Proc. of the 2nd International Conference on Electrical, Communication and Computer Engineering (ICECCE)
14-15 April 2020, Istanbul, Turkey

Forecasting of Water Flow in a Hydroelectric Power Plant Using LSTM Recurrent Neural Network

Arlindo Rodrigues Galvão Filho
Scientific Computing Lab
Pontifical Catholic University of Goiás
Goiânia, Brazil
https://orcid.org/0000-0003-2151-8039

Diogo Fernandes Costa Silva
Scientific Computing Lab
Pontifical Catholic University of Goiás
Goiânia, Brazil
difercosil@gmail.com

Rafael Viana de Carvalho
Scientific Computing Lab
Pontifical Catholic University of Goiás
Goiânia, Brazil
https://orcid.org/0000-0001-8656-5472

Filipe de Souza Lima Ribeiro
Operational department
Jirau Hidroeletric Power Plant
Energia Sustentável do Brasil
Porto Velho, Brazil
filipe.ribeiro@esbr.com.br

Clarimar José Coelho
Scientific Computing Lab
Pontifical Catholic University of Goiás
Goiânia, Brazil
https://orcid.org/0000-0002-5163-2986

***Abstract*—The water flow is an important information for the study of the energy production as well as management and hydropower control. The forecast of future information scenarios allows to take advanced actions in order to optimize electricity generation. This work proposes a model to generate a forecast of water flow based on a Recurrent Neural Network, more specifically the long short-term memory (LSTM) type. The dataset used to validate the LSTM model is obtained from flow history of Jirau Hydroelectric Power Plant, installed on the Madeira River in the state of Rondônia, Brazil. The model was trained and tested on one-day time-step resolution. Experimental results shows a lower error according its predictive capacity. The resulted forecast was evaluated by operating staff ex-perts, from Jirau Hydroelectric Power Plant, which attested the results can be used in real operation scenario. It is concluded that LSTM model is a good strategy for the forecast of water flow for the study of hydroelectric turbine efficiency.**

***Index Terms*—Water Flow Forecasting, Energy, Recurrent Neural Network, RNN, Long Short-Term Memory, LSTM.**

I. Introduction

Hydroelectric production is normally scheduled with respect to demand in power network at any time [1]. At hydroelectric power plant, regulating streamflow is an important strategy to optimize energy production. Electricity is supplied to system according to prompt demand [2]. In this way, it is possible to control flow of water through the spillway and turbine in order to maximize power in long run [3]. This entire system is dependent on the water flow of the river. Therefore, future predictions of this information are necessary to optimize streamflow control of reservoir and generation of electricity [4]–[8].

Large benefits in forecasting water flow is to reduce the risk in decision making [9]. Information regarding stream flow, at any given point of interest, is necessary in analysis and operation of reservoir [10], [11]. However, reservoir operation is not an universal system, each hydroelectric installation has its specific restriction. Therefore, it is necessary to understand a particular system to determine optimal reservoir operation for each hydroelectric installation [12], [13]. Moreover, a successful water management requires accurate streamflow forecasting.

Several studies have demonstrated effectiveness of neural network to forecast floods and water flow at hydroeletric installation. Pierini et al. [14] proposed predictions using autoregressive an neural network models of water flows in Colorado River, Argentina. Chang et al. [15] proposed a recurrent learning strategy to forecast a two-step-ahead real-time streamflow of Da-Chia River in Taiwan. Kumar et al. [16] compared two different architectures for Hemavathi river in India. Wu et al. [17] proposed an context-aware long-term memory (CALSTM) to predict the flow rate at the river Changhua, China. Yang et al. [18] use a Recurrent Neural Network (RNN) for inflow reservoir forecast of Chao Phraya River in Thailand. Aljahdali et al. [19] compared a feedfoward and RNN for Black and Gila rivers in USA.

Le et al. [20], proposed a long short-term memory (LSTM) RNN model for flood forecasting, where the daily discharge and rainfall were used as input data. The authors combined two different input data sets from the Hoa Binh dam, located at Da River, in Vietnam. The model resulted in one-day, two-day, and three-day flowrate forecasting ahead at Hoa Binh Station. The Nash–Sutcliffe efficiency (NSE) reached 99%, 95% e 87% corresponding to three forecasting cases, respectively. The findings of this study suggest the model is a viable option for flood forecasting on the Da River in Vietnam.

In this paper it is proposed a LSTM RNN model to forecast the streamflow per day. The method uses a decomposition techniques to preprocessing data and a LSTM RNN to forecast 30 days ahead streamflow. The data used were obtained from Jirau Hydroelectric Power Plant, installed on Madeira River in state of Rondônia, Brazil. The Jirau power plant is managed by Energia Sustentável do Brasil (ESBR) consortium.

978-1-7281-7116-6/20/$31.00 © 2020 IEEE

II. Long Short Term Memory

Long-short-term memory (LSTM) is a recurrent artificial neural network (RNN) architecture used in context of deep learning. LSTM has feedback connections to data streams such as non-segmented handwriting recognition and speech recognition. The LSTM remembers values at arbitrary intervals and is suitable for predicting time series with time intervals of unknown duration [21].

The LSTM network model correspond to memory cells composed by self-loops and three regulators of the information flow inside the cell. The self-loops are responsible to store temporal information encoded on the cell state. The three regulators: input gate (i_g), forget gate (f_g), and output gate (o_g), are responsible for writing, erasing and reading information from the cells memory state through network. The LSTM architecture of memory cell is depicted in Fig 1. One cell operation is expressed by (1) through (6).

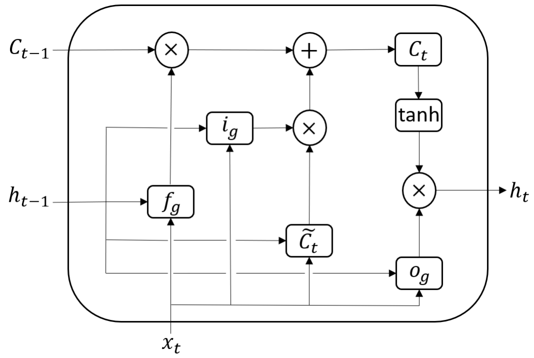

Fig. 1: LSTM cell according to [22]

$$i_g = \mathrm{sigm}(x_t V_i + h_{t-1} W_i + b_i) \quad (1)$$
$$f_g = \mathrm{sigm}(x_t V_f + h_{t-1} W_f + b_f) \quad (2)$$
$$o_g = \mathrm{sigm}(x_t V_o + h_{t-1} W_o + b_0) \quad (3)$$
$$\tilde{C}_t = \tanh(x_t V_c + h_{t-1} W_c + b_u) \quad (4)$$
$$C_t = i_g \odot \tilde{C}_t + f_g \odot C_{t-1} \quad (5)$$
$$h_t = o_g \circ \tanh(C_t) \quad (6)$$

In (1) through (6), h_t is a vector which denotes the hidden state of the cell. Likewise, C_t is the cell state and $\bar{C}_t$ is the candidate cell state at time step t which captures the important information to be persisted through to future. Meanwhile, W_i, W_o, W_f, W_c denote the weight matrices of the input gate, output gate, forget gate and the cell state, respectively. Similarly, V_i, V_o, V_f, V_c and b_i, b_o, b_f, b_c denote, respectively weight matrices and bias vectors for the current input x_t.

In LSTM, input gate (i_g) uses a sigmoid function as a switch, whose off/on state depends on the current input and previous output, as defined in (1). If the i_g is close to zero, the update signal is multiplied by zero, and the state will not be affected by update as in (5) The output gate (o_g), defined in (3), and forget gate works (f_g), defined in in (2), operates in similar way.

III. Case Study

This study uses data from Madeira River, provided by Jirau Hydroelectric Power Plant. The Madeira River is located in north of Brazil. This river has enormous hydroelectric potential, with flow rates reaching 60,000m^3/s. Due to local geography, being predominantly plain, dams built on this river have a low nominal fall, approximately 15 meters. A creative solution to take advantage of the river's hydroelectric potential was to place a large number of turbines with lower power. In case of Jirau Plant, there are 50 generating units. The Energia Sustentável of Brasil (ESBR) consortium is responsible for manage Jirau Plant and has provided dataset used for testing, model configuration and evaluation.

A. Dataset

The dataset consists of 1350 measurements of water flow history from 01/01/2016 to 12/09/2019 with one-day resolution. The data were preprocessed with transformation to a logarithmic scale and normalization. This process provide to structure data in a scale of common magnitude and stabilize variance as data has high value [22], [23]. Equation (7) correspond o transformation step,

$$z_t = \begin{cases} log(d_t), & min(d) > 0; \\ log(d_{t+1}), & min(d) = 0; \end{cases} \quad (7)$$

where d is dataset; $\min(d)$ is lower value from dataset; d_t is a sample from dataset at time t, and z_t is sample at time t after variance stabilization. Wile, (8) correspond to normalization preprocessing step,

$$i_t = \frac{(z_t - \bar{z})}{\sigma(z_t)} \quad (8)$$

where i_t is normalized sample at time t from dataset, z_t is measured sample with variance stabilized on time t; $\bar{z}$ is measured sample, and $\sigma(z_t)$ is standard deviation from samples.

Fig. 2 shows original and processed dataset. As depicted in Fig. 2(a), dataset has high magnitude ($\times 10^4$) and significant differences in variance over time horizon. The resulted preprocessing dataset is depicted in Fig. 2(b). One can notice that data contains more uniform measures, contributing to time series forecasting.

After the preprocessing steps, the overall dataset of 1350 measurements was divided in two sets. First $n = 1320$ measurements for a training set and last 30 for testing set.

B. LSTM proposed model

The training set of proposed model uses a moving window (MW) approach to sample time series dataset, as depicted in Fig. 3. The MW strategy transforms observations of entire time series into pairs of input (x_t) and output (y_t) samples of LSTM cell. For the size of MW, it is subsample three data measures for input (x_t) to generate a predicted output ($\hat{y}_t$) to be evaluated by real measurement (y_t).

Since training dataset has 1320 measurements, there will be a total of 439 MW. This configuration was empirically defined,

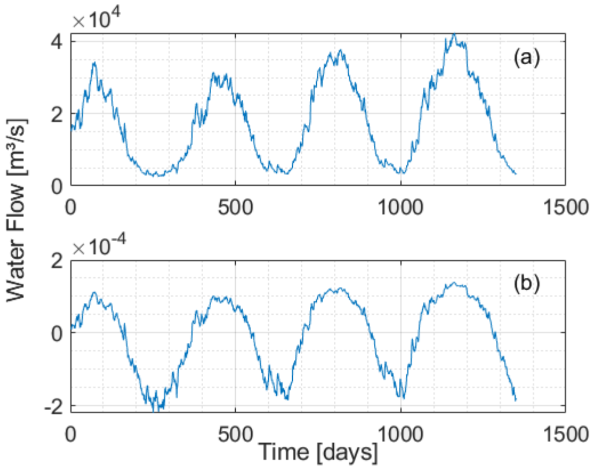

Fig. 2: Original (a) and processed (b) water flow dataset

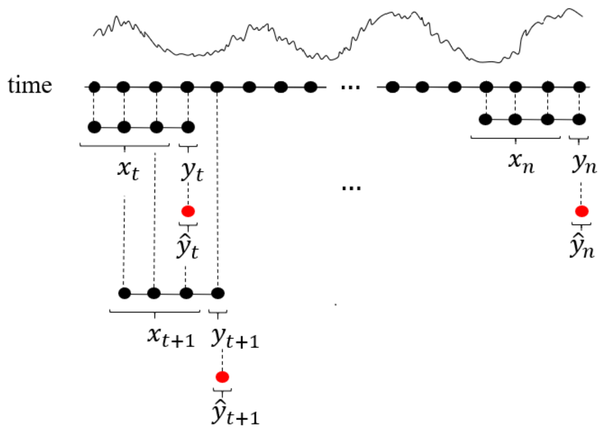

Fig. 3: Moving window of the time series dataset

in order to best fit this set of data. These frames are generated based in multi-input multi-output (MIMO) principle used in multi-step forecasting, which predicts all future observations up to intended forecasting horizon [22], [23]. To this work, a variation of MIMO model was used with multiple inputs, tree steps of temporal series x_t and single output, y_t, as shown in (9).

$$\begin{aligned} x_t &= \{i_t \quad i_{t+1} \quad i_{t+2}\} \\ y_t &= \{i_{t+3}\} \end{aligned} \tag{9}$$

For training model, it was used optimizer algorithm Adam [24], due to its better performance according to literature [22], [25]. Initially it was empirically set 250 epochs and a learning rate starting at 0.005, reducing to 0.001 after 125 epochs. This reduction together with gradient limit set to 1 was necessary in order to prevent gradient explosion.

Fig. 4 depicts a schematic representation for testing the proposed model. In first iteration (t_1), last two measures of training dataset are used to obtain two predicted outputs. The second iteration (t_2) uses last training measure as an input together with the first value predicted in previous iteration, to predict a new two outputs. From third iteration (t_3) onwards, first value predicted in first iteration will be used together with first value of second iteration to predict a new two outputs. In summary, to predict a new output, first predictions from two past iterations are used. The second prediction from the two past iterations are always discarded. This procedure was performed to entire testing dataset, resulting in a 30 days forecast.

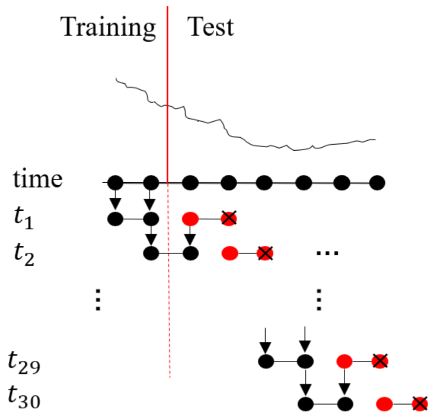

Fig. 4: Windows moving to make predict

C. *Evaluation of LSTM models*

Predictive ability of proposed LSTM model was evaluated by three different criteria. They were root mean square error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}, \tag{10}$$

Mean absolute error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i|, \tag{11}$$

and determination coefficient (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}, \tag{12}$$

where N is the amount of data, $\hat{y}_i$ predicted value, y_i measured value, $\bar{y}_i$ average of measured values.

The magnitudes of errors are aggregated with RMSE, MAE and R^2 as a single measure of predictive power. They were used to measure accuracy for forecasting errors of different LSTM units for training and testing for the water flow datasets. One value of RMSE equal to zero indicates a perfect fit for the data. However, a lower RMSE is better than a higher one. On other hand, MAE was used to measure forecast error that contributes in proportion with the absolute value of error. Absolute value is important because it doesn't allow for any form of cancellation of error values. MAE will lead to forecasts of median while RMSE will lead to forecasts of mean. R^2 is a coefficient that shows the relationship between the variance of prediction and total variance of data. The closer to one, more correlated variances are, which indicates that the model is valid for prediction.

IV. Experimental Results

The proposed model was evaluated by training and test in nine different architecture. Table I shows average results of 30 realizations on training and testing datasets. The RMSE, MAE and R^2 it present for nine different numbers of LSTM units.

TABLE I: Average results of 30 realizations on training and testing datasets for 30 days forecast

Dataset	Units	RMSE	MAE	R^2
Training	5	0.1303	0.0963	0.7656
	10	0.1260	0.0928	0.7972
	25	0.1214	0.0896	0.8022
	50	0.1203	0.0888	0.8110
	75	0.1200	0.0882	0.8078
	100	0.1223	0.0896	0.8076
	150	0.1217	0.0898	0.8084
	250	0.1208	0.0892	0.8259
	500	0.1227	0.0905	0.8196
Testing	5	0.3547	0.2945	0.2512
	10	0.2320	0.1956	0.5524
	25	0.1040	0.0858	0.8282
	50	**0.0962**	**0.0792**	**0.8519**
	75	0.0997	0.0818	0.8407
	100	0.1152	0.0946	0.7913
	150	0.2441	0.2041	0.6157
	250	0.2318	0.1914	0.5638
	500	0.2130	0.1793	0.5480

One can notice that, the best values of RMSE, MAE and R^2 for training dataset were obtained using between 50 and 75 LSTM units. By increasing or decreasing this amount, there is a worsening of such errors. For training set, the best result found for three errors simultaneously was using 50 LSTM units, as shown in Table I (bold). Therefore, 50 units were considered a good choice for having obtained the best results in the prediction on the test data, and thus greatest potential for generalization of tested scenarios is expected.

Fig. 5 and 6 shows prediction on training and testing dataset, respectively using proposed model. The results were obtained using best prediction among 30 realization for fifty LSTM units. This number of units was chosen because it provides best performance in three errors simultaneously, as shown in Table I. It can be seen that proposed model is able to produce results close to real measurements on training and testing dataset.

The forecast obtained with proposed LSTM model fits with real water flow curve. Parameters tuning during model training provided adjustments of model to produce results compatible with real dataset curve. From magnitude of RMSE, MAE and R^2 errors, it is possible to expect that forecast curve will keep its excursion within of variation limits of real water flow curve. The resulted forecast was evaluated by operating staff experts, from Jirau Hydroelectric Power Plant, which attested the accuracy of result from proposed model and its value for analysis of turbine performance, energy production as well as management and hydropower control.

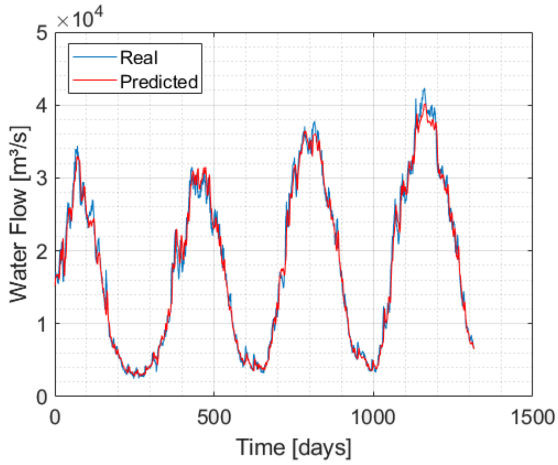

Fig. 5: Prediction results for the training dataset

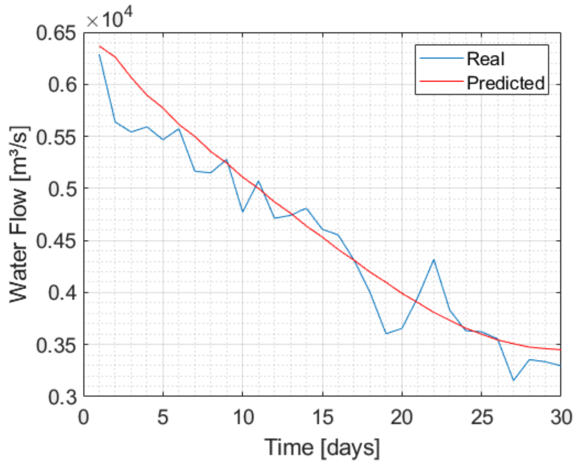

Fig. 6: Prediction results for the testing dataset

V. Conclusion

This paper presents an LSTM based neural network for building a water flow forecasting of a hydroelectric power plant. This model was trained and tested using a dataset with one-day resolution obtained from flow history of Madeira River. The predictive capacity of proposed model was tested in terms of RMSE, MAE and R^2, resulting in low errors for training and test sets. The number of LSTM units was tested in different configurations, and best results were obtained with fifty LSTM units. It can be seen that the average results obtained with the best configuration of the LSTM model here proposed can predict values similar to the real values of the Madeira River. Therefore, it can be concluded that LSTM based neural network model is a promising approach for water flow forecasting of Madeira River, and can collaborate for efficiency studies of Jirau Hidroeletric Power Plant and energy production as well as hydropower control and management. For future works, it is necessary to study other configurations of the neural network, such as the use of multiple LSTM layers and other deep learning strategies to improve the generalization of the model.

Acknowledgment

Authors thank Energia Sustentável do Brasil for their support in conducting this study "*Projeto regulamentado pela ANEEL e desenvolvido no âmbito do Programa de P&D da Energia Sustentável do Brasil S.A. (PD-06631-0007/2018)*"

References

[1] Singh, K. V., Singal, S. K., Operation of hydro power plants-a review, Renewable and Sustainable Energy Reviews, n. 69, p. 610–619, 2017.

[2] Kishor, N., Saini, R.P., Singh, S.P.: A review on hydropower plant models and control. Renew. Sustain. Energy Rev. 11, 776–796 (2007).

[3] Bakken, T. H., Killingtveit, A., Engeland, K., Alfredsen, K., Harby, A., Water consumption from hydropower plants – review of published estimates and an assessment of the concept, Hydrol. Earth Syst. Sci., v. 17, p. 3983–4000,

[4] Georgakakos, A.P., The Value of Streamflow FOrecasting in Reservoir Operation. JAWRA Journal of the American Water Resources Association, 25: 789-800.

[5] Ahmed, J. Sarma, A., Genetic Algorithm for Optimal Operating Policy of a Multipurpose Reservoir, An International Journal, Published for the European Water Resources Association, Springer;European Water Resources Association (EWRA), v. 19, n. 2, pp. 145-161, 2005.

[6] Xu, B., Zhong, P.-A., Wan, X., Zhang, W., Chen, X., Dynamic Feasible Region Genetic Algorithm for Optimal Operation of a Multi-Reservoir System, Energies, v. 5, n. 8, pp. 1-17, 2012.

[7] Li, G., Sun, Y., He, Y., Li, X., Tu, Q., Short-Term Power Generation Energy Forecasting Model for Small Hydropower Stations Using GA-SVM, Hindawi Publishing Corporation, Mathematical Problems in Engineering, v. 2014, Article ID 381387, 9 pages, http://dx.doi.org/10.1155/2014/381387, 2014.

[8] Marino, D. L., Amarasinghe, K., Manic, M., Building energy load forecasting using deep neural networks. In: IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society. IEEE, 2016. p. 7046-7051.

[9] Unes, F., Demirci, M., Taşar, B., Kaya, Y. Z., Varçin, H., Estimating Dam Reservoir Level Fluctuations Using Data-Driven Techniques, Pol. J. Environ. Stud. v. 28, n. 5, pp. 3451-3462, 2019.

[10] Khai, W. J., Alraih, M., Ahmed, A. N., Fai, C. M., El-Shafie, El-Shafie, A., Daily forecasting of dam water levels using machine learning, International Journal of Civil Engineering and Technology (IJCIET) v. 10, n. 06, pp. 314-323, 2019.

[11] Chen, C.-T., Wang, W.-C., Xu, D.-M., Chau, K. Optimizing Hydropower Reservoir Operation Using Hybrid Genetic Algorithm and Chaos," Water Resources Management: An International Journal, Published for the European Water Resources Association (EWRA), Springer; European Water Resources Association (EWRA), v. 22, n. 7, pp. 895-909, 2008.

[12] Ehsani, N., Vörösmarty, C. J., Balázs M. Fekete, Eugene Z. Stakhiv, Reservoir operations under climate change: Storage capacity options to mitigate risk, Journal of Hydrology, v. 555, p. 435–446, 2017.

[13] Altunkaynak, A., Predicting Water Level Fluctuations in Lake Van Using Hybrid Season-Neuro Approach, Journal of Hydrologic Engineering, v. 24, n. 8, 2019.

[14] Pierini, J. O., Gómez, E. A., Telesca, L., Prediction of water flows in Colorado River, Argentina. Latin American Journal of Aquatic Research, v. 40, n. 4, p. 872-880, 2012.

[15] Chang, L.-C.; Chang, F.-J.; Chiang, Y.-M,. A two-step-ahead recurrent neural network for stream-flow forecasting, Hydrological Processes, v. 18, n. 1, p. 81-92, 2004.

[16] Kumar, D.N., Raju, K.S.; Sathish, T., River flow forecasting using recurrent neural networks, Water resources management, v. 18, n. 2, p. 143-161, 2004.

[17] Wu, Y., Liu, Z., Xu, W., Feng, J, Palaiahnakote, S., Lu, T., Context-Aware Attention LSTM Network for Flood Prediction, 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, pp. 1301-1306, 2018.

[18] Yang, S., Yang, D., Chen, J., Zhao, B., Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. Journal of Hydrology, v. 579, p. 124229, 2019.

[19] Aljahdali, S., Sheta, A., Turabieh, H., River Flow Forecasting: A Comparison Between Feedforward and Layered Recurrent Neural Network. In: International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning. Springer, Cham, 2019. p. 523-532.

[20] Le, X., Ho, H.V., Lee, G, and Jung, S, Application of long short-term memory (LSTM) neural network for flood forecasting, Water, v. 11, n. 7, p. 1387, 2019.

[21] Kelleher, J. D., Deep Learning, The MIT Press, 2019.

[22] Hewamalage, H., Bergmeir, C., Bandara, K. Recurrent neural networks for time series forecasting: Current status and future directions. arXiv preprint arXiv:1909.00590, 2019.

[23] Bandara, K., Bergmeir, C., Hewamalage, H. LSTM-MSNet: Leveraging Forecasts on Sets of Related Time Series with Multiple Seasonal Patterns. arXiv preprint arXiv:1909.04293, 2019.

[24] Kingma, D. P., Ba, J., Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

[25] Kong, W., et al. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Transactions on Smart Grid, v. 10, n. 1, p. 841-851, 2019.