	K_{total}	1024	1024
# of pilot subcarriers	K_{pilot}	352	256
symbol duration	T	104.86 ms	170.7 ms
subcarrier spacing	$\Delta f := 1/T$	9.54 Hz	5.86 Hz
zero guard interval	T_g	24.6 ms	79.3 ms

per OFDM block based on measurements at pilot subcarriers using a sparse channel estimator which exploits the multipath sparsity in the delay and the Doppler domain [29]. Four types of slowly-varying channel parameters derived from the estimated CIRs are examined in this section, including the average channel SNR, the Nakagami-m fading parameter, the average RMS delay spread, and the average RMS Doppler spread (c.f. Section II-A). While many environmental parameters have impact on UWA channels, the wind speed and temperature are chosen in this work to evaluate the proposed algorithms based on their availability and low acquisition cost. In addition, noticing that the water condition in KW-AUG14 was calm with negligible Doppler effect, we skip the analysis of the average RMS Doppler spread in this experiment.

For performance comparison, we introduce a recursive linear regression (LR) method where the time sequence of a slowly-varying channel parameter is modeled as the summation of a time-invariant component γ_0 and a process $\{g_k\}$ described by environmental measurements defined as in (47); see *Remark 3* in Section III-B.

A. SPACE08 WITH NON-SEASONAL CHANNEL VARIATIONS

In SPACE08, we consider the signals received by a 12-element hydrophone array, which was vertically mounted on a fixed tripod 200 meters away from the source. The adjacent elements have a 12 cm spacing and the top element is 3.25 meters above the sea floor. The water depth is about 15 meters. The source transducer was mounted 4 meters above the bottom. The average channel SNR scaled by

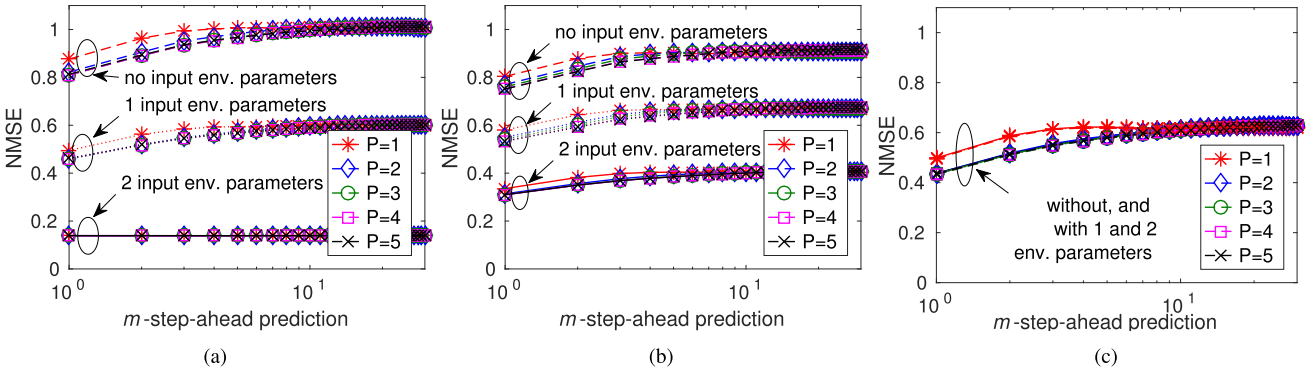


FIGURE 9. Seasonal channels: Prediction performance without the channel generation knowledge. $P = 2$. (a) energy ratio $\eta = 0$. (b) energy ratio $\eta = 0.5$. (c) energy ratio $\eta = 1$.

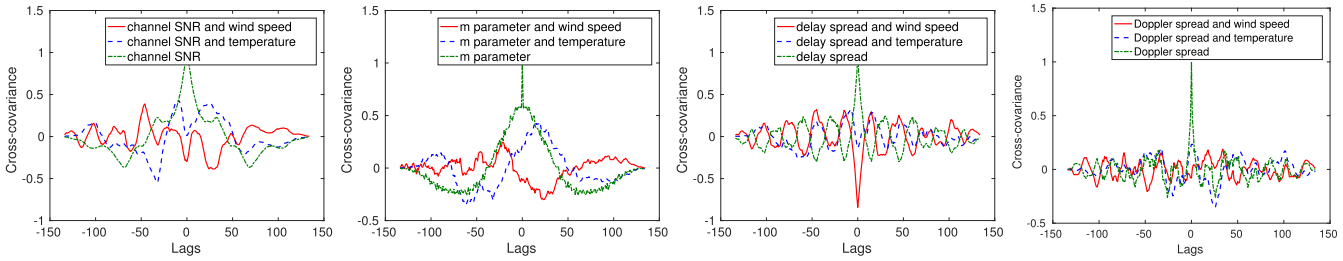


FIGURE 10. SPACE08: Autocorrelation of slowly-varying channel parameters and their correlation with environmental measurements. (a) average channel SNR. (b) Nakagami-*m* fading parameter. (c) average RMS delay spread. (d) average RMS Doppler spread.

the transmission power, the Nakagami-*m* fading parameter, the average RMS delay spread, and the average RMS Doppler spread within each transmission and over the 12 hydrophones are shown in Fig. 2(a), along with the mean wind speed and the mean air temperature measurements measured respectively by a 3-axis sonic anemometer and a VaiPTU located at 12.5 meters above the mean sea level on the meteorological mast of the Martha's Vineyard Coastal Observatory (MVCO) [30]. The autocorrelation of the slowly-varying channel parameters and their correlation with environmental measurements are depicted in Fig. 10. One can see that the average channel SNR and the average RMS delay spread are negatively correlated with the wind speed and their correlation with the temperature are not obvious. The Nakagami-*m* fading parameter exhibits high inherent temporal correlation, and slight positive correlation with the temperature and slight negative correlation with the wind speed. The average RMS Doppler spread shows slight positive correlation with the temperature and negligible correlation with the wind speed.

In the proposed algorithm, we set the forgetting factor $\lambda = 0.96$ for the average channel SNR sequence, $\lambda = 0.92$ for the Nakagami-*m* parameter sequence, and $\lambda = 1$ for the sequences of other slowly-varying channel parameters. According to the MDL criterion in (46), the optimal order of the latent process is chosen as $P = 1$ for the sequences of the average channel SNR and the average RMS delay spread, $P = 2$ for the sequence of the average RMS Doppler spread, and $P = 4$ for the sequence of the Nakagami-*m* fading parameter. In addition, the sequences of the wind speed and the temperature are normalized individually to have a unit power, and a linear combination of the two types of environmental parameters will be used for modeling the process $\{g_k\}$ (c.f. (8)). With the incorporation of both types of environmental measurements into the modeling, the model parameters estimated by the proposed algorithm are listed in Table 2. The estimated coefficients in $\hat{\mathbf{b}}$ indicate the amount of contribution from each type of environmental parameters, and the value of η reveals the energy ratio between the estimated latent process $\{\hat{x}_k\}$ and the summed process $\{\hat{x}_k + \hat{g}_k\}$.

The prediction performance of the proposed algorithm and the recursive LR are shown in Fig. 11. Specifically, the proposed algorithm with the incorporation of both types of environmental measurements achieves the best performance for all the four types of slowly-varying channel parameters. For the average channel SNR and the Nakagami-*m* fading parameter, the proposed algorithm outperforms considerably the recursive LR by introducing the latent process to model the temporal variation caused by unknown physical mechanisms. About the average RMS delay spread, thanks to its high correlation with the wind speed, the recursive LR achieves a good performance and outperforms the proposed algorithm without the incorporation of environmental measurements. About the average RMS Doppler spread, due to its fast decaying autocorrelation and limited correlation with environmental measurements, its prediction performance is not as good as the other three types of channel parameters.

B. KW-AUG14 WITH SEASONAL CHANNEL VARIATIONS

In KW-AUG14, the transmission waveform was received by an acoustic modem located 312 meters away from the source. The acoustic modem has 4 hydrophones which are fixed at the vertexes of a horizontal square with 7 cm side length. The water depth of the experimental area varies from 3 to 6 meters. The average channel SNR scaled by the transmission power, the Nakagami-*m* fading parameter, and the average RMS delay spread within each transmission and over 4 hydrophones are depicted in Fig. 2(b), along with the wind speed and the temperature measurements obtained from the Weather Underground [27]. The autocorrelation of those slowly-varying channel parameters and their correlation with environmental measurements are shown in Fig. 12. It can be seen that both the average channel SNR and the Nakagami-*m* fading parameter have high negative correlation with both the wind speed and the temperature, while the average RMS delay spread exhibits positive correlation with both types of environmental measurements. In addition, the sequences of all the three types of slowly-varying channel parameters exhibit a seasonal cycle of 96 (24 hours).

In the proposed algorithm, we set the forgetting factor $\lambda = 1$ for all the three types of slowly-varying channel parameters. According to the MDL criterion in (46), the optimal orders of the latent process are chosen as $P = 1$ and $P_{\text{se}} = 1$ for the average channel SNR and the Nakagami-*m* fading parameter, and $P = 2$ and $P_{\text{se}} = 1$ for the average RMS delay spread. In addition, the sequences of the wind speed and the temperature are normalized individually to

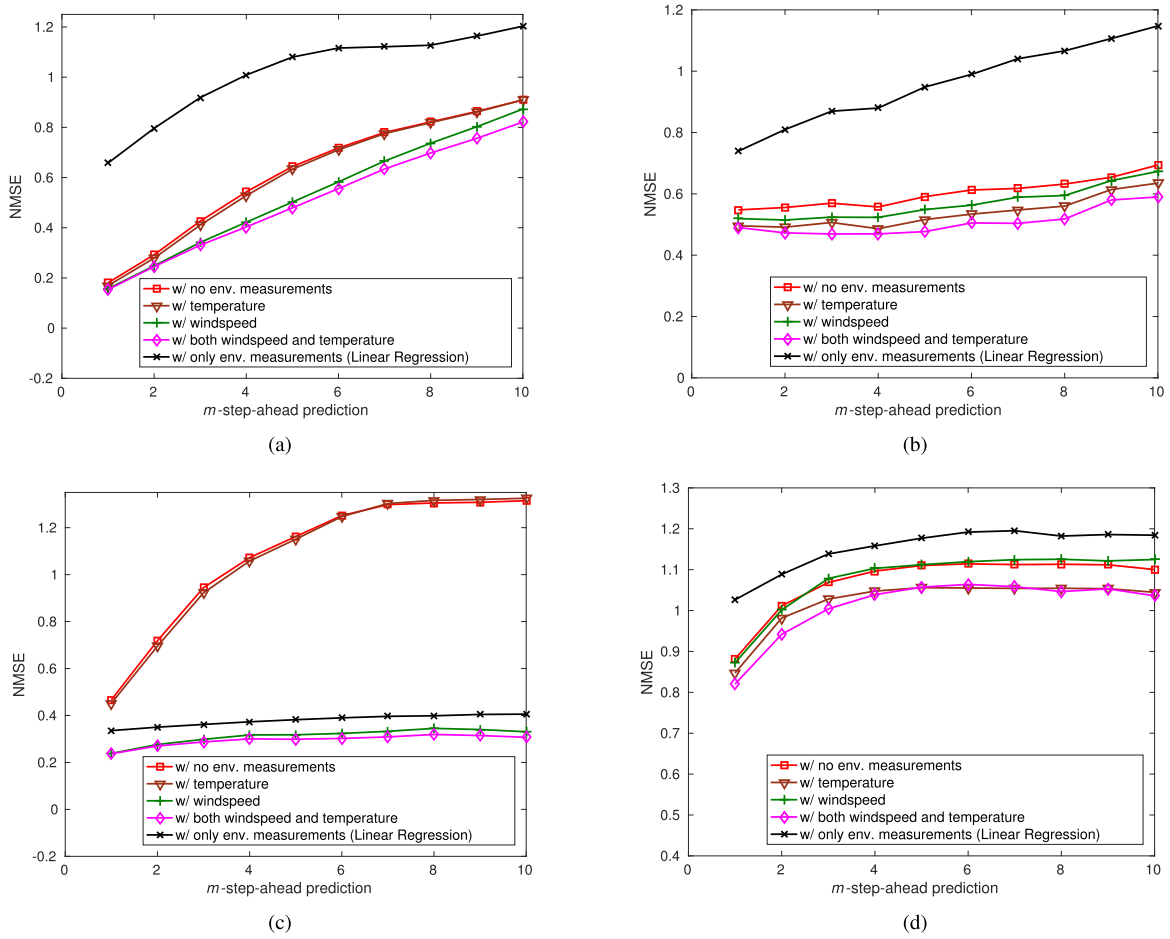


FIGURE 11. SPACE08: Prediction performance of the proposed algorithm and the recursive linear regression in non-seasonal channels. (a) average channel SNR. (b) Nakagami-*m* fading parameter. (c) average RMS delay spread. (d) average RMS Doppler spread.

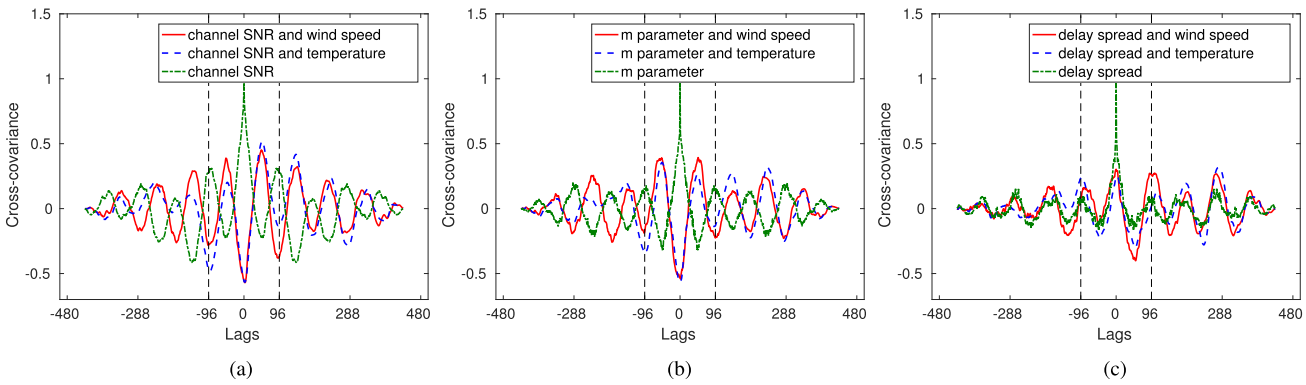


FIGURE 12. KW-AUG14: Autocorrelation of slowly-varying channel parameters and their correlation with environmental measurements. (a) average channel SNR. (b) Nakagami-*m* fading parameter. (c) average RMS delay spread.

have a unit power, and a linear combination of the two types of environmental parameters will be used for modeling the process $\{g_k\}$ (c.f. (8)). With the incorporation of both types of environmental measurements into the modeling, the estimated model parameters are listed in Table 2.

The prediction performance of the average channel SNR, the Nakagami-*m* fading parameter, and the average RMS delay spread are shown in Fig. 13. For comparison, the algorithm proposed for non-seasonal channels is also evaluated, where the latent process is modeled as an AR(P) process without considering the seasonality. It can be observed that the proposed algorithm for seasonal channels achieves the best performance when both types of environmental measurements are incorporated. Additionally, as the wind speed and the temperature have similar cross-correlation with the sequences of the three slowly-varying channel parameters (c.f. Fig. 12), similar performances are obtained when either type of the environmental measurements is incorporated into

TABLE 2. The estimated model parameters.

Slowly-varying channel parameter	$\hat{\mathbf{a}}$	$\hat{\xi}$	$\hat{\mathbf{b}}$ ([wind speed, temperature])	$\hat{\gamma}_0$	η
SPACE08: Channel SNR	0.967	-	[-1.189, -0.206]	12.503	0.828
SPACE08: m fading parameter	[0.153,0.197,0.251,0.189]	-	[-0.0004, 0.0005]	7.99	0.989
SPACE08: RMS delay spread	0.759	-	[-0.411, -0.039]	2.204	0.235
SPACE08: RMS Doppler spread	[0.267,0.130]	-	[0.010, 0.013]	0.748	0.952
KW-AUG14: Channel SNR	0.908	0.068	[-0.098, -0.145]	23.486	0.513
KW-AUG14: m fading parameter	[0.593,0.148,0.145]	-0.013	[-0.193, -0.142]	5.96	0.391
KW-AUG14: RMS delay spread	[0.259,0.224]	0.061	[0.006, 0.006]	1.296	0.976

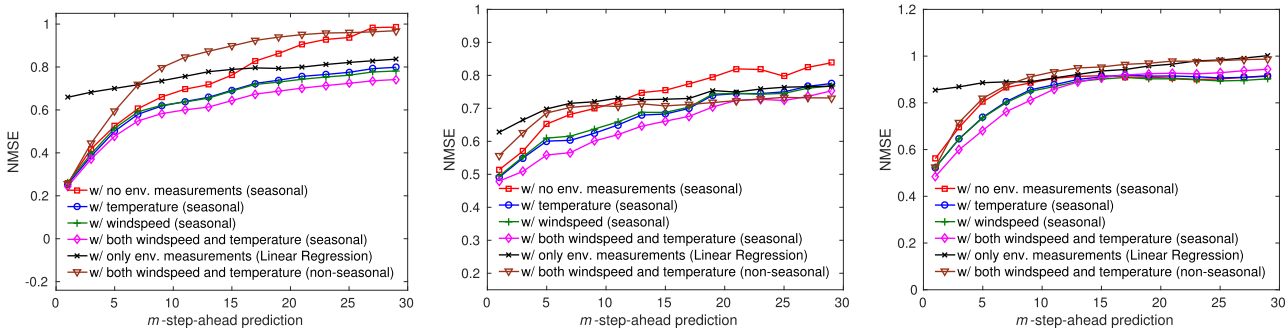


FIGURE 13. KW-AUG14: Prediction performance of several algorithms in seasonal channels. (a) average channel SNR. (b) Nakagami-*m* fading parameter. (c) average RMS delay spread.

the modeling. Furthermore, compared to the model and the algorithm proposed for non-seasonal channels, the proposed model and algorithm for seasonal channels achieve superior performance by explicitly modeling the channel seasonality and correspondingly exploiting the seasonality for prediction.

VIII. CONCLUSIONS

This work studied the online modeling and prediction of slowly-varying locally-averaged channel parameters over a long term, by exploiting their inherent temporal correlation and correlation with environmental conditions. From a data-driven perspective, the temporal evolution of a slowly-varying channel parameter of interest was modeled as the summation of a time-invariant component, a process that can be explicitly represented by available and relevant environmental parameters, and a Markov latent process that describes the contribution from unknown or unmeasurable physical mechanisms. A recursive algorithm was developed to estimate the unknown model parameters based on sequentially collected channel measurements and environmental parameters during real-time system operations. The updated model allows multiple-step-ahead prediction of the slowly-varying channel parameter, which could then guide higher-level proactive adaptation of communication strategies to the channel dynamics. The proposed model and the recursive algorithm were extended to seasonal channels by introducing a multiplicative seasonal AR process to model the channel seasonal correlation. Simulations and data sets from two shallow-water experiments were used to validate the effectiveness of the proposed models and algorithms. The experimental data processing revealed that the average channel SNR, the Nakagami-m fading parameter, and the average RMS delay spread can be reasonably well predicted. In addition, superior modeling and prediction performance can be achieved by exploiting the seasonal correlation in seasonal channels.

APPENDIX

DERIVATION FROM EQ. (21) to Eqs. (22) and (23)

According to (21), we have the batched representation of $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}})$,

$$Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}) = \mathbb{E}[\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})] + \sum_{k'=0}^{k-1} \lambda^{k-k'} \mathbb{E}[\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})] + \ln f(\mathbf{x}_{-1}|\boldsymbol{\Theta}), \tag{52}$$

where the expectation of $[\ln f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})]$ is performed with respect to $\tilde{f}(x_k, \mathbf{x}_{k-1}|y_k, \mathbf{y}_0^{k-1}, \hat{\boldsymbol{\Theta}})$, and the expectation of $[\ln f(x_{k'}, y_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})]$, $k' < k$ is performed with respect to $\tilde{f}(x_{k'}, \mathbf{x}_{k'-1}|y_{k'}, \mathbf{y}_0^{k'-1}, \hat{\boldsymbol{\Theta}}_{k'})$. Note that $f(x_k, y_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta}) = f(y_k|x_k, \boldsymbol{\Theta}) f(x_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})$. We have,

$$Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}) = \mathbb{E}[\ln f(x_k|\mathbf{x}_{k-1}, \boldsymbol{\Theta})] + \mathbb{E}[\ln f(y_k|x_k, \boldsymbol{\Theta})] + \sum_{k'=0}^{k-1} \lambda^{k-k'} \{\mathbb{E}[\ln f(x_{k'}|\mathbf{x}_{k'-1}, \boldsymbol{\Theta})] + \mathbb{E}[\ln f(y_{k'}|x_{k'}, \boldsymbol{\Theta})]\} + \ln f(\mathbf{x}_{-1}|\boldsymbol{\Theta}). \tag{53}$$

Substitute $f(y_k|x_k, \boldsymbol{\Theta}) = \mathcal{N}(\gamma_0 + x_k + \mathbf{b}^{\mathrm{T}}\mathbf{u}_k, \sigma_v^2)$ and $f(x_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{a}^{\mathrm{T}}\mathbf{x}_{k-1}, \sigma_w^2)$ into (53), and set the partial derivative of $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}})$ with respect to each unknown parameter in the set $\boldsymbol{\Theta} = \{\gamma_0, \mathbf{a}, \mathbf{b}, \sigma_w^2, \sigma_v^2\}$ to zero. One can

obtain the batched estimation of the unknown parameters. The recursive estimation can then be derived based on the batched estimation. Next, we take $\mathbf{a}$ as an example to illustrate the derivation of the recursive estimation in (22a). The recursive estimation of all the other unknown parameters can be similarly derived.

Substitute $f(x_{k'}|\mathbf{x}_{k'-1}) = \mathcal{N}(\mathbf{a}^{\mathrm{T}}\mathbf{x}_{k'-1}, \sigma_w^2)$ into (53). We have

$$\begin{aligned} -Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}) = \mathbb{E}\left[\frac{1}{2\sigma_w^2}(x_k - \mathbf{a}^{\mathrm{T}}\mathbf{x}_{k-1})^2\right] \\ + \sum_{k'=0}^{k-1} \lambda^{k-k'} \mathbb{E}\left[\frac{1}{2\sigma_w^2}(x_{k'} - \mathbf{a}^{\mathrm{T}}\mathbf{x}_{k'-1})^2\right] \\ + \text{others}. \end{aligned} \tag{54}$$

Set the partial derivative of $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}})$ with respect to $\mathbf{a}$ to zero,

$$\begin{aligned} -\frac{\partial Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}})}{\partial \mathbf{a}} = \mathbb{E}\left[\frac{1}{2\sigma_w^2}(x_k - \mathbf{a}^{\mathrm{T}}\mathbf{x}_{k-1})\mathbf{x}_{k-1}^{\mathrm{T}}\right] \\ + \sum_{k'=0}^{k-1} \lambda^{k-k'} \mathbb{E}\left[\frac{1}{2\sigma_w^2}(x_{k'} - \mathbf{a}^{\mathrm{T}}\mathbf{x}_{k'-1})\mathbf{x}_{k'-1}^{\mathrm{T}}\right] = 0. \end{aligned} \tag{55}$$

We obtained the batched estimation of $\mathbf{a}$ at time k,

$$\hat{\mathbf{a}}_k = \mathbf{M}_{k-1}^{-1}\boldsymbol{\pi}_k, \tag{56}$$

where the matrix $\mathbf{M}_{k-1}$ and the vector $\boldsymbol{\pi}_k$ are defined, respectively, as

$$\begin{aligned} \mathbf{M}_{k-1} &:= \mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}] + \sum_{k'=0}^{k-1} \lambda^{k-k'} \mathbb{E}[\mathbf{x}_{k'-1}\mathbf{x}_{k'-1}^{\mathrm{T}}], \\ \boldsymbol{\pi}_k &:= \mathbb{E}[x_k\mathbf{x}_{k-1}] + \sum_{k'=0}^{k-1} \lambda^{k-k'} \mathbb{E}[x_{k'}\mathbf{x}_{k'-1}]. \end{aligned}$$

which can be recursively represented as,

$$\begin{aligned} \mathbf{M}_{k-1} &= \lambda\mathbf{M}_{k-2} + \mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}], \\ \boldsymbol{\pi}_k &= \lambda\boldsymbol{\pi}_{k-1} + \mathbb{E}[x_k\mathbf{x}_{k-1}]. \end{aligned}$$

According to the Woodbury matrix identity [26], we have

$$\mathbf{M}_{k-1}^{-1} = \lambda^{-1}\mathbf{M}_{k-2}^{-1} - \lambda^{-1}\mathbf{M}_{k-1}^{-1}\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\mathbf{M}_{k-2}^{-1}. \tag{57}$$

The recursive representation of (56) can then be derived as,

$$\begin{aligned} \hat{\mathbf{a}}_k &= \mathbf{M}_{k-1}^{-1}\mathbb{E}[x_k\mathbf{x}_{k-1}] \\ &\quad + \left(\lambda^{-1}\mathbf{M}_{k-2}^{-1} - \lambda^{-1}\mathbf{M}_{k-1}^{-1}\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\mathbf{M}_{k-2}^{-1}\right)\lambda\boldsymbol{\pi}_{k-1} \\ &= \mathbf{M}_{k-1}^{-1}\mathbb{E}[x_k\mathbf{x}_{k-1}] \\ &\quad + \left(\mathbf{M}_{k-2}^{-1}\boldsymbol{\pi}_{k-1} - \mathbf{M}_{k-1}^{-1}\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\mathbf{M}_{k-2}^{-1}\boldsymbol{\pi}_{k-1}\right) \\ &= \mathbf{M}_{k-1}^{-1}\mathbb{E}[x_k\mathbf{x}_{k-1}] + \left(\hat{\mathbf{a}}_{k-1} - \mathbf{M}_{k-1}^{-1}\mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\hat{\mathbf{a}}_{k-1}\right) \\ &= \hat{\mathbf{a}}_{k-1} + \mathbf{M}_{k-1}^{-1}\left(\mathbb{E}[x_k\mathbf{x}_{k-1}] - \mathbb{E}[\mathbf{x}_{k-1}\mathbf{x}_{k-1}^{\mathrm{T}}]\hat{\mathbf{a}}_{k-1}\right). \end{aligned} \tag{58}$$

For the proposed recursive and iterative algorithm in Section III-B, corresponding to the parameter set estimation $\hat{\boldsymbol{\Theta}}_k^{(i)}$ in the ith iteration, the result in (58) can be generalized to (22a) which is obtained by maximizing $Q_k(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}_k^{(i)})$.

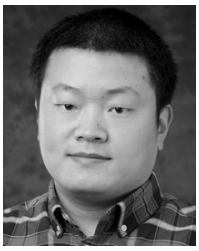
ACKNOWLEDGMENT

This paper was partially presented at the MTS/IEEE OCEANS Conference, Shanghai, China, in April 2016 [1].

REFERENCES

[1] W. Sun and Z. Wang, "Modeling and prediction of large-scale temporal variation in underwater acoustic channels," in *Proc. MTS/IEEE OCEANS Conf.*, Shanghai, China, Apr. 2016, pp. 1–6.
[2] N. M. Carbone and W. S. Hodgkiss, "Effects of tidally driven temperature fluctuations on shallow-water acoustic communications at 18 kHz," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 84–94, Jan. 2000.
[3] A. Song, M. Badiey, H. C. Song, W. S. Hodgkiss, and M. B. Porter, "Impact of ocean variability on coherent underwater acoustic communications during the Kauai experiment (KauaiEx)," *J. Acoust. Soc. Amer.*, vol. 123, no. 2, pp. 856–865, Feb. 2008.
[4] P. A. van Walree, "Propagation and scattering effects in underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 614–631, Oct. 2013.
[5] D. B. Kilfoyle and A. B. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 4–27, Jan. 2000.
[6] M. Chitre, S. Shahabudeen, and M. Stojanovic, "Underwater acoustic communications and networking: Recent advances and future challenges," *Marine Technol. Soc. J.*, vol. 42, no. 1, pp. 103–116, 2008.
[7] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 701–717, Oct. 2013.
[8] A. Caiti *et al.*, "Linking acoustic communications and network performance: Integration and experimentation of an underwater acoustic network," *IEEE J. Ocean. Eng.*, vol. 38, no. 8, pp. 758–771, Oct. 2013.
[9] B. Tomasi, J. Preisig, G. B. Deane, and M. Zorzi, "A study on the wide-sense stationarity of the underwater acoustic channel for non-coherent communication systems," in *Proc. Eur. Wireless Conf.*, Apr. 2011, pp. 1–6.
[10] M. Chitre and K. Pelekanakis, "Channel variability measurements in an underwater acoustic network," in *Proc. Underwater Commun. Netw. (UComms)*, Sestri Levante, Italy, Sep. 2014, pp. 1–4.
[11] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Hoboken, NJ, USA: Wiley, 2008.
[12] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. Berlin, Germany: Springer Science & Business Media, 2011.
[13] M. A. Ainslie and J. G. McColm, "A simplified formula for viscous and chemical absorption in sea water," *J. Acoust. Soc. Amer.*, vol. 103, no. 3, pp. 1671–1672, 1998.
[14] D. E. Weston and P. A. Ching, "Wind effects in shallow-water acoustic transmission," *J. Acoust. Soc. Amer.*, vol. 86, no. 4, pp. 1530–1545, 1989.
[15] D. Jackson, "APL-UW high-frequency ocean environmental acoustic models handbook," Appl. Phys. Lab., Univ. Washington, Seattle, WA, USA, Tech. Rep. 9407, 1994.
[16] B. B. Ma, J. A. Nystuen, and R.-C. Lien, "Prediction of underwater sound levels from rain and wind," *J. Acoust. Soc. Amer.*, vol. 117, no. 6, pp. 3555–3565, May 2005.
[17] M. Chitre, "A high-frequency warm shallow water acoustic communications channel model and measurements," *J. Acoust. Soc. Amer.*, vol. 122, no. 5, pp. 2580–2586, 2007.
[18] P. A. van Walree and R. Otnes, "Ultrawideband underwater acoustic communication channels," *IEEE J. Ocean. Eng.*, vol. 38, no. 4, pp. 678–688, Oct. 2013.
[19] J. Llor and M. P. Malumbres, "Statistical modeling of large-scale signal path loss in underwater acoustic networks," *Sensors*, vol. 13, no. 2, pp. 2279–2294, 2013.

[20] P. Qarabaqi and M. Stojanovic, "Modeling the large scale transmission loss in underwater acoustic channels," in *Proc. Allerton Conf. Commun., Control Comput.*, Monticello, IL, USA, Sep. 2011, pp. 445–452.
[21] C. M. Bishop, *Pattern Recognition and Machine Learning*, 6th ed. New York, NY, USA: Springer-Verlag, 2006.
[22] E. Demirors, G. Sklivanitis, G. E. Santagati, T. Melodia, and S. N. Batalama, "A high-rate software-defined underwater acoustic modem with real-time adaptation capabilities," *IEEE Access*, vol. 6, pp. 18602–18615, 2018.
[23] Z.-H. Wang, C. Wang, and W. Sun, "Adaptive transmission scheduling in time-varying underwater acoustic channels," in *Proc. MTS/IEEE OCEANS Conf.*, Washington, DC, USA, Oct. 2015, pp. 1–6.
[24] B. Tomasi, L. Toni, P. Casari, L. Rossi, and M. Zorzi, "Performance study of variable-rate modulation for underwater communications based on experimental data," in *Proc. MTS/IEEE OCEANS Conf.*, Seattle, WA, USA, Sep. 2010, pp. 1–8.
[25] H. Hashemi and D. Tholl, "Statistical modeling and simulation of the RMS delay spread of indoor radio propagation channels," *IEEE Trans. Veh. Technol.*, vol. 43, no. 1, pp. 110–120, Feb. 1994.
[26] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. Hoboken, NJ, USA: Wiley, 2001.
[27] The Weather Channel. (Jun. 2015). *Weather History and Data Archieve*. [Online]. Available: http://www.wunderground.com/history/
[28] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
[29] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.
[30] The Woods Hole Oceanographic Institution. (Aug. 2016). *The Martha's Vinyard Coastal Observatory*. [Online]. Available: http://www.whoi.edu/mvco



WENSHENG SUN (S'15) received the B.S. degree in electrical engineering from Yanshan University, Qinhuangdao, China, in 2010 and the M.S. degree in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 2012. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, USA.

His research interests lie in the areas of wireless communications, statistical signal processing, and machine learning, with the recent focus on online learning and inversion for underwater acoustic communications.

ZHAOHUI WANG (S'10–M'13) received the B.S. degree from the Beijing University of Chemical Technology in 2006, the M.S. degree from the Institute of Acoustics, Chinese Academy of Sciences, Beijing, China, in 2009, and the Ph.D. degree from the University of Connecticut, Storrs, in 2013, all in electrical engineering.

She has been with the Department of Electrical and Computer Engineering, Michigan Technological University, Houghton, as an Assistant Professor, since 2013. Her research interests lie in the areas of wireless communications, networking, and statistical signal processing, with a recent focus on signal processing and machine learning techniques for wireless communications and networking in underwater acoustic environments.

Dr. Wang was honored with the Women of Innovation Award by the Connecticut Technology Council in 2013. She was a recipient of the NSF CAREER Award in 2017. She served as a technical reviewer for many premier journals and conferences. She was recognized as an Outstanding Reviewer by the IEEE JOURNAL OF OCEANIC ENGINEERING from 2012 to 2016.