TWMS J. App. Eng. Math. V.13, N.1, 2023, pp. 373-385

PREDICTING THE OCEAN CURRENTS USING DEEP LEARNING

C. BAYINDIR[1], §

Abstract. In this paper, we analyze the predictability of the ocean currents using deep learning. More specifically, we apply the Long Short Term Memory (LSTM) deep learning network to a data set collected by the National Oceanic and Atmospheric Administration (NOAA) in Massachusetts Bay between November 2002-February 2003. We show that the current speed in two horizontal directions, namely u and v, can be predicted using the LSTM. We discuss the effect of training data set on the prediction error and on the spectral properties of predictions. Depending on the temporal or the spatial resolution of the data, the prediction times and distances can vary, and in some cases, they can be very beneficial for the prediction of the ocean current parameters. Our results can find many important applications including but are not limited to predicting the statistics and characteristics of tidal energy variation, controlling the current induced vibrations of marine structures and estimation of the wave blocking point by the chaotic oceanic current and circulation.

Keywords: Oceanic current and circulations, deep learning, long short term memory, predictability of oceanic circulations, spectral properties of oceanic current.

AMS Subject Classification: 68Q32, 68T01, 68T07, 94A12.

1. Introduction

Ocean currents and circulations are one of the very important phenomena observed in ocean hydrodynamics and they have attracted a lot of attention throughout the historical development of ocean science [1, 2]. These currents and circulations can be generated by many different mechanisms. These mechanisms include but are not limited to winds [3, 4], tides [5], waves [6], changes in underwater topography [7] and buoyancy fluxes [6, 8, 9, 10]. There are different ways of classifying the currents. According to their forcing mechanisms, they may be classified as wind driven, tidal or thermohaline. One other possible classification is according to the depth of occurrence, currents may be classified as the surface or the subsurface currents. Majority of the currents occurring on the ocean surface are wind-driven, however thermohaline currents are driven by the temperature and salinity. The sinking of denser saltier water at colder parts of the ocean

[1] Engineering Faculty, Istanbul Technical University, 34469, Maslak, İstanbul.
Engineering Faculty, Boğaziçi University, 34342, Bebek, İstanbul.
e-mail: cbayindir@itu.edu.tr; ORCID: http://orcid.org/0000-0002-3654-0469.

§ Manuscript received: December 19, 2020; accepted: February 01, 2021.

drives subsurface thermohaline currents. Yet another possible classification is to classify the currents according to their occurrence locations, i.e. the currents flowing some distance away from the shore can be classified as an offshore current and the ones closer to shore can be classified as an inshore current. From a time series analysis point of view, a more appropriate classification is to classify them as periodic or mean currents. The currents which have changing speed and directions cyclically at regular intervals can be classified as periodic currents. However, mean circulations in the ocean have mean parts that experience no or relatively little changes in time and space.

Ocean currents and circulations have very important effects on the marine environment. They are one of the driving mechanisms of the mass and heat transfer on earth. They can heat up or cool down the oceans and can transport biological agents, nutrients, and chemicals. They may lead to many engineering problems for offshore engineering structures and marine travel, such as excessive vibrations, an increase of transit times and accident risks in strategic waterways, such as the Bosphorus [11, 12]. Additionally, they may lead to the formation of rogue waves due to their wave blocking effect [13]. They are the main driving mechanism of the tidal energy converters [14, 15, 16].

Due to the reasons, some of which are summarized above, the prediction of ocean currents and circulation in the marine environment is a vital problem for the safety of marine operations and structures. Various attempts are made for the prediction of the different ocean processes including the ocean currents and circulations. Various wave models and time-series methods are utilized for the prediction of the ocean wave height and energy in [17] and [18]. The predictability of the near-surface winds using a Kalman filtering technique is discussed in [19]. A transfer function based method was proposed in [20] for the wave height forecast. The usage of artificial neural networks are discussed for the prediction of waves is studied in [21]. A genetic programming approach is used for the real-time wave forecasting in [22]. More recently, the prediction of wave conditions using a machine learning approach is proposed in [23]. Again a machine learning approach is proposed in [24] for the prediction of thermocline parameters. A statistical machine learning approach is utilized in [25] for the prediction of ocean processes. Spatio-temporal prediction of the ocean currents using Gaussian processes is discussed in [26]. Another approach was to apply deep learning to ocean data inference and subgrid parameterization [27].

In this study, we discuss and analyze the predictability of the time-varying ocean currents using a deep learning approach [28]. To be more specific, we apply the LSTM deep learning network to an ocean current data set collected by NOAA in Massachusetts Bay between November 2002-February 2003. We discuss the performance of the LSTM for the prediction of the ocean current speed data and discuss the root-mean-square (RMS) error and spectral properties of predictions. We also investigate the effect of the length of the training data set on predictions. We discuss our findings and comment on our results.

2. Methodology

2.1. **Review of the Long Short Term Memory.** One of the most commonly used deep learning networks is the LSTM network, which was introduced in [29]. One of the possible uses of the LSTM is the prediction of the various time series data [29, 30]. Using the dependencies between the time steps of the sequences of data, the LSTM learns long-term dependencies. In order to predict the values of the data at the future time steps, a sequence-to-sequence regression LSTM network can be trained. For the LSTM network, training sequences with one-time step shifted values are used as the responses [29], thus

the LSTM network learns to predict the values of the next time step at every time step of the training sequence [29]. The LSTM layer architecture is shown in Fig. 1.

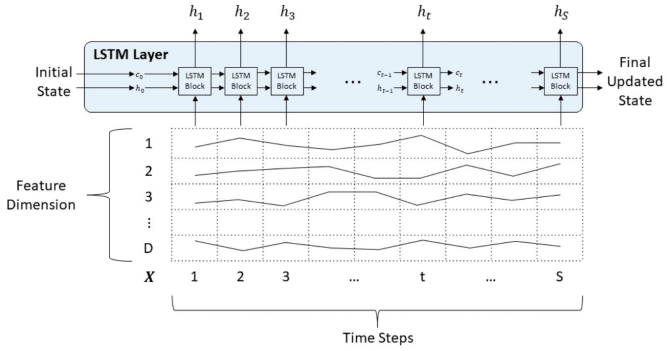

FIGURE 1. LSTM layer architecture [29].

In the LSTM layer architecture, the flow of a time series having D features through an LSTM layer is illustrated. In this architecture, h denotes the output and c denotes the cell state [29]. In LSTM architecture the first LSTM block depicted in Fig. 1 uses the initial state and the first time step of the sequence [29]. Then, the first output and the updated cell state are computed. At an arbitrary time step t, the LSTM block uses the current state and the next time step of the input sequence and computes the output as well as the updated cell state [29]. The output of the LSTM layer is contained in the hidden state at a given time step t. The learned information at the previous time steps is contained in the cell state. The information is either added or removed from the cell state at each time step of the input sequence. The gates shown in Fig. 2, controls these updates.

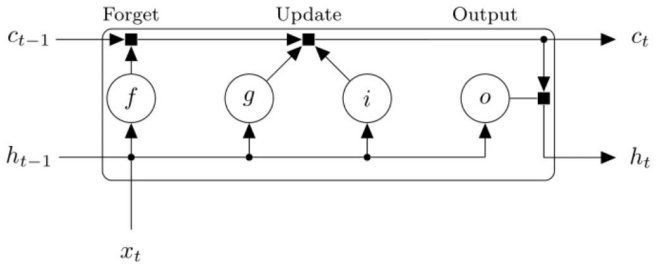

FIGURE 2. LSTM gates [29].

The letters i, f, g, o in this figure denote the input, forget gates, cell candidate and the output gate, respectively [29]. At time step t, the cell state is computed by

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g_t \tag{1}$$

In here, $\otimes$ is the Hadamard product (element-wise multiplication) [29]. At time t, the hidden state is computed using

$$h_t = o_t \otimes \sigma_c(c_t) \tag{2}$$

where σ_c is the state activation function. Among different possibilities, the tanh function is used as the state activation function throughout this study due to its faster convergence and to overcome vanishing gradient computations efficiently. Some other possible alternatives which give values between 0 and 1 include sigmoid function, hard tanh function,

Component	Formula
Input gate	$i_t = \sigma_g(W_i x_t + \mathrm{R}_i h_{t-1} + b_i)$
Forget gate	$f_t = \sigma_g(W_f x_t + \mathrm{R}_f h_{t-1} + b_f)$
Cell candidate	$g_t = \sigma_c(W_g x_t + \mathrm{R}_g h_{t-1} + b_g)$
Output gate	$o_t = \sigma_g(W_o x_t + \mathrm{R}_o h_{t-1} + b_o)$

FIGURE 3. LSTM components and formula [29].

ReLU and leaky ReLU functions. The gates and their formula are shown in Fig. 3. In here, W shows the input weight, R shows the recurrent weight and b shows the bias [29]. σ_g shows the gate activation function. A sigmoid function, $\sigma(x) = (1 + e^{-x})^{-1}$ is used as the gate activation function since 3 gates are involved and flow of information between gates need to be controlled [29]. In this approach, the training data is standardized to have zero mean and unit variance for a better fit. Then, the LSTM layer is specified to have 200 hidden units. The number of training epochs is selected as 250 and a gradient threshold value of 1 is used to avoid the divergence of the gradients. The predictions are started with an initial learning rate of 0.005, and after 200 epochs, it is dropped to 0.004. Then the predicted time series is unstandardized using the parameters discussed earlier. Additionally, in the LSTM network approach one can also update the network state with observed values instead of predictions by utilizing the time steps between predictions which generally results in better prediction performance. All of the LSTM network-based predictions in this work are performed using the parameters discussed above. Our paper focuses on discussing the usage of deep learning for the prediction of ocean currents and circulations. Thus, the reader is referred to references [29, 30] for a more comprehensive discussion and the details of the LSTM network.

3. Properties and Review of the Ocean Current Data

In this study, we use an experimental data set to test the performance of the LSTM deep learning network for the ocean current and circulation predictions. The data set is recorded by NOAA in North Atlantic in the Massachusetts Bay at the location $42°22'38.3''N$, $70°46'54.8''W$ $(42.3773N, 70.7819W)$. The experimental data were recorded using a VMCM type moored current meter between the dates 2002/10/24 and 2003/02/12. The seafloor depth at this specific location was $33.7m$ and the current meter was deployed at a depth of $23.5m$. The horizontal current velocities, namely u and v are measured at every 3 minutes 44 seconds during this experiment. The original data set and further information can be seen at NOAA's website `https://www.nodc.noaa.gov/gocd/data/a0060062/info/gocd{_}a0060062{_}6964vm-a.html`. The time series of the 2D horizontal current velocities, u and v are depicted in Fig. 4. As Fig. 4 confirms, horizontal current velocity in both directions exhibit an oscillatory behavior, however the x-component of the horizontal velocity exhibits a more steady-in-the-mean trend compared to the y-component v.

4. Results and Discussion

In this section, we assess the performance of the LSTM deep learning network for the prediction of the ocean current speed data depicted in Fig. 4 in the preceding section.

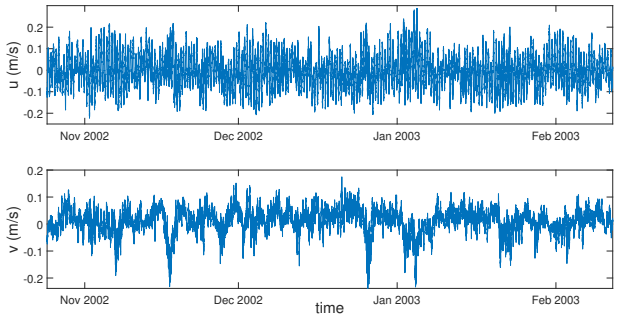

FIGURE 4. Ocean current speed time series a) u (m/s) b) v (m/s).

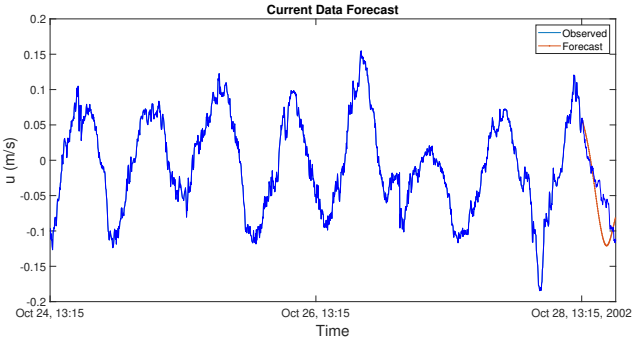

FIGURE 5. Observed and predicted time series of the first component of the current velocity (u) obtained using the initial 95 % as the training sequence.

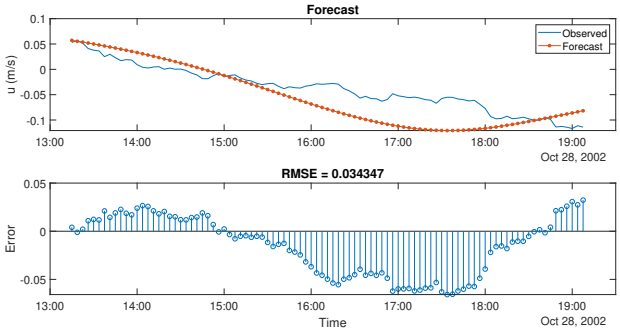

FIGURE 6. a) Comparisons between the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 95 % as the training sequence, b) the RMS error of predictions.

In Fig. 5, the initial part of the first component of the original current speed data (u) between the dates October 24-October 28, 2002 is shown. In this prediction, we use the initial 95 % of this part of the current speed (u) data set as the training sequence. The training data set (in blue) and the predicted time series (in red) are depicted in Fig. 5. We compare the predicted and observed time series and depict the root-mean-square (RMS) error in Fig. 6. As one can realize from these figures, the LSTM deep learning network

performs quite well for the prediction of the first component of ocean current speed, u. The predictions are better for the first few time steps. As depicted in Fig. 6, the absolute value of the RMS error remains less than $0.05m/s$ for the data having an absolute peak value of approximately $0.15m/s$. It is useful to note that, the sampling interval of this data set is 3 minutes 44 seconds. An early prediction time on the order of a few time steps can be very beneficial for many practices in marine science such as for tidal energy harvesting and vibration control, just to name a few. Depending on the temporal resolution of the data and with the advance of algorithms, this prediction times can vary and may be substantially improved.

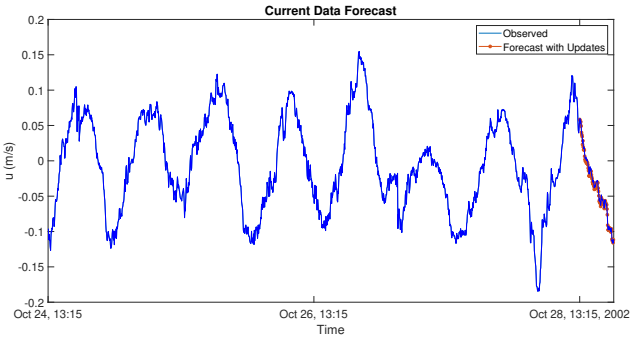

FIGURE 7. Observed and predicted time series of the first component of the current velocity (u) obtained using the initial 95 % as the training sequence and using the updates.

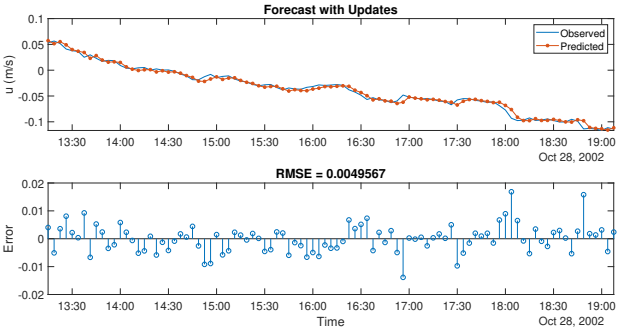

FIGURE 8. a) Comparisons between the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 95 % as the training sequence and using the updates, b) the RMS error of predictions.

Next, we turn our attention to discuss the effects of using the observed values instead of predicted ones which are used for updating the LSTM network. We plot the first component of the predicted ocean current speed (u) time series and the RMS error for this case in Fig. 7 and Fig. 8. The same data set depicted Fig. 5 is used and its initial 95 % is used as the training sequence, as before. Checking Fig. 7 and Fig. 8, it is possible to conclude that the LSTM prediction with updates leads to better predictions and RMS error significantly reduces. For the same data set, the peak absolute value of the RMS error becomes less than $0.02m/s$ for this case.

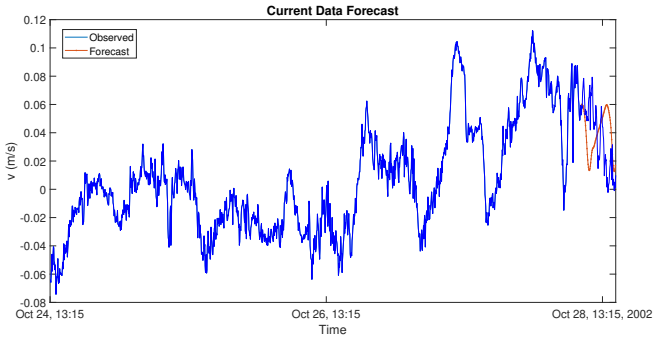

FIGURE 9. Observed and predicted time series of the second component of the current velocity (v) obtained using the initial 95 % as the training sequence.

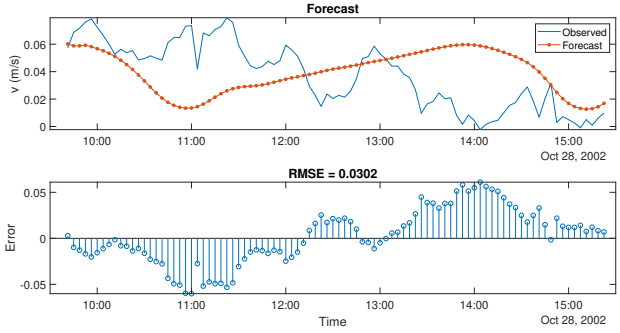

FIGURE 10. a) Comparisons between the observed and predicted time series of the second component of the current velocity (v) obtained using the initial 95 % as the training sequence, b) the RMS error of predictions.

Focusing on the second component of the horizontal current velocity, namely v, we plot the predicted v time series and the RMS error LSTM in Fig. 9 and Fig. 10. The training sequence is selected as the initial 95 % of this part of the v time series. Checking these figures it is possible to realize that LSTM performs worse compared to the predictions of u. While u has a periodic steady-in-mean behavior, v exhibits a tendency of increase and does not have a steady-in-mean behavior, which makes the LSTM network predictions poorer. The absolute value of the RMS error is on the order of $0.05m/s$ for the data having an absolute peak value of approximately $0.10m/s$.

It is known that if there is access to the actual values of time steps between predictions, one can update the network state with the observed values instead of using the predictions [29, 30]. In order to discuss the effects of updating the LSTM network with such observed values, we repeat the same simulations for the case with updates and depict the related results in Fig. 11 and Fig. 12. Comparing Fig. 9 and Fig. 10 with Fig. 11 and Fig. 12 respectively, it is possible to argue that using the observed values to update the LSTM network significantly improves the prediction performance. The peak absolute value of the RMS error becomes less than $0.04m/s$ for this case. It is also useful to note that using the observed values as updates imposes a limit on the prediction time and limits it to a one time step. The value of the time step is 3 min 44 seconds for the data set used, however, depending on the size of the sampling interval the prediction time can vary and can still

be very beneficial. Additionally, it is possible to upsample or downsample the data and adaptive time-stepping can be utilized to enhance the prediction time scales.

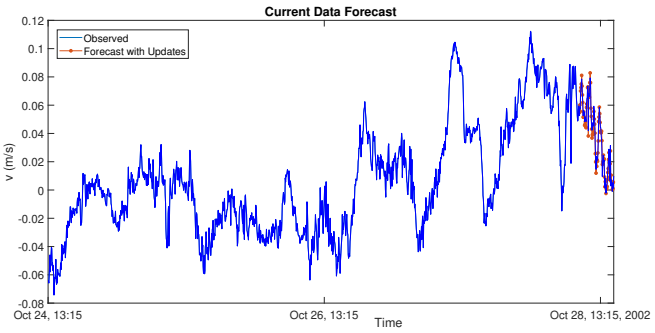

FIGURE 11. Observed and predicted time series of the second component of the current velocity (v) obtained using the initial 95 % as the training sequence and using the updates.

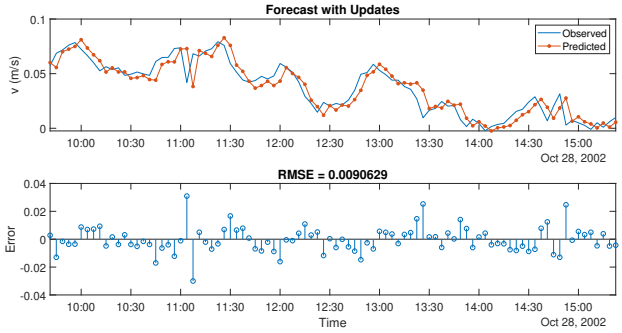

FIGURE 12. a) Comparisons between the observed and predicted time series of the second component of the current velocity (v) obtained using the initial 95 % as the training sequence and using the updates, b) the RMS error of predictions.

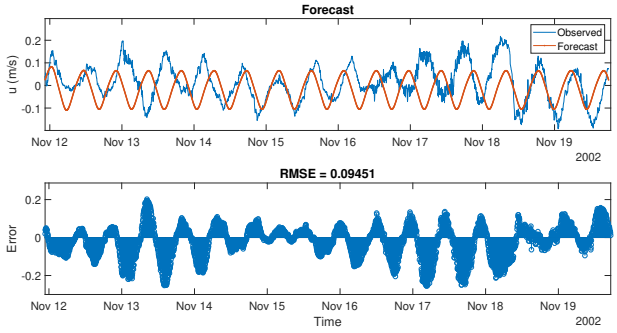

FIGURE 13. a) Comparisons between the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 70 % as the training sequence, b) the RMS error of predictions.

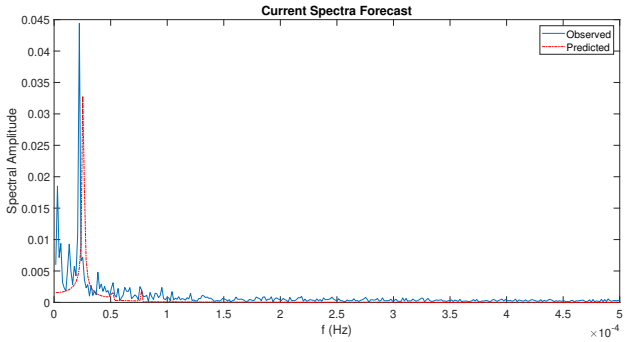

FIGURE 14. Comparison of the Fourier spectra of the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 70 % as the training sequence.

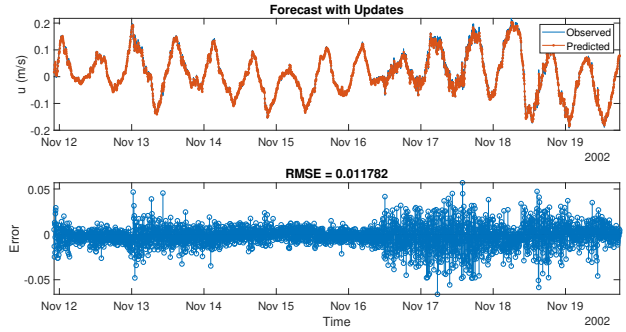

FIGURE 15. a) Comparisons between the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 70 % as the training sequence and using the updates, b) the RMS error of predictions.

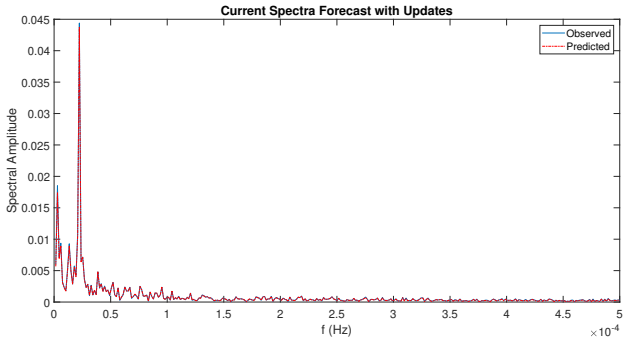

FIGURE 16. Comparison of the Fourier spectra of the observed and predicted time series of the first component of the current velocity (u) obtained using the initial 70 % as the training sequence and using the updates.

Next, we investigate the effects of the size of the LSTM training data set and the spectral properties of predictions obtained by the LSTM network. With this motivation, we use the initial part of the time series between the dates October 24-November 19, 2002. We select the initial % 70 of this part of the time series data as the training data set and

predict the remaining part using the LSTM network. Results for the first component of the current speed data (u) with no updates are depicted in Fig. 13 and its spectrum is depicted in Fig. 14. The results for u with updates are depicted in Fig. 15 and its spectrum is depicted in Fig. 16.

Checking Figs. 13-16, it is possible to argue that LSTM with no updates produces quite sufficient results for the prediction of the ocean current speed. When the observed values are used as updates, the results significantly improve and predictions become very successful for many marine operations and marine engineering purposes. The depicted Fourier spectra in Fig. 15 and Fig. 16 clearly show that the predicted time series with the LSTM network has a higher peak frequency and its bandwidth is significantly limited compared to the observations when no observations are used as updates. Again, when the observations are used as updates, the Fourier spectrum of the predicted time series by the LSTM network is in very good agreement with the spectrum of the observations. Comparing Fig. 4 and Fig. 13, it is also possible to state that when the training sequence is longer, the prediction performance of the LSTM network significantly improves as expected.

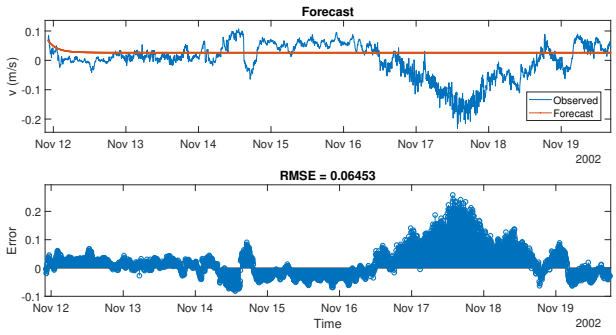

FIGURE 17. a) Comparisons between the observed and predicted time series of the second component of the current velocity (v) obtained using the initial 70 % as the training sequence, b) the RMS error of predictions.

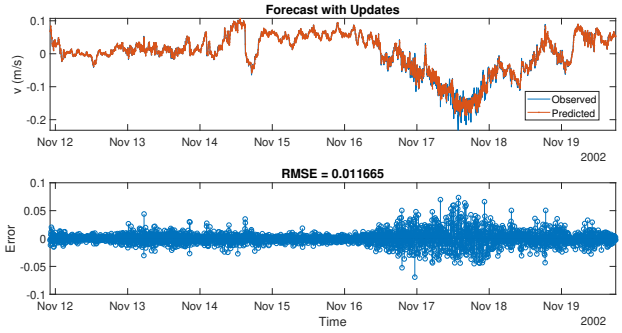

FIGURE 18. a) Comparisons between the observed and predicted time series of the second component of the current velocity (v) obtained using the initial 70 % as the training sequence and using the updates, b) the RMS error of predictions.

Lastly, we turn our attention again to the second component of the current speed time series, namely v. Again, we use the initial part of v time series between the dates October 24-November 19, 2002 and as before we select the initial % 70 of this part of the data

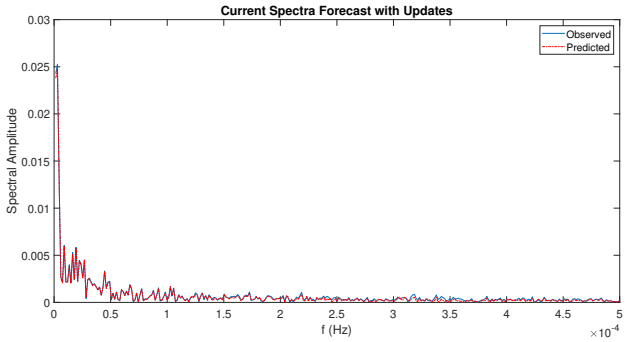

FIGURE 19. Comparison of the Fourier spectra of the observed and predicted time series of the second component of the current velocity (v) obtained using the initial 70 % as the training sequence and using the updates.

set as our training sequence. The predicted v time series with no updates is depicted in Fig. 17 and the predicted v time series with updates is depicted in Fig. 18. Checking these figures, it is possible to argue that the prediction with no updates is poor and can not represent the data after a few time steps. This mismatch is a result of using a shorter training sequence as well as the unsteady-in-mean nature of the v time series. However, when the observed values are used as updates in the LSTM network, the predicted time series and its spectrum become excellent for many possible engineering uses as depicted in Figs. 18 and 19.

5. Conclusion and Future Work

In this paper, we have discussed the predictability of the ocean current speed time series by deep learning. Specifically, we have investigated the applicability of the LSTM deep learning network to the ocean current speed time series and discussed its prediction performance. The experimental data set used in our study was collected by NOAA in Massachusetts Bay between the dates November 2002-February 2003. Implementing the LSTM network-based deep learning approach on this data set, we have analyzed the RMS error and the Fourier spectra of the predicted time series. We showed, depending on the steady-state-in-the-mean behavior of the predicted time series, the LSTM network with no updates can be very successful in the prediction of the ocean current speed data, and may allow for accurate predictions at least a few time steps in advance. We showed that the LSTM network predictions with no updates have a higher peak frequency compared to the observed spectra. However, when the observed values are used as updates in the LSTM network, the prediction accuracy significantly develops. We also showed that longer training sequences result in better prediction accuracy of the ocean current time series. Depending on the sampling interval of the data, the prediction times can significantly vary. Thus, the LSTM network-based deep learning approach proposed in this paper can be used to predict the ocean current time series data with very beneficial prediction times. Our results can be generalized for the prediction of other oceanic and atmospheric time-series data including but are not limited to Gulf Stream, North Equatorial, Agulhas currents and East Wind drift. Possible usage areas include but are not limited to predicting the tidal energy variation, controlling the current induced vibrations of marine structures and estimation of the wave blocking point by the chaotic oceanic current and circulation.

References

[1] Richardson, P., (1980), Benjamin Franklin and Timothy Folger's first printed chart of the Gulf Stream, Science, 207, pp. 643-645.

[2] Peterson, R. G., Stramma, L. and Kortum, G., (1996), Early concepts and charts of ocean circulation, Prog. Oceanogr., 37, pp. 1-115.

[3] Munk, W. H., (1950), On the wind driven ocean circulation, J. Meteor., 7, pp. 79-93.

[4] Sverdrup, H. U., (1947), Wind Driven Currents in a Baroclinic Ocean; With Application to the Equatorial Currents of the Eastern Pacific, Proc. Natl. Acad. Sci. USA, 33, pp. 318-326.

[5] Parker, B. B., (1991), Tidal Hydrodynamics, John Wiley and Sons.

[6] Stommel, H., (1957), A survey of ocean current theory, Deep Sea Res., 4, pp. 149-184.

[7] Tomczak, M. and J. S. Godfrey, (1994), Regional Oceanography: an Introduction, Pergamon Press.

[8] Wust, G., (1935), Schichtung und Zirkulation des Atlantisches Ozeans, Die Stratosphare, Wissenschaftliche Ergebnisse der Deutschen Atlantischen Expedition auf dem Forschungs-und Vermessungsschiff 'Meteor' 1925–1927, 6.

[9] Sverdrup, H. U., Johnson, M. W. and Fleming, R. H., (1942), The Oceans: Their Physics, Chemistry and General Biology, Prentice Hall.

[10] Siedler, G., Church, J. and Gould, J., (2001), Ocean Circulation and Climate: Observing and Modelling the Global Ocean, Academic Press.

[11] Oğuz, T., Özsoy, E., Latif, M. A., Sur, H. I. and Ünlüata, Ü., (1990), Modeling of hydraulically controlled exchange flow in the Bosphorus Strait, J. Phys. Oceanogr., 20, pp. 945-965.

[12] Jarosz, E., Teague, W. J., Book, J. W. and Beşiktepe, Ş., (2011), On flow variability in the Bosphorus Strait, J. Geophys. Res., 116, C08038.

[13] Bayındır, C., (2016), Shapes and statistics of the rogue waves generated by chaotic ocean current, Proc. of the 26th ISOPE, Rhodes, Greece.

[14] Magagna, D. and Uihlein, A., (2015), Ocean energy development in Europe: current status and future perspectives, Int. J. Mar. Energy, 11, pp. 84-104.

[15] Uihlein, A. and Magagna, D., (2016), Wave and tidal current energy – A review of the current state of research beyond technology, Renew. Sust. Energ. Rev., 58, pp. 1070-1081.

[16] Lynn, P. A., (2013), Electricity from wave and tide: an introduction to marine energy, John Wiley and Sons.

[17] Roulston, M. S., Ellepola, J., Von Hardenberg, J. and Smith, L. A., (2005), Forecasting wave height probabilities with numerical weather prediction models, Ocean. Eng., 32, pp. 1841-1863.

[18] Reikard, G., Pinson, P. and Bidlot, J-R., (2011), Forecasting ocean wave energy: the ECMWF wave model and time series methods, Ocean. Eng., 38, pp. 1089-1099.

[19] Malmberg, A., Holst, U., Holst, J., (2005), Forecasting near-surface ocean winds with Kalman filter techniques, Ocean. Eng., 32, pp. 273-291.

[20] Ho, P. C. and Yim, J. Z., (2006), Wave height forecasting by the transfer function model, Ocean. Eng., 33, pp. 1230-1248.

[21] Londhe, S. N. and Panchang, V., (2006), One-day wave forecasts based on artificial neural networks, J. Atmos. Ocean. Technol., 23, pp. 1593-1603.

[22] Gaur, S. and Deo, M. C., (2008), Real-time wave forecasting using genetic programming, Ocean. Eng., 35, pp. 1166-1172.

[23] James, S. C., Zhang, Y. and O'Donncha, F., (2018), A machine learning framework to forecast wave conditions, Coast. Eng., 137, pp. 1-10.

[24] Jiang, Y., Gou, Y., Zhang, T., Wang, K. and Hu, C., (2017), A machine learning approach to Argo data analysis in a thermocline, Sensors, 17, 2225.

[25] Hollinger, G., Pereira, A., Ortenzi, V. and Sukhatme, G., (2012), Towards improved prediction of ocean processes using statistical machine learning, Proc. Robotics: Science and Systems Workshop on Robotics for Environmental Monitoring (RSS), Sydney, Australia.

[26] Sarkar, D., Osborne, M. A. and Adcock, T. A., (2019), Spatio-temporal prediction of tidal currents using Gaussian processes, J. Geophy. Res.: Oceans, 124, 2697-2715.

[27] Bolton, T. and Zanna, L., (2019), Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization, J. Adv. Model. Earth Sys., 11, 376-399.

[28] MacKay, D. J. C., (2003), Information theory, inference and learning algorithms, Cambridge University Press.

[29] Hochreiter, S. and Schmidhuber, J., (1997), Long Short-Term Memory, Neur. Comp., 9, 1735-1780.

[30] Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R. and Schmidhuber, J., (2016), LSTM: A search space odyssey, IEEE Trans. on Neur. Netw. and Learn. Sys., 28, 2222-2232.

Cihan Bayındır for the photography and short autobiography, see TWMS J. App. and Eng. Math. V.5, N.2.