Modeling realistic underwater acoustic networks using experimental data

Mandar Chitre
Department of Electrical and Computer Engineering,
and ARL, Tropical Marine Science Institute,
National University of Singapore.
mandar@arl.nus.edu.sg

Gabriel Chua
ARL, Tropical Marine Science Institute,
National University of Singapore.
gabriel@arl.nus.edu.sg

***Abstract*—Since underwater network experiments are logistically challenging and expensive to conduct, many researchers evaluate the performance of their protocols through simulation. The validity of simulation results strongly depends on the accuracy of the channel model used. Most network simulators either neglect time variability of links, or model the variability to be independent across links. Both approaches are questionable approximations of reality, since link performance strongly depends on environmental factors that may vary with space and time. We use data from underwater network experiments to measure and model the spatiotemporal variability of network performance. This approach allows researchers to test protocols in realistic simulation environments driven by representative experimental datasets, long after the experiments are conducted.**

I. Introduction

Underwater acoustic channels exhibit a high degree of temporal and spatial variability, especially in shallow coastal areas. Recent years have seen a growing interest in understanding, modeling and simulating this variability [1]–[7]. Analysis of data from several experiments has led to statistical models that efficiently simulate time-varying channel impulse responses [4]. While these models accurately predict the statistics of individual multipath arrivals, they do not model the correlation between multiple arrivals or across multiple network links. Such correlations can arise from common environmental factors that directly affect multiple paths or links.

Since underwater networking experiments are logistically challenging and expensive to conduct, researchers often evaluate the performance of their protocols through simulation. The validity of simulation results strongly depends on the accuracy of the channel model used. Most network simulators either neglect time variability of links, or model the variability to be independent across links. Since link performance strongly depends on environmental factors that may vary with space and time, both approaches could potentially lead to a poor match between simulation and reality.

In order to understand time-variability of the underwater acoustic channel across multiple links, and its impact on network performance, the MISSION 2012 and 2013 experiments were conducted in Singapore waters in October 2012 and November 2013 respectively. The experiments involved multiple underwater networks with static and mobile nodes operating concurrently. Several environmental sensors were also deployed in the area, so that a rich communication and environmental dataset was obtained for statistical analysis. In [8], we presented preliminary analysis of channel impulse response variability measurements made as part of the experiment. In this paper, we use data from the MISSION 2013 experiment to measure and model the spatio-temporal variability of the network performance. This allows testing of protocols in realistic simulation environments driven by representative experimental datasets, long after the experiments are conducted.

We first examine the prevalence of spatial and temporal correlation using experimental data. We consider the packet transmission success as a non-independent Bernoulli random process, and characterize it though the use of the auto mutual information function. We find that while spatial correlation between network links can usually be neglected, temporal dependence must be correctly modeled. We use a Markov model for this, and present a method to easily estimate model parameters. Finally, we present a method to extend the Markov model to account for long-term variability of the channel.

The paper is organized as follows. In section II, we outline the experiment that provided the data used in this work. We present our estimates of spatial and temporal network performance variability in section III. In section IV, we present a Markov model that explains the observations from section III. We summarize the resulting simulation method in section V, and summarize our conclusions in section VI.

II. Experimental Setup

The MISSION 2013 experiment was conducted at Selat Pauh in Singapore waters from 15$^{\text{th}}$ to 29$^{\text{th}}$ November 2013. The experiment involved multiple underwater networks with static and mobile nodes operating concurrently. Several environmental sensors were also deployed in the area, so that a rich communication and environmental dataset was obtained for statistical analysis. Several underwater communication techniques, network protocols, and localization algorithms were tested during the experiment. In this paper, we focus our attention on data collected over a 3-day period from 26$^{\text{th}}$ to 28$^{\text{th}}$ November.

The UNET network deployed during the experiment consisted of a UNET-2 modem [9] (node 21) mounted 6 meters

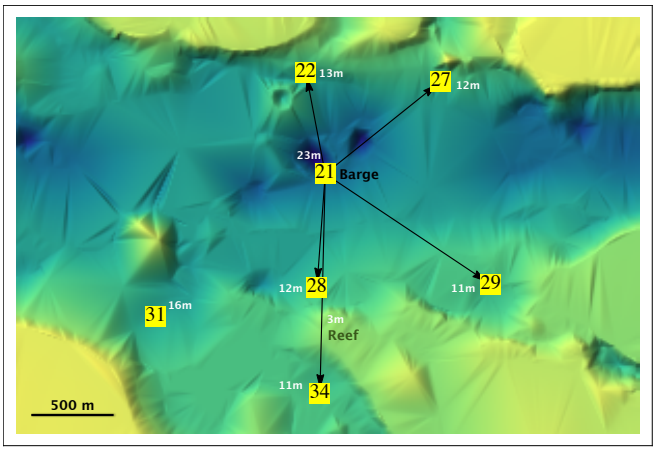
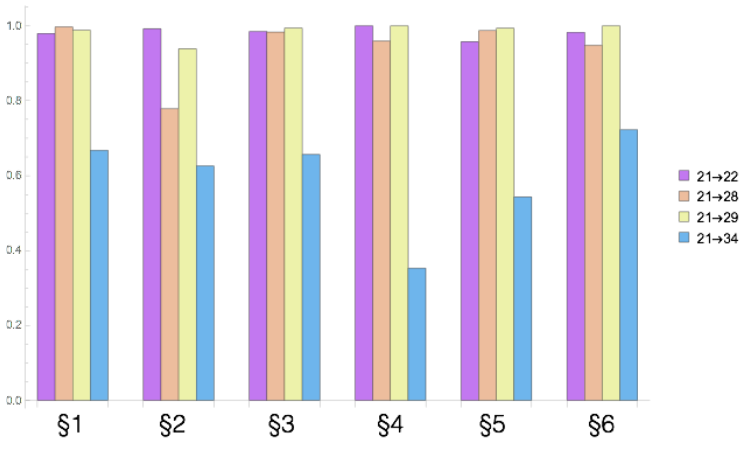


Fig. 1. UNET network node locations during the MISSION 2013 experiment (deployment #2). Yellow squares are network nodes. Water depths are marked in white at selected areas.



	21→22	21→28	21→29	21→34
§1 :: 0.1-1.6 hrs, 343 pkts	0.9796	0.9971	0.9883	0.6676
§2 :: 20.6-21. hrs, 131 pkts	0.9924	0.7786	0.9389	0.6260
§3 :: 22.4-24.9 hrs, 664 pkts	0.9849	0.9834	0.9940	0.6566
§4 :: 44.5-45.1 hrs, 221 pkts	1.000	0.9593	1.000	0.3529
§5 :: 45.6-46. hrs, 784 pkts	0.9579	0.9872	0.9936	0.5434
§6 :: 46.4-47.1 hrs, 173 pkts	0.9827	0.9480	1.000	0.7225

Fig. 2. Coarse-scale PDR for various links and data segments (§1–§6). Data segments are spread across 3 days. The time, relative to the start of the experiment, and the number of packet transmissions for each data segment are shown in the left column of the table. The y-axis of the graph and the right column of the table shows the PDR.

below a barge, and six UNET-PANDA nodes [5] (nodes 22, 27, 28, 29, 31 and 34) deployed at various locations within a 2×2 km area around the barge. The water depth in the area ranged from 23 m near the barge to about 3 m at the reef. Most of the UNET-PANDA nodes were in about 10-15 m water depth, deployed about 3 m above the seabed. The network geometry is depicted in Fig. 1.

III. Network Performance Variability

A. Coarse-scale variability

At a coarse-scale, we measure the average packet delivery ratio (PDR) for various links at various time intervals during the experiment. PDR estimates for links originating at the central node 21 for various data segments (§1-§6) are shown in Fig. 2. We see that the PDR varies significantly across links, and across data segments. The time variability for different links is different, and not very correlated. While we do not, as yet, have a clear conclusion on what factors control link performance at a coarse scale, it is suspected that the performance is dependent on environmental factors such as wind, currents, etc. We are currently exploring the relationship of link performance with various environmental factors.

B. Fine-scale variability

At a finer scale, we consider each packet transmitted from a specific node (say node 21) and estimate the probability of successful reception at each of the other nodes. We model the process as a Bernuolli random process. While a conventional Bernuolli random process assumes independent arrivals, we do not require that the arrivals across links or over time be independent random variables.

We measure the uncertainty of whether a packet would be successfully received using an entropy measure $\mathcal{E}(X)$ where X denotes the random process with 0 for a successful delivery, and 1 for a failed delivery:

$$\mathcal{E}(X) = -\sum_{x\in\{0,1\}} \mathcal{P}_X(x)\log_2 \mathcal{P}_X(x). \tag{1}$$

Since two processes are not necessarily independent, we measure the dependency between the processes by computing the mutual information:

$$\begin{aligned}\mathcal{I}(X,Y) &= \mathcal{E}(X)+\mathcal{E}(Y)-\mathcal{E}(XY) \\ &= \sum_{x,y\in\{0,1\}} \mathcal{P}_{XY}(x,y)\log_2 \frac{\mathcal{P}_{XY}(x,y)}{\mathcal{P}_X(x)\mathcal{P}_Y(y)}.\end{aligned} \tag{2}$$

The values of $\mathcal{P}_X(x)$ and $\mathcal{P}_{XY}(x,y)$ are easily estimated from experimental data using the success/failure outcomes for each transmitted packet on each link.

1) Spatial dependence: We check if pairs of links are independent by estimating the mutual information between the random processes for each of the links. The results are shown in Fig. 3. Table (a) shows the dependence between two links that have very little spatial overlap, while table (b) shows the results for two links which pass through the same region in space for some distance. We see that the mutual information $\mathcal{I}$, in both cases, is generally about two orders of magnitude lower than the individual link entropies $\mathcal{E}_1$ and $\mathcal{E}_2$. This suggests the the spatial dependence is small enough to be ignored for most practical purposes, and that a model/simulation does not need to take it into account explicitly.

2) Temporal dependence: The temporal dependence can be characterized using the mutual information between the random process for a link, and a delayed version of it. In a

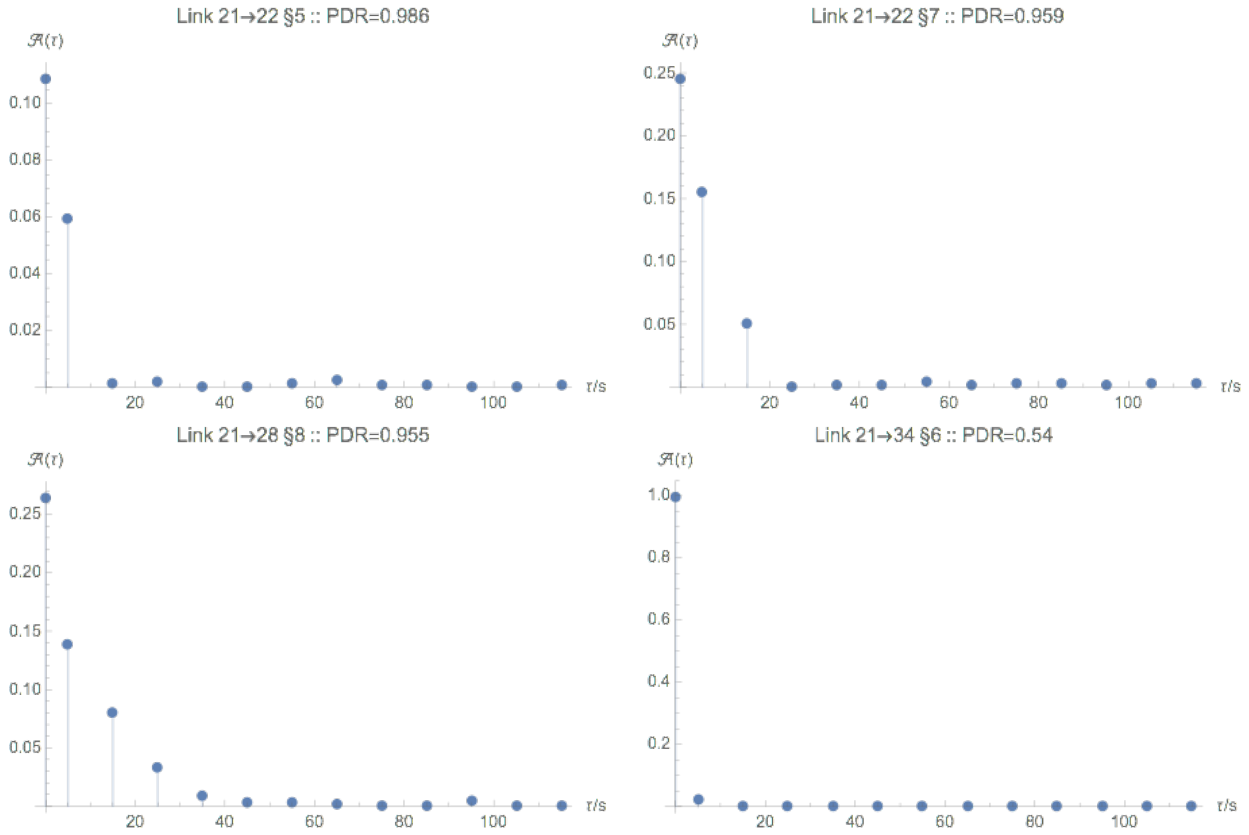


Fig. 4. AMIF $\mathcal{A}(\tau)$ estimated for various links and data segments.

Time (hrs)	pkts	$\mathcal{I}$	$\mathcal{E}_1$	$\mathcal{E}_2$
0.1-1.6	410	0.0001	0.14	0.04
20.6-21.	146	0.0025	0.06	0.76
22.4-25.5	923	0.0011	0.11	0.13
44.5-45.4	348	0.	0.	0.26
45.6-46.	824	0.0007	0.25	0.09
46.4-47.1	181	0.0014	0.12	0.31
47.4-49.4	570	0.0015	0.1	0.24

(a) Links between node 21→22 and 21→28.

Time (hrs)	pkts	$\mathcal{I}$	$\mathcal{E}_1$	$\mathcal{E}_2$
0.1-1.6	363	0.0044	0.03	0.92
20.6-21.	135	0.0018	0.76	0.96
21.4-22.1	192	0.0082	0.05	0.92
22.4-25.	714	0.0055	0.14	0.92
44.5-45.2	261	0.0236	0.22	0.95
45.6-46.	793	0.003	0.11	0.99
46.4-49.4	824	0.0007	0.23	0.89

(b) Links between node 21→28 and 21→34.

Fig. 3. Mutual information $\mathcal{I}$ and entropies $\mathcal{E}_1$, $\mathcal{E}_2$ for pairs of links.

similar vein as an auto-correlation function (ACF), this leads to a concept of an auto mutual information function (AMIF):

$$\mathcal{A}_X(\tau) = \mathcal{I}(X_t, X_{t-\tau}). \tag{3}$$

The AMIF at lag zero is simply the entropy of the link, i.e., $\mathcal{A}_X(0) = \mathcal{E}(X)$. The AMIF is a symmetric function: $\mathcal{A}_X(\tau) = \mathcal{A}_X(-\tau)$, and its value goes to zero for large lags: $\lim_{\tau\to\infty} \mathcal{A}_X(\tau) = 0$. The rate at which the AMIF decays to zero characterizes the temporal coherence scale of the process.

Fig. 4 shows the AMIF for various links and data segments. These AMIFs are representative of the entire dataset, and show how the typical AMIFs decay with time lag. Any model which accurately characterizes the temporal dependence of the random Bernoulli process for a link must model this decay well.

IV. Packet Delivery Model

The AMIF for a random independent Bernoulli process is zero for all non-zero lags. Given the measured estimates in Fig. 4, the random independent Bernoulli process is obviously a poor model of reality. More complex models such as Markov models or hidden Markov models (HMM) can more accurately produce AMIF such as those estimated from data. We focus on Markov models due to their parsimony and simplicity in estimation. We find that they adequately fit our experimental observations, and that we do not need to resort to more complex models (such as HMMs). Other researchers have previously reported that while Markov models were sometimes sufficient, at other times they required HMMs to accurately model the data [10]. However, the Markov models and HMMs used in [10] are based on discrete time modeling with transmissions at regular intervals, are hence somewhat

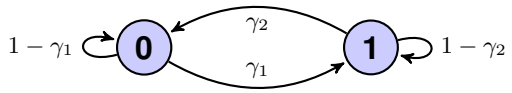


Fig. 5. Markov model for packet success (0) / failure (1).

different from the ones proposed here.

A. Markov model

The Markov model we propose is represented schematically in Fig. 5. The model is a simple 2-state Markov process with a 1-step transition matrix

$$\Gamma = \begin{pmatrix} 1-\gamma_1 & \gamma_1 \\ \gamma_2 & 1-\gamma_2 \end{pmatrix} \tag{4}$$

where γ_1 and γ_2 are the parameters of the model. We start by assuming a discrete time step of 1 second, and later generalize to continuous time. A k-step transition matrix (representing a k-second step) is then simply Γ^k. An eigensystem analysis of the matrix allows us to write the matrix power in terms of a scalar power:

$$\Gamma^k = \begin{pmatrix} \delta_1 & \delta_2 \\ \delta_1 & \delta_2 \end{pmatrix} + (1-\gamma_1-\gamma_2)^k \begin{pmatrix} \delta_2 & -\delta_2 \\ -\delta_1 & \delta_1 \end{pmatrix} \tag{5}$$

where $\delta_1 = \gamma_2/(\gamma_1+\gamma_2)$ and $\delta_2 = \gamma_1/(\gamma_1+\gamma_2)$. By allowing k to take on non-integer (fractional) positive values, we generalize our model to allow continuous time.

B. Parameter estimation

The PDR for the Markov model is easily computed:

$$\text{PDR} = \delta_1 = \frac{\gamma_2}{\gamma_1+\gamma_2}. \tag{6}$$

The theoretical AMIF can also be computed from the model:

$$\mathcal{A}(\tau) = \sum_{i,j\in\{1,2\}} (\Gamma^\tau)_{i,j}\,\delta_i \log_2 \frac{(\Gamma^\tau)_{i,j}}{\delta_j} \tag{7}$$

where $(\Gamma^\tau)_{i,j}$ represents the i^{th} row and j^{th} column entry in the matrix Γ^τ.

If we have a measured AMIF $\hat{\mathcal{A}}(\tau)$, we estimate the parameters γ_1^* and γ_2^* through the following one-dimensional minimization:

$$\gamma_1^* = \arg\min_{\gamma_1} \sum_{0<\tau\le\tau_{\max}} \left|\mathcal{A}(\tau;\gamma_1,\gamma_2) - \hat{\mathcal{A}}(\tau)\right| \tag{8}$$

subject to PDR $p = \gamma_2/(\gamma_1+\gamma_2)$ and $0 \le \gamma_1 \le 1$. Once γ_1^* is estimated, γ_2^* is trivial to find:

$$\gamma_2^* = \frac{p}{1-p}\gamma_1^*. \tag{9}$$

The 1-norm is used in the minimization in preference to the more common Euclidean 2-norm to ensure that the fitting is not dominated by the high-entropy region near zero-lag. We choose $\tau_{\max}$ to be 45 seconds, since most of our observed AMIFs go to zero by then.

C. Model validation

Using the method outlined above, we estimate the Markov model parameters for various links and data segments. We then plot the theoretical AMIF (solid line) against the measured AMIF (solid dots) in Fig. 6. We see that the model fits our data accurately.

V. Modeling & Simulation Methodology

Based on the results in the section IV, we next summarize our proposed method to model and simulate underwater networks based on experimental data.

A. Quasi-static channels

For models and simulations spanning a relatively short time (less than few hours), it is reasonable to assume that the channel is stationary and can be represented by a Markov model with fixed parameters. We call this channel a *quasi-static* channel.

To model the quasi-static channel, we use experimental data to estimate the average PDR and AMIF for each link in the network. We then use the parameter estimation method in section IV-B to find Markov model parameters γ_1 and γ_2 for each link. This then provides us a probability model that can be used for protocol design and performance evaluation. To simulate networks with this model, we simulate each link as a Bernoulli Markov process with the estimated parameters[1].

B. Modeling longer timescale variability

While Markov models may be able to model the network at all times, the parameters for the model may change over longer timescales. To model and simulate networks over these longer timescales, we use experimental data to determine model parameters at various time points over a long experiment. For this, the experimental data is divided into data segments spanning time $[t_i, t_{i+1})$. Each data segment yields one set of model parameters $(\gamma_1,\gamma_2)_i$ at time point $(t_i+t_{i+1})/2$. We then interpolate (e.g. using cubic splines) the estimated set of parameters to yield functions $\gamma_1(t)$ and $\gamma_2(t)$ for any arbitrary time t within the experiment. The model and simulation then simply use time-varying parameters $\gamma_1(t)$ and $\gamma_2(t)$ with the Markov model to effectively model long term channel variability accurately.

VI. Conclusions

Analysis of the MISSION 2013 packet delivery success/failure data suggests that a continuous time Markov model adequately represents the average PDR link performance and the measured link AMIF. The spatial dependence between links is small, and can be ignored for most practical purposes. The resulting model is simple and parsimonious, and useful in design and simulation of underwater networking protocols.

[1] An implementation of this model for use with UnetStack [11], based on the MISSION 2013 experiment, is available for download (www.unetstack.net).

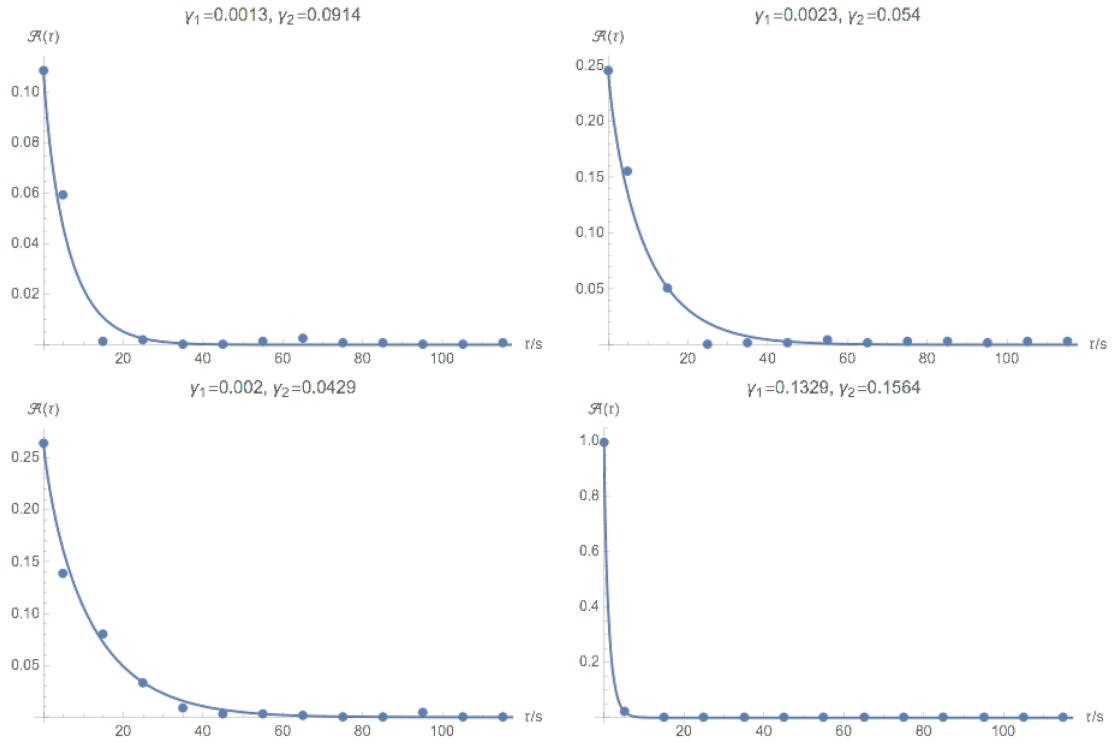


Fig. 6. Theoretical AMIF $\mathcal{A}(\tau)$ fitted to data for various links and data segments. The fitted model parameters γ_1 and γ_2 are shown for each fit.

ACKNOWLEDGEMENT

We are grateful to Temasek Defence Systems Institute, which provided partial funding for this work. We thank Dr. Venugopalan Pallayil for his leadership in organizing the MISSION 2013 experiment, and also thank Iulian Topor, Mohan Panayamdam, Manu Ignatius, Shanmugam Mpl and Sivaraman Selvakumar, who helped with instrumentation, deployment and logistics.

REFERENCES

[1] M. Badiey, Y. Mu, J. Simmen, and S. Forsythe, "Signal variability in shallow-water sound channels," *Oceanic Engineering, IEEE Journal of*, vol. 25, no. 4, pp. 492–500, Oct 2000.

[2] A. Song, M. Badiey, H. C. Song, W. S. Hodgkiss, M. B. Porter, and KauaiEx-Group, "Impact of ocean variability on coherent underwater acoustic communications during the kauai experiment (kauaiex)," *The Journal of the Acoustical Society of America*, vol. 123, no. 2, pp. 856–865, 2008.

[3] C. T. Tindle, G. B. Deane, and J. C. Preisig, "Reflection of underwater sound from surface waves," *The Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 66–72, 2009.

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

[5] M. Chitre, I. Topor, R. Bhatnagar, and V. Pallayil, "Variability in link performance of an underwater acoustic network," in *Proceedings of IEEE OCEANS 2013 Bergen*, June 2013.

[6] H. Dol, M. Colin, M. Ainslie, P. van Walree, and J. Janmaat, "Simulation of an underwater acoustic communication channel characterized by wind-generated surface waves and bubbles," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 642–654, Oct 2013.

[7] P. van Walree, "Propagation and scattering effects in underwater acoustic communication channels," *Oceanic Engineering, IEEE Journal of*, vol. 38, no. 4, pp. 614–631, Oct 2013.

[8] M. Chitre and K. Pelekanakis, "Channel variability measurements in an underwater acoustic network," in *Underwater Communications Networking (UComms 2014)*, Italy, 2014.

[9] M. Chitre, I. Topor, and T.-B. Koay, "The UNET-2 modem – an extensible tool for underwater networking research," in *Proceedings of IEEE OCEANS'12 Yeosu*, May 2012.

[10] B. Tomasi, P. Casari, L. Finesso, G. Zappa, K. McCoy, and M. Zorzi, "On modeling JANUS packet errors over a shallow water acoustic channel using Markov and hidden Markov models," in *Proceedings of Military Communications Conference 2010 (MILCOM 2010), IEEE*, November 2010, pp. 2406 –2411.

[11] M. Chitre, R. Bhatnagar, and W.-S. Soh, "UnetStack: an agent-based software stack and simulator for underwater networks," in *Proceedings of IEEE OCEANS'14 St John's*, September 2014.