International Journal of Civil and Structural Engineering Research ISSN 2348-7607 (Online)
Vol. 9, Issue 2, pp: (83-91), Month: October 2021 - March 2022, Available at: www.researchpublish.com

# Prediction of oceanographic data using LSTM network: A case study for FSRU in Arabian Sea

A. Basu[1], A.A. Purohit[2], K.A. Chavan[3]

[1]Scientist-C, [2]Scientist-E, [3]Research Assistant

[1,2,3] Central Water and Power Research Station, Pune, India

***Abstract:*** **Information on site specific measured oceanographic data such as wind generated waves, currents is seldom available and plays an important role in offering solution to various coastal engineering problems. Due to unanticipated difficulties during field measurement campaigns, it has been observed that measured data for some specific duration is missing or lost. In such circumstances, data driven methods are essential to predict the missing data and its use to calibrate the physical model and to offer engineering solution to the project. Present study investigates the applicability of one of the popular recently developed artificial intelligence (AI) model named as long short-term memory (LSTM) in predicting current data in a macro tide dominated Thane creek and the significant wave height ($H_s$) for Floating Storage & Regasification Unit (FSRU) structure. The results indicate that LSTM networks formed based on data of 7/15 days of current/tide, can predict current data for 8/10 days with root mean square error (RMSE) of 0.091 and 0.09 respectively. Study also reveals that for location in Thane creek, LSTM prepared on the basis of 8 days of $H_s$ data is able to predict $H_s$ for 33 hours with RMSE value of 0.14. Based on the predicted data, current data is used to calibrate the physical tidal model to determine flow conditions for proposed FSRU in Thane creek and also operable condition (Hs). The study reveals that FSRU needs to be aligned at 33° N and hence LSTM network was found to be useful in design of waterfront structures.**

***Keywords:*** **Artificial intelligence, current, LSTM, oceanographic data, significant wave height.**

## I. INTRODUCTION

The importance of measured oceanographic data such as wind generated waves, currents, tides, suspended sediment concentration, salinity etc. plays an important role for understanding various nearshore coastal processes such as coastal erosion/accretions, shoreline changes; determination of design/operational conditions for coastal structures; in planning schedule for safe navigation of ships; in assessing the trajectory of movement of dredged materials/oil spills; in validation/calibration of physical/numerical models on tidal/wave hydrodynamics, sedimentation/wave transformation etc. Presently various organisations such as Indian National Centre for Ocean Information Services (INCOIS) deployed buoys all over the Indian Ocean to measure different metocean parameters and the organisations like European Centre for Medium Range Weather Forecasts (ECMWF), National Oceanic and Atmospheric Administration (NOAA) forecast various metocean parameters at various grid resolutions all over the globe. However, the collection of site specific measured metocean data even although a costly affair and involves high risk of human life, measuring instruments; is still inevitable in solving various site specific coastal engineering problems. Due to the various difficulties during the field measurement campaign, many times it has been observed that the collected field data for some specific duration is either missing or is lost and, in such scenario, to predict the missing data, data driven methods are essential to predict the missing data and its use to calibrate physical/numerical model will offer engineering solution to the project. The development of Floating Storage and Regasification Unit (FSRU) near the entrance of Thane creek, Mumbai for availing storage facility for LNG (1,70,000 cum) and berthing of LNG tanker was under consideration. In order to provide design basis for this waterfront facility in macro tidal region (tidal range of 5 m), its alignment needs to be finalised in such a way that irrespective of phase of tide (flood/ebb), flow direction at berth of LNG tanker/FSRU should remain parallel to the prevailing flow direction to minimise the undesirable forces on mooring ropes and also to determine operational conditions. Thus to study these aspects, the application of well calibrated hydrodynamic/wave transformation model/data driven method and reliable field data is essential.

In recent years with rapid development in the field of artificial intelligence (AI), its applicability in solving diversified complex non-linear system is widely increasing. Many researchers have applied AI to predict oceanographic parameters such as sea surface temperature, salinity, significant wave height, current etc. Feed forward artificial neural network with various training algorithm was applied by Deo MC and Naidu CS [1] for predictions of wave height in offshore region of Bay of Bengal for the duration of 3 to 24 hrs. Park et. Al. [2] adopted feed forward neural network (FNN) and LSTM to predict Hs for 48 hrs in Korea Straight and it was observed that predicted $H_s$ by LSTM matches well with the observed data as compared to feed forward neural network. The main limitation in using feed forward neural network in predicting Hs is due to its less temporal dependencies and it makes the model more sensitive to noise. Nonlinear auto-regressive network with eXogenous inputs (NARX) was applied by Zubier [3] wherein wind shear velocity and wave direction were used as input to improve the prediction of wave height in the Eastern Central Red Sea for 3, 6, 12, 24 hrs durations. Delong Chen et. al. [4] had integrated wavelet transformation and graph neural network to form Wavelet Graph Neural Network (WGNN) to predict $H_s$ and the study indicates that WGNN performs better than numerical model, machine learning model such as ANN, SVM etc. The main hindrance in accurate prediction of significant wave height is caused due to the randomness and instability of sea wave (Hou et al., 2012 [5]). Nitsure et. al., 2012 [6] used future wind field information to improve the prediction accuracy of $H_s$. Peres et al. (2015) [7] used the future wind field forecasted by the numerical model to predict $H_s$. However, in the present study futuristic $H_s$ were predicted (inside of a tide dominated creek) through LSTM model with the help of only measured $H_s$.

As compared to the prediction of Hs through AI, the application of AI in prediction of current is limited. Garg et al. (2006) [8] proposed application of ANN model to predict alongshore current and ARIMA models for prediction of cross-shore current. Dauji S. et. al. (2015) [9] had applied ANN to predict short term current data (1 hr to 24 hr) in North Atlantic Ocean and North Pacific Ocean. Dauji S. et. al. (2016) [10] had applied combined numerical and neural technique (feed forward neural network) for short term (1 day to 5 day) prediction of ocean currents in the Indian Ocean. In the present study LSTM model was used to predict current data for short term (8 days) to medium term (10 days) with the help of measured current and tidal data inside of a macro tide dominated creek (Thane). The current/wave measurement locations are shown in Fig. 1a.

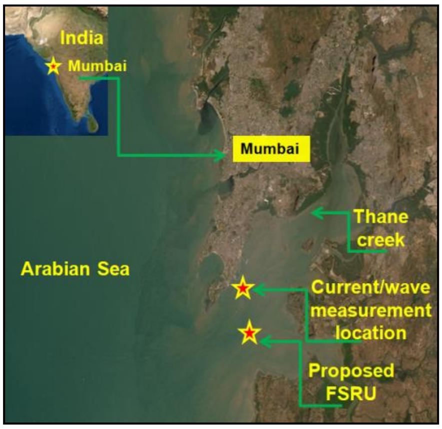


**Fig 1a. Current and wave measurement location**

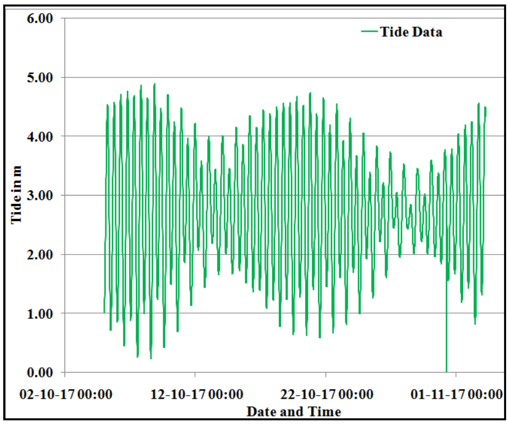


**Fig 1b. Plot of measured tide data**

## II. OCEANOGRAPHIC DATA AND LSTM NETWORK

### *A. Oceanographic Data*

In the present study the current/wave data measurement location being situated on the leeside of Mumbai island at about 5 km north from the entrance of Thane creek, the wave action is relatively insignificant however due to the presence of macro type of semi diurnal tidal phenomena (maximum tidal range about 5 m), there is huge exchange of tidal flux through the 10 km wide entrance of creek. The tidal excursion in Thane creek is about 35 km on the north from the entrance of the creek. The exchange of huge amount of tidal flux creates strong tidal current inside of the creek and it governs the flow regime of the region. The current data was collected to calibrate the tidal model (physical) for finalising the alignment of proposed Floating Storage and Regasification Unit (FSRU) and its berthing facility near the entrance of Thane creek. The bottom mounted ADCP was used to measure the current data at 10-minute interval over the entire water depth at 0.5 m interval starting from 1 m above the sea bed. The ADCP was installed at 8 m water depth below CD (chart datum). The weighted average current strength over the entire water depth was calculated to compute the depth average current strength at every time step and the analysis of data shows that maximum depth average current strength of 0.9 m/sec and 1.37 m/sec during flood/ebb phase of tide prevails at site. The direction of current during flood varies between 16°-20° N, while during ebb it reverses and happens to be 200°-220° N. The information on measured tide data for the same duration at 10 minute interval was also available near the current location i.e. at Apollo Bunder. The analysis of tidal data carried out reveal that maximum tidal range is 4.8m, while minimum is 0.59m. The tides are semi-diurnal in nature with diurnal inequality. The plots of tidal data, measured current strength and direction are shown in Fig. 1b, 2a and 2b respectively. The Hs data was recorded at 1 hr. interval at the same location where the current data was collected and the record shows that it's magnitude varies randomly within a small range i.e. from about 0.1 m to 0.85 m. The current/Hs data was collected during September-October 2017. The purpose of selecting two different oceanographic parameters viz. wave and current is primarily to check the suitability of LSTM model in predicting randomly varied data (Hs) and data (tidal current) representing gradually varying flow phenomenon. The plot of measured Hs inside the creek is shown in Fig. 3.

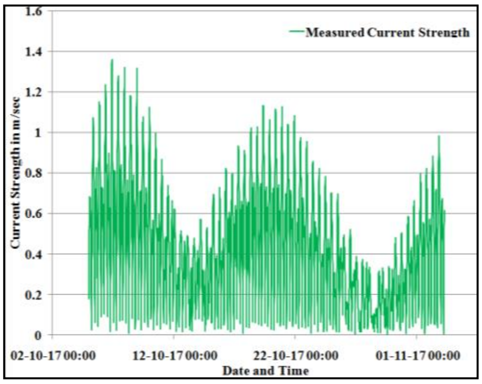


Fig 2a. Measured current strength inside the creek

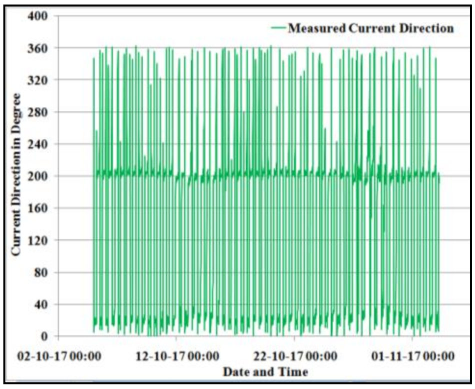


Fig 2b. Measured current direction inside the creek

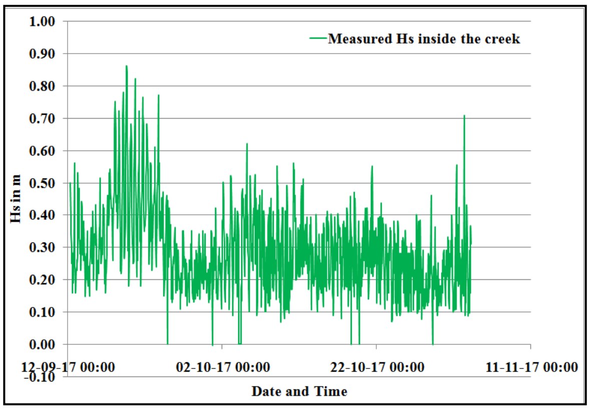


Fig 3. Measured wave height inside the creek

***B. LSTM Network***

The LSTM is one kind of recurrent neural network (RNN) which is recently being widely used in time series forecasting, text summarization, learning the behavior of how people write and talk etc. The basic structure of LSTM network is shown in Fig 4.

The LSTM network consists of forget gate ($f_t$); cell state i.e. the state of the current moment ($c_t$), the state of the previous moment ($c_{t-1}$); input gate ($i_t$) and output gate ($o_t$), ($h_{t-1}$)is the previous cell state. The "Sig" and "tanh" are smooth step functions and hyperbolic tangent functions respectively. In LSTM network the input gate, output gate and forget gate controls the information of input and output. The main function of the input gate is to update the input value, the output gate controls the information to be transmitted to the next cell, and the forget gate decides which information to be retained and which information to forget and this feature of the LSTM helps in selectively remembering the patterns in long sequences of historical data over conventional feed-forward neural networks and other RNNs. As the forget gate controls the past information to be retained or to be forgotten, it used the "sigmoid" function with the value ranges from 0 to 1 wherein if the value of the function is 0, information of the previous state is completely forgotten, and if the value is 1, information is completely retained. The information which is to be retained gets stored in input gate and it processes the values with "tanh" function and Hadamard product operator to transfer it further.

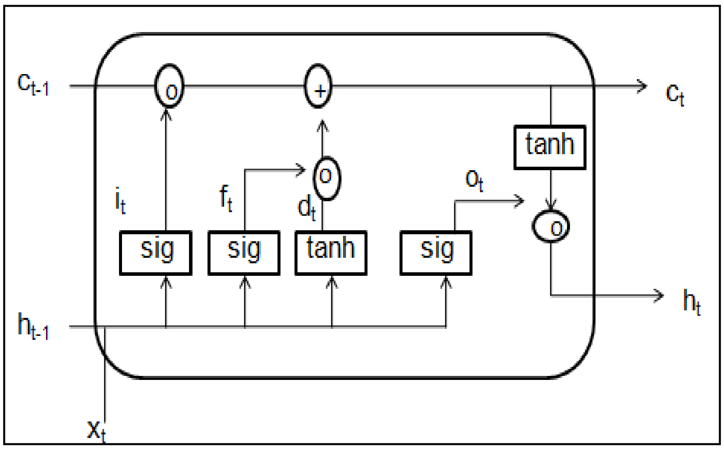


**Fig 4. Architecture of LSTM network**

In LSTM the information state in any cell gets transferred as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \; o \; c_{t-1} + b_i)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \; o \; c_{t-1} + b_f)$$

$$c_t = f_t \; o \; c_{t-1} + i_t \; o \; tanh(W_{xc}x_t + W_{hc}h_{t-1} + W_{hc}oh_{t-1} + b_c)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \; o \; c_t + b_o)$$

$$h_t = o_t \; o \; tanh(c_t)$$

Wherein "$f_t$" is forget gate, "$c_t$" is cell state i.e. the state of the current moment (ranges from -1 to 1), "$c_{t-1}$" is the state of the previous moment, "$i_t$" is input gate (ranges from 0 to 1),"$o_t$" is output gate, "$h_{t-1}$"is the previous cell state, W is weight co-efficient for a given state, "b" represents corresponding bias coefficient, "**σ**" is sigmoid function and "o" is Hadamard product operator.

In the present study two types of the LSTM networks were formed for the prediction of current (strength/direction) and Hs. The normalized current data (derived based on its mean and standard deviation), corresponding tidal range, rate of increase/decrease in water level were used to form the LSTM network to predict current data. The LSTM network was trained with the 7/15 days of current/tide data and the trained networks were used to predict futuristic current data. LSTM network was also formed based on the normalized Hs which was derived from its mean and standard deviation and for inside creek location, the network was trained with 8 days of Hs at 1hr interval to predict futuristic Hs. The performance of the network as well as its prediction accuracy is discussed in following section.

## III. DATA PREDICTION AND MODELING FOR FSRU

### *A. Data Prediction*

The LSTM networks were formed to train the network based on the tide/current data near the proposed FSRU for the duration of 7/15 days i.e from 05.10.2017 to 11.10.2017 and 04.10.2017 to 18.10.2017 respectively. The tidal range, rate of increase of tidal level and normalized depth averaged current data (strength/direction) at every 10 minutes interval were used to form the network. The depth average current strength was obtained as follows:

$$Depth\ average\ current\ strength\ at\ any\ time\ step = \frac{(v_1 z_1 + v_2 z_2 + \cdots v_n z_n)}{(z_1 + z_2 + \cdots z_n)}$$

Wherein $v_1, v_2 \ldots. v_n$ are measured current strength at $z_1, z_2 \ldots\ldots z_n$ distance from the sea bed.

The normalized current strength/directions were computed as follows:

$$Normalised\ Data = \frac{(Measured\ data - mean)}{StandaredDeviation}$$

As such the LSTM network was trained with three input vectors (i.e. tidal range, rate of increase of tidal level and normalized current data) and the targeted output was current data (strength/direction). Once the network was trained with desired accuracy, it was used to predict futuristic the current strength/direction for the duration 8/10 days. The performance of the LSTM networks was evaluated by computing the error statistics of correlation coefficient (R), root mean square error (RMSE) and from the computed RMSE value it reveals that the current strength/direction can be predicted with desired accuracy for 8/10 days ahead after training the LSTM with 7/15 days duration respectively. The prediction plots of the network for 7/15 days trained network are given in Fig. 5a to 5d. The performance plots of the network are also given in Fig. 5e and 5f. The RMSE values of the predicted data are given in Table I. As the measured current strength/direction in Thane creek is primarily governed by the presence of macro type of tidal phenomena, the network formed with the tidal range, rate of increase/decrease of water level and the corresponding normalized current data provides more opportunity for the LSTM network to understand the pattern of change of current strength/direction. This helps in predicting relatively longer duration of current strength/direction with desired accuracy provided the information on the measured tidal data is available.

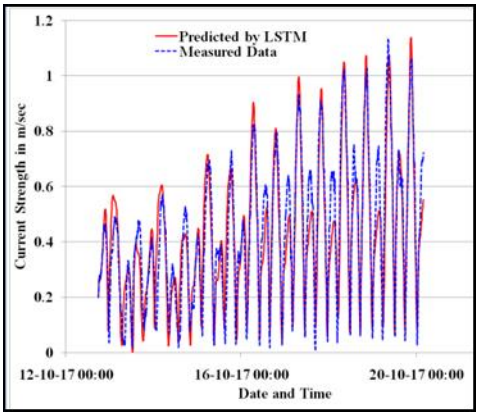


**Fig.5a. Typical plot of predicted current strength after training the network for 7 days of current strength**

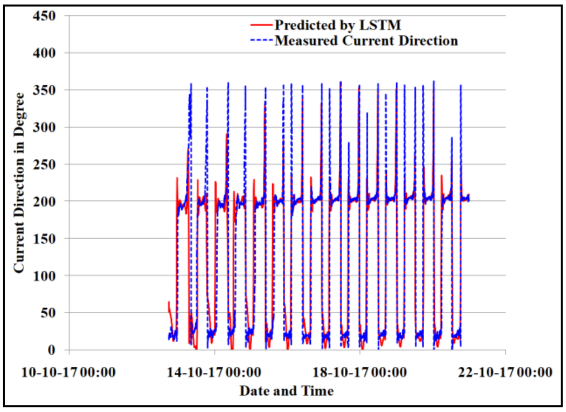


**Fig. 5b. Typical plot of predicted current strength after training the network for 7 days of current direction**

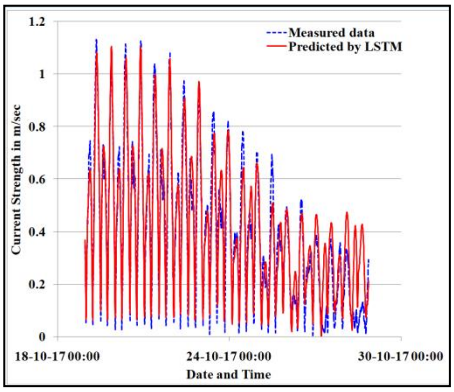


**Fig. 5c. Typical plot of predicted current strength after training the network for 15 days of current strength**

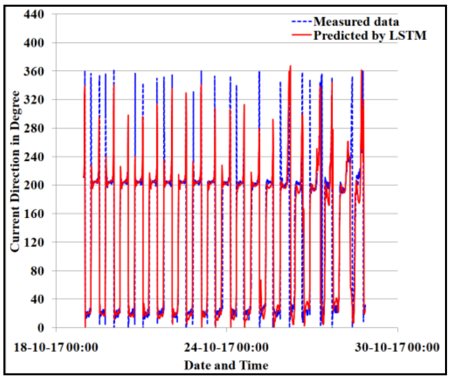


**Fig 5d. Typical plot of predicted current strength after training the network for 15 days of current direction**

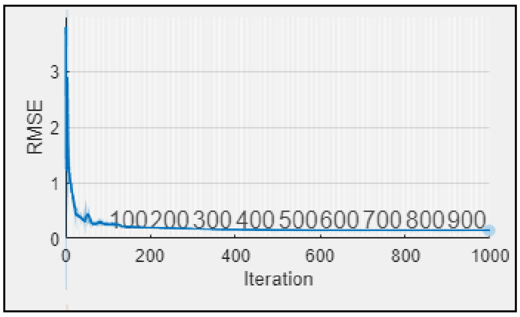


**Fig 5e. Performance plot of LSTM after training the network for 7 days of current strength**

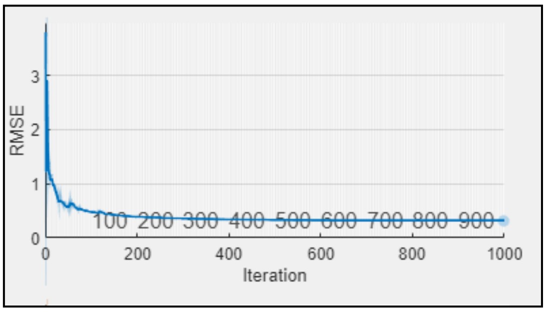


**Fig 5f. Performance plot of LSTM after training the network for 7 days of current direction**

**TABLE I: RMSE VALUES OF THE PREDICTED CURRENT STRENGTH/DIRECTION**

| Duration for which the LSTM was trained | Duration for which data was predicted | RMSE for current strength prediction | R- value for Current Strength Prediction | R- value for Current Direction Prediction |
|---|---|---|---|---|
| 7 days | 8 days | 0.091 | 0.93 | 0.8 |
| 15 | 10 days | 0.09 | 0.94 | 0.83 |

The LSTM networks were also formed to train the network based on the measured narrow range (0.75 m) of Hs at 1 hr interval for the location inside the Thane creek near the proposed FSRU for the duration of 8 days i.e from 12.09.2017 to 20.09.2017. The measured Hs were normalized as follows:

$$Normalised\ Hs = \frac{(Measured\ Hs - mean)}{Standard\ Deviation}$$

The sequential normalized Hs at every 1 hr interval were used to form the network. As such the LSTM network was trained with an input vectors (i.e. normalized Hs) and the targeted output was normalized Hs. Once the network was trained with desired accuracy, it was used to predict futuristic the Hs. The performance of the LSTM networks was evaluated by computing the root mean square error (RMSE) and prediction plots of the network and its performance plot for inside creek location are given in Fig. 6a and 6b respectively. The plot indicates that LSTM can predict Hs for the duration of 33 hrs with RMSE of 0.14. The RMSE values of the predicted data for the inside creek location are given in Table II. The accuracy of the predicted Hs entirely depends on the randomness of waves and its pattern of repeatability which also depends on the local wind. As the network was only formed with the help of measured $H_S$ only, its futuristic prediction duration is also limited.

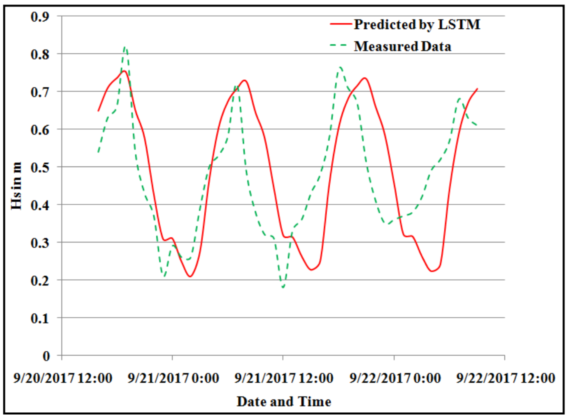


**Fig. 6a. Typical plot of predicted Hs (inside the creek) after training the network for 8 days of Hs**

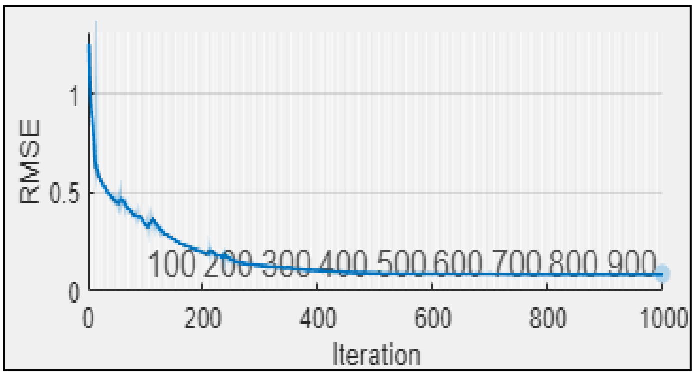


**Fig 6b. Performance plot of LSTM after training the network for 8 days of $H_s$ measured inside the creek**

**TABLE II: RMSE VALUES OF THE PREDICTED HS FOR THE INSIDE CREEK LOCATION**

| **Duration for which the LSTM was trained** | **Duration for which data was predicted** | **RMSE for predicted Hs** | **R- value for Hs Prediction** |
|---|---|---|---|
| 8 days | 33 hrs | 0.14 | 0.7 |

### *B. Modelling for FSRU*

The development of FSRU for providing temporary storage facility for LNG was proposed in wide estuarine harbour near the entrance of Thane creek (10 km wide) wherein water spreads far inside up to 35 km on north. The FSRU project consists of storage facility and berthing of LNG carrier for which the ships will have access through main navigational channel of Mumbai port. The tides being macro type, the design consideration is based on flow conditions (current strength/direction) prevailing at site. The missing data predicted using LSTM network is of more significance to calibrate the physical tidal model of Mumbai port (1:400 Horz. & 1:80 Vert.) available at CWPRS and determine the flow conditions at proposed FSRU location (Lat. 18° 54' N and Long. 72°49' E)for different tidal scenarios. The bathymetry simulated on physical model for FSRU is shown in Fig. 7a, the calibration of tidal pattern for typical spring tide is shown in Fig. 7b.

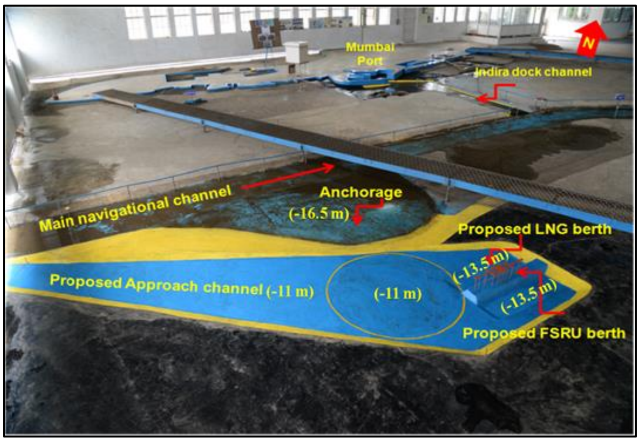


**Fig.7a. Physical model for FSRU in Thane creek**

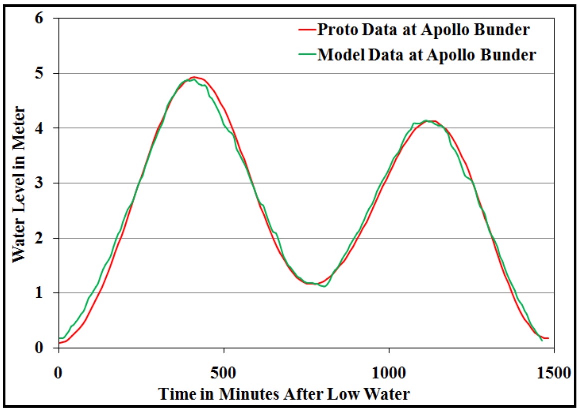


**Fig 7b. Current strength & direction at FSRU in model**

Based on this the current strength and direction plot for typical spring & neap tide is shown in Fig. 8a and Fig. 8b

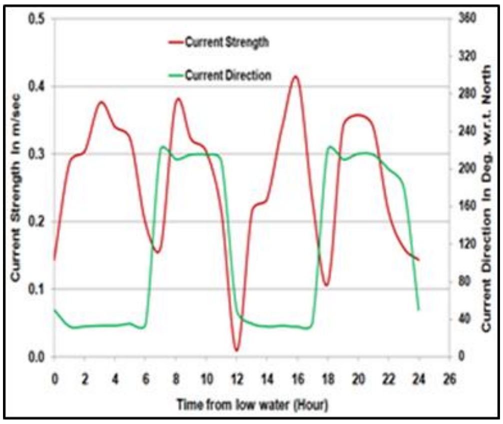


**Fig.8a. Current Strength/Direction at FSRU for spring tide**

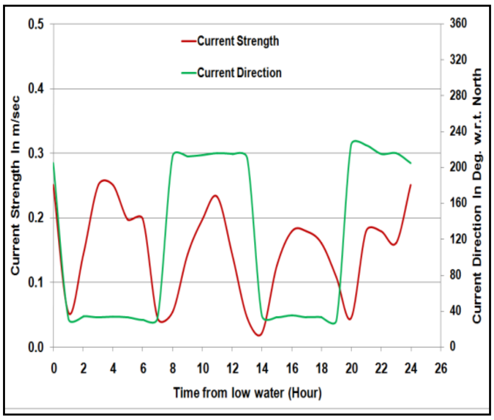


**Fig 8b. Current Strength/Direction at FSRU for neap tide**

The flow conditions observed on model for proposed FSRU during flood tide is shown in Fig. 9a while during ebb tide is in. Fig. 9b

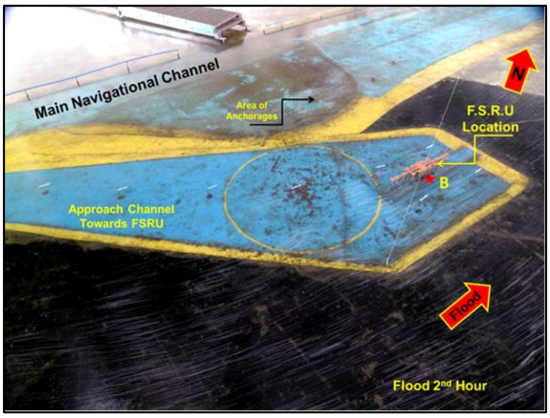


**Fig.9a. Flood flow at FSRU in Thane creek**

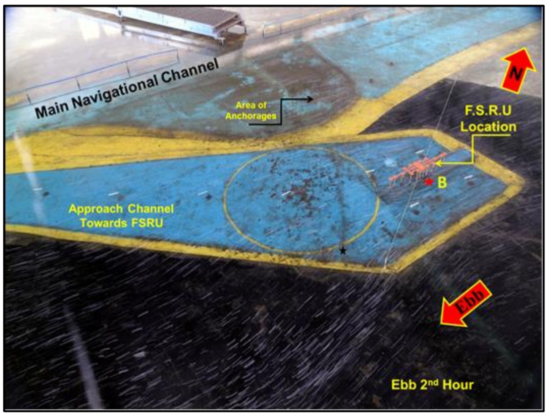


**Fig 9b. Ebb flow at FSRU in Thane creek**

Thus model studies carried out for the proposed layout of FSRU, wherein deepening to the desired depths of 11 m (below CD) in approach channel towards FSRU, turning circle and also 13.5 m (below CD) in the berth pocket reveals that along the proposed location of FSRU during flood phase of spring tide current strength varies from 0.4 - 0.43 m/s while during neap tide it varies from 0.23 - 0.25 m/s. Whereas the maximum current strength along the proposed location of FSRU during ebb phase of spring tide varies from 0.38 - 0.43 m/s while during neap tide it varies from 0.23 - 0.27 m/s. The current direction varies between 30˚- 36˚ N during flood [11] whereas 205˚ - 210˚ N during for most of the duration of ebb phase for spring/neap tide, however for short duration it deviates from exact reversal. However, it is observed that no eddies are seen to be formed at the location of proposed FSRU/LNG berthing facility Thus FSRU is proposed to be aligned at 33°N.

## IV. CONCLUSION

- The site specific measured oceanographic data is seldom available for current strength/direction, significant wave height and is essential in evolving design of marine structures. Due to various unpredictable situations during field measurement campaign, many times the measured data is either missed or lost and cannot be measured again being a costly affair and also as the same phenomenon does not repeat in the nature.

- Important valuable data missed/lost during measurement can be predicted after having known information on the available data by use of recently developed data driven method named as LSTM. This Artificial Intelligence (AI) technique is used to predict short term "Hs" and current strength/direction for relatively longer duration.

- The measured current strength/direction in macro tide dominated creek is primarily governed by the presence of tidal phenomena, the network formed with the tidal range, rate of increase/decrease of water level and the corresponding normalized current data helps the LSTM network to understand the pattern of change of current strength/direction. It helps in predicting relatively longer duration of current data with desired accuracy. The study reveals that the LSTM networks formed based on the information of 7, 15 days of current/tide, can predict current data for 8/10 days of current strength/direction with RMSE varies from 0.091- 0.09.

- The significant wave height “Hs” was also predicted and the study reveals that for the narrow range of wave dataset, the LSTM prepared on the basis of 8 days of measured data is able to predict the Hs for 33 hours with RMSE value of 0.14. The accuracy of the predicted Hs entirely depends on the randomness of waves and its pattern of repeatability which also depends on the local wind.

- The application of such data was used to calibrate the physical model of Mumbai port and finalising the alignment of FSRU proposed in Thane creek. Based on the studies carried out the alignment of FSRU was finalised and recommended as33°N along with proposed deepening of approach channel and turning circle at 11 m below CD and berth pocket at 13.5 m below CD. Also the predicted wave data (Hs) is useful in assessing its suitability in determining operable condition (tranquillity condition) at FSRU.

- The LSTM network has been found to be promising technique for carrying out calibration of physical model for evolving alignment of FSRU under various tidal conditions and also to determine operable condition toarrive at design basis for the waterfront structure.

## ACKNOWLEDGMENTS

The authors are thankful to Director, Central Water and Power Research Station, Pune (India) for his continuous encouragement for carrying out the research work. The Authors are also thankful to Dr. J.D. Agrawal, Ex. Scientist-“E” for his valuable suggestions during the course of the studies.

## REFERENCES

[1] M.C. Deo and C.S. Naidu, “Real Time Forecasting Using Neural Networks,” Ocean Eng. Vol. 26, pp. 191–203, 1998.

[2] S.B. Park, S.Y. Shin, K.H. Jung and B.G. Lee, “Prediction of Significant Wave Height in Korea Strait Using Machine Learning,” Journal of Ocean Engineering and Technology. Vol. 35, no. 5, pp. 336-346, October 2021.

[3] K.M. Zubier, “Using an Artificial Neural Network for Wave Height Forecasting in the Red Sea,” Indian J. Geo Mar. Sci. Vol. 49, pp. 184–191, 2020.

[4] Delong, L. Fan, Z. Zheqi, L. Xiaomin and L. Zewen, “Significant Wave Height Prediction based on Wavelet Graph Neural Network,” IEEE 4th International Conference on Big Data and Artificial Intelligence. 2021.

[5] Y.J. Hou, Y.L. Duan, G.X. Chen, Q.I. Peng and G.C. Song, “Statistical distribution of nonlinear random wave height in shallow water,” Journal of Xiamen University (Natural Science). vol. 53, no. 2, pp. 267–273, 2012.

[6] S.P. Nitsure, S.N. Londhe, and K.C. Khare, “Waveforecasts using wind information and genetic programming,” Ocean Engineering. vol. 54, no. 1, pp. 61–69, 2012.

[7] Peres, C. Iuppa, L. Cavallaro, A. Cancelliere and E. Foti, “Significant wave height record extension by neural networks and reanalysis wind data,” Ocean Modelling. vol. 94, no. 7, pp. 128–140, 2015.

[8] N.K. Garg, M.C. Deo and V.S. Kumar, “Forecasting coastal currents with Genetic Programming, Indian National Conference on Advances in Hydraulic Engineering,” Conference on Hydraulics, Water Resources, Coastal and Environmental Engineering Hydro 2006. pp 549-55.

[9] S. Dauji, M.C. Deo, and K. Bhargav, “Prediction of ocean currents with artificial neural networks,” ISH Journal of Hydraulic Engineering. vol. 21, no. 1, pp. 14-27, 2015.

[10] S. Dauji, M.C. Deo and B. Kapilesh, “A combined numerical and neural technique for short term prediction of ocean currents in the Indian Ocean,” Environ System. vol. 5, no. 4, 2016.

[11] Central water & power research station, “Physical hydraulic model studies for the development of proposed FSRU in Mumbai harbour,” Tech. Rep. TR-5804, March 2020.