

Deep Learning Prediction of Time-Varying Underwater Acoustic Channel Based on LSTM with Attention Mechanism

Zhengliang Zhu^{1,2}, Feng Tong^{1,2}, Yuehai Zhou^{1,2}, Ziqiao Zhang³ and Fumin Zhang⁴

Received: 14 September 2022 / Accepted: 22 November 2022
© Harbin Engineering University and Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

This paper investigates the channel prediction algorithm of the time-varying channels in underwater acoustic (UWA) communication systems using the long short-term memory (LSTM) model with the attention mechanism. AttLstmPreNet is a deep learning model that combines an attention mechanism with LSTM-type models to capture temporal information with different scales from historical UWA channels. The attention mechanism is used to capture sparsity in the time-delay scales and coherence in the gap-time scale under the LSTM framework. The soft attention mechanism is introduced before the LSTM to support the model to focus on the features of input sequences and help improve the learning capacity of the proposed model. The performance of the proposed model is validated using different simulation time-varying UWA channels. Compared with the adaptive channel predictors and the plain LSTM model, the proposed model is better in terms of channel prediction accuracy.

Keywords Long short-term memory (LSTM); Attention mechanism; Underwater acoustic communication; Underwater acoustic channel; Channel prediction

1 Introduction

The underwater acoustic (UWA) channel is recognized as one of the most challenging communication channels due to its adverse time-frequency selectivity, limited bandwidth, and random noise (Jiang et al., 2022; Jiang and Diamant, 2023; Zhu et al., 2023; Zhu et al., 2021; Song et al.,

2019). The time variation of the UWA channel is generally coupled with the complicated marine environment at different time scales (Oliveira et al., 2021; Zhou et al., 2015; Yang et al., 2017). For example, the scattering of surface waves may lead to small-scale phenomena. Meanwhile, the sound speed profile, surface height, and the motion of the transmitter/receiver will cause large-scale phenomena (Qarabaqi and Stojanovic, 2013). Substantial research has conducted in the past years, overcoming the impact of the underwater environment and improving the UWA communication quality. A consensus indicates that the throughput and robustness of the UWA communication link will be improved further if the UWA channel properties are maximized and an adaptive UWA communication method is utilized.

The quality of channel information, which is feedback from the receiver, determines the performance of the adaptive UWA communication systems. Supported by a proper channel, the adaptive UWA communication system can determine the physical layer parameters and the selected environment match modulation schemes, such as using a low-modulation scheme and spread spectrum technology under the multipath channel. The channel received by the transmitter is frequently outdated due to the slow acoustic propagation and the fast time-varying UWA channel. Therefore, a channel prediction model that does not rely on extra training sequences in the receiver is required.

In terrestrial wireless communication, channel prediction has been widely used in beam-forming, physical secu-

Article Highlights

- LSTM model with attention mechanism is proposed for deep learning prediction of time-varying channels.
- The proposed AttLstmPreNet model combines an attention mechanism with LSTM-type models to exploit and capture temporal information at different scales from historical UWA channels.
- The simulation results demonstrate that the proposed AttLstmPreNet model exhibits better performance than both adaptive channel predictors and the plain LSTM model.

✉ Feng Tong
ftong@xmu.edu.cn

¹ National and Local Joint Engineering Research Center for Navigation and Location Service Technology, Xiamen University, Xiamen 361005, China

² College of Ocean and Earth Sciences, Xiamen University, Xiamen 361005, China

³ School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta 30314, USA

⁴ Cheng Kar-Shun Robotics Institute, Hong Kong University of Science and Technology, Hong Kong 00852, China

rity, and fifth-generation usage scenarios. However, in the UWA communication society, the UWA channel prediction has only seen application in recent years. The channel prediction approach can be mainly divided into two categories: model-dependent and model-independent (Liu et al., 2021; Zhang et al., 2020a). The former approach assumes that the previous knowledge of the UWA channel is known, and the prediction performance can then be improved when the correctly defined model matches the realistic channel model. The UWA channel evolution in Fuxjaeger and Iltis (1994) was modeled as an autoregressive process, and the channel parameters were predicted by extended Kalman filter because of the assumption of the uncorrelation channel tap and independent transitions. Nadakuditi and Preisig (2004) assumed that the multipath is adopted to optimize the postfilter coefficients for tracking the channel. Thus, the model-dependent approach could attain no more than the comprehensive of the researchers. That is, if the natural UWA channel does not match the assumed model, then the prediction performance based on the model-dependent approaches cannot be effectively presented.

In contrast to the model-dependent approach, the model-independent method disregards previous knowledge of the channel, such as the adaptive and deep learning channel predictions. This type of approach can track the channel per tap based on the history of channels. The adaptive algorithms, namely the least means square (LMS) and RLS, are widely used to update the channel. Similar to many adaptive processing applications, the RLS has superior accuracy with high computational complexity compared with LMS. A channel prediction network that comprises a one-dimensional convolutional neural network and a long short-term memory (LSTM) model is designed in Liu et al. (2021). However, the existing UWA channel predictor occasionally considers the inherent physical sparsity and the evolutionary correlation of the UWA channel. A channel prediction model based on LSTM integrated with the attention mechanism is proposed in this paper, and the performance is analyzed under different UWA channels to exploit their characteristics.

The key idea of the attention mechanism comes from the human visual attention mechanism. When people perceive things, they often provide increased attention to specific parts based on their needs. Motivated by this natural phenomenon of humans, the attention mechanism is widely applied in several fields with remarkable results. A recurrent neural network (RNN) using an attention mechanism was proposed by Mnih et al. (2014) in computer vision for image classification. In Bahdanau et al. (2016), an attention mechanism with RNN was first used in machine translation work. The attention mechanism can generally be regarded as a resource allocation scheme, which is the primary means to solve the problem of information overload (Niu et al., 2021).

The significant contributions of this paper are presented as follows. A learning model named AttLstmPreNet is designed for UWA channel prediction, which is integrated

with an attention mechanism using an LSTM network. The proposed AttLstmPreNet can exploit the sparsity and time evolution of the UWA channel to conduct the channel prediction effectively. The prediction performance was evaluated using a set of simulation channels. Notably, the proposed AttLstmPreNet is a hybrid channel prediction approach between model-dependent and model-independent because this approach is driven by the history of channel data and the previous knowledge of channel data.

2 UWA channel characteristics

Owing to the physical characteristics of the underwater environment, namely the dynamic of the sea environment, the UWA signal will be reflected and scattered, resulting in the receipt of signal along with multiple propagation paths, and this phenomenon is known as the multipath effect. Theoretically, the number of propagation paths was infinite. However, most paths can be ignored with the transmission loss, and a few dominant paths with high energy should be considered; thus, the UWA channel will almost emerge in a sparse structure.

UWA channels can be regarded as a function of time and depth. A coherent function was proposed to describe the temporal coherence of UWA channels (Yang, 2012), which can be expressed as follow:

$$\Gamma(\tau) \equiv \left\langle \frac{[h^*(t)h(t+\tau)]}{\sqrt{[h^*(t)h(t)][h^*(t+\tau)h(t+\tau)]}} \right\rangle \quad (1)$$

where the $h(t)$ is the reference estimated channel at geotime t , and $h(t+\tau)$ is the channel arriving with a latency τ , $[pq]$ denotes the maximum valued of the correlation between the p and q signal. $(\cdot)^*$ is the conjugation operation, and the $\langle \cdot \rangle$ is the ensemble average function over geotime t (Huang et al., 2013). The temporal coherence of the UWA channel was computed in different time scales (i.e., interpacket, intrapacket) under different sea states using Equation (1). The results show that the channel coherence is high during the intrapacket and the interpacket in the calm sea. The correlation of channels is poor under the rough sea. Figure 1 shows the intrapacket temporal coherence under different sea conditions.

The following observations can be drawn from the discussion above,

- Most UWA channels have a sparse multipath arrival structure, which is caused by the propagation environment.
- Under different environmental conditions, the evolution of the UWA channel has varying temporal coherence. In particular, the UWA channel exhibits remarkable temporal coherence in stationary environments (such as in a calm sea or protected ocean). By contrast, the temporal correlation of the channel is poor in nonstationary environments (such as in a rough sea or open ocean).

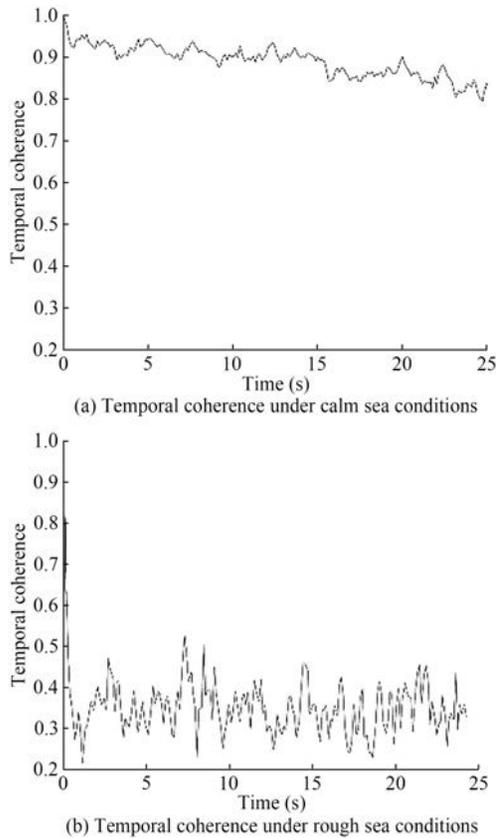


Figure 1 Temporal coherence of different UWA channels

The paper aims to provide an LSTM-based channel prediction model with an attention mechanism to improve adaptive UWA communication by exploiting the channel properties (sparsity and temporal coherence).

3 UWA channle prediction model based on LSTM with attention mechanism

3.1 UWA single-carrier communication system

Consider a single-input-single-output single-carrier acoustic communication system in a dynamic ocean environment. The binary information bit is divided into P groups, in which P denotes the number of bits per symbol, and each group is mapped to one of the 2^P -ary symbols of the map alphabet $A = \{ \alpha_p \}_{p=1}^P$, where α_p can be a complex or real value. The transmitted symbol vector is constructed by concatenating a training symbol sequence of length N_t with the information symbol sequence. That is, the transmitted symbol vector can be represented as $\mathbf{x} = \{ x_n \}_{n=1}^{N_s}$, where the N_s denotes the length of the transmitted symbols.

The transmitted symbols are filtered via a raised-cosine pulse-shaping filter with an impulse response $g(t)$ to form the baseband signal $u(t)$, which can be written as follows:

$$s(t) = \sum_{n=1}^{N_s} x_n g(t - nT_s) \tag{2}$$

where $g(t)$ is a raised-cosine pulse-shaping filter with a roll-off factor γ , and T_s is the symbol interval. The pass-band transmitted signal $u(t)$ is produced by the $s(t)$ modulated with a carrier frequency f_c , namely:

$$u(t) = s(t)e^{j \cdot 2\pi f_c \cdot t} \tag{3}$$

Assume that the channel delay length is L . The received baseband signal is distorted by multipath spread and noise can then be expressed as shown below:

$$y(t) = \sum_{l=0}^{L-1} h_l(t)s(t - \tau_l(t)) + \text{noise}(t) \tag{4}$$

where the $h_l(t)$, τ_l is the time-varying fading factor and delay time corresponding to the l th propagation path, and $\text{noise}(t)$ is the equivalent baseband noise at the receiver, which is independent of the $s(t)$.

The baseband channel information can be estimated by using the knowledge of the transmitted and the received baseband signals, that is, $s(t)$ and $y(t)$. Many scholars have developed UWA channel estimation algorithm in the past years, such as maximum likelihood, least squares, and compressed sensing. Herein, the estimated channel in baseband can be expressed as shown below:

$$h(\tau, t) = \sum_{l=0}^{L-1} h_l(t)\delta(\tau - \tau_l(t)) \tag{5}$$

where $h_l(t)$ and $\tau_l(t)$ denote the fading amplitude and the time delay of the discrete channel path, respectively.

3.2 Brief review of the UWA channel prediction method

The traditional model-independent channel predictor, such as the LMS and RLS, are highly suited for channel prediction in real-world communication systems (Lin et al., 2015; Ma et al., 2019; Radosevic et al., 2011). The general channel prediction structure can be modeled as follows:

$$\tilde{\mathbf{h}}[m + 1, l] = \mathbf{w}^T[m, l] \hat{\mathbf{h}}[m, l] \tag{6}$$

where l denotes the channel tap index corresponding to the time delay, m denotes the channel geotime, $\mathbf{w}[m, l] = [w_0[m, l], w_1[m, l], \dots, w_{n-1}[m, l]]^T$ represents the prediction coefficients, $\hat{\mathbf{h}}[m, l] = [h[m, l], h[m - 1, l], h[m - 2, l], \dots, h[m - n + 1, l]]^T$ represents the estimated value at l th channel tap in the historical moment, and n is the order of the predictor.

From the above description of Equation (6), the channel prediction approach works by the per channel tap, implying that the computational complexity rises with the channel length. The channel prediction algorithm can impose several significant channel taps to eliminate the noise perturbation and improve the channel prediction accuracy due to the sparsity of the UWA channel; however, determining the critical channel tap should be carefully studied (Lin et al., 2015; Radosevic et al., 2011).

The goal for adaptive predictor based on LMS or RLS is solving the prediction coefficients $w[m, l]$ of Equation (6). The two approaches are briefly described below. Assuming at the geotime $(m - 1)$ th, the channel prediction error at geotime m th can be computed by the Equation (6) and be written as follows:

$$e[m, l] = \hat{h}[m, l] - \tilde{h}[m, l] = \hat{h}[m, l] - w^T[m - 1, l] \hat{h}[m - 1, l] \tag{7}$$

In the core idea of LMS, the update of $w[m, l]$ lies in the criterion of the least mean square error, namely the minimization of the cost function $J_w = E\{|e[m, l]|^2\} \approx |e[m, l]|^2$. Thusm the prediction coefficients can be updated below:

$$\begin{aligned} w[m, l] &= w[m - 1, l] - \mu \nabla J_w \\ &= w[m - 1, l] - \mu \frac{\partial |e[m, l]|^2}{\partial w[m - 1, l]} \\ &= w[m - 1, l] + 2\mu e[m, l] \hat{h}[m - 1, l] \end{aligned} \tag{8}$$

where μ is the step size and $0 < \mu < 1$, which controls the tracking performance of w . According to the Equation (8), the previous coefficients and the past channel information with step size are managed to form the new prediction coefficients. The step size μ must compromise the prediction performance and convergence speed of the LMS algorithm.

Unlike the LMS prediction algorithm, the goal of RLS is to minimize the sum of the error squares between the actual and predicted value. The cost function can be expressed as

$$J_w = \sum_{j=1}^n \lambda^{n-j} |e[m, l]|^2, \text{ where } \lambda \text{ is the forget factor } 0 < \lambda < 1, \text{ which is used to balanced the old and present data.}$$

The selection of λ should match the emporal coherence of UWA channels. The prediction step of RLS is as follows:

$$w[m, l] = w[m - 1, l] + k[m - 1, l] e[m, l] \tag{9}$$

where $e[m, l] = \hat{h}[m, l] - w^T[m - 1, l] \hat{h}[m - 1, l]$ is the prediction error, and $k[m - 1, l]$ is the RLS gain vector:

$$k[m - 1, l] = \frac{P[m - 2, l] \hat{h}[m - 1, l]}{\lambda + \hat{h}^T[m - 1, l] P[m - 2, l] \hat{h}[m - 1, l]} \tag{10}$$

where the matrix P can be computed by the recursion:

$$P[m - 1, l] = \frac{1}{\lambda} (I - k[m - 1, l] \hat{h}^T[m - 1, l]) P[m - 2, l] \tag{11}$$

where the $P[0, l] = \delta^{-1} I$ and δ is a tiny number.

3.3 Proposed UWA channel prediction model

3.3.1 Channel prediction based on LSTM

The LSTM is a variant of an RNN and can solve the problems of exploding and vanishing gradients, which is suitable for time series prediction (Zhu et al., 2021; Zhang, 2020b; Zhang et al., 2019; Greff et al., 2017).

In the application of channel prediction, the channel data (measured from the real world or generated from the UWA channel simulator (Qarabaqi and Stojanovic, 2013)) are applied to train the LSTM model. The update of the well-trained LSTM network for channel prediction can be expressed as follows:

$$h_t = g_{LSTM}(H_t, W_{LSTM}) \tag{12}$$

where H_t is the history channels, h_t is the predicted output channel, and W_{LSTM} is the parameters of LSTM that can be trained. Contatenting LSTM, a regression layer is used to make a regresion of the prediction channel as follows:

$$\hat{h}_t = W_{reg} h_t \tag{13}$$

where W_{reg} is the paramters of the regression layer that can be trained.

3.3.2 Attention mechanism

Under the same principle of the attention mechanism, many scholars have developed some modifications and improvements in the attention mechanism for adapting to a specific task. The variants of attention mechanism can be categorized into two types: soft attention, and hard attention, under the softness of attention criterion (Niu et al., 2021). Compared with soft attention, hard attention requires minimal computation because computing all attention weights for all elements is unnecessary. The operation of hard attention can be regarded as making a hard decision for each input and is thus non-differentiable and difficult to optimize (Xu et al., 2015).

On the contrary, soft attention aims to compute the probability for each input sequence element when calculating the attention distribution. Comprehensively, he weights on each element are usually calculated by a SoftMax function, and the output is an operation of weights with an input sequence. The entire computing process can be described as a differentiable function (Xu et al., 2015). Thus, the attention mechanism can be trained jointly with the rest of the networking using backpropagation method.

Assuming the input sequences, $X = [x_1, x_2, \dots, x_N]$, $x_n = [x_{n,1}, x_{n,2}, \dots, x_{n,L}]^T$, and $n \in [1, N]$. For each element of X , the attention mechanism generates a positive weight $\alpha_{i,j}$,

$i \in [1, N]$, and $j \in [1, L]$. The weight $\alpha_{i,j}$ is computed by an attention model $f_{att}(\cdot)$ and can be expressed as follows:

$$\begin{aligned} [e_1, e_2, \dots, e_N] &= f_{att}(X) \\ [\alpha_1, \alpha_2, \dots, \alpha_N] &= \text{Softmax}[e_1, e_2, \dots, e_N] \end{aligned} \tag{14}$$

where $\alpha_{i,j}$ is the element of α_i . While the refined sequences by the attention mechanism can be written as:

$$X_{att} = X \odot [\alpha_1, \alpha_2, \dots, \alpha_N] \tag{15}$$

where \odot represents element-wise multiplication.

Soft attention is adopted in this paper. A simple but effective soft attention mechanism is practiced in the designed mode via a dense connection layer with a Softmax function.

3.3.3 Designed attention LSTM for UWA channel prediction

Following the brief introduction of LSTM and the attention mechanism, a learning model called Attention LSTM is developed for the UWA channel (AttLstmPreNet). Figure 2 shows the full architecture of the proposed network, which comprises an attention mechanism and LSTM model. The attention mechanism is adapted to focus on the characteristic of the history channels, which are constructed by a dense layer connected with the input of the network.

Consider a set of channels $H \in R^{M \times N}$, where M and N denote the geotime and channel tap (i.e., the time delay), respectively. After processing by the attention mechanism, the focused dataset H^* can be computed as follows:

$$H^* = H \odot [\alpha_1, \alpha_2, \dots, \alpha_N] \tag{16}$$

where $[\alpha_1, \alpha_2, \dots, \alpha_N]$ can be learned from the attention mechanism following Equation (14). An LSTM model is then adopted to predict channels under the learned long-term dependencies. The output channel of the LSTM model can be obtained on the basis of Equation (12) and (13). The attention mechanism can provide corresponding attention to the input channel set with the natural properties of UWA channels. Therefore, the proposed AttLstmPreNet incorporates the advantages of LSTM in the temporal information processing and the benefits of the attention mechanism in the feature selection.

The performance of the proposed AttLstmPreNet is evaluated in Section 4. In particular, the application of the atten-

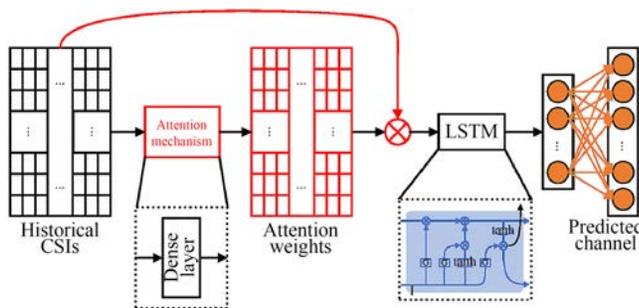


Figure 2 Structure of AttLstmPreNet

tion mechanism will be imposed on the following three aspects: channel tap (AttLstmPreNet-Type1), prediction time step (AttLstmPreNet-Type2), and jointly with channel tap and prediction time step (AttLstmPreNet-Type3).

4 Numerical simulation

The performance of the aforementioned channel prediction models, namely, the type of AttLstmPreNet-based, the plain LSTM-based, the LMS-based, and the RLS-based, is compared using the simulation channel measurements. The simulation channel measurement data are generated from the time-varying UWA channel simulator (Qarabaqi and Stojanovic, 2013), which also considers physical aspects of acoustics propagation as the inevitable random channel variations. Zhang et al. (2020a) used the simulator to approximate the real-world channel measurements with different degrees of mobility, such as the Surface Processes and Acoustic Communications Experiment in 2009 (SPACE'08) and the Mobile Acoustic Communications Experiment in 2010 (MACE'10). Two UWA channel measurements were simulated (Cases A and B) in the current study. The system parameters are provided in Table 1 as follows.

Table 1 Simulation parameters used in channel simulation (Qarabaqi and Stojanovic, 2013)

Simulation parameters	Case A	Case B
Carrier frequency (kHz)	10	15.5
Bandwidth (kHz)	10	5
Depth of water (m)	100	10
Transmitter height (m)	20	4
Receiver height (m)	50	4
Distance between Tx and Rx (km)	1	1
Sound source angle (°)	[-60, 60]	[-89, 89]
Sound speed (m/s)	1 530.00	1 511.50

4.1 Simulaiton channel analysis

Following Table 1, Figure 3 shows the ensembles of simulated time-varying UWA channels over the measurement duration of 180 s. The vertical and horizontal axes denote the geotime and time delay, respectively. Figure 3 shows that the channel is time-varying, especially in the channel amplitude and delay. Meanwhile, Figure 3 also indicates the sparse structure of the UWA channel. It should be noted that the horizontal axes in Figure 3 represent the time-delay axis with the unit ms, and the vertical axis represents the geo-time axis and its unit is s. Following Equation (1), the temporal coherence of the mentioned simulation UWA channels was computed, as shown in Figure 4. The temporal coherence in Case B is higher than that in Case A, which is consistent with the simulation settings. The similarity definition (Yang, 2012; Zhou et al., 2017;

Tu et al., 2021) revealed that the channel whose temporal coherence is larger than 0.85 is defined as a slowly varying channel; otherwise, the channel is fast varying.

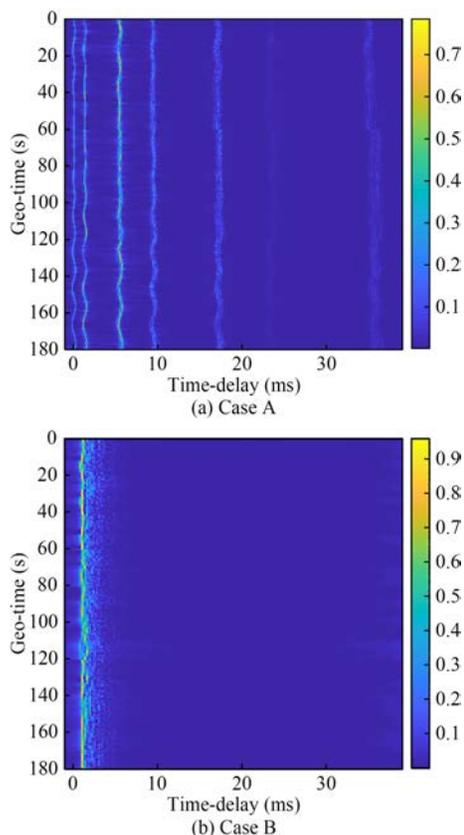


Figure 3 Simulation time-varying channel impulse responses

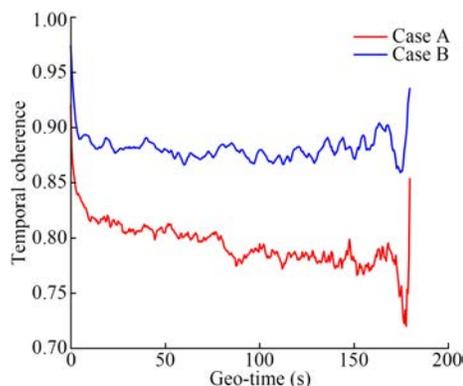


Figure 4 Temporal coherence of Case A and Case B

4.2 Prediction performance comparison

The aforementioned channel predictors are deployed for training and analysis. All channel predictors were trained on the simulation channel measurements of Cases A and B, comprising 361 samples (80% of the channel set were used for training, and the rest were used for testing). The AttLstmPreNet model comprises attention, LSTM, and

dense models, while the plain LSTM is constructed by two LSTM layers and a dense layer. The loss function of the LSTM is the mean absolute error (MAE). All LSTM models were implemented with Keras running on top of TensorFlow, which is performed on a computer with 2.5 GHz Intel (R) Core (TM) i7-11700K CPU (64 G RAM) and GeForce GTX 1050 Ti GPU. The prediction order for LMS and RLS channel predictors is set 8. The step size μ of LMS is to set 0.7 and 0.3 for Cases A and B channel sets, respectively. The forget factor λ of RLS is set to 0.85 and 0.75 for Cases A and B channel sets, respectively (the parameter set referenced from Liu et al. (2021)). Notably, the parameters of LMS and RLS are empirical values depending on the UWA channel properties. The training loss of the mentioned LSTM-based models in different channel sets is given in Figure 5. The loss value for each model decreased as the training epoch progressed. Figures 5(a) and (b) show that the AttLstmPreNet-Type2 achieves low loss value curves with a faster rate than other LSTM-based models in Cases A and B channel sets, respectively, thereby imposing that the attention mechanism on the time step is

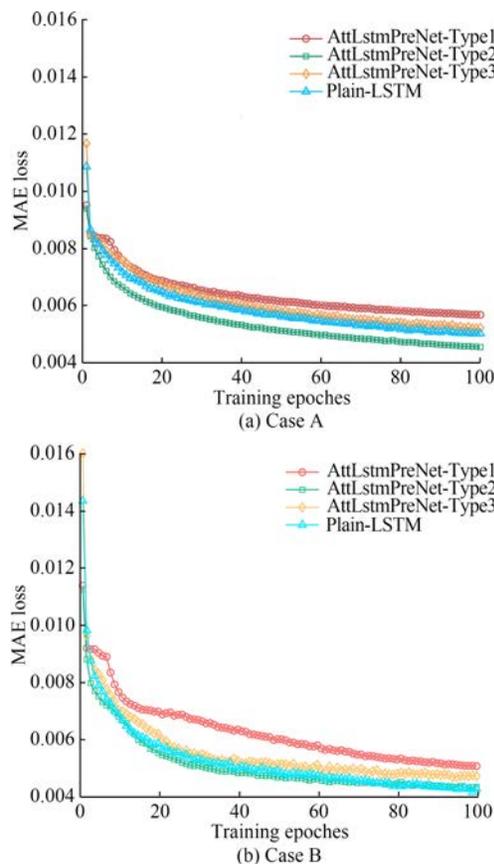


Figure 5 MAE loss curve of the mentioned prediction models during the training phase

beneficial to reducing the training loss. In particular, in the fast time-varying UWA channel set (i.e., Case A), AttLstmPreNet-Type2 has a faster convergence speed compared with the plain LSTM.

Figures 6 and 7 show the prediction channel by applying the well-trained methods on Cases A and B, respectively, namely AttLstmPreNet-Type1, AttLstmPreNet-Type2, AttLstmPreNet-Type3, plain LSTM, LMS, and RLS. Similar to the training phase, the MAE loss function is used to measure the error of prediction results, as shown in Table 2. From the MAE loss, the mentioned predictors work well in the slow time-varying channels. The RLS predictor achieved lower MAE values among the fast and slow time-varying channels than the LMS

predictor. Notably, the parameters of RLS should be finetuned compared with the LSTM-based predictor. For LSTM-based predictors, the AttLstmPreNet-Type2 and AttLstmPreNet-Type3 work slightly better than plain LSTM in different time-varying channels. Considering model simplification, the AttLstmPreNet-Type2 should be encouraged. In a large number of multipath conditions (i.e., Case A), the AttLstmPreNet-Type1 showed similar performance to plain LSTM, while the performance of AttLstmPreNet-Type1 is worse than the plain

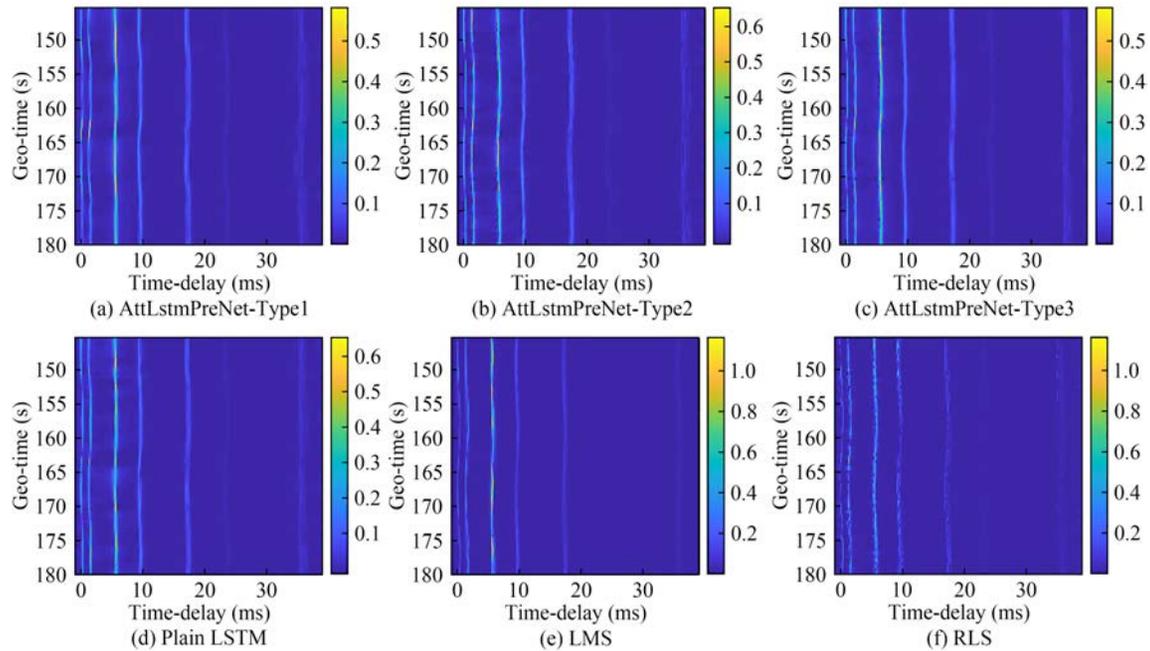


Figure 6 Prediction results for Case A using the mentioned predictors

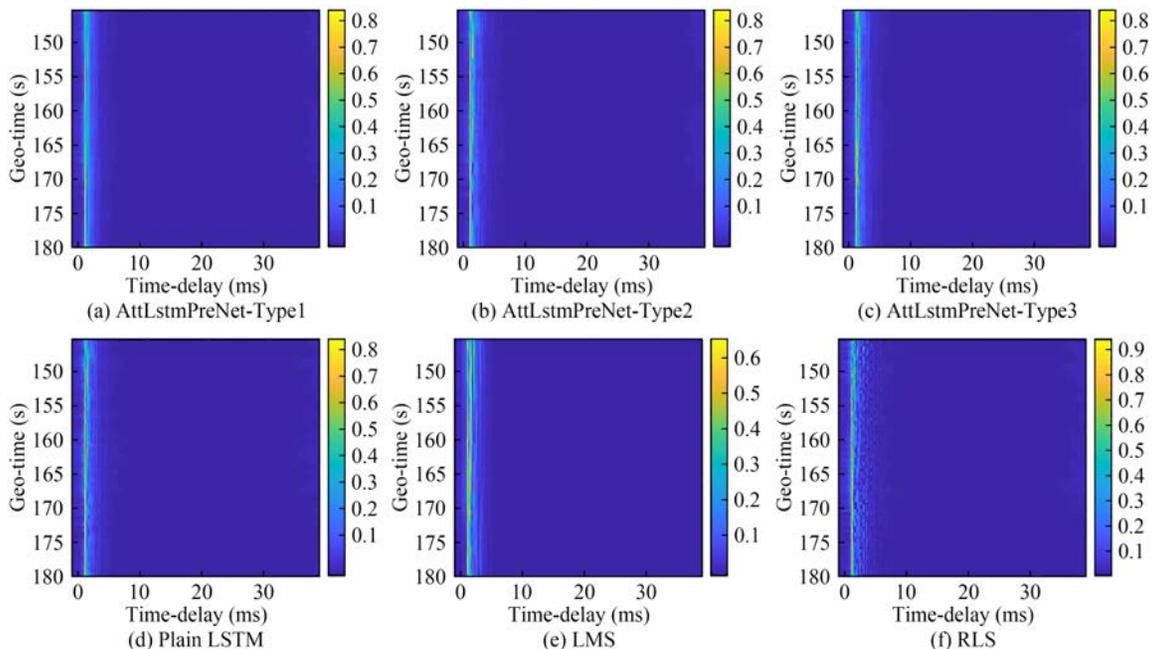


Figure 7 Prediction results for Case B using the mentioned predictors

LSTM in Case B. The result indicates that the AttLstmPreNet-Type1 is suitable for a channel with more cluster multipath than plain LSTM with a simple structure.

Table 2 MAE comparison of mentioned prediction models

MAE	AttLstmPreNet			Plain LSTM	LMS	RLS
	Type1	Type2	Type3			
Case A	0.009 2	0.008 7	0.008 2	0.009 1	0.011 7	0.008 8
Case B	0.006 2	0.005 1	0.005 2	0.005 5	0.010 0	0.006 2

5 Conclusions

A deep learning approach is applied in this paper to time-varying UWA channel prediction. A novel deep learning model named AttLstmPreNet is designed on the basis of LSTM with an attention mechanism. In the proposed model, AttLstmPreNet exploits the attention mechanism and LSTM network to capture temporal information from the historical UWA channels. A soft attention mechanism is introduced before the LSTM to support the model to focus on the features of input sequences and help improve the learning capacity of the proposed model. Experimental results on the simulated dataset demonstrate the effectiveness of the proposed model. The results show that the proposed model is better than the adaptive channel predictors (e.g. LMS or RLS predictor).

The authors believe that this paper is the first to design an LSTM model with an attention mechanism for UWA channel prediction with a simulation channel dataset. Applying the attention mechanism to real-world UWA communication systems will be an essential issue. Thus, highly advanced attention mechanisms will be introduced in the future. Meanwhile, the parameter optimization of the adaptive channel predictors must be studied jointly with the UWA properties.

Acknowledgement The authors would like to thank Dr. Xingbin Tu of the Ocean College, Zhejiang University, for his assistance in computing the temporal coherence of the UWA channel.

Funding Supported by the National Key Research and Development Program of China (No. 2018YFE0110000), the National Natural Science Foundation of China (No. 11274259, 11574258), and the Science and Technology Commission Foundation of Shanghai (21DZ1205500).

Competing interest Feng Tong is an editorial board member for the Journal of Marine Science and Application and was not involved in the editorial review, or the decision to publish this article. All authors declare that there are no other competing interests.

References

Bahdanau D, Cho K, Bengio Y (2016) Neural machine translation by

jointly learning to align and translate. arXiv:1409.0473 [cs, stat]. Available from <http://arxiv.org/abs/1409.0473> [Accessed March 29, 2022]

- Fuxjaeger AW, Iltis RA (1994) Adaptive parameter estimation using parallel Kalman filtering for spread spectrum code and doppler tracking. *IEEE Transactions on Communications* 42(6): 2227-2230. DOI: 10.1109/26.293672
- Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J (2017) LSTM: a search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems* 28(10): 2222-2232. DOI: 10.1109/TNNLS.2016.2582924
- Huang SH, Yang TC, Huang CF (2013) Multipath correlations in underwater acoustic communication channels. *The Journal of the Acoustical Society of America* 133(4): 21802190. DOI: 10.1121/1.4792151
- Jiang W, Diamant R (2023) Long-range underwater acoustic channel estimation. *IEEE Transactions on Wireless Communications*, Early Access. DOI: 10.1109/TWC.2023.3241230
- Jiang W, Tong F, Zhu Z (2022) Exploiting rapidly time-varying sparsity for underwater acoustic communication. *IEEE Transactions on Vehicular Technology* 71(9): 9721-9734. DOI: 10.1109/TVT.2022.3181801
- Lin N, Sun H, Cheng E, Qi J, Kuai X, Yan J (2015) Prediction based sparse channel estimation for underwater acoustic OFDM. *Applied Acoustics* 96: 94-100. DOI: 10.1016/j.apacoust.2015.03.018
- Liu L, Cai L, Ma L, Qiao G (2021) Channel state information prediction for adaptive underwater acoustic downlink OFDMA system: deep neural networks based approach. *IEEE Transactions on Vehicular Technology* 70(9): 9063-9076. DOI: 10.1109/TVT.2021.309979
- Ma L, Xiao F, Li M (2019) Research on time-varying sparse channel prediction algorithm in underwater acoustic channels. 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE), 2014-2018
- Mnih V, Heess N, Graves A, Kavukcuoglu K (2014) Recurrent models of visual attention. *Proceedings of the 27th International Conference on Neural Information Processing Systems*, 2204-2212
- Nadakuditi R, Preisig JC (2004) A channel subspace post-filtering approach to adaptive least-squares estimation. *IEEE Transactions on Signal Processing* 52(7): 1901-1914. DOI: 10.1109/TSP.2004.828926
- Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. *Neurocomputing* 452: 48-62. DOI: 10.1016/j.neucom.2021.03.091
- Oliveira TCA, Lin YT, Porter MB (2021) Underwater sound propagation modeling in a complex shallow water environment. *Frontiers in Marine Science* 8: 751327. DOI: 10.3389/fmars.2021.751327
- Qarabaqi P, Stojanovic M (2013) Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels. *IEEE Journal of Oceanic Engineering* 38(4): 701-717. DOI: 10.1109/JOE.2013.2278787
- Radosevic A, Duman TM, Proakis JG, Stojanovic M (2011) Channel prediction for adaptive modulation in underwater acoustic communications. *OCEANS 2011 IEEE-Spain, 1-5*. DOI: 10.1109/Oceans-Spain.2011.6003438
- Song A, Stojanovic M, Chitre M (2019) Editorial underwater acoustic communications: where we stand and what is next? *IEEE Journal of Oceanic Engineering* 44(1): 1-6. DOI: 10.1109/JOE.2018.2883872
- Tu XB, Xu XM, Song AJ (2021) Frequency-domain decision feedback equalization for single-carrier transmissions in fast time-varying underwater acoustic channels. *IEEE Journal of Oceanic Engineering*

- 46(2): 704-716. DOI: 10.1109/JOE.2020.3000319
- Xu K, Lei J, Kiros R, Cho K, Courville A, Salakhutdinov R, Zemel RS, Bengio Y (2015) Show, attend and tell: neural image caption generation with visual attention. Proceedings of the 32nd International Conference on Machine Learning, 2048-2057
- Yang Mei, Li XK, Yang Y, Meng XX (2017) Characteristic analysis of underwater acoustic scattering echoes in the wavelet transform domain. *Journal of Marine Science and Application* 16(1): 93-101. DOI: 10.1007/s11804-017-1398-6
- Yang TC (2012) Properties of underwater acoustic communication channels in shallow water. *The Journal of the Acoustical Society of America* 131(1): 129-145. DOI: 10.1121/1.3664053
- Zhang Y, Venkatesan R, Dobre OA, Li C (2020a) Efficient estimation and prediction for sparse time-varying underwater acoustic channels. *IEEE Journal of Oceanic Engineering* 45(3): 1112-1125. DOI: 10.1109/JOE.2019.2911446
- Zhang Z, Hou M, Zhang F, Edwards CR (2019) An LSTM based Kalman filter for spatio-temporal ocean currents assimilation. WUWNET'19: International Conference on Underwater Networks & Systems, Atlanta, 1-7
- Zhang ZQ, Al-Abri S, Wu WC, Zhang FM (2020b) Level curve tracking without localization enabled by recurrent neural networks. 5th International Conference on Automation, Control and Robotics Engineering (CACRE), 759-763
- Zhou Y, Song A, Tong F (2017) Underwater acoustic channel characteristics and communication performance at 85 kHz. *The Journal of the Acoustical Society of America* 142(4): EL350-EL355. DOI: 10.1121/1.5006141
- Zhou YH, Cao XL, Tong F (2015) Acoustic MIMO communications in a very shallow water channel. *Journal of Marine Science and Application* 14(4): 434-439. DOI: 10.1007/s11804-015-1323-9
- Zhu Z, Tong F, Jiang W, Zhang F, Zhang Z (2021) Evaluating underwater acoustics sensor network based on sparse LMS algorithm driven physical layer. WUWNet'21: The 15th International Conference on Underwater Networks & Systems, Shenzhen, 1-8.
- Zhu Z, Tong F, Zhou Y, Wu F (2023) Dual parameters optimization lp-LMS for estimating underwater acoustic channel with uncertain sparsity. *Applied Acoustics* 202: 109150. DOI: 10.1016/j.apacoust.2022.109150