

Article

Water Flow Forecasting Based on River Tributaries Using Long Short-Term Memory Ensemble Model

Diogo F. Costa Silva ¹, Arlindo R. Galvão Filho ^{1,2} , Rafael V. Carvalho ^{2,*} , Filipe de Souza L. Ribeiro ³ and Clarimar J. Coelho ² 

¹ Institute of Informatics, Federal University of Goiás, Goiânia 74690-900, Brazil; difercoasil@gmail.com (D.F.C.S.); arlindogalvao@ufg.br (A.R.G.F.)

² Master's School of Production and Systems Engineering (MEPROS), Pontifical Catholic University of Goiás, Goiânia 74605-220, Brazil; clarimarc@gmail.com

³ Operational Department, Jirau Hydroelectric Power Plant, Energia Sustentável do Brasil, Porto Velho 76840-000, Brazil; filipe.ribeiro@esbr.com.br

* Correspondence: rvcarvalho@gmail.com

Abstract: Water flow forecasts are an essential information for energy production, management and hydropower control. Advanced actions to optimize electricity production can be taken based on predicted information. This work proposes an ensemble strategy using recurrent neural networks to generate a forecast of water flow at Jirau Hydroelectric Power Plant (HPP), installed on the Madeira River in Brazil. The ensemble strategy consists of combining three long short-term memory (LSTM) networks that model the Madeira River and two of its tributaries: Mamoré and Abunã rivers. The historical data from streamflow of the Madeira river and its tributaries are used to validate the ensemble LSTM model, where each time series of river tributaries are modeled separated by LSTM models and the result used as input for another LSTM model in order to forecast the streamflow of the main river. The experimental results present low errors for training and test sets for individual LSTM networks and ensemble model. In addition, these results were compared with the operational forecasts performed by Jirau HPP. The proposed model showed better accuracy in four of the five scenarios tested, which indicates a promising approach to be explored in water flow forecasting based on river tributaries.

Keywords: water flow forecasting; energy; ensemble model; long short-term memory; LSTM



Citation: Costa Silva, D.F.; Galvão Filho, A.R.; Carvalho, R.V.; de Souza L. Ribeiro, F.; Coelho, C.J. Water Flow Forecasting Based on River Tributaries Using Long Short-Term Memory Ensemble Model. *Energies* **2021**, *14*, 7707. <https://doi.org/10.3390/en14227707>

Academic Editors: Michele La Rocca and Paul Christodoulides

Received: 5 October 2021

Accepted: 10 November 2021

Published: 17 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water flow forecasts evaluate streamflow in terms of lead time. Prediction is based on probability of the streamflow and its historical records. The prediction measures can be used to understand the complexity of water resource management, in order to deal with uncertainty of climate and to support decision and management in hydroelectric power plant [1]. The streamflow forecasts are performed in short-term and long-term to flood management, water supply and in the analysis and operation of reservoirs in hydroelectric power plants [2]. Despite the climate uncertainty that influences the streamflow, traditional models of forecasts are based on statistics of stationary historical flows. The statistics of non-stationary series increases the uncertainty for investments and water resource planning. Therefore, many models have been explored in order to reduce the uncertainty in planning of water resources uses [3].

Short-term forecasting (real time) is performed continuously or after some warnings condition. Generally, short-term forecasting is provided for operational purposes when required by hydroelectric power plant and navigation. In hydroelectric power plant systems usually the planning is based in flow statistics and adjusted on monthly, weekly or daily data bases [4]. When a forecast is used for flood control and power production, an expected volume is used in planning. Flood forecast is performed during flood season,

after a flood warning in a river basin. It could be the level of the basin, rainfall or weather condition. The classification is based on the required lead time or waiting time of the basin level in relation to rainfall. Floods can be sudden, medium (basin flood) and large floods [5].

A forecast short-time of 5–10 days is ideal to increase flood response for large river basins [6]. Moreover, it can be used to regulate the streamflow of a hydroelectric power plant in a river system (basin) since it is an important strategy to optimize energy production. Information regarding streamflow is necessary in analysis and operation of reservoir. Therefore, it is very important to study the flow pattern to support decision and management of hydroelectric power plant [7].

The major benefits of river water flow forecast in the context of a hydroelectric power plant is to reduce risks in decision making, to short-term action planning in order to minimize impact of disasters and to improve energy production [8]. The reservoir operation is particular to each hydroelectric installation and it is necessary to know the characteristics of the river basins to determine proper reservoir operation [9,10].

Studies have used several models to develop short-time water flow forecasting in order to increase accuracy in prediction. Stochastic models such as autoregressive (AR) [11], and AR with moving average with exogenous inputs (ARMAX) [12] have been used for short-time flow prediction based on the time series. These models analyze time series datasets in a method that simulates water flow using classical statistics models. Nevertheless, these models have limitations to capture nonlinear characteristics of data. However, machine learning (ML) based data-driven models [13] such as fuzzy neural network (FNN) [14], support vector machine (SVM) [15], artificial neural network (ANN) [16], extreme learning machine (ELM) [17], and genetic programming (GP) [18], have proven to have the best results in modeling processes compared to the stochastic model.

Kratzert et al. [19] proposed an approach based on long short term memory (LSTM) for modeling rainfall-runoff of catchments with snow influence. The results of the LSTM model were better than reached with traditional models. Fu et al. [20] developed a model based on LSTM and a classic backpropagation neural network model for predicting water flow using historical data from a specific period of time. The results showed that performance of LSTM was superior to traditional model in different situations. Zaini, et al. [21] developed a forecast daily time series for Malaysia's rivers water level based on LSTM. The forecasting models were named $LSTM_{t-1}$, $LSTM_{t-2}$ and $LSTM_{t-3}$ and corresponding to $1-h$ ahead of time at multiple lag time which are $1-h$, $2-h$ and $3-h$ lag time. Ha and collaborators [22] propose three methods using deep neural network based on a monthly streamflow data of Yangtze river (from 1952 to 2018) to predict monthly streamflow of Yantze River in extreme flood years and small flood year. The proposed models used stacked LSTM, Conv LSTM encoder-decoder LSTM and Conv LSTM encoder-decoder gate recurrent unit. The results confirm that Conv LSTM is more stable than traditional models for prediction of Yangtze River streamflow.

Liu et al. [23] proposed a real-time rolling forecast short-term model based on LSTM to predict high uncertainty of water level in urban river in Fuzhou City, China. The results shown that LSTM is feasible method to real-time forecasting river water level. Ghimire et al. [24] combined two deep neural network to make an integrated model to predict hourly short-term at Brisbane and Teewah Creek rivers in Australia. The convolutional neural network (CNN) integrated with LSTM model were named CNN-LSTM model. The results of CNN-LSTM model were compared with standalone CNN model, LSTM models and with conventional artificial models. In all cases, prediction with CNN-LSTM shows better results than standalone models and conventional artificial models. Le et al. [25] proposed six supervised learning models to evaluate the performance of deep learning models to streamflow forecasting. The deep learning models include a feed-forward neural network, a CNN and four LSTM models. Two LSTM models with just one hidden layer and gated recurrent unit that are used in two more complex models: stacked

LSTM model and bidirectional LSTM model. According to the authors that LSTM-based models provided a better result.

In common with these studies, the authors used streamflow and rainfall data, comparing different neural networks methods presenting LSTM as a method with better accuracy. In a big river basin such as Madeira River, there are several tributaries with different flows characteristics that influence the streamflow of the main river. In this context, this paper proposes an ensemble LSTM model to forecast the streamflow of Madeira River using data only from the streamflow of two of its tributaries: Mamoré and Abunã rivers as input. Meteorological data is not considered in this model where each time series of river tributaries (Madeira and Mamoré) are modeled separated by LSTM models and the result used as input for another LSTM model in order to forecast the streamflow of the main River. The dataset used as a case study was provided by the Jirau HPP, installed on the Madeira River, in the state of Rondônia, Brazil. The Jirau power plant is managed by the Consortium Energia Sustentável do Brasil (ESBR).

Five scenarios were tested in order to compare the accuracy of the ensemble model with the statistical model used by Jirau Hydroelectric Power Plant. The tested scenarios were strict to a limited period of time in order to compare the models. The ensemble LSTM model outperformed the statistical method in four of five scenarios tested. The findings show that is possible to use ensemble LSTM models for water flow forecast on Madeira River based only on the streamflow from its tributaries. Therefore, the proposed method can contribute to the Jirau HPP to manage and plan decision and processes over data from 5 days in advance with high accuracy.

2. Materials and Methods

This section describes the related background of LSTM, the case study and the characteristics of the dataset and the description of the methodology of the proposed method.

2.1. Long Short Term Memory

Long short term memory is a recurrent artificial neural network (RNN) architecture generally applied in deep learning forecasting problems [26,27]. They are composed of LSTM cells capable of capturing long-term dependencies in sequences while attenuate gradient vanishing/exploding problem [28]. This capacity is achieved by the use of forget and update gates to modify memory cell state that allow gradients to also flow unchanged [29,30]. The LSTM memory cells are composed by self-loops that encoded temporal information in the cell states, and three regulators gates that operate the flow of information within each cell. Figure 1 presents a schematic representation of an LSTM memory cell. Self-loops are responsible for storing encoded temporal information from the past, in the state of the cell. The three gates are called: forget gate f_g , input gate i_g and output gate o_g , which operate the information flow by erasing, writing and reading, respectively. Therefore, LSTM models memorize information at different intervals and are suitable to predict time series with a certain duration interval [30,31].

The cell operation is expressed by Equations (1)–(6)

$$f_g = \text{sigm}(X_t V_f + h_{t-1} W_f + b_f) \quad (1)$$

$$i_g = \text{sigm}(X_t V_i + h_{t-1} W_i + b_i) \quad (2)$$

$$o_g = \text{sigm}(X_t V_o + h_{t-1} W_o + b_o) \quad (3)$$

$$\tilde{C}_t = \tanh(X_t V_c + h_{t-1} W_c + b_c) \quad (4)$$

$$C_t = i_g \odot \tilde{C}_t + f_g \odot C_{t-1} \quad (5)$$

$$h_t = o_g \odot \tanh(C_t) \quad (6)$$

where, h_t is a vector that represents the hidden state of cell, corresponding to short term memory. Likewise, C_t is the cell state that corresponds to long-term memory, and \tilde{C}_t is candidate for cell state in time step t , responsible to select possible important information

to be stored over time. The weight matrices of forget gate (f_g), input gate (i_g), output gate (o_g) and cell state (\tilde{C}_t) are denoted as W_f , W_i , W_o , W_c , respectively. The weight matrices and the bias for current entry X_t are denoted as V_f , V_i , V_o , V_c , and b_f , b_i , b_o , b_c , respectively.

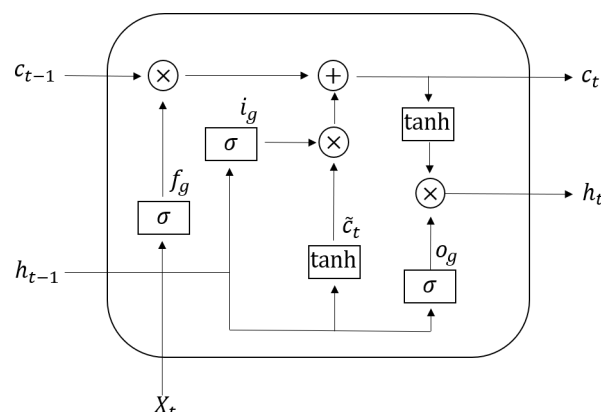


Figure 1. LSTM cell partially adapted from [32].

The forget gate uses the sigmoid activation function, generating values in the range between 0 and 1, depending on current input and previous output of LSTM cell according to Equation (1). A value of 0 in the forget gate means that all information in the state of the previous cell must be erased, and consequently will not continue to persist over time. Already a value of 1 in forget gate means that the previous information in cell state must be completely maintained. The input gate works in a similar way, in which values between 0 and 1 can control writing of new information in cell state, according to information of current input, cell candidate and previous state. Similarly, output gate can control output of information that must be read in cell state, also according to values between 0 and 1.

2.2. Case Study

For the case study it was used data from Madeira river and two of its tributaries: Mamoré and Abunã rivers, provided by Jirau HPP. The Madeira River basin is depicted in Figure 2.

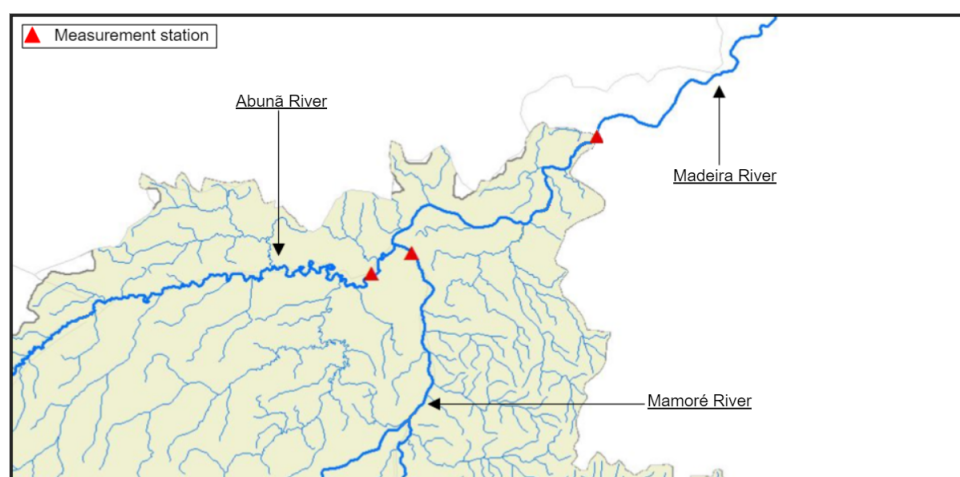


Figure 2. Madeira river basin.

The Madeira River basin is located in the north of Brazil with a big hydroelectric potential, where the flow rates of the Madeira River can reach $60,000 \text{ m}^3/\text{s}$. As the geography of the river is predominantly plain, dams built on this river have approximately 15 m of nominal fall, which can be considered low for standard dams. To take advantage of the hydroelectric potential of the river, dams at this river uses a large number of turbines with

lower power. Particularly in case of Jirau plant, there are 50 generating units in Madeira River. The Energia Sustentável do Brasil (ESBR) consortium is responsible to manage Jirau HPP and has provided dataset used for construct and evaluation of the proposed model.

2.3. Dataset

The dataset consists of three time series, containing 2069 daily measurements of flow history from 28 May 2014 to 22 January 2020. Figure 3 shows the original time series from Mamoré (a), Abunã (b) and Madeira (c) rivers.

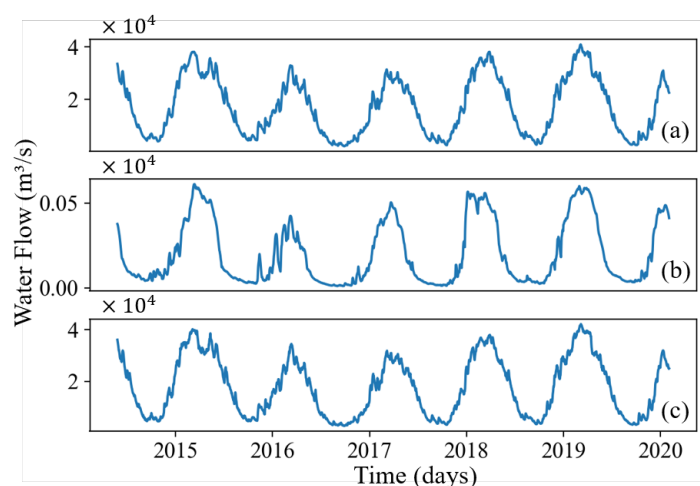


Figure 3. Original time series dataset of Mamoré (a), Abunã (b) and Madeira (c) rivers.

In order to make the data to fit the LSTM model, a preprocessing step is necessary to normalize raw data between 0 and 1 [33]. This process allows to adjust data on a common magnitude scale, providing a more effective weights adjustments for the neural networks [32]. The normalization is performed by Equation (7)

$$X_t = \frac{x_t - \min(x)}{\max(x) - \min(x)}, \quad (7)$$

where x_t represents time series sample at time step t , while X_t expresses sample at time step t after normalization step. $\min(x)$ and $\max(x)$ are lowest and highest value in time series, respectively.

The size of the training and test dataset varies according to the scenarios for each problem [34]. For this experiment, the three time series were individually normalized and divided into three sets: first 1382 measures for training set, 682 measures in sequence for validation set and last four measures for test set, except for Madeira River time series, which has last five measurements for test set. This setup was choose in order to compare with the forecast of the statistical model provided by Jirau HPP, considering five days ahead forecast.

2.4. Ensemble LSTM Model

The proposed LSTM ensemble model is divided into two stages, as depicted in Figure 4. First stage has two univariate LSTM networks, called LSTM 1 and LSTM 2, which should generate a 4-day forecast for Mamoré and Abunã rivers time series, respectively. The second stage has a multivariate LSTM network, called LSTM 3, which uses the results of first stage forecasts as an input to forecast 5-days of Madeira River time series [35].

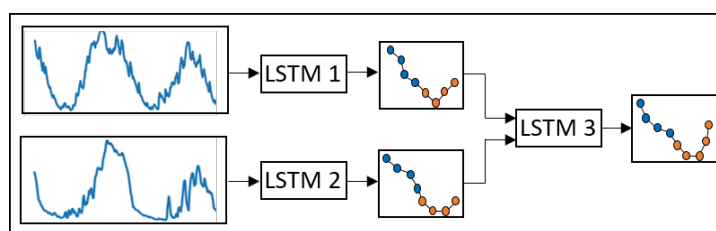


Figure 4. Schematic representation of forecast by the proposed ensemble LSTM model.

For the training set, a moving window (MW) strategy is used in order to sample the time series dataset as show in Figure 5. The MW strategy convert the entire time series observations into pairs of input (x_t) and output (y_t) samples of LSTM cell. A sample of time series X_t can be observed in time step t , with total time steps of dataset $n + 1$ and total number of MW m used. Each LSTM network has a specific MW that subsample the data measures for input empirically defined. LSTM 1 uses subsample of a single measure for i_t and two outgoing measures y_t , which must be evaluated by two real measurements o_t . For LSTM 2, the size of MW is a subsample with two sequential measurements for input i_t and one output y_t , which will be evaluated with actual measurement o_t . Furthermore, for LSTM 3 the MW consists of a subsampling of three sequential measures of each time series for input i_t , and one output y_t , which will be evaluated by one measure o_t of the time series of Madeira River. It is important to mention that LSTM 1 and LSTM 2 networks are univariate, this means that i_t and o_t samples belong to same time series. Moreover, LSTM 3 uses data from Rio Mamoré and Rio Abunã simultaneously as the two time series input.

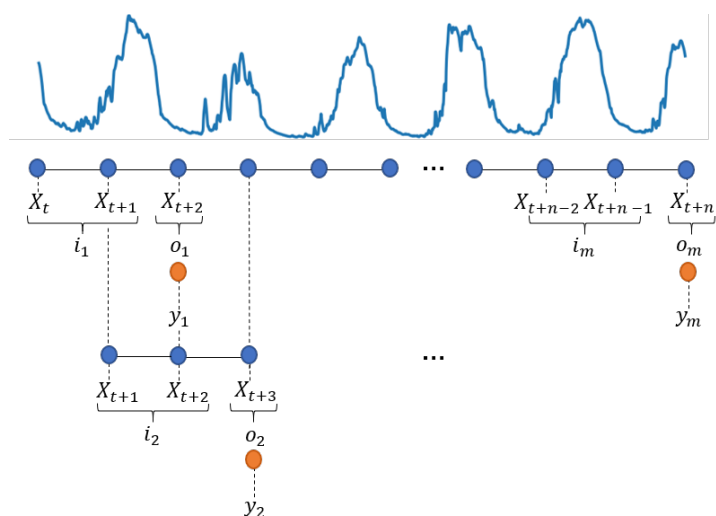


Figure 5. Example of LSTM 2 training.

The training set has 1382 samples, total of MW will result in $m = 1380$ MWs for LSTM 1 and LSTM 2 and $m = 1379$ MWs for LSTM 3. The principle of multi-input multi-output (MIMO) was used in order to generate the windows. This is a strategy in multi-step forecasting that predicts all future observations up to intended forecasting horizon [36,37]. The LSTM parameters are empirically defined according to an extensive hyperparameter optimization using grid search to set the length of the input, learning rate, number of LSTM units, number of layers and maximum training of model (number of epochs). To train the model, ADAM optimization algorithm was used with a learning rate of 0.0001. The ADAM optimizer is generally expected to perform better than other optimizers [38]. The LSTM 1 and LSTM 2 networks were configured with only one LSTM layer with 250 LSTM units. LSTM 3 also has only one layer, but it contains 100 LSTM units. These were best settings found for these datasets.

Figure 6 shows the test of the first stage of the proposed model. In the first iteration t_1 of LSTM 2, last two measurements in training dataset are used to predict an output from

next time step. The second iteration t_2 uses the last measure of training set as an input, along with the first value predicted in previous iteration in order to predict a new output. From the third iteration t_3 onwards, predictions from the last two iterations are used as an input to predict next output. This process is carried out up to 4 days. This procedure is similar to LSTM 1, using the result of the last iteration as an input to generate the output of two sequential forecasts.

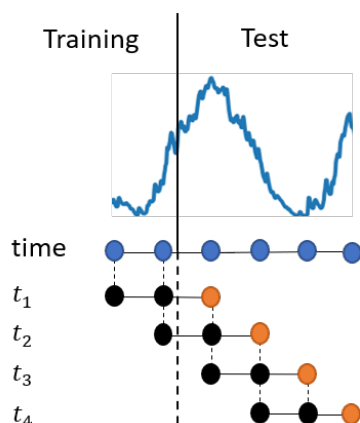


Figure 6. Example of individual LSTM test.

LSTM 3 uses the same process, but in the first iteration t_1 the last three measurements of the training dataset are used to predict an output. In the second iteration t_2 , the last two measurements of the training set are used together with LSTM 1 and LSTM 2 predictions of time step t_1 , to predict the next output of the Madeira River dataset. In the third iteration t_3 , the last measurement of training set is used, and LSTM 1 and LSTM 2 forecasts of time steps t_1 and t_2 are used to predict an output. From the fourth iteration t_4 onwards, predictions from step 1 and those prior to the time step of the current iteration are used as an input to predict an output. This process is carried out until reaching a 5-day forecast.

2.5. Evaluation Method of Ensemble LSTM Model

To evaluate the predictive ability of the proposed LSTM model, the root mean square error (RMSE) was used, as well as the mean absolute error (MAE) criteria:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2}, \quad (8)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i|, \quad (9)$$

where N correspond to the amount of data, \hat{y}_i is the predicted value and y_i is the measured value.

Currently, the Jirau HPP uses its own models to generate predictions of the flow of the Madeira River. Such models were generated in order to assist in optimization of electric energy production. In this way, the forecasts obtained by the proposed model is compared with the forecasts provided by Jirau Hydroelectric Power Plant. This comparison aims to verify, in terms of RMSE and MAE, applicability of proposed model in a real practical use scenario.

3. Results

All the experimental results were generated by using TensorFlow [39] in version 2.10, Keras 2.3.1 and Python 3.7. Table 1 shows average results of 30 realizations with the respective standard deviation and the best realization on train and test datasets. The experimental

results are presented in terms of RMSE and MAE, for Mamoré, Abunã and Madera rivers streamflow datasets with LSTM 1, LSTM 2 and LSTM 3, respectively.

Table 1. Average, standard deviation and best results of 30 realizations on train and test datasets, for four days forecast to Mamoré and Abunã rivers, and five days forecast to Mareda river.

Dataset	Time Series	RMSE		MAE	
		Mean \pm Std	Best	Mean \pm Std	Best
Train	Mamoré River	0.0074 \pm 0.0061	0.0221	0.0073 \pm 0.0063	0.0167
	Abunã River	0.0211 \pm 0.0136	0.0244	0.0207 \pm 0.0140	0.0158
	Madeira River	0.0215 \pm 0.0209	0.0208	0.0155 \pm 0.0150	0.0150
Test	Mamoré River	0.0072 \pm 0.0064	0.0065	0.0070 \pm 0.0066	0.0058
	Abunã River	0.0327 \pm 0.0021	0.0035	0.0326 \pm 0.0021	0.0030
	Madeira River	0.0034 \pm 0.0027	0.0027	0.0028 \pm 0.0024	0.0023

One can notice that RMSE and MAE results from test dataset in Table 1 were significantly lower compared to results obtained with train dataset. That occurs due to difference in size between datasets, since test dataset were fixed in a shorter time window in order to compare results obtained with forecast provided by Jirau HPP. Therefore, lowest RMSE and MAE values of test dataset occur due to evaluation of these data not having accumulated error of a long test history. Figure 7 depicts this behavior. The results obtained are close to the actual measurements, but Figure 7a,c,e have more accumulated error compared to Figure 7b,d,f.

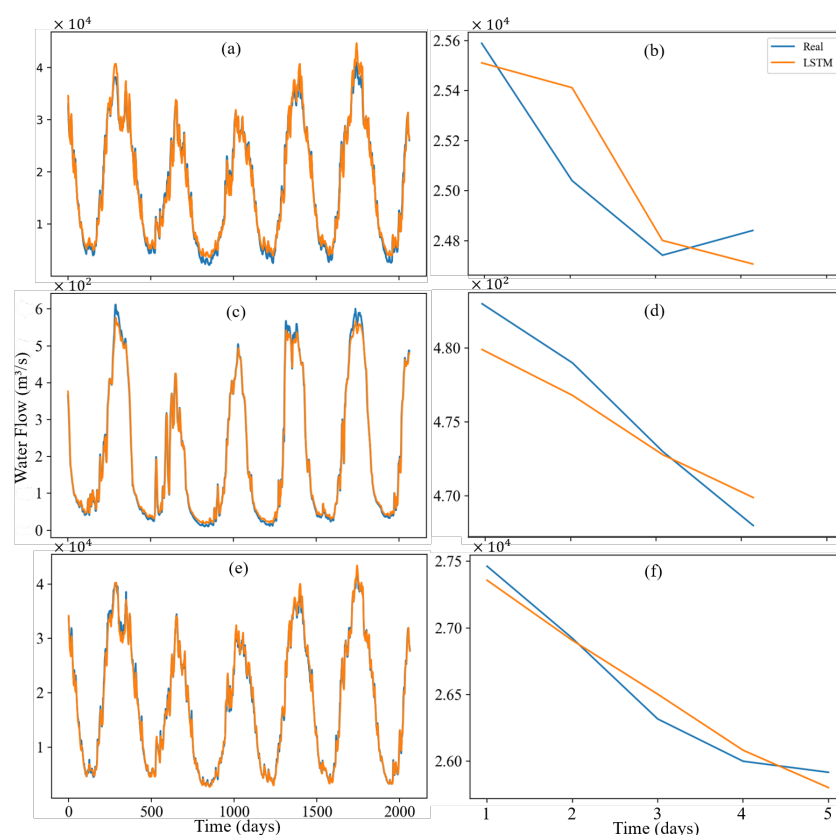


Figure 7. Prediction results of Mamoré River for (a) training and (b) testing dataset, Abunã River (c) training and (d) testing dataset and Madera River (e) training and (f) testing dataset.

Figure 7 shows the prediction results for the train and test dataset obtained using the best prediction among 30 realizations, compared to the real streamflow measurements.

Figure 7a shows the prediction results of the Mamoré River train dataset, while Figure 7b shows the forecast with its four days test dataset. Similarly, Figure 7c shows the prediction results of the Abunã River train dataset, and Figure 7d shows the forecast with its four days test dataset. Finally, Figure 7e shows the prediction results of the Madera River train dataset, and Figure 7f shows the forecast with its five days test dataset. As one can notice, the numerical results of all proposed LSTMs methods (LSTM 1, LSTM 2 and LSTM 3) are close to real the measurements of each river on train and test datasets.

In order to compare results, the forecast from the proposed LSTM ensemble model were compared to forecasts generated by the Jirau HPP. Moreover, two standard LSTM multivariate models were used in the experiments to compare the accuracy: Vanilla LSTM with 100 units and a stacked LSTM with 4 layers and 250 units in each other layer [40]. Five different scenarios were tested, providing for five days ahead on different dates in January 2020. Scenario 1 consists of forecast from 23 January 2020 to 27 January 2020, scenario 2 from 17 January 2020 to 21 January 2020, scenario 3 from 20 January 2020 to 24 January 2020, scenario 4 from 21 January 2020 to 25 January 2020 and finally scenario 5 from 22 January 2020 to 26 January 2020. The scenarios were chosen according dataset provided by Jirau HPP related to their model prediction.

Table 2 shows the numerical results in terms of RMSE and MAE for five scenarios. The smallest errors obtained are highlighted in bold. As one can notice, scenarios 1, 3, 4 and 5 present considerably lower values of both errors for forecasts obtained with the proposed ensemble model. For scenario 2, the Jirau HPP strategy was better in both errors, but with a very close margin. For the other LSTM models (Vanilla and Stacked), it can be noted that the approaches obtained considerably high values in both errors for the five scenarios tested. Numerically, errors can seem low in both forecasts, but when denormalized, the difference is significant in relation to m^3/s of water. Such an amount is relevant to the generation of electric energy in the long-term, according to the Jirau HPP.

Table 2. Results of five different scenarios comparing the forecast errors of Jirau HPP, Vanilla LSTM, Stacked LSTM, and ensemble LSTM proposed.

Model	Scenario 1		Scenario 2		Scenario 3		Scenario 4		Scenario 5	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Jirau HPP	0.0169	0.0151	0.0088	0.0083	0.0165	0.0148	0.0123	0.0096	0.0154	0.0123
Vanilla LSTM	0.0550	0.0500	0.0491	0.0476	0.0393	0.0340	0.0425	0.0356	0.0521	0.0459
Stacked LSTM	0.0296	0.0263	0.0462	0.0455	0.0162	0.0115	0.0218	0.0179	0.0270	0.0216
Ensemble LSTM	0.0029	0.0025	0.0093	0.0084	0.0083	0.0079	0.0081	0.0070	0.0073	0.0065

Figure 8 shows the forecasts of five scenarios, comparing the real value of Madeira River streamflow with the forecasts of the proposed LSTM ensemble, Jirau HPP, Vanilla LSTM and Stacked LSTM. The forecast from the proposed ensemble LSTM model fits the real measures of water flow. By tuning the parameters of the model, the training provides results compatible to the real dataset curve. Considering the RMSE and MAE magnitudes, it is expected that the forecast curve will keep its line within the variation limits of the real water flow curve.

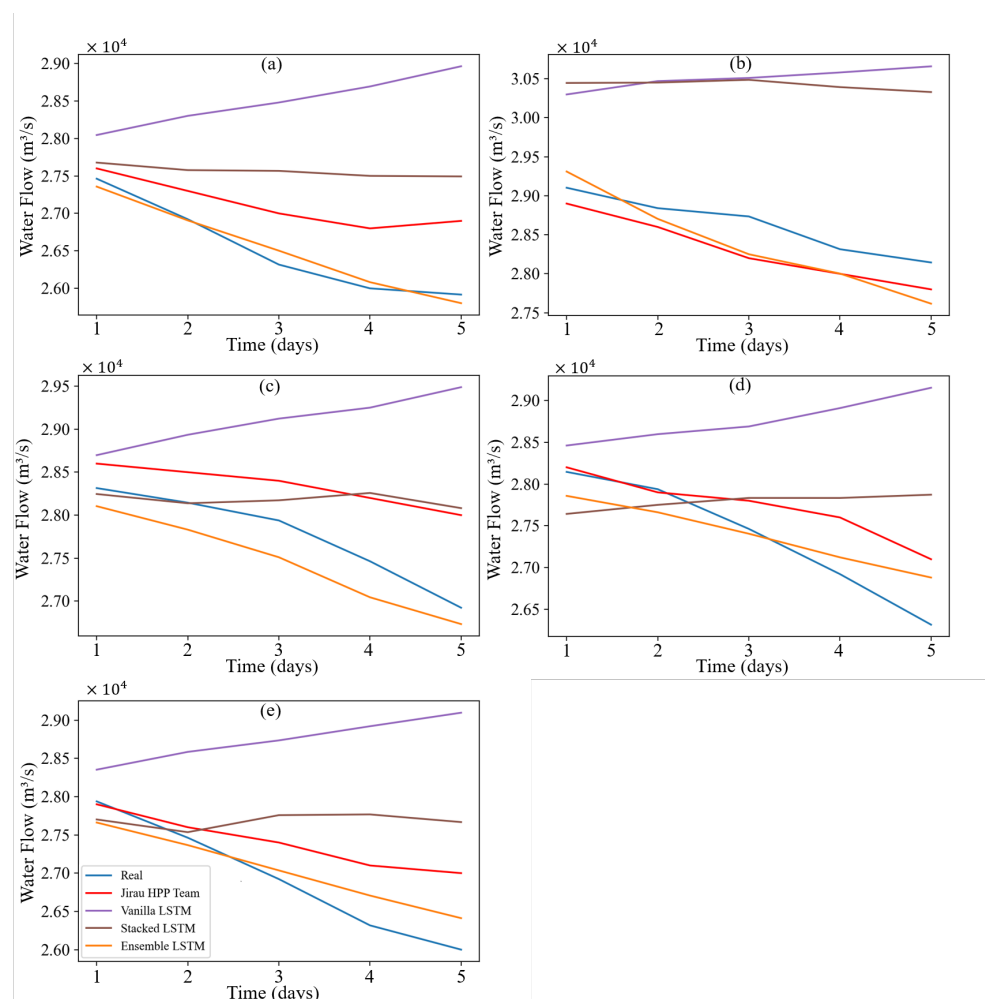


Figure 8. Forecasts of (a) scenario 1, (b) scenario 2, (c) scenario 3, (d) scenario 4 and (e) scenario 5, comparing the real value of the Madeira River streamflow, the forecasts of the Jirau HPP, Vanilla LSTM, Stacked LSTM, and proposed LSTM ensemble.

4. Discussions and Conclusions

This work presents an approach to forecast the water flow based on river tributaries using ensemble long short-term memory network. As a case study, it used data from the Madeira River and two of its tributaries: Madeira and Mamoré Rivers, located in Brazil. The dataset used for training and testing the model correspond to studies resolution obtained from the water flow history of Madeira River and its tributaries, provided by the Jirau HPP. Rainfall data was not considered in order to check the forecast accuracy only with inflow measurements. The predictive capacity of the proposed model was tested in terms of RMSE and MAE and compared to individual LSTM models (i.e., Vanilla and Stacked LSTM), and also compared to the statistical model used by the Jirau HPP to forecast water flow and HPP control.

Five different scenarios considering five-day ahead prediction was used in order to compare the accuracy of the models tested. The ensemble LSTM model resulted in better accuracy compared to the other LSTM models. Moreover, when the model is compared to the statistical model of the Jirau HPP, the ensemble LSTM model outperformed in four of five scenarios.

Despite finding similar works in the literature, there is no reference so far on water flow forecasting based only on river tributaries. Therefore, the ensemble LSTM based neural network model is a promising approach to be explored in water flow forecasting based on river tributaries. For the specific case study of the Madeira River and its tributaries, this

work can collaborate with the Jirau HPP to support decision making for the management of efficient energy production and control.

Author Contributions: Conceptualization, D.F.C.S. and C.J.C.; Data Curation, D.F.C.S. and F.d.S.L.R.; Formal Analysis, A.R.G.F.; Funding Acquisition, C.J.C.; Investigation, D.F.C.S. and A.R.G.F.; Methodology, A.R.G.F.; Project Administration, F.d.S.L.R., R.V.C. and C.J.C.; Resources, F.d.S.L.R.; Software, D.F.C.S.; Validation, A.R.G.F.; Visualization R.V.C.; Writing—Original Draft, D.F.C.S., A.R.G.F., R.V.C. and C.J.C.; Writing—Review and Editing, R.V.C. and C.J.C. All authors have read and agreed to the published version of the manuscript.

Funding: Programa de P&D da Energia Sustentável do Brasil S.A. (PD-06631-0007/2018).

Acknowledgments: 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)”.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Paiva, R.C.D.; Chaffe, P.L.B.; Anache, J.A.A.; Fontes, A.S.; Araujo, L.M.N.; Araujo, A.N.; Bartiko, D.; Bleninger, T.; Amorim, P.B.; Buarque, D.C.; et al. Advances and challenges in the water sciences in Brazil: A community synthesis of the XXIII Brazilian Water Resources Symposium. *Braz. J. Water Resour.* **2020**, *25*, e50. [\[CrossRef\]](#)
2. Raff, D.; Brekke, L.; Werner, K.; Wood, A.; White, K. Short-Term Water Management Decisions: User Needs for Improved Climate, Weather, and Hydrologic Information, Technical Report. 2013. Available online: <https://www.usbr.gov/research/st/roadmaps/WaterSupply.pdf> (accessed on 3 June 2021).
3. Loucks, D.P.; van Beek, E. Water Resources Planning and Management: An Overview. In *Water Resource Systems Planning and Management*; Springer: Cham, Switzerland, 2017. [\[CrossRef\]](#)
4. Rai, S.; De, M. Analysis of classical and machine learning based short-term and mid-term load forecasting for smart grid. *Int. J. Sustain. Energy* **2021**, 1–19. [\[CrossRef\]](#)
5. Belvederesi, C.; Dominic, J.A.; Hassan, Q.K.; Gupta, A.; Achari, G. Short-Term River Flow Forecasting Framework and Its Application in Cold Climatic Regions. *Water* **2020**, *12*, 3049. [\[CrossRef\]](#)
6. Palash, W.; Jiang, Y.; Akanda, A.S.; Small, D.L.; Nozari, A.; Islam, S. A Streamflow and Water Level Forecasting Model for the Ganges, Brahmaputra, and Meghna Rivers with Requisite Simplicity. *J. Hydrometeorol.* **2018**, *19*, 201–225. [\[CrossRef\]](#)
7. Hussain, D.; Khan, A.A. Machine learning techniques for monthly river flow forecasting of Hunza River. *Earth Sci. Inform.* **2020**, *13*, 939–949. [\[CrossRef\]](#)
8. Unes, F.; Demirci, M.; Taşar, B.; Kaya, Y.Z.; Varçin, H. Estimating Dam Reservoir Level Fluctuations Using Data-Driven Techniques. *Pol. J. Environ. Stud.* **2019**, *28*, 3451–3462. [\[CrossRef\]](#)
9. Ehsani, N.; Vörösmarty, C.J.; Fekete, B.M.; Stakhiv, E.Z. Reservoir operations under climate change: Storage capacity options to mitigate risk. *J. Hydrol.* **2017**, *555*, 435–446. [\[CrossRef\]](#)
10. Altunkaynak, A. Predicting Water Level Fluctuations in Lake Van Using Hybrid Season-Neuro Approach. *J. Hydrol. Eng.* **2019**, *24*, 04019021. [\[CrossRef\]](#)
11. Chen, Y.; Koch, T.; Lim, K.G.; Xu, X.; Zakiyeva, N. A review study of functional autoregressive models with application to energy forecasting. *Wiley Interdiscip. Rev. Comput. Stat.* **2021**, *13*, e1525. 2020. [\[CrossRef\]](#)
12. Banaś, J.; Utnik-Banaś, K. Evaluating a seasonal autoregressive moving average model with an exogenous variable for short-term timber price forecasting. *For. Policy Econ.* **2021**, *131*, 102564. [\[CrossRef\]](#)
13. Bata, M.; Carriveau, R.; Ting, D.S.K. Short-term water demand forecasting using hybrid supervised and unsupervised machine learning model. *Smart Water* **2020**, *5*, 1–18. [\[CrossRef\]](#)
14. Bakirtzis, A.G.; Bakirtzis, J.B.; Kiartzis, S.J.; Satsios, K.J. Short term load forecasting using fuzzy neural networks. *IEEE Trans. Power Syst.* **1995**, *10*, 1518–1524. [\[CrossRef\]](#)
15. Zaini, N.; Malek, M.A.; Yusoff, M.; Mardi, N.H.; Norhisham, S. Daily River Flow Forecasting with Hybrid Support Vector Machine—Particle Swarm Optimization. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *140*, 012035. [\[CrossRef\]](#)
16. Hu, Y.; Yan, L.; Hang, T.; Feng, J. Stream-Flow Forecasting of Small Rivers Based on LSTM. *arXiv* **2020**, arXiv:2001.05681v1.
17. Acikgoz, H.; Yildiz, C.; Sekkeli, M. An extreme learning machine based very short-term wind power forecasting method for complex terrain. *Energy Sources Part A Recover. Util. Environ. Eff.* **2020**, *42*, 2715–2730. [\[CrossRef\]](#)
18. Wu, Z.Y.; Yan, X. Applying Genetic Programming Approaches to Short-Term Water Demand Forecast for District Water System. Water Distribution Systems Analysis 2010. In Proceedings of the 12th International Conference, Tucson, AZ, USA, 12–15 September 2010; pp. 1498–1506. [\[CrossRef\]](#)
19. Kratzert, F.; Klotz, D.; Brenner, C.; Schulz, K.; Herrnegger, M. Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, *22*, 6005–6022. [\[CrossRef\]](#)

20. Fu, M.; Fan, T.; Ding, Z.; Salih, S.Q.; Al-Ansari, N.; Yaseen, Z.M. Deep Learning Data-Intelligence Model Based on Adjusted Forecasting Window Scale: Application in Daily Streamflow Simulation. *IEEE Access* **2020**, *8*, 32632–32651. [\[CrossRef\]](#)
21. Zaini, N.; Abdul, M.; Shuhairy, M.; Mardi, N.H. Deep Learning Neural Network for Time Series Water Level Forecasting. In *Lecture Notes in Civil Engineering, Proceedings of the International Conference on Civil, Offshore and Environmental Engineering, Kuching, Malaysia, 13–15 June 2020*; Springer: Singapore, 2020; pp. 22–29.
22. Ha, S.; Liu, D.; Mu, L. Prediction of Yangtze River streamflow based on deep learning neural network with El Niño–Southern Oscillation. *Sci. Rep.* **2021**, *11*, 11738. [\[CrossRef\]](#)
23. Liu, Y.; Wang, H.; Feng, W.; Huang, H. Short Term Real-Time Rolling Forecast of Urban River Water Levels Based on LSTM: A Case Study in Fuzhou City, China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 9287. [\[CrossRef\]](#)
24. Ghimire, S.; Yaseen, Z.M.; Farooque, A.A.; Deo, R.C.; Zhang, J.; Tao, X. Streamflow prediction using an integrated methodology based on convolutional neural network and long short-term memory networks. *Sci. Rep.* **2021**, *11*, 17497. [\[CrossRef\]](#)
25. Le, X.-H.; Nguyen, D.-H.; Jung, S.; Yeon, M. Comparison of Deep Learning Techniques for River Streamflow Forecasting. *IEEE Access* **2021**, *9*, 71805–71820. [\[CrossRef\]](#)
26. Jozefowicz, R.; Zaremba, W.; Sutskever, I. An Empirical Exploration of Recurrent Network Architectures. In Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 6–11 July 2015; Volume 37.
27. Salehinejad, H.; Sankar, S.; Barfett, J.; Colak, E.; Valaee, S. Recent Advances in Recurrent Neural Networks. *arXiv* **2018**, arXiv:1801.01078.
28. Mali, A.; Ororbia, A.; Kifer, D.; Giles, C.L. Recognizing Long Gramatical Sequences Using Recurrent Networks Augmented with an External Differentiable Stack. *arXiv* **2020**, arXiv:2004.07623v2.
29. Tsang, G.; Deng, J.; Xie, X. Recurrent Neural Networks for Financial Time-Series Modelling. In Proceedings of the International Conference on Pattern Recognition, Beijing, China, 20–24 August 2018; pp. 892–897.
30. Staudemeyer, R.C.; Morris, E.R. Understanding LSTM—A tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv* **2019**, arXiv:1909.09586.
31. Maulik, R.; Egele, R.; Lusch, B.; Balaprakash, P. Recurrent Neural Network Architecture Search for Geophysical Emulation. *arXiv* **2020**, arXiv:2004.10928.
32. Hewamalage, H.; Bergmeir, C.; Bandara, K. Recurrent neural networks for time series forecasting: Current status and future directions. *arXiv* **2019**, arXiv:1909.00590.
33. Yunpeng, L.; Di, H.; Junpeng, B.; Yong, Q. Multi-step Ahead Time Series Forecasting for Different Data Patterns Based on LSTM Recurrent Neural Network. In Proceedings of the Web Information Systems and Applications Conference, Liuzhou, China, 11–12 November 2017; pp. 305–310.
34. Le, X.; Ho, H.V.; Lee, G.; Jung, S. Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting. *Water* **2019**, *11*, 1387. [\[CrossRef\]](#)
35. Choi, J.Y.; Lee, B. Combining LSTM Network Ensemble via Adaptive Weighting for Improved Time Series Forecasting. *Math. Probl. Eng.* **2018**, *2018*, 2470171. [\[CrossRef\]](#)
36. Sahoo, B.B.; Jha, R.; Singh, A.; Kumar, D. Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. *Acta Geophys.* **2019**, *67*, 1471–1481. [\[CrossRef\]](#)
37. Bandara, K.; Bergmeir, C.; Hewamalage, H. LSTM-MSNet: Leveraging Forecasts on Sets of Related Time Series with Multiple Seasonal Patterns. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *32*, 1586–1599. [\[CrossRef\]](#)
38. Kingma, D.P.; Ba, J.L. ADAM: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
39. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* **2016**, arXiv:1603.04467.
40. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D Nonlinear Phenom.* **2020**, *404*, 132306. [\[CrossRef\]](#)