

# Short-Term Water Demand Forecasting Based on LSTM Using Multi-Input Data <sup>†</sup>

Dingtong Wang, Yanning Li, Benwei Hou <sup>\*</sup> and Shan Wu

College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100124, China; wangdington20010601@outlook.com (D.W.); liyn137@emails.bjut.edu.cn (Y.L.); wushan@bjut.edu.cn (S.W.)

<sup>\*</sup> Correspondence: benweiyou@bjut.edu.cn; Tel.: +86-13488674786

<sup>†</sup> Presented at the 3rd International Joint Conference on Water Distribution Systems Analysis & Computing and Control for the Water Industry (WDSA/CCWI 2024), Ferrara, Italy, 1–4 July 2024.

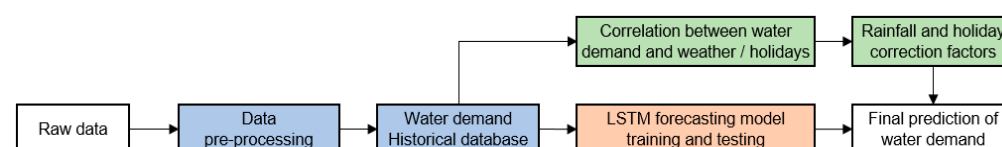
**Abstract:** This study presents a forecasting framework for the hourly water demand of district metered areas (DMAs) based on a bidirectional long short-term memory (LSTM) model which is a participant in the Battle of Water Demand Forecasting (BWDF) during the 3rd WDSA/CCWI Joint Conference. The framework consists of three portions: raw data preprocessing, initial forecasting model establishment based on LSTM, and a correction based on weather and holiday factors. The application results in the DMAs provided by the BWDF show that the proposed framework demonstrates reasonable forecasting of water demands.

**Keywords:** short-term water demand; water demand forecasting; long short-term memory; weather factor correction

## 1. Introduction

In the field of time series forecasting by deep learning models, the models based on long short-term memory (LSTM) have demonstrated a significant advantages [1,2]. In recent years, LSTM has been utilized in water demand forecasting due to its powerful performance [3,4]. This paper uses the LSTM as the initial prediction model for hourly water demand forecasting. Then, a correction model that considers the impacts of weather factors and holidays was developed based on statistical analysis.

As illustrated in Figure 1, this research encompasses three main portions: (1) raw data processing; necessary procedures were performed to complete missing data and identify and correct anomalies. Moreover, the impact of weather factors on water demand was analyzed. (2) The water demand forecasting model was developed using LSTM, considering the dual input parameters of temperature and water demand and yielding initial predictive results. (3) A correction model for water demand was established based on holiday and weather parameters and incorporating the impact of rainfall on initial water demand predictions.



**Figure 1.** The hourly water demand forecasting framework proposed in this study.

## 2. Data Analysis

### 2.1. Water Demand Features of DMAs

Figure 2a illustrates the hourly water demand data for the ten DMAs provided by the Battle of Water Demand Forecasting (BWDF) from 1 January 2021 00:00 to 5 March 2023



**Citation:** Wang, D.; Li, Y.; Hou, B.; Wu, S. Short-Term Water Demand Forecasting Based on LSTM Using Multi-Input Data. *Eng. Proc.* **2024**, *69*, 103. <https://doi.org/10.3390/engproc2024069103>

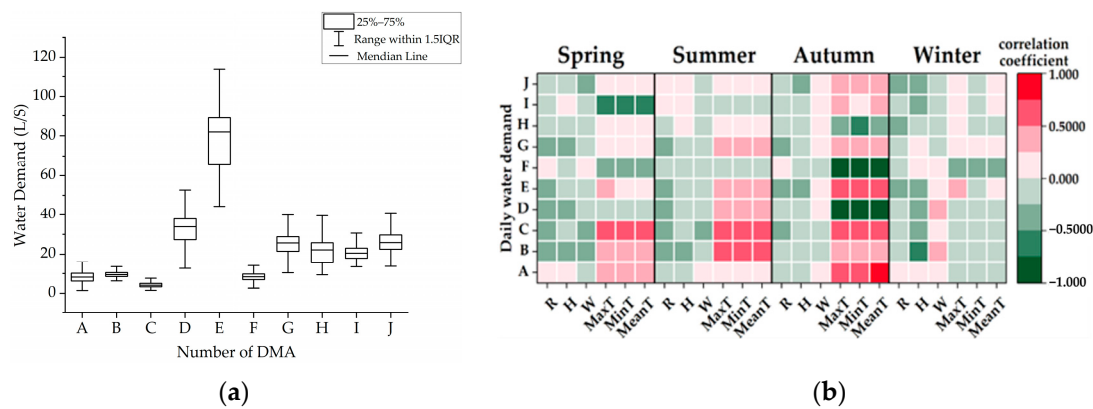
Academic Editors: Stefano Alvisi, Marco Franchini, Valentina Marsili and Filippo Mazzoni

Published: 10 September 2024



**Copyright:** © 2024 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/>).

23:00. The flow data reveal significant variability and diversity across the DMAs. Based on these observations, this study establishes an individual forecasting model for each DMA to enhance prediction accuracy.



**Figure 2.** The data features of water demand in the DMAs provided by BWDF. (a) Statistics of water demand in ten DMAs; (b) Correlation between DMA water demand and weather factors.

The historical water demand data for each DMA provided by the BWDF exhibit various levels of missing data and anomalies. This research deals with the raw data of water demand through two procedures: anomaly identification and data correction. Anomalies are categorized into three types: missing values, unusual peak values, and sudden changes. For the data anomalies, a moving average method is employed for data completion and correction.

## 2.2. Correlations between Water Demand and Weather Factors

The relationship between various weather factors and daily water demand was preliminarily assessed according to the Spearman correlation coefficients. Figure 2a depicts the correlation of weather factors with water demand across different seasons in the various DMAs. The impact of weather temperature is most significant during the spring and autumn, followed by summer, with virtually no effect in winter. Consequently, the LSTM-based forecasting model was established with two modules: one for water demand data and the other for temperature data (Figure 2b).

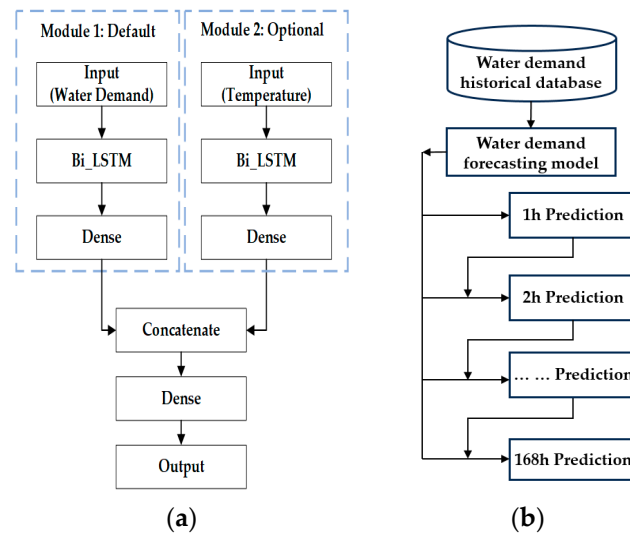
Beyond temperature, rainfall has the greatest impact on water demand, followed by humidity, while wind speed has almost no effect. Rainfall is particularly influential during the spring and summer seasons. Considering the correlation between humidity and rainfall, this study performs statistical analysis to calculate correction coefficients for water demand based on rainfall, which are used to correct the initial water demand forecasting by the LSTM model.

## 3. Water Demand Forecasting Model

### 3.1. Initial Water Demand Forecasting Based on LSTM

The long short-term memory (LSTM) model is a type of recurrent neural network (RNN) equipped with a gating mechanism [5]. Compared to traditional RNNs, LSTMs incorporate three different gate structures: the input gate, the forget gate, and the output gate. These gates effectively control the updating and discarding of data, overcoming inherent drawbacks of RNN.

Figure 3a presents the architecture of the LSTM model developed in this study. Module 1, a default component, processes water demand data, consisting of water demand data from a DMA over previous  $n$  time steps. These data are then fed into an LSTM layer to extract and learn the raw information and then into a dense layer to adjust the data to the desired output (i.e., the prediction of water demand at the next time step).



**Figure 3.** The water demand forecasting model was established by LSTM method. (a) Illustration of LSTM model with multi-inputs of water demand and temperature; (b) procedures for recursive multi-time-step prediction.

In addition to water demand data, if weather temperature significantly affects water demand for a particular DMA, an optional Module 2 can be added. This module uses temperature data corresponding to the time steps of water demand in the DMA as inputs. Then, the concatenation layer merges the data and information from Modules 1 and 2. It is important to note that the concatenation layer performs the merging operation without additional processing. The merged data are subsequently processed through another dense layer to reach the desired output.

The BWDF requires water demand forecasting of each DMA for one week with a time step of 1 h. Therefore, in this study, the recursive multi-step prediction procedure is used, and the iterative method can easily generalize the one-step-ahead prediction model to multi-step prediction [2]. As shown in Figure 3b, this method allows the extrapolation of the 1 h water demand forecasting model to 168 h.

### 3.2. Water Demand Correction Considering Weather and Holiday Factors

According to related research, rainfall impacts water demand on the rainy day ( $d$ ) and the following day ( $d + 1$ ). To ease the statistical analysis, the rainfall periods are segmented in 6 h intervals, and the water demand data are presented in 3 h intervals. Total amount of rainfall and water demand for their respective intervals are calculated, followed by the computation and fitting of water demand correction coefficients  $R_{d,t}$  and  $R_{d+1,t}$  for the corresponding rainfall periods. Statistical analysis revealed varying impacts of rainfall on DMA water demand; this study only took into consideration rainfall events with more than 5 mm and applied corrections to water demand from the start of the rainfall event until the end of the following day.

There are dramatic differences between water demand patterns on holidays and workdays. This study calculates holiday correction coefficients  $R_d$  (with  $R_d = 1$  for non-holidays) based on the water demand characteristics of the workdays adjacent to the holiday and the corresponding days of the adjacent weeks. Seasonal variations are considered to derive proportionate reductions for workdays and holidays.

Finally, the initially predicted water demand ( $\hat{Q}_{d,initial}$ ) by LSTM was adjusted by the rainfall corrector ( $R_d$ ) and holiday corrector ( $H_d$ ) to obtain the finally predicted water demand ( $\hat{Q}_{d,final}$ ), according to Equation (1):

$$\hat{Q}_{d,final} = R_d \times H_d \times \hat{Q}_{d,initial} \quad (1)$$

#### 4. Application Results

The water demand data of DMA\_D and DMA\_E during a period of rainfall from 6 June to 10 June 2022 were used to assess the effectiveness of rainfall correction. Similarly, the water demand data of DMA\_H and DMA\_J for the week of 1 November to 7 November 2021, which include holidays on Monday and Wednesday, were used to assess the effectiveness of holiday corrections. Table 1 shows the metrics provided by the BWDF, detailing the performance before and after these corrections. The corrections for rainfall and holidays mitigated their impacts on water demand to some extent, thereby enhancing the accuracy of water demand forecasts.

**Table 1.** Water demand forecasting results by the initial LSTM model and correction.

| Performance Indicator | Initial Water Demand Forecasting by LSTM |      |      |      | Water Demand Correction Considering Rainfall and Holiday |      |      |      |
|-----------------------|------------------------------------------|------|------|------|----------------------------------------------------------|------|------|------|
| DMA                   | D                                        | E    | H    | I    | D                                                        | E    | H    | I    |
| PI_1                  | 2.00                                     | 1.87 | 2.30 | 2.36 | 2.00                                                     | 1.86 | 1.37 | 0.80 |
| PI_2                  | 8.41                                     | 6.38 | 6.46 | 5.76 | 8.41                                                     | 6.38 | 5.22 | 3.45 |
| PI_3                  | 2.35                                     | 1.45 | 0.62 | 1.11 | 2.04                                                     | 1.41 | 0.55 | 0.91 |

#### 5. Conclusions

This study proposes a water demand forecasting model that incorporates LSTM-based forecasting and adjustments for holidays and weather factors, which effectively account for both historical and future trends in water demand data. A statistical analysis-based model has been developed to obtain the correction coefficients of water demand on holidays and rainy days, thereby minimizing the impact of these variables on the accuracy of water demand predictions.

**Author Contributions:** Data analysis, Y.L.; writing—original draft preparation, D.W. and Y.L.; writing—review and editing, B.H.; conceptualization and supervision, S.W.; project administration, S.W. and B.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data are available for downloading through the following link <https://wdsa-ccwi2024.it/battle-of-water-networks/> (accessed on 19 January 2024).

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

#### References

1. Han, Z.; Zhao, J.; Leung, H.; Ma, K.F.; Wang, W. A Review of Deep Learning Models for Time Series Prediction. *IEEE Sens. J.* **2021**, *21*, 7833–7848. [CrossRef]
2. Lim, B.; Zohren, S. Time-Series Forecasting with Deep Learning: A Survey. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2021**, *379*, 20200209. [CrossRef] [PubMed]
3. Zanfei, A.; Brentan, B.M.; Menapace, A.; Righetti, M. A Short-Term Water Demand Forecasting Model Using Multivariate Long Short-Term Memory with Meteorological Data. *J. Hydroinformatics* **2022**, *24*, 1053–1065. [CrossRef]
4. Mu, L.; Zheng, F.; Tao, R.; Zhang, Q.; Kapelan, Z. Hourly and Daily Urban Water Demand Predictions Using a Long Short-Term Memory Based Model. *J. Water Resour. Plan. Manag.* **2020**, *146*, 05020017. [CrossRef]
5. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.