

# Predicting underwater acoustic network variability using machine learning techniques

Vignesh Kalaiarasu, Hari Vishnu, Ahmed Mahmood, Mandar Chitre

Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore

e-mail: vigneshkalai@outlook.com, {hari, ahmed, mandar}@arlab.nus.edu.sg

**Abstract**—Predicting the performance of an underwater acoustic network (UAN) is a challenging task due to the spatio-temporal variability of the links and its complicated dependence on multiple factors. We present a machine-learning model based on logistic regression (LogR) to capture the spatio-temporal variation in the performance of a UAN. The model captures the effect of environmental factors such as wind speed, tide, current velocity etc., and modem-specific factors, on the performance of the UAN, which can be quantified by the packet success rate (PSR). As the PSR is a complicated non-linear function of environmental/modem-specific factors, developing a forward model in this regard is a difficult task, motivating our data-driven model. Our results indicate that LogR can quantify UAN performance with fair accuracy.

**Index Terms**—underwater acoustics, communication network, variability, logistic regression, environmental factors

## I. INTRODUCTION

Underwater acoustic networks (UANs) comprise of several nodes communicating acoustically with each other via communication links. UANs have garnered interest due to their potential for use in applications such as large-scale monitoring [1]. Experiments using UANs are expensive and not easy to conduct. Thus, in order to evaluate communication protocols and assess capabilities of systems being developed, researchers often resort to simulations based on models of the UAN. Some simulators used for this purpose can be found in the literature [2], [3].

An underwater acoustic channel is severely band-limited, offers significant delay spread and is time-varying [4]. The spatio-temporal variability of the underwater acoustic channel and its effect on UAN performance have been studied previously [5]–[7]. Analysis of data collected from experiments in Singapore waters in 2012 showed that link performance was affected by underwater currents which caused vibrations of the modems [6]. In [7], the authors studied the short-term and long-term variability of the multipath arrivals in an acoustic channel using experimental data, and suggested that the short-term variability of multipath arrivals can be modeled as Gaussian or Rician distributed. In [5], the performance in terms of packet success rate (PSR) was modeled as a non-independent Bernoulli random process and a Markov model was used to estimate model parameters.

The variability of acoustic channels causes individual link performances in a UAN to vary significantly [5]. There is a multitude of interconnected factors such as local environmental conditions and communication equipment-related factors

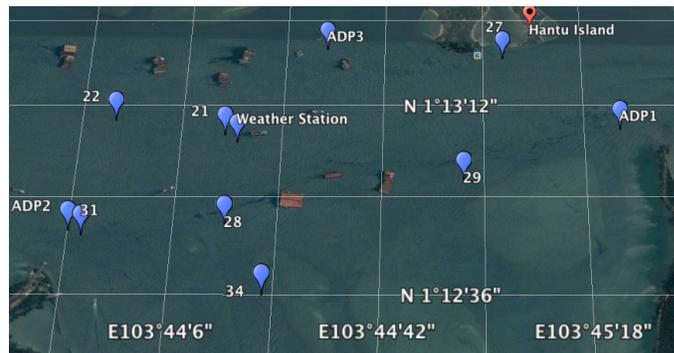


Fig. 1. Layout of the MISSION 2013 experiment

that can affect UAN performance. Some of these factors may not be fully understood or quantifiable. Thus, it is difficult to model and predict the UAN's performance as it formulates into a complicated non-linear forward model. In this scenario, machine-learning techniques are ideal tools to model the underlying dependencies of the UAN performance on various factors, in a data-driven fashion [8]. We can also incorporate our existing understanding of the factors that affect UANs into the training of this model.

We use a data-driven approach to model the performance variability of UANs based on fluctuations in environmental conditions and communication equipment-related factors. We employ logistic regression (LogR) to model the performance in terms of PSR [9]. Experimental data acquired by us at the MISSION 2013 experiment [7] is used for the modeling. Furthermore, we interpret the physical implications of the modeled variations. This paper is organized as follows: In Section II, we provide details on the experiment whose data is used in training and testing our LogR model. This is followed by a discussion on the features utilized in our model in Section III. Details of the modeling procedure are discussed in Section IV. Section V discusses the results of our modeling and our conclusions are presented in Section VI.

## II. THE MISSION 2013 EXPERIMENT

MISSION 2013 was a UAN experiment conducted by the Acoustic Research Laboratory of the National University of Singapore [6]. It was conducted at Selat Pauh in Singapore waters in 2013. The UAN consisted of seven nodes (numbered 21, 22, 27, 28, 29, 31 and 34) [6]. The UAN geometry is

TABLE I  
LEGEND OF PACKET OUTCOME STATES

Encoding	Outcome
R	Packet successfully received and decoded
X	Packet successfully received but decoding failed
U	Packet was not received
V	Packet was overheard (successfully received and decoded)
D	Packet was dropped

highlighted in Fig. 1, showing the locations of these nodes. Node 21 was mounted six meters below a barge. All remaining nodes consisted of an anchor, an underwater modem, batteries and an acoustic release buoy. When dropped, the anchor sinks to the seabed, while the modem-battery-buoy assembly floats several meters from the bottom. Three acoustic Doppler current profilers (ADPs) were deployed to measure the prevailing current profile and monitor tide changes. The ADPs are labeled as ADP1, ADP2 and ADP3, and their locations during the experiment are shown in Fig. 1. Additionally, a weather station was deployed in the vicinity of node 21.

In the experiment, packets were sent and received between multiple modems in a half-duplex manner and the outcome of each received packet was classified into five states denoted as R, X, U, V, D. These are summarized in Table I. We consider the performance of the UAN in terms of success in decoding a received packet. The Rs and Vs constitute successful packet decodings (denoted as 1) whereas Xs and Us constitute unsuccessful decodings (denoted as 0). The Ds are not considered in our dataset as they are not unbiased indicators of bad link quality. This is because packet drops could have occurred not only due to bad link quality, but also because the receiver node under consideration was scheduled to transmit at the same time. The timeseries of the packet outcomes is generated from the transmission logs of the experiment. We use the data collected over the 3-day period of the experiment in training and testing the LogR model.

### III. FEATURE SELECTION AND PREPROCESSING

#### A. Environmental factors measured

The environmental factors measured at the experiment were:

- **Current** - Measured by three ADPs in centimeters per second (cm/s) at a sampling interval of 5 minutes.
- **Wind** - Measured by the weather station in meters per second (m/s) once every 15 seconds.
- **Tide** - Measured in meters (m) via a pressure sensor on each ADP.
- **Range** - Range between the transmitter and receiver, measured in kilometers (km).
- **Depth** - Depth at which each node was located, measured in meters.

TABLE II  
FEATURES FOR THE MODEL

Feature class	Features
Along-current	AC1, AC2, AC3
Cross-current	CC1, CC2, CC3
Wind	AW, CW
Tide	Tide
Range	Range between transmitter and receiver
Depth	Receiver depth, transmitter depth
Modem-specific	Transmitter and receiver ID one-hot labels

#### B. Preprocessing and insight behind features

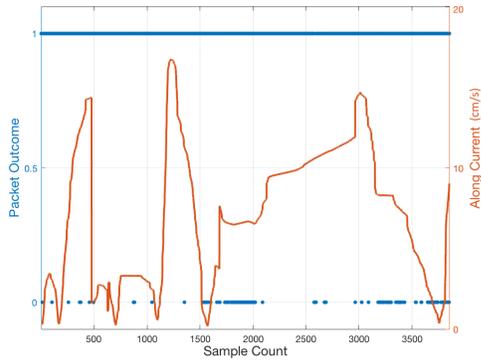
The current velocities measured at each ADP are resolved into two components: one in the direction of each link and the other perpendicular to it. We term these as along-current ( $ACx$ ) and cross-current ( $CCx$ ), respectively, where  $x$  refers to the ADP being used. The reason for doing so stems from the fact that each component contributes to a different effect on the transmitting/receiving modems. The AC component moves nodes along the direction of the link, thus introducing time-varying multipath. On the other hand, CC can lead to performance loss by introducing vortex induced vibrations of the modems.

Similarly, the wind speeds measured by the weather station are also resolved into along-wind (AW) and cross-wind (CW) components. Wind can affect the communication links by introducing ambient noise. The AW and CW components can also affect the PSR of a UAN via different mechanisms. More precisely, AW leads to surface-reflected rays undergoing magnification (due to focusing of rays by the water surface) or attenuation. On the other hand, we expect that CW can scatter surface-reflected rays away from the link direction, thus adversely affecting UAN performance. We also note that tide can essentially change the multipath structure of the link, which may lead to better or worse performance.

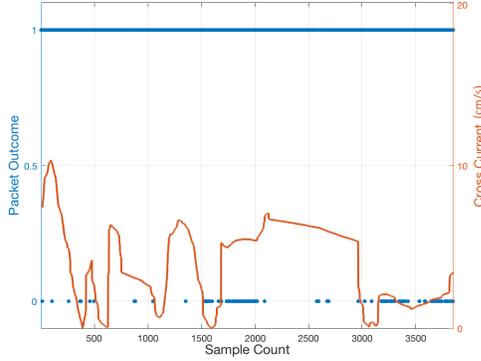
The depth of each node is included as an input feature as it allows us to capture some multipath-related performance variations. Similarly, the range between a transmitter and receiver captures some of the multipath-related variation and also characterizes geometric spreading loss incurred by sound traveling through the medium.

The identity of each transmitter and receiver is encoded as a feature in our learning model. This captures link performance variations that arise due to modem-specific factors such as hardware. Though the modems are labeled with numerical IDs, we cannot use these as features as we want node identities to be treated as categorical variables and not numerical quantities. Thus, we encode the IDs by employing one-hot labels [10]. As seven modems were deployed in the experiment, this results in 14 one-hot labels.

In summary, the environmental features employed in our model are  $ACx$ ,  $CCx$  ( $x \in \{1, 2, 3\}$ ), AW, CW, tide, modem

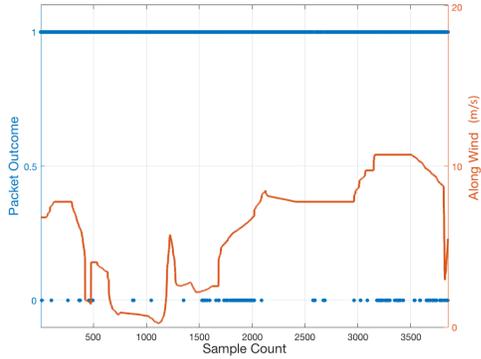


(a) AC1

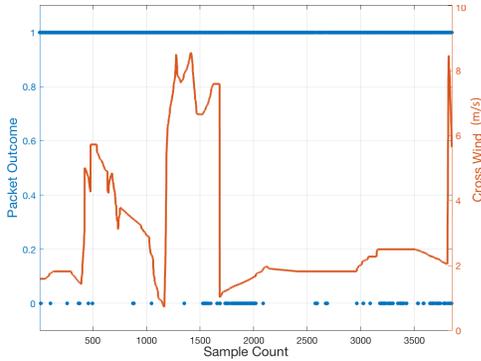


(b) CC1

Fig. 2. Packet outcomes for link 21 to 22 and current measurements.



(a) AW



(b) CW

Fig. 3. Packet outcomes for link 21 to 22 and wind measurements.

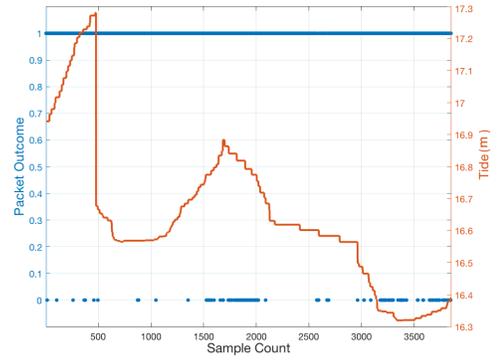


Fig. 4. Packet outcomes for link 21 to 22 and tide measurements.

depths, range between transmitter-receiver pairs and the ID of each transmitter and receiver modem as a one-hot label. The features used in the model are summarized in Table II.

### C. Observations

To visualize the experimental data, we plot the current variations recorded by ADP1 and packet outcomes for link 21 to 22 in Fig. 2. The current profile is resolved into AC and CC components in Figs. 2 (a) & 2 (b), respectively. Moreover, we also plot the wind measurements and tide profiles in Figs. 3 & 4, respectively. These again are accompanied with packet outcomes for link 21 to 22. The wind measurements are divided into AW and CW components with respect to link 21 to 22 in Figs. 3 (a) & 3 (b). Blue dots in Figs. 2 - 4 refer to packet outcome in terms of 1s and 0s (to be read against the primary y-axis). The orange line shows the feature (to be read against secondary y-axis). Link 21 to 22 was a 1.6 km long surface to seabed link. From visual inspection, it is not easy to draw straightforward correlations between PSR performance and the features. However, we note that packet successes and failures seem to occur in bursts and are temporally correlated.

## IV. PSR MODELING

The objective of our work is to model the PSR by learning the dependence of packet outcomes on the environmental factors. The packet outcome may be 1 or 0, indicating success or failure in decoding the transmitted packet, respectively. Treating this as a binary classification problem, we employ a LogR network to model the packet outcome. LogR belongs to a class of linear models that models the probability of the target belonging to a particular category given an observation [9]. In this model, the probabilities describing the possible outcomes of a single trial are modeled using the logistic function. Training of a binary class  $\ell_2$ -penalized LogR network using  $N$  training data points involves the following cost function minimization

$$\arg \min_{\mathbf{w}, c} \frac{1}{2} \mathbf{w}^T \mathbf{w} + \lambda \sum_{i=1}^N \log(\exp(-y_i(\mathbf{x}_i^T \mathbf{w} + c)) + 1)$$

where  $\mathbf{w}$  refers to the vector of (real) weights used in the network,  $\lambda$  is the inverse of the regularization parameter,  $y_i$  is the  $i^{th}$  target label,  $\mathbf{x}_i$  refers to the input feature vector

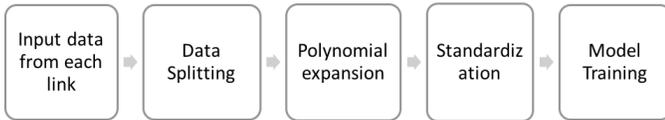


Fig. 5. Model training pipeline

corresponding to the  $i^{th}$  observation, and  $c$  is the bias [9]. Our dataset consists of 39,942 transmitted packets between all link pairs, of which 60% belong to class 1 and 40% to class 0.

### A. Performance metric

The performance metric used to gauge the model is the Matthews correlation coefficient (MCC) which is a balanced measure that is not affected by biased datasets [11]. The MCC is measure of the quality of binary classification which takes into account true and false positives and negatives, and can be used even if the classes are of very different sizes. It is, in essence, a correlation coefficient value between -1 and +1. A coefficient of +1 represents a perfect prediction, 0 an uninformed random prediction and -1 an inverse prediction [12]. The MCC is evaluated by

$$MCC = \frac{T_P T_N - F_P F_N}{\sqrt{(T_P + F_P)(T_P + F_N)(T_N + F_P)(T_N + F_N)'}}$$

where  $T_P$  is the count of true positives,  $T_N$  is the count of true negatives,  $F_P$  is the count of false positives and  $F_N$  is the count of false negatives.

### B. Modeling technique and hyperparameters

The LogR network was trained based on the performance metric detailed above, using the pipeline described in Fig. 5. Each of the steps in the pipeline is explained below:

- **Data splitting** - The data from each link is divided into 70% for training, 15% for validation and 15% for testing.
- **Standardization** - The data was standardized, which means that it was shifted such that each of its features has a distribution with a zero mean and a standard deviation of one. This ensures that the input data does not saturate the logistic nonlinearity of the neurons in the input layer, but rather keeps it in the active linear range.
- **Polynomial expansion** - We expand the initial feature space by additionally using polynomial powers of the features. This allows us to model more nonlinear and non-monotonic dependencies on the features [13]. We observe that though the features summarized in subsection III-B (which are essentially degree-1) were able to capture the variability in PSR performance based on the environmental features, better predictions can be made on enhancing the space to degree-2 and degree-3 features. All features are expanded to polynomial orders except for the one-hot labels, because the latter consists of ones and zeros and cannot provide any additional information to the LogR network on expansion.

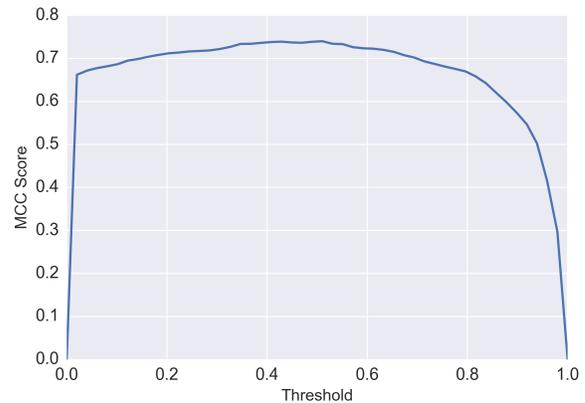


Fig. 6. Variation of MCC score of output predicted by LogR network using validation set, with threshold value.

Our model uses a LogR network with one hidden layer. We use  $\ell_2$ -regularization during training to prevent any single neuron weight becoming too large [14]. This ensures the network learns smooth variations with respect to features and does not over-fit the small random fluctuations in the data. We adopt the grid search cross-validation technique to fix the model’s hyperparameters such as the regularization strength.

The LogR network ends with a *softmax* output layer that yields a soft-value prediction which can be treated as a probability. This probability is then thresholded to give a hard decision on the predicted outcome (1 or 0). The threshold at which the decision boundary is set is a hyperparameter which is selected based on the validation set. The threshold is set to a value that yields the best MCC on validation data. Figure 6 shows one instance of the variation of the MCC of outputs predicted using the validation set when different thresholds are used. The prediction is done using a LogR network with a degree-3 feature space. In Fig. 6, we see that the maximum MCC is obtained at a threshold of 0.42 for the example considered. In general, we observed that the threshold values selected via the validation set were in the range of 0.4-0.6.

## V. RESULTS

In this section, we gauge the performance of the trained LogR network and study its learnt dependencies.

### A. Test performance

The performance of our LogR network models using different combinations of features is shown in Table III. The performance refers to the MCC obtained with the held-out test dataset. We note that the model using wind, current, tide and range features can yield an MCC of upto 0.50. With the addition of receiver depth, the MCC increases to 0.57. Adding one-hot labels, range and transmitter depth, increases the MCC by 12.5% over this network, thus indicating that these features provide information which aids in modeling. With polynomial expansion of the feature space, the MCC improves from 0.64 to 0.73, an enhancement of 14.2%. We did not expand the

TABLE III  
LOGR NETWORK PERFORMANCE COMPARISON WITH TEST DATA.

Features/modeling technique used	Test MCC
Wind, current, tide and range between transmitter and receiver	0.50
Wind, current, tide, range and receiver depth	0.57
Wind, current, tide, range, receiver depth and one-hot labels for transmitter and receiver	0.64
Wind, current, tide, receiver and transmitter depth, one-hot labels for transmitter and receiver (degree 1)	0.64
Wind, current, tide, receiver and transmitter depth, one-hot labels for transmitter and receiver (degree 2)	0.69
Wind, current, tide, receiver and transmitter depth, one-hot labels for transmitter and receiver (degree-3)	0.73

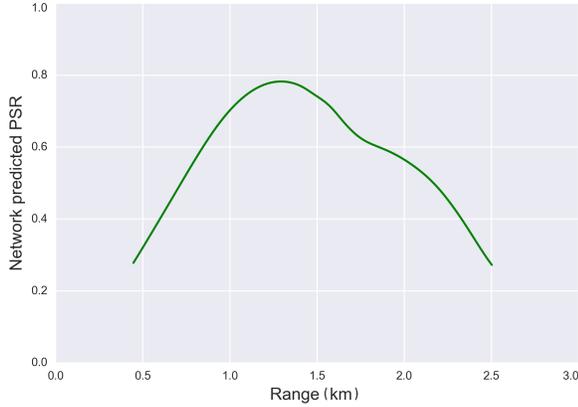


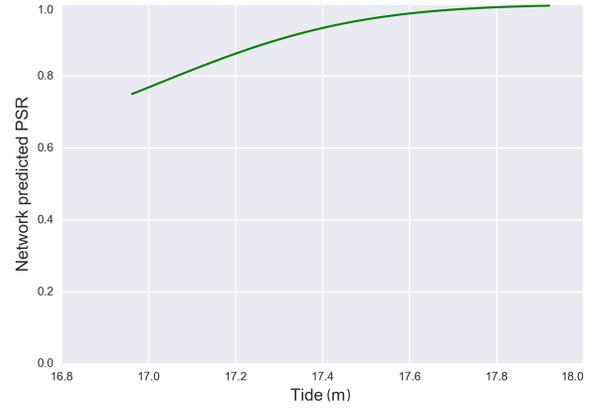
Fig. 7. PSR variation with range.

feature space to polynomial orders beyond degree-3 as the ratio of the number of data points to number of features was too low and could result in over-fitting.

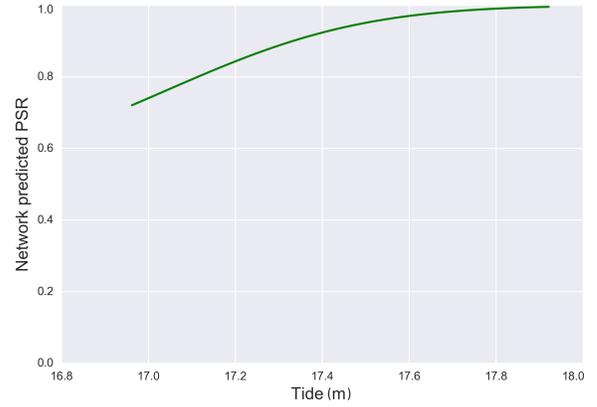
### B. Modeled variation and insights

We now examine the relationships that the trained model has learnt between the predicted outcome and individual features. We do this by making predictions using the trained LogR network on a prediction dataset where only one feature is varied within a relevant range and all other features are kept at their median values. The predictions are based on the degree-3 feature space.

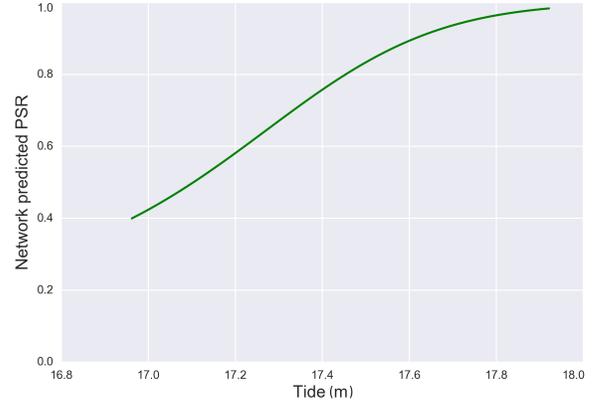
1) *Variation with range:* The dependence of performance on range learnt by the LogR model is shown in Fig. 7. The model shows a non-monotonic variation. At large ranges the model predicts that the UAN performs poorly, which is due to geometric spreading of sound waves from the transmitter that leads to low SNR at the receiver. The model also shows a degradation in performance at shorter ranges which could be due to the higher delay spread exhibited by raypath arrivals at short ranges which can degrade communication performance.



(a) Link 21 to 22.



(b) Link 21 to 29.



(c) Link 21 to 34.

Fig. 8. PSR variation with tide for 3 different links.

An interplay of the contrasting effects of delay spread at short range and low SNR at large range could lead to the existence of an intermediate ‘optimal range’ of communication performance. The LogR network seems to be capturing this in its non-monotonic variation. It is also possible that the network was over-fitting due to data points from some short-range links that performed poorly due to other factors, like link 27 to 29, link 22 to 31 and link 28 to 31.

2) *Variation with tide:* Figure 8 shows the variation in performance of three links with variation in tide as learnt

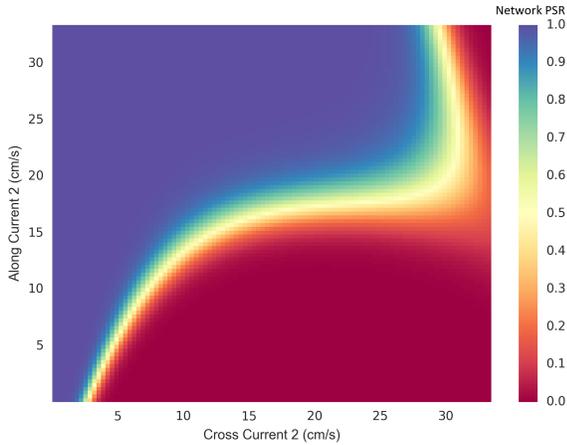


Fig. 9. Predicted PSR for link 21 to 22 against ADP2.

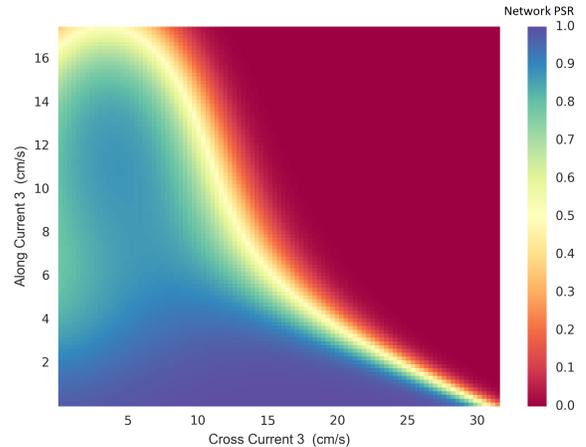


Fig. 10. Predicted PSR for link 21 to 22 against ADP3.

by the model. The model indicates that an increase in tide *increases* the PSR. This is especially noticeable for link 21 to 34. On examining the bathymetry of this link, we found that the direct path between the nodes was blocked by a reef at low tide. An increase in tide provided a way for the packets to overcome the reef and reach the receiver, leading to an increase in the PSR [7].

3) *PSR variation with current and wind*: Figures 9 & 10 show the variations learnt by the network for link 21 to 22 with respect to along/cross components of currents measured at ADPs 2 and 3, respectively, which were closest to the link. The heat map displays values between 0 to 1 indicating the predicted softmax probability output.

In general, the learnt variation indicates that the performance of the link decreases when the CC increases. As highlighted in subsection III-B, this decrease could be because an increase in CC leads to vortex-induced vibrations of the modem. These vibrations lead to Doppler distortion in the transmitted signals, thus resulting in a performance drop. The performance drop could also be attributed to CCs moving scatterers such as bubbles into the path of the link which would otherwise be stable. Figures 9 & 10 also indicate that the performance dependence on  $AC_x$  is not consistent for measurements from the two ADPs.

Figure 11 highlights the performance variation with wind learnt by the model. It indicates that the performance of the link improves for large AW. This could be because the AW led to a change in the surface reflected rays which improved link quality. The decreasing variation with CW is likely because cross-winds lead to wave patterns perpendicular to the link direction which scatter the acoustic energy in directions away from the link.

### C. Short-term PSR variability predictions

The PSR predicted by the LogR model for each transmitter-receiver pair is shown as a colour-coded table in Fig. 12, after setting all the environmental features and depths at their median values. We compare this against the median PSR computed from the groundtruth data, shown as a color-coded

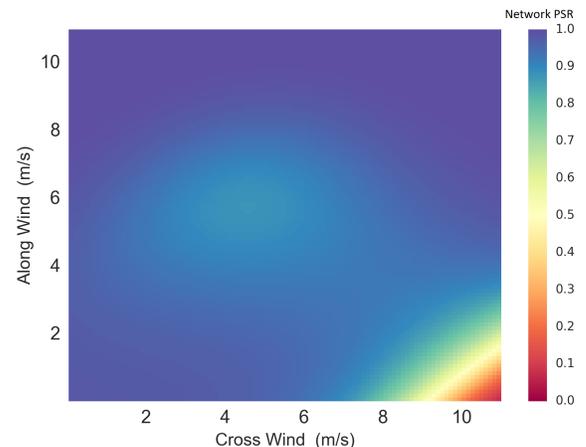


Fig. 11. Predicted PSR for link 21 to 22 against wind.

table in Fig. 13. Note that the LogR network captures the trends in the median link performance with fair accuracy. It indicates that certain links and modems, such as modem 27, performed very poorly on an average. We corroborated this with experimental logs and found that modem 27 faced hardware issues during the experiment. Thus, the one-hot labels are able to capture some of the modem and link-specific information in the modeling.

In Figs. 14 & 15, we plot the short-term average PSR for two different links from the ground-truth data. The average is taken over a window of 6 minutes so that it would be larger than the ADP's sampling rate of 5 minutes. We compare the ground-truth against the predicted softmax probability outputs of the LogR model, which are an indicator of the PSR. We observe that the predictions from the network are almost in line with the ground-truth data. The variations in PSR are captured well by the LogR model.

## VI. CONCLUSION

We have developed a LogR model that can predict, with some consistency, the UAN performance given some information on prevailing environmental conditions. The model



Fig. 12. Predicted PSR for each transmitter-receiver pair.

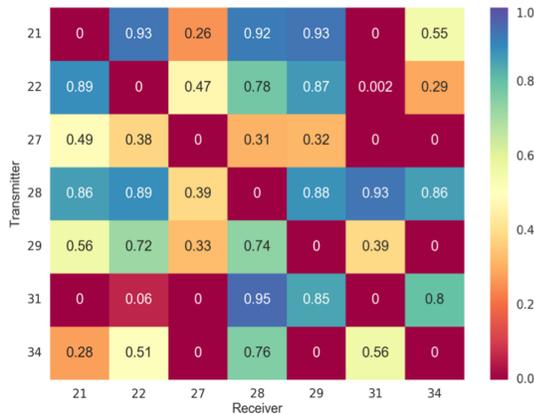


Fig. 13. Median ground-truth PSR for each transmitter-receiver pair.

was able to learn the dependence of performance on the environmental factors, from the data. It also captured the short-term variability of the UAN performance with fair accuracy. The variations learnt by the model with respect to the environmental factors seemed to corroborate with known physical phenomena. However, more exploration needs to be undertaken to understand the learnt variations better. Our predictive model enables us to simulate, predict and plan more effectively for future UAN deployments.

#### ACKNOWLEDGEMENTS

We thank the MISSION 2013 team including Dr. Venugopalan Pallayil, Iulian Topor, Mohan Panayamdram, Manu Ignatius, Shanmugam Mpl and Sivaraman Selvakumar for their efforts in collecting the experimental data used in this paper. We also thank Mr. Gabriel Chua for his valuable inputs on the data and its processing.

#### REFERENCES

[1] "Connecting the ocean to the internet," December 2016. [Online]. Available: <http://hosting.fluidbook.com/open-resource-magazine-03/#/36>

[2] R. Masiero, S. Azad, F. Favaro, M. Petrani, G. Toso, F. Guerra, P. Casari, and M. Zorzi, "DESERT Underwater: An NS-Miracle-based framework to design, simulate, emulate and realize test-beds for underwater network protocols," in *Oceans 2012 - Yeosu*. IEEE, may 2012, pp. 1–10.

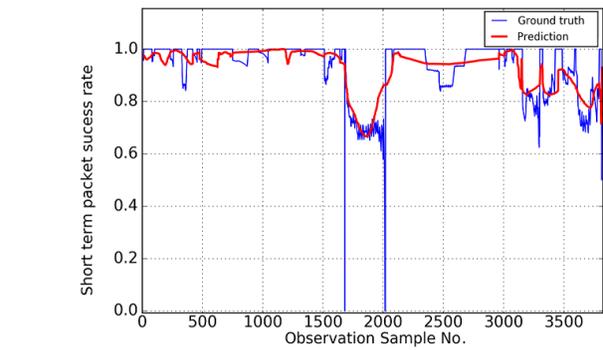


Fig. 14. Comparison of ground-truth and predicted short-term averaged PSRs for transmissions between link 21 to 22.

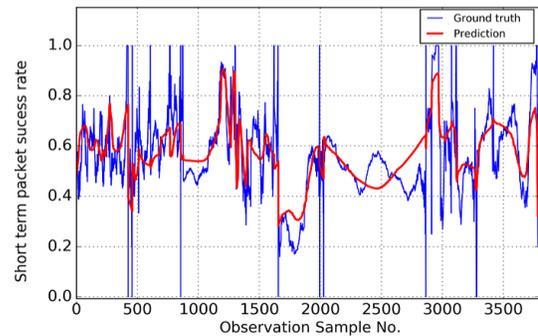


Fig. 15. Comparison of ground-truth and predicted short-term averaged PSRs for transmissions between link 21 to 34.

[3] M. Chitre, R. Bhatnagar, and W.-S. Soh, "UnetStack: An agent-based software stack and simulator for underwater networks," in *Oceans 2014 - St. John's*. IEEE, sep 2014, pp. 1–10.

[4] M. Chitre, S. Shahabudeen, and M. Stojanovic, "Underwater acoustic communications and networking: Recent advances and future challenges," *Mar. Technol. Soc. J.*, vol. 42, no. 1, pp. 103–116, 2008.

[5] M. Chitre and G. Chua, "Modeling realistic underwater acoustic networks using experimental data," *48th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA*, pp. 39–43, 2014.

[6] M. Chitre, I. Topor, R. Bhatnagar, and V. Pallayil, "Variability in link performance of an underwater acoustic network," *2013 MTS/IEEE OCEANS - Bergen*, pp. 1–7, 2013.

[7] M. Chitre and K. Pelekanakis, "Channel variability measurements in an underwater acoustic network," *Underwater Communications and Networking (UComms), Sestri Levante*, pp. 1–4, 2014.

[8] M. Kanevski, A. Pozdnoukhov, and V. Timonin, *Machine Learning for Spatial Environmental Data - Theory, Applications, and Software*. EFPL Press, 2009.

[9] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, ser. Springer Texts in Statistics, 2013.

[10] A. Daly, T. Dekker, and S. Hess, "Dummy coding vs effects coding for categorical variables: Clarifications and extensions," *Journal of Choice Modelling*, vol. 21, pp. 36–41, December 2016.

[11] D. Powers, "Evaluation: From precision, recall and f- measure to roc, informedness, markedness and correlation," *J. Mach. Learn. Technol*, vol. 2, pp. 37–63, 2011.

[12] P. Baldi, S. Brunak, Y. Chauvin, C. A. F. Andersen, and H. Nielsen, "Assessing the accuracy of prediction algorithms for classification: an overview," *Bioinformatics -Oxford*, vol. 16, pp. 412–424, 2000.

[13] K. Yao, W. Lu, S. Zhang, H. Xiao, and Y. Li, "Feature expansion and feature selection for general pattern recognition problems," *International Conference on Neural Networks and Signal Processing*, vol. 1, pp. 29–32, 2003.

[14] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.