Hourly and Daily Urban Water Demand Predictions Using a Long Short-Term Memory Based Model

Mu, Li; Zheng, Feifei; Tao, Ruoling; Zhang, Qingzhou; Kapelan, Zoran

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Hourly and daily urban water demand predictions using a long short-term memory based model

Li Mu, Feifei Zheng*, Ruoling Tao, Qingzhou Zhang and Zoran Kapelan

Li Mu: Deputy Chief Engineer, Hefei Water Supply Group Co., Ltd, muli@hfwater.cn

Feifei Zheng: Corresponding author, Professor, College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, China 310058. feifeizheng@zju.edu.cn

Ruoling Tao: Master Student, College of Civil Engineering and Architecture, Zhejiang University, taoruoling@zju.edu.cn

Qingzhou Zhang: Postdoctoral fellow, College of Civil Engineering and Architecture, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, China 310058. wdswater@gmail.com

Zoran Kapelan, Professor, Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Water Management, Stevinweg 1, 2628 CN Delft, Netherlands. z.kapelan@tudelft.nl

17   **Abstract:**

18   This case study uses a long short-term memory (LSTM) based model to predict

19   short-term urban water demands for the Hefei City of China. The performance of the

20   LSTM based model is compared with autoregressive integrated moving average

21   (ARIMA) model, the support vector regression (SVR) model and the random forests

22   (RF) model based on data with time resolutions ranging from 15-minute to 24-hour.

23   Additionally, this paper investigates the performance of the LSTM based model in

24   predicting multiple successive data points. Results show that the LSTM based model

25   can offer predictions with improved accuracy than the other models when dealing

26   with data with high time resolutions, data points with abrupt changes and data of a

27   relatively high uncertainty level. It is also observed that the LSTM based model

28   exhibit the best performance in predicting multiple successive water demands with

29   high time resolutions. In addition, the inclusion of external parameters (e.g.,

30   temperature) cannot enhance the performance of the LSTM based model, but it can

31   improve ARIMAX's prediction ability (ARIMAX is the ARIMA with variables).

32   These obtained insights based on the Hefei case study provide additional and

33   improved knowledge as well as evaluations regarding the LSTM based models used

34   for short-term urban water demand forecasting, thereby enabling their wider take-ups

35   in practical applications.

36   **Key words:** Water demand prediction; long short-term memory; data-driven models;

37   ARIMA models

38

Introduction

Urban water demand predictions are often important to the sustainable management of water supply systems for a range of purposes, including system design, maintenance and operation (Billings and Jones, 2008; Zheng et al. 2016, 2017; Qi et al., 2018). Accurate urban demand forecasts have become even more vital for many cities in recent years due to the emerged water crisis as a result of rapid urbanization and climate change, as well as driven by the need of real-time system operation (Hutton and Kapelan, 2014; Pacchin et al., 2019). This, consequently, has motived intensive studies to develop models for urban demand prediction, thereby enabling an effective water usage planning and scheduling (Pacchin et al., 2019).

A number of models are available for urban water demand forecasts with different prediction periodicity and forecast horizon (Donkor et al., 2014). More specifically, long-term forecasts usually focus on time periods more than ten years, often providing guidance for city planning and development (Levin et al., 2006). Medium-term forecasts often predict demands at a monthly or yearly resolution, and these predictions are mainly used to develop strategies for water usages (Ghiassi et al.,

55   2008). Short-term forecasts at hourly or daily resolutions are generally employed to

56   enable the effective operations of water treatment plants or pumping stations,

57   typically aimed to provide sufficient demands for urban users with the lowest

58   operation cost (Guo et al., 2018).

59   Traditionally, urban demand forecast models are generally developed based on

60   statistical methods (Howe and Linaweaver, 1967). This is because demand variations

61   are often driven by a group of factors including meteorological parameters and

62   socioeconomic elements (Arbués et al., 2003). Therefore, various linear regression

63   models are used to reveal the underlying relationships between urban water demands

64   and the external affecting parameters, thereby providing long-term demand forecasts

65   based on the projections of the external parameters (e.g., populations, Jain et al.,

66   2001). However, the accuracies of these simple linear regression models are often

67   unsatisfactory, especially in the case of predicting short-term urban water demands

68   (e.g., daily, Wong et al., 2010).

69   In recognizing the potential limitation of simple linear regression models, many

70   data-driven models have been developed to improve demand forecast accuracy

71 (Donkor et al., 2014). Autoregressive models, one type of data-driven models, have

72 been widely used in both the academic field and engineering community, in which a

73 time series analysis is often used to analyze the historical data (Chen and Boccelli,

74 2018). It has been widely demonstrated that these autoregressive models, such as

75 autoregressive integrated moving average (ARIMA) model, can exhibit better

76 performance than traditional linear regression models in predicting short-term urban

77 water demands (Chen and Boccelli, 2018).

78   In parallel to the development of the autoregression models, many other

79 data-driven models are also proposed to predict urban water demands

80 (Ghalehkhondabi et al., 2017). These include artificial neural networks (ANNs) that

81 have been broadly used for urban water demand forecasts (Ghiassi et al., 2008), the

82 support vector regression (SVR, Bai et al., 2015) model and the random forests (RF,

83 Chen et al., 2017) model that also show great merits for demand predictions. These

84 advanced data-driven models have shown improved performance than many

85 traditional prediction methods, such as autoregressive models (Villarin and

86 Rodriguez-Galiano, 2019).

87   In recent years, a type of recurrent neural networks named as the long short-term

88   memory (LSTM) based model has been emerged as an important prediction tool (Guo

89   et al., 2018). Compared to traditional ANNs, the LSTM based model is better suited

90   for time-series predictions as they possess the ability to preserve previous information

91   through learning time series data, thereby improving the accuracy of predictions

92   (Mikolov et al., 2010, Zhang et al., 2018). While the LSTM based models have been

93   broadly used in the area of artificial intelligence, such as language processing

94   (Sundermeyer et al., 2012), speech recognition (Graves and Jaitly, 2014), and image

95   captioning (Wang et al., 2016). To our best knowledge, only limited studies have been

96   undertaken so far to apply the LSTM based models to predict short-term urban water

97   demands. Guo et al. (2018) have made the first attempt to implement the LSTM

98   method for urban water demand predictions. In the study of Guo et al. (2018), the

99   performance of the LSTM based model has been compared with ARIMA and ANNs

100   based on data with 15-minute resolution, and results showed that the LTSM based

101   models exhibited better capacity than the other two methods in predicting accurate

102   water demands.

103   Given that the LSTM has only been investigated in Guo et al. (2018), there is

104   therefore a lack of sufficient case study application experience as well as

105   comprehensive understanding on its performance in dealing with short-term urban

106   water demand forecasts. These include how the LSTM based models perform (i)

107   when handling urban water demand predictions with various time resolutions as only

108   15-minute resolution data were considered in Guo et al. (2018), (ii) when predicting

109   inflection data points that have abrupt changes relative to their corresponding

110   nerbouring demand values, as well as data with a relatively high uncertainty level, (iii)

111   when comparing with other advanced data-driven models such as SVR and RF

112   models, in addition to the traditional ARIMA model considered in Guo et al. (2018),

113   and (iv) when predicting data with a 24-hour time resolution with the aid of external

114   covariates (such as temperature and rainfall). The present case study paper aims to

115   provide additional and improved knowledge as well as evaluations regarding the

116   LSTM' performance in predicting short-term urban water demands, thereby enabling

117   the wider up-takes of the LSTM based models for real-world applications.

118

Short-term urban water demand prediction models

As previously stated, the ARIMA, SVR and RF models are selected to enable the

performance comparison with the LSTM based models. The ARIMA is chosen due to

its wide applications in both the academic and industry fields, representing a standard

urban water demand prediction model (Guo et al., 2018). The SVR and RF models are

selected because they are advanced data-driven models that have shown great merits

for urban water demand forecasts (Bai et al., 2015, Chen et al., 2017), and hence it is

interested to demonstrate whether the LSTM based model (also a type of data-driven

model) can outperform the SVR and RF models or not (this comparison has not been

done in the area of the urban water demand prediction).

The long short-term memory (LSTM) based model

A recurrent neural network (RNN) model is a specific kind of artificial neural

networks (ANNs), where the network of a RNN typically has connections between

neurons and form a directed cycle (Sutskever et al., 2014). This type of structure

creates an internal self-looped cell, which allows dynamic temporal behavior. The

gradients of RNNs can be computed via Backpropagation Through Time (BPTT)

135   algorithm (Gers et al., 2000), but this method is inefficient when learning patterns

136   from long-term dependency. To solve this problem, a long-short term memory

137   (LSTM) has been developed, where it is featured by that it can bring information

138   crossing several time steps, and hence prevent early signals from fading away (Zhang

139   et al., 2018). The main structure of the LSTM network is illustrated in Figure 1 (Gers,

140   2001), stressing the importance of three gates within the algorithm structure. These

141   are input gate, forget gate and output gate, with each gate represented by a sigmoid

142   neural network layer (σ) and a multiplicative unit ($\times$). These components allow the

143   weights converge dynamically, even though the model parameters are fixed.

144   The LSTM network computes a mapping from an input sequence to an output

145   sequence by calculating network unit activations using the equations as follows (Gers

146   et al., 2000):

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{1}$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{2}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{3}$$

$$g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g) \tag{4}$$

$$s_t = g_t \otimes i_t + s_{t-1} \otimes f_t \tag{5}$$

$$h_t = \tanh(s_t) \otimes o_t \tag{6}$$

147   where $\otimes$ denotes element-wise multiplication of two vectors; t denotes the current

148   time; W_i, W_f, W_o, W_g, U_i, U_f, U_o and U_g denote the weights; b_i, b_f, b_o and b_g denotes the

149   bias; σ and tanh are the sigmoid functions; x_t is the input vector; i_t refers to the

150   input threshold; f_t is the forget threshold; o_t refers to the output threshold; g_t is the

151   candidate cell state generated by the tanh neural network layer; s_t is the cell state at

152   time t; h_t is the output vector. Specifically, the forget gate controls whether the cell

153   state of previous time is forgotten or not (Equation 2) and the input gate is responsible

154   for the input series at the current time (Equations 1). The two gates act on the

155   updating of current cell state (Equation 5) and then generate the output with the

156   output gate (Equations 3 and 6). One output h_t is the input of the recurrent procedure

157   as shown in Figure 1. Consequently, the LSTM method can prevent the gradient

158   explosion or vanishing issues during error back flow, and predict the output with

159   updated index.

160

161   *Autoregressive integrated moving average (ARIMA)*

162   The development of ARIMA model can be dated back to 1976 by Box and

163   Jenkins (1976), and this model describes data sequence using linear functions of

164   previous data and random errors. The ARIMA is featured by its great ability to

165   capture the trend, seasonality and randomness of time series (Williams, 2001).

166   Generally, an ARIMA model consists of an autoregressive (AR) model, a difference

167   process that deals with non-stationary data, and a moving average (MA) model, with

168   details presented in Hao et al., (2013).

169   *Support vector regression (SVR) models*

170   The core concept of the support vector regression (SVR) model is that it uses a

171   relatively small number of support vectors to represent the entire sample set and then

172   figures out a curve that can minimize the residual error for the data (Rasouli et al.,

173   2011). Given a set of l samples [(x_1, y_1), …, (x_l, y_l)], where x_i are the input vectors and

174   y_i are the corresponding output values (i=1, 2, …, l), a group of functions $f(x, \alpha)$ can

175   be formulated to approximate the relationship between the x_i and y_i, where α is the

176   parameter vector of the function. Generally, a nonlinear decision function of an SVR

177   model ($f(w, b)$) can be expressed as:

$$f(w, b) = w \cdot \phi(x) + b \tag{7}$$

178   where w and b are the parameter vectors of the function; x is the input vector; $\phi(x)$

179   is a nonlinear function. The objective of the SVR model is to select a function from

180   the group of $f(x, \alpha)$ that can predict the output value as accurately as possible, which is

181   obtained by the minimization of the empirical risk R_{emp} as shown below,

$$R_{emp} = \frac{1}{N} \sum_{i=1}^{N} L_{\varepsilon}(y - f(x)) \tag{8}$$

182   where L_{ε} is the loss function between the observations (y) and model predictions ($f(x)$),

183   with details given in Gunn (1998). To solve the objective function in Equation (8), a

184   standard quadratic programming algorithm with a dual set of Lagrange multipliers is

185   often adopted (Yu et al., 2006), which is

$$\min_{w, b, \xi, \xi^*} \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle x_i \cdot x_j \rangle + \varepsilon \sum_{i=1}^{l} (\alpha_i + \alpha_i^*) - \sum_{i=1}^{l} y_i(\alpha_i - \alpha_i^*) \tag{9}$$

186   with constraints

$$\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) = 0 \tag{10}$$

$$0 \leq \alpha_i, \alpha_i^* \leq C, i=1, 2, \ldots, l \tag{11}$$

187   where C is the error penalty factor; l is the length of the training data; $\langle x_i \cdot x_j \rangle$ is the

188   inner product of x_i, x_j; α_i and α_i^* are the Lagrange multipliers for the ith data point; ε is

189   the error tolerance which is specified by the users (ε=0.1 is often used). To deal with

190   nonlinear regressions, $\langle x_i \cdot x_j \rangle$ in Equation (9) is replaced by the computation of

191    $\langle \phi(x_i) \cdot \phi(x_j) \rangle$ often using a radial basis function (RBF, Yu et al., 2006) as shown

192   below,

$$\langle \phi(x_i) \cdot \phi(x_j) \rangle = e^{-\gamma|x_i - x_j|^2} \tag{12}$$

193   where γ is a user-defined parameter. In this study, the value of C and γ are determined

194   based on a grid search method as described in Cherkassky and Ma (2004).

195   ***Random forests (RF)***

196   Given an input vector X and the corresponding output Y, the random forests (RF)

197   model builds a number of q regression trees formed as $\hat{h}(X, S_n^{\theta_q})$ followed by

198   averaging the results, which can be presented as (Villarin and Rodriguez, 2019)

$$Y = \frac{1}{q} \sum_{l=1}^{q} \hat{h}(X, S_n^{\theta_l}) \tag{13}$$

199   Where S_n is the training set; n is the number of observations; the bagging method

200   selects several bootstrap samples $(S_n^{\theta_1}, ..., S_n^{\theta_q})$, and accordingly a set of trees

201 $(\hat{h}(X, S_n^{\theta_1}), \dots, \hat{h}(X, S_n^{\theta_q}))$; θ is the independent identically distributed random

202 variables representing the random selection.

203   Generally, two parameters need to be pre-specified for a RF model, that is, the

204 number of decision trees to be generated (q) and the number of selected input

205 variables m_t for each split θ. Since a RF model is often computationally efficient and

206 does not overfit, q can be set to a relatively large value (Guan et al., 2013). The

207 selection of m_t is based on the following equation (Were et al., 2015),

$$m_t = \left\lceil \sqrt{m} \right\rceil \tag{14}$$

208 where m is the total number of input variables (covariates), $[x]$ denotes the ceiling

209 function of x.

210 ***Benchmarking metrics***

211   Four metrics are considered in this study to enable the statistical analysis of the

212 model performance. These are the mean absolute percentage error (*MAPE*), the

213 Nash-Sutcliffe model efficiency (*NSE*), the coefficient of determination (R^2) and the

214 root mean square error (*RMSE*). Lower values of *MAPE* and *RMSE* indicate better fits

215 of the models, and larger values of *NSE* (the best value is 1) and R^2 (the best value is

216 1) represent better model performance These four metrics are selected due to their

217 wide applications in the area of urban water demand forecasts (Chen et al., 2017,

218 Zhang et al., 2018). The *MAPE* is defined as

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100\% \tag{15}$$

219 where Y_i represents the ith observed value, and $\hat{Y}_i$ is the ith prediction value; N is the

220 total number of data points being predicted; $\left|\frac{Y_i-\hat{Y}_i}{Y_i}\right|$ is the absolute relative error.

221   The NSE is defined as

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{16}$$

222 where $\bar{Y}$ is the mean of the observations. The R^2 is defined as

$$R^2 = \frac{\left(\sum_{i=1}^{n}(Y_i - \tilde{Y})(Y_i - \bar{Y})\right)^2}{\sum_{i=1}^{n}(Y_i - \tilde{Y})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{17}$$

223 where $\tilde{Y}$ is the mean of the predictions. The *RMSE* is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)}{n}} \tag{18}$$

224 **Case study**

225 ***Case study description***

226   The LSTM based model has been validated and its performance has been

227   compared to other three models on water demand records with a 15-minute resolution

228   in the city of Hefei, China. This city has a population of approximately eight million,

229   and the total water demands were approximately 0.59 billion m^3 per year. As shown in

230   Figure 2, a total of seven water treatment plants (WTPs) are used to supply water to

231   this city. Such a large number of WTPs induces high operational complexities for this

232   system, and hence short-term water demand forecasts are important to enable an

233   effective operation of this system, thereby saving the clean water production and

234   operational cost. More specifically, the demand predictions of the 15-min resolution

235   can greatly facilitate the real-time modelling of this water supply system, which can

236   be accordingly used to, for example, enable the leakage and energy analysis (Creaco

237   et al. 2017). The 1-hour demand predictions are often utilized to determine optimal

238   scheduling strategies for the pump stations in the WTPs, thereby reducing the

239   operation cost (Guo et al. 2018).

240   A total of 70,080 records at a 15-min resolution from May 2016 to May 2018

241   have been collected from the local water utility in the city of Hefei. These demand

242   records are the total readings from the outflow meters at the water treatment plants as

243   there are no tanks in this water supply system. Figure 3(a) shows one-week records

244   with 15-min resolution for the total demands (TD), and Figure 3(b) presents one-week

245   demands with 15-minute resolution for a district metering area (DMA) within this

246   water supply system. It is seen that the demands of this DMA are very small relative

247   to the total demands of the entire city (TD), implying that this DMA only provides

248   water for a very small population size. Consequently, the demands of this DMA are

249   significantly more variable than the total demands as visualized in Figure 3,

250   representing a dataset with a relatively high uncertainty level.

251   ***Computational experiments and model parameterizations***

252   A number of R and Python packages were used to develop the prediction models

253   applied to the case study. More specifically, the LSTM models were developed in the

254   python environment, with the aid of the functions from Keras library (Chollet, 2015).

255   R packages of "TSA", "e1071" and "randomForest" were used to develop the

256   ARIMA, SVM and RF models respectively (Chang and Lin, 2001; Breiman, 2001).

257   The inputs of the LSTM based models were determined based on a comprehensive

258   sensitivity analysis, following the method outlined in Guo et al. (2018). More

259   specifically, for the LSTM based model applied to data with 15-min and 1-hour

260   resolutions, the timeline of the inputs was divided into three fragments, the current

261   day, the previous day and the day before yesterday. In each time fragment, a certain

262   number of data points between zero and ten have been tried to identify the inputs that

263   have the best performance. For the LSTM based model applied to data with 24-hour

264   resolution, one to ten previous consecutive days were tried as the inputs. The selected

265   inputs with the best model performance were presented in Table 1. As shown in this

266   table, to predict the data with the 15-min resolution at time t of the current day (Q_t^0),

267   the inputs were the demands of previous three time steps at the current day (Q_{t-3}^0,

268   Q_{t-2}^0, Q_{t-1}^0), demands of five consecutive time steps centered at time t at the previous

269   day ($Q_{t-2}^{-1}, Q_{t-1}^{-1}, Q_t^{-1}, Q_{t+1}^{-1}, Q_{t+2}^{-1}$), and demands of five consecutive time steps centered at

270   time t at the day before yesterday ($Q_{t-2}^{-2}, Q_{t-1}^{-2}, Q_t^{-2}, Q_{t+1}^{-2}, Q_{t+2}^{-2}$). In a similar way, the

271   inputs of the 1-hour and 24-hour resolutions for the LSTM based models, as well as

272   the inputs for the SVR and RF models were outlined in Table 1. For the ARIMA

273   model with 15-minute and 1-hour resolution at time t, the inputs were their

274   corresponding previous 672 consecutive data points as presented in Table 1, and the

275   previous 56 consecutive data points with 24-hour resolution were used to predict the

276   24-hour demand at time t.

277   A sensitivity analysis was conducted to determine the appropriate architecture

278   for the LSTM model, and the number of layers was 2 with the number of nodes being

279   128 and 16 respectively, the learning rate was 0.002, tanh and ReLU were used as the

280   activation functions, the number of epochs was 100 and the batch size was 60 (Guo et

281   al., 2018). The ARIMA parameters were automatically determined after model

282   calibrations. For the SVR models, the range of the C parameters was integer numbers

283   between 1 and 10, and potential γ values were 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.15

284   and 0.20 following the approach outlined in Friedrich and Igel (2005). Finally, C=1

285   and γ=0.06 were selected using the grid search method as this parameter combination

286   exhibited the best model performance (Cherkassky and Ma, 2004). For the RF models,

287   the number of decision trees q=1000 and m_t=4 based on the method described in Guan

288   et al. (2013). It is noted that the ARIMA models needed to be re-calibrated for each

289   new set of inputs, while RNNs, SVR and RF models only calibrated once using the

290   training data set. The training dataset were records of the first 21 months and data of

291   the last three months were used for model validations.

292   **Results and Discussions**

293   *Performance comparisons of models applied to total water demands*

294   Figure 4 presents the predictions versus the observations for the four models

295   applied to the total water demands (TD) with different time resolutions. All the four

296   models were able to capture the overall trend of the observations, with errors mainly

297   produced at the extreme values of the observations. The detailed comparisons of these

298   four models are given below.

299   Boxplots in Figure 5 show the absolute relative errors of the predictions

300   generated by the four models applied to the total water demands (TD). It is noted that

301   these results were produced using the validation dataset. It is seen that the LSTM

302   based model exhibited moderately better performance than the other three models for

303   data with 15-minute and 1-hour resolutions, while the four models performed overall

304   similarly when dealing data with the 24-hour resolution. The LSTM's better

305   performance relative to its counterparts can also be supported by the statistics of the

306   prediction errors in Table 2. As shown in this table, the *MAPE* value of the LSTM

307 based models for the 15-minute and 1-hour resolution data were 1.40% and 2.56%

308 respectively, which were lower than those provided by other models. For all different

309 time resolutions, the values of *NSE* and R^2 of the LSTM based models were

310 consistently higher than the other models as shown in Table 2. For the *RMSE* values,

311 the LSTM based model also showed better performance than the other three models

312 for 15-min and 1-hour time resolutions, but it performed similarly with the ARIMA

313 for the 24-hour resolution as shown in Table 2. It is noted that the extreme values of

314 the absolute relative errors are not presented in Figure 5 for the sake of easy

315 comparisons of the overall results.

316 ***Model comparisons for predicting multiple successive data points***

317 It is practically meaningful to predict multiple successive high time resolution

318 data as these predictions can be used to facilitate the decision-making regarding the

319 operation strategies for water production and pumping. Following the method used in

320 Guo et al. (2018), the prediction at time *t* was used as the potential inputs to predict

321 water demands at time *t*+1, thereby predicting multiple successive data points (the

322 number is referred as *k*). For instance, *k*=4 indicated that four successive data points

323   were generated using the model, and the *MAPE*, *NSE*, R^2 and *RMSE* values were

324   computed based on successive data predictions relative to their corresponding

325   observations.

326   In this study, the data with the 15-minute resolution were employed for model

327   developments, aimed to predict *k*=4 (1-hour time period) and 96 (24-hour time period)

328   successive data points, with results given in Figure 6. It is seen that while all models

329   exhibited deteriorated prediction accuracy as the number of *k* increased, the LSTM

330   based model performed significantly better than the ARIMA, SVR and RF models,

331   with advantages being more noticeable for a larger value of *k*. For instance, the *MAPE*

332   values of the LSTM based model were 2.21% and 5.23% for *k*=4 and *k*=94

333   respectively as shown in Table 3, which were appreciably lower than the other three

334   models. Similar observations can be made for the *NSE*, R^2 and *RMSE* values as

335   outlined in Table 3.

336   It is observed from Figure 6 and Table 3 that the performance of the ARIMA

337   model deteriorated in a significantly quicker rate compared to the other three models

338   when the value of *k* increased. This can be also supported by the results shown in

339   Figure 7, where large deviations were observed for the ARIMA predictions relative to

340   the observations, especially for k=96. The performance variation between the LSTM

341   based models (also the SVR and RF models) and the ARIMA model in predicting

342   multiple successive data points was caused by the differences of their model

343   structures. More specifically, the inputs of the LSTM based models (also SVR and RF

344   models) were formed by some records in the current day and some data points taken

345   from previous days (see Table 1), while the inputs of the ARIMA model were many

346   successive records before the prediction time. This, consequently, leads to that a

347   larger number of inputs of the ARIMA model would be replaced by the forecasts

348   compared to the LSTM based models, SVR and RF models when predicting multiple

349   successive data points ahead, resulting in larger accumulative errors within the

350   predictions.

351

352   ***Model comparisons for data points with abrupt changes***

353   The data points with abrupt changes are often difficult to predict, and hence they

354   can be used to demonstrate the ability of the prediction models. In this study, a new

355   dataset was extracted from the original observations using the following procedures.

356   Firstly, each data point was compared with its first previous data point and first data

357   point behind in terms of relative errors, followed by the identification of inflection

358   points based on the signs of the relative errors. Secondly, these inflection data points

359   were ranked based on their mean of the absolute relative errors in a descending order,

360   and finally a new dataset was formed by the first 10% of the ranked data points.

361   Within practical applications, these data points were often referred as "abrupt points",

362   which were of great interest as many models often failed to produce accurate

363   predictions for them. In this study, the dataset with abrupt changes was respectively

364   extracted from the original 15-minute and 1-hour observations to enable the

365   prediction analysis, as shown in Table 4.

366   Interestingly, the LSTM based model exhibited significantly better performance

367   than the other three models when applied to datasets with abrupt changes as shown in

368   Table 4. This was supported by that the *MAPE* values of the LSTM based models

369   were lower than 3% for both datasets with 15-minute and 1-hour time resolutions,

370   while *MAPE* values of the other models were all around 5%. We also compared the

371   *MAPE* values of the four models used to produce multiple successive data points for

372   the dataset with abrupt changes extracted from 15-minute observations, with results

373   given in Table 4. Clearly, the LSTM based models also appreciably outperformed the

374   ARIMA, SVR and RF models, with similar observations when measured using *NSE*,

375   R^2 and *RMSE* metrics. Combining the results (Table 2 and 3) that the four models

376   applied to the full dataset, it can be deduced that the advantage of the LSTM based

377   models relative to the other three models can be more prominent when applying to

378   data with abrupt changes.

379   ***Model comparisons for data with a relatively high uncertainty level***

380   Table 5 shows the validation results measured by four statistic metrics of the four

381   models applied to the DMA demands with different time resolutions. As shown in this

382   table, the overall performances of the four models for this DMA demands were worse

383   than those from the total demands of the water supply system (see Table 2), especially

384   for the 15-min and 1-hour resolutions. This was expected as the DMA demands were

385   quite small relative to the total demands of this supply system and hence its demand

386   uncertainty was higher, resulting in challenges for the prediction models.

387   It is seen from Table 5, the LSTM based models consistently outperformed the

388   ARIMA, SVR and RF models for the dataset from the DMA demands. For instance,

389   for the LSTM applied to this dataset with 15-min resolution, $MAPE$=11.77%,

390   NSE=0.924, R^2=0.935, and $RMSE$=0.74 m^3 were achieved, which were better than

391   those from the other three models. Same observations can be made for the four

392   models applied to DMA demands with 1-hour and 24-hour time resolutions.

393   ***Model comparisons when accounting for external parameters***

394   To examine the influence of external parameters on the models' performance, a

395   range of parameters were considered as the covariates to develop the models for the

396   total water demands with the 24-hour resolution. These include daily maximum

397   temperature (T_{max}), the daily average of the temperature (T_{avg}), and the accumulative

398   daily rainfall (R_c) as these external parameters have been demonstrated to be

399   important influential factors that could affect the prediction accuracy of the models

400   (Bai et al., 2015).

401   Figure 8 presents the results of the four models with external parameters

402   considered as covariates for model calibrations and validations, where NC indicated

403   that no external parameter were used. It was observed that external parameters had

404   limited impacts on the performances of the LSTM based models, but they can slightly

405   enhance the prediction accuracy of the ARIMA, SVR and RF models, especially

406   when the daily maximum temperature (T_{max}) was used as the covariate. Similar

407   observations can be made based on *MAPE*, *NSE*, *R^2* and *RMSE* metric values.

408   **Conclusions**

409   This case study paper proposed the use of the long short-term memory (LSTM)

410   network for short-term urban water demand predictions, motivated by that the LSTM

411   networks have already been demonstrated to be an effective forecast tool in many

412   other research fields. To systematically demonstrate the performance of the LSTM

413   based models, the autoregressive integrated moving average (ARIMA) model that has

414   been widely used so far, as well as the support vector regression (SVR) model and the

415   random forest (RF) model that have shown great potentials for urban demand

416   predictions were also implemented in this study. These four models were applied to

417   urban demand predictions with different time resolutions ranging from 15-minute to

418   24-hour for the Hefei City of China. The main observations based on the case study

419   results obtained are as follows,

420   (i) The LSTM based models exhibited better performance than the ARIMA,

421   SVR and RF models in predicting data with high time resolutions (e.g., 15-minute and

422   1-hour), with merits being more significant when handling data points with abrupt

423   changes and data with a relatively high uncertainty level. When predicting data with

424   relatively low time resolutions (e.g., 24-hour), the four models performed overall

425   similarly in terms of prediction accuracy. These observations are practically

426   meaningful as they can be used to facilitate the selection of the appropriate models for

427   real-world problems based on the data properties. In addition, it was found that the

428   LSTM based model showed the significantly improved performance when predicting

429   multiple successive high time-resolution demands, with advantage being more

430   noticeable for the larger number of successive data points. Such ability is of great

431   importance as it is often very important to predict a series of successive demands with

432   a high time resolution, thereby enabling the optimal decision regarding real-time

433   operation strategies.

434   (ii) External parameters such as temperature and rainfall had limited impacts on

435 the performance of the LSTM based models in predicting data with 24-hour

436 resolution, indicating that the performance of the LSTM based model was dominated

437 by its great ability in capturing the underlying relationships within the data

438 themselves. This is also a great merit of the LSTM based models for practical

439 applications as collecting external parameters in a high time resolution is often

440 time-consuming and costly.

441   The observations mentioned above based on the Hefei Case study provide

442 important additional experiences and evaluations regarding the applications of the

443 LSTM based models for short-term urban demand forecasts. These knowledge go

444 beyond the findings reported in Guo et al (2018) as in their study only data with

445 15-min resolution were considered (no covariates), as well as that the LSTM based

446 models were only compared with ARIMA and ANN models. In addition, this study

447 demonstrated that the LSTM based models can exhibit significantly better

448 performance than other models in predicting data points with abrupt changes as well

449 as data with a high uncertainty level, which have not been considered in Guo et al.

450 (2018).

Data Availability Statement

All data, models, or code generated or used during the study are available from the

corresponding author by request (feifeizheng@zju.edu.cn).

References

Adamowski, J. (2008). Peak Daily Water Demand Forecast Modeling Using Artificial Neural Networks. *Journal of Water Resources Planning and Management, 134*(2), 119-128. doi:10.1061//ASCE/0733-9496/2008/134:2/119.

Arandia, E., Ba, A., Eck, B., & McKenna, S. (2016). Tailoring Seasonal Time Series Models to Forecast Short-Term Water Demand. *Journal of Water Resources Planning and Management, 142*(3), 04015067. doi:10.1061/(asce)wr.1943-5452.0000591.

Bai, Y., Li, C., & Xie, J. (2015). Dynamic Forecast of Daily Urban Water Consumption Using a Variable-Structure Support Vector Regression Model. *Journal of Water Resources Planning and Management, 141*(3). doi:0.1061/(ASCE)WR.1943-5452.000045710.1061/(ASCE).

472   Bakker, M., Vreeburg, J. H. G., van Schagen, K. M., & Rietveld, L. C. (2013). A fully
473   adaptive forecasting model for short-term drinking water demand. *Environmental*
474   *Modelling Software, 48*(5), 141-151.

475   Bhanja, S. N. , Mukherjee, A. , Saha, D. , Velicogna, I. , & Famiglietti, J. S. . (2016).
476   Validation of grace based groundwater storage anomaly using in-situ groundwater
477   level measurements in india. *Journal of Hydrology, 543*, 729-738.

478   Billings, R. B., & Jones, C. V. (2008). *Forecasting Urban Water Demand* (2nd ed.). Denvor,
479   CO: American Water Works Association.

480   Bougadis, J., Adamowski, K., & Diduch, R. (2005). Short-term municipal water demand
481   forecasting. *Hydrological Processes, 19*(1), 137-148. doi:10.1002/hyp.5763.

482   Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-23.

483   Chang, C. C., & Lin, C. J. (2001). *Libsvm: a library for support vector machines*.

484   Chen, G., & Boccelli, D. L. (2018). Forecasting Hourly Water Demands With Seasonal
485   Autoregressive Models for Real-Time Application. *Water Resources Research, 54*(2),
486   879-894.

487   Chen, G., Long, T., Xiong, J., & Bai, Y. (2017). Multiple Random Forests Modelling for
488   Urban Water Consumption Forecasting. *Water Resources Management, 31*(15),
489   4715-4729.

490   Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise
491   estimation for SVM regression. *Neural Networks, 17*(1), 113-126.

492   Chollet, F. (2015). Keras. Retrieved from https://github.com/fchollet/keras.

493   Creaco, E., Pezzinga, G., and Savic, D. (2017). "On the choice of the demand and hydraulic modeling
494   approach to WDN real-time simulation." *Water Resources Research*, 53(7), 6159-6177.

495   Donkor, E. A., Mazzuchi, T. A., Soyer, R., & Roberson, J. A. (2014). Urban Water Demand
496   Forecasting- Review of Methods and Models. *Journal of Water Resources Planning*
497   *and Management, 140*(2), 146-156. doi:0.1061/(ASCE)WR.1943-5452.

498   Gers, F. A. (2001). *Long Short-Term Memory in Recurrent Neural Networks.* (Doctor),

499   University of Hannover.

500   Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: continual prediction
501   with LSTM. *Neural Computation, 12*, 2451-2471.

502   Ghalehkhondabi, I., Ardjmand, E., Young, W. A., 2nd, & Weckman, G. R. (2017). Water
503   demand forecasting: review of soft computing methods. *Environ Monit Assess, 189*(7),
504   313. doi:10.1007/s10661-017-6030-3.

505   Ghiassi, M., Zimbra, D. K., & Saidane, H. (2008). Urban Water Demand Forecasting with a
506   Dynamic Artificial Neural Network Model. *Journal of Water Resources Planning and*
507   *Management, 134*(2), 138-146. doi:10.1061/(asce)0733-9496(2008)134:2(138).

508   Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z., & Yang, X. (2013). Integration of
509   orthoimagery and lidar data for object-based urban thematic mapping using random
510   forests. *International Journal of Remote Sensing, 34*(14), 5166-5186.

511   Guo, G., Liu, S., Wu, Y., Li, J., Zhou, R., & Zhu, X. (2018). Short-Term Water Demand
512   Forecast Based on Deep Learning Method. *Journal of Water Resources Planning and*
513   *Management, 144*(12), 04018076. doi:10.1061/(asce)wr.1943-5452.0000992.

514   Hao, Y., Wu, J., Sun, Q., Zhu, Y., Liu, Y., Li, Z., & Yeh, T. J. (2013). Simulating effect of
515   anthropogenic activities and climate variation on Liulin Springs discharge depletion
516   by using the ARIMAX model. *Hydrological Processes, 27*, 2605-2613.

517   Howe, C. W., & Linaweaver, F. P. (1967). The Impact Price on Residential Water Demand
518   and Its Relation to System Design and Price Structure. *Water Resources Research,*
519   *3*(1), 13-32.

520   Hyndman, R. J., & Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast
521   Package for R. *Journal of Statistical Software, 26*(3).

522   Jain, A., Varshney, A., & Joshi, U. (2001). Short-term water demand forecast modelling at
523   IIT Kanpur using artificial neural networks. *Water Resources Management, 15*(5),
524   299-321.

525   Kwiatkowski, D., Philips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null
526   hypothesis of stationarity against the alternative of a unit root. *Journal of*

527 *Econometrics, 54*, 159-178.

528 Levin, E., Maddaus, W., Sandkulla, N., & Pohl, H. (2006). Forecasting wholesale demand
529 and conservation savings. *Journal of American Water Works Association, 98*(2),
530 102-111.

531 Mikolov, T., Karafiat, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). *Recurrent*
532 *neural network based language model*. In: Proceedings of the 11th Annual
533 Conference of the International Speech Communication Association, Makuhari,
534 Chiba, Japan, pp. 1045–1048.

535 Pacchin, E., Gagliardi, F., Alvisi, S., & Franchini, M. (2019). A Comparison of Short-Term
536 Water Demand Forecasting Models. *Water Resources Management, 33*(4), 1481-1497.
537 doi:10.1007/s11269-019-02213-y

538 Prosdocimi, I., Kjeldsen, T. R., & Svensson, C. (2013). Non-stationarity in annual and
539 seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth*
540 *System Sciences, 1*(5), 5499-5544.

541 Qi, Z., Zheng, F., Guo, D., Zhang, T., Shao, Y., Yu, T. Maier, H. R. (2018). A comprehensive
542 framwork to evaluate hydraulic and water quality impacts of pipe breaks on water
543 distribution systems. *Water Resources Research, 54*(10), 8174-8195.

544 Rasouli, K., Hsieh, W. W., & Cannon, A. J. (2011). Daily streamflow forecasting by machine
545 learning methods with weatherand climate inputs. *Journal of Hydrology, 414-415*,
546 284-293.

547 Romano, M., & Kapelan, Z. (2014). Adaptive water demand forecasting for near real-time
548 management of smart water distribution systems. *Environmental Modelling Software,*
549 *60*(7), 265-276.

550 Sardinha-Lourenço, A., Andrade-Campos, A., Antunes, A., & Oliveira, M. S. (2018).
551 Increased performance in the short-term water demand forecasting through the use of
552 a parallel adaptive weighting strategy. *Journal of Hydrology, 558*, 392-404.
553 doi:10.1016/j.jhydrol.2018.01.047

554 Sundermeyer, M., Schluter, R., & Ney, H. (2012). *LSTM neural networks for language*

modeling. In: Proceedings of the 12th Annual Conference of the International Speech Communication Association, Portland, Oregon, USA, pp. 601–608.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). *Sequence to sequence learning with neural networks*. In: Proceedings of the 28th Conference on Neural Information Processing Systems, Montreal, Canada, pp. 3104–3112.[SEP]

Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer.

Villarin, M. C., & Rodriguez-Galiano, V. F. (2019). Machine Learning for Modeling Water Demand. *Journal of Water Resources Planning and Management, 145*(5). doi:0.1061/(ASCE)WR.1943-5452.0001067

Were, K., Bui, D. T., Dick, O. B., & Singh, B. R. (2015). A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators, 52*, 394-403.

Williams, B. M. (2001). Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling. *Transportation Research Record Journal, 1776*(1), 194-200.

Wong, J. S., Zhang, Q., & Chen, G. (2010). Statistical modeling of daily urban water consumption in Hong Kong- Trend, changing patterns, and forecast. *Water Resources Research, 46*. doi:0.1029/2009WR008147

Zhang, D., Lindholm, G., & Ratnaweera, H (2018). Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. Journal of Hydrology, 556, 409-418.

Zhang, J., Zhu, Y., Zhang, X., Ye, M., & Yang, J. (2018). Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. *Journal of Hydrology, 561*, 918-929.

Zheng, F., Zecchin, A., Maier, H., and Simpson, A. (2016). "Comparison of the Searching Behavior of NSGA-II, SAMODE, and Borg MOEAs Applied to Water Distribution System Design Problems." *Journal of Water Resources Planning and Management*, 142(7), 04016017.

583   Zheng, F., Zecchin, A., Newman, J., Maier, H., and Dandy, G. (2017). "An Adaptive
584   Convergence-Trajectory Controlled Ant Colony Optimization Algorithm with
585   Application to Water Distribution System Design Problems." *IEEE Transactions on*
586   *Evolutionary Computation*, 21(5), 773-791.

587   Zhou, S. L., McMahon, T. A., Walton, A., & Lewis, J. (2002). Forecasting operational
588   demand for an urban water supply zone. *Journal of Hydrology, 259*, 189-202.

589

590

Table 1 Inputs of the four models

Mode types	Time resolutions	Inputs and outputs
LSTM	t=15-minute	$Q_t^0 = f(Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_{t-2}^1, Q_{t-1}^1, Q_t^1, Q_{t+1}^1, Q_{t+2}^1, Q_{t-2}^2, Q_{t-1}^2, Q_t^2, Q_{t+1}^2, Q_{t+2}^2)$
	t=1-hour	$Q_t^0 = f(Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_{t-1}^1, Q_t^1, Q_{t+1}^1, Q_{t-1}^2, Q_t^2, Q_{t+1}^2)$
	t=24-hour	$Q_t^0 = f(Q_t^{-1}, Q_t^{-2}, Q_t^{-3})$
ARIMA	t=15-minute	$Q_t = f(Q_{t-1}, Q_{t-2}, ..., Q_{t-671}, Q_{t-672})$
	t=1-hour	$Q_t = f(Q_{t-1}, Q_{t-2}, ..., Q_{t-671}, Q_{t-672})$
	t=24-hour	$Q_t = f(Q_{t-1}, Q_{t-2}, ..., Q_{t-55}, Q_{t-56})$
SVR	t=15-minute	$Q_t^0 = f(Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_{t-2}^1, Q_{t-1}^1, Q_t^1, Q_{t+1}^1, Q_{t+2}^1, Q_{t-2}^2, Q_{t-1}^2, Q_t^2, Q_{t+1}^2, Q_{t+2}^2)$
	t=1-hour	$Q_t^0 = f(Q_{t-5}^0, Q_{t-4}^0, Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_t^{-1}, Q_t^{-2})$
	t=24-hour	$Q_t^0 = f(Q_t^{-1}, Q_t^{-2}, Q_t^{-3}, Q_t^{-4}, Q_t^{-5})$
RF	t=15-minute	$Q_t^0 = f(Q_{t-5}^0, Q_{t-4}^0, Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_t^{-1}, Q_t^{-2})$
	t=1-hour	$Q_t^0 = f(Q_{t-5}^0, Q_{t-4}^0, Q_{t-3}^0, Q_{t-2}^0, Q_{t-1}^0, Q_t^{-1}, Q_t^{-2})$
	t=24-hour	$Q_t^0 = f(Q_t^{-1}, Q_t^{-2}, Q_t^{-3})$

591

592

593 **Table 2 Statistics of the model prediction errors for the total water demands**

Time resolutions	Models	*MAPE*	*NSE*	R^2	*RMSE* (m^3)
15-minute	*LSTM*	*1.40%*	*0.991*	*0.991*	*315*
	ARIMA	2.14%	0.974	0.975	551
	SVR	2.01%	0.985	0.986	421
	RF	2.03%	0.984	0.984	425
1-Hour	*LSTM*	*2.56%*	*0.978*	*0.981*	*1976*
	ARIMA	4.26%	0.937	0.937	3367
	SVR	3.40%	0.963	0.966	2587
	RF	3.70%	0.945	0.945	3153
24-Hour	*LSTM*	*2.89%*	*0.820*	*0.822*	*55,605*
	ARIMA	2.94%	0.811	0.821	55,463
	SVR	3.82%	0.680	0.769	74,181
	RF	3.08%	0.816	0.821	56,179

594

595

596   **Table 3 Statistics of prediction errors for models used for multiple successive**

597   **data forecasts**

No. of successive predictions (k)	Models	*MAPE*	*NSE*	R^2	*RMSE* (m³)
k=4	**LSTM**	**2.21%**	**0.980**	**0.981**	**475**
	ARIMA	3.19%	0.954	0.954	728
	SVR	3.05%	0.970	0.973	591
	RF	3.11%	0.959	0.959	685
k=96	**LSTM**	**5.23%**	0.899	0.909	1075
	ARIMA	16.28%	0.206	0.348	3018
	SVR	7.41%	0.832	0.836	1390
	RF	8.19%	0.751	0.754	1692

598

599

600 **Table 4 Statistics of model prediction errors for data with abrupt changes**

Time resolutions	Models	*MAPE*	*NSE*	R^2	*RMSE* (m^3)
15-minute	**LSTM**	**2.96%**	**0.961**	**0.962**	**596**
	ARIMA	5.58%	0.897	0.909	967
	SVR	4.56%	0.939	0.940	744
	RF	5.49%	0.916	0.916	873
1-Hour	**LSTM**	**2.89%**	**0.979**	**0.982**	**2111**
	ARIMA	5.75%	0.913	0.983	4307
	SVR	4.94%	0.956	0.974	3057
	RF	6.95%	0.884	0.973	4973
k=4*	**LSTM**	**3.56%**	**0.962**	**0.963**	**588**
	ARIMA	5.33%	0.929	0.936	803
	SVR	4.69%	0.933	0.938	780
	RF	4.76%	0.920	0.923	853
k=96*	**LSTM**	7.19%	0.821	0.862	1274
	ARIMA	15.69%	0.315	0.368	2492
	SVR	9.57%	0.688	0.731	1681
	RF	9.36%	0.678	0.732	1708

601 *k*=4 and 96 represents 4 and 96 successive predictions with 15-min resolution.

602

Table 5 Statistics of prediction errors for models used for data with a relatively uncertainty level

Time resolutions	Models	*MAPE*	*NSE*	*R^2*	*RMSE* (m³)
15-minute	**LSTM**	**11.77%**	**0.924**	**0.935**	**0.74**
	ARIMA	19.94%	0.843	0.843	0.94
	SVR	17.78%	0.856	0.861	0.90
	RF	18.95%	0.856	0.856	0.90
1-hour	**LSTM**	**10.29%**	**0.942**	**0.942**	**2.18**
	ARIMA	19.14 %	0.860	0.859	3.39
	SVR	14.59 %	0.898	0.905	2.92
	RF	13.90%	0.899	0.900	2.86
24-hour	**LSTM**	**1.36%**	**0.878**	**0.895**	**11.23**
	ARIMA	1.86%	0.811	0.852	13.99
	SVR	7.66%	-1.704	0.280	52.92
	RF	2.64%	0.425	0.642	24.39

Figure 1: The structure of a long-short term memory (LSTM) network, where the dotted lines represent the recurrent procedure
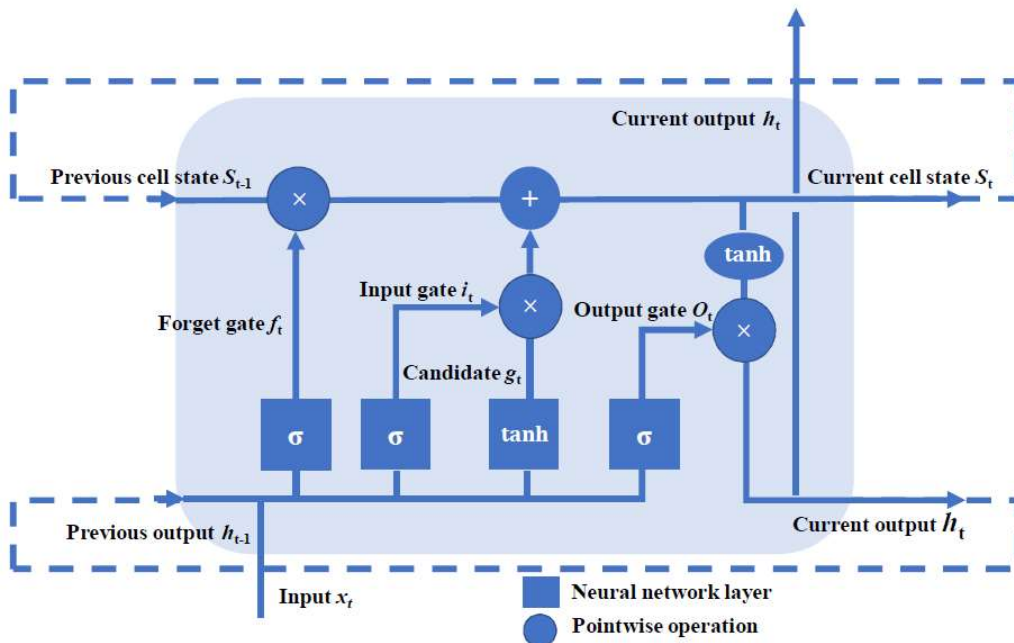
Figure 2: Water treatment plants (WTPs) distributed in the city of Hefei, China, with green liens representing the water distribution pipelines.
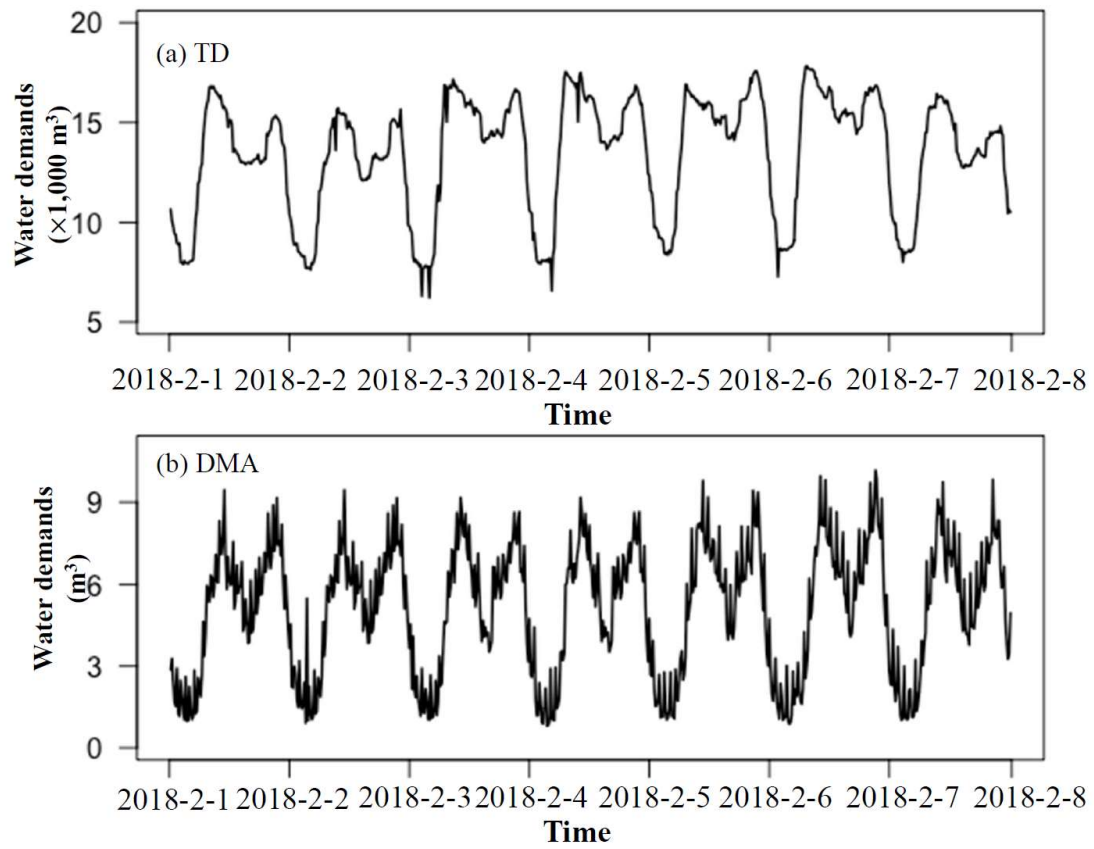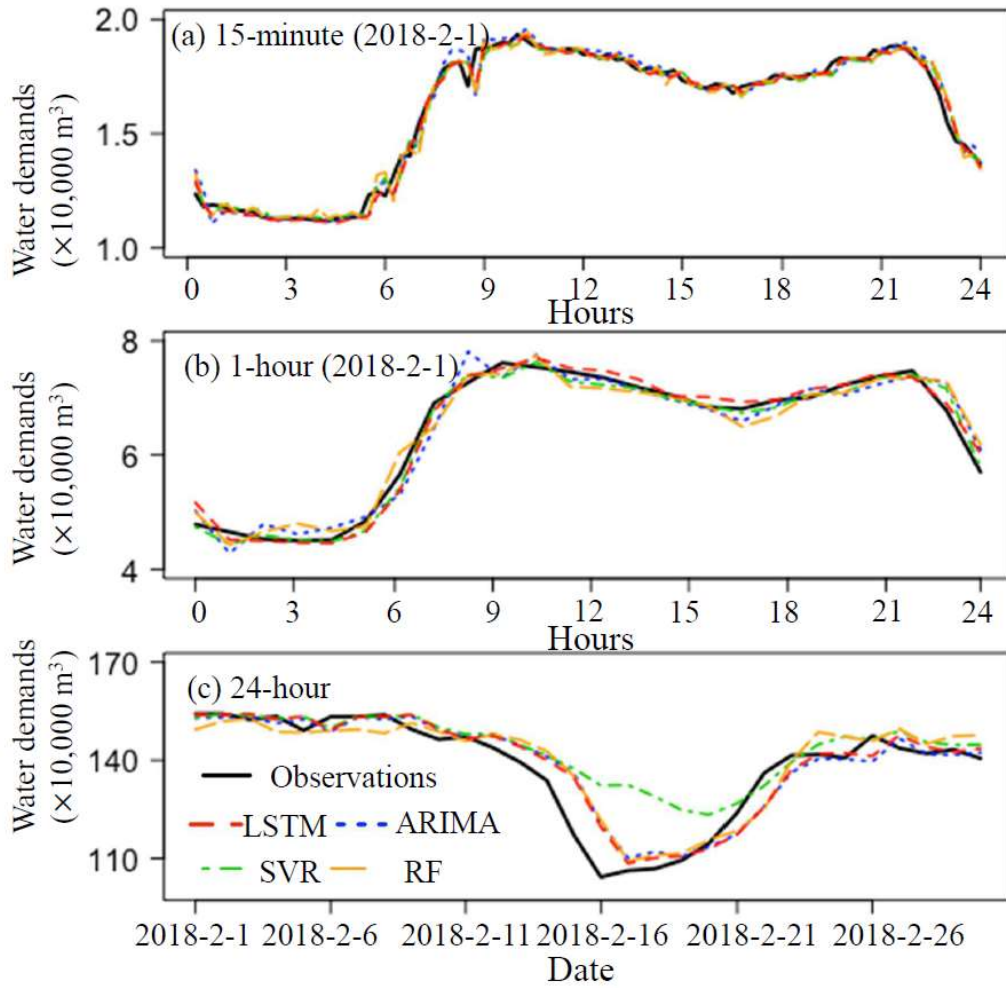
615

616 **Figure 3: Records of total water demands (TD) and from a DMA with 15-min**
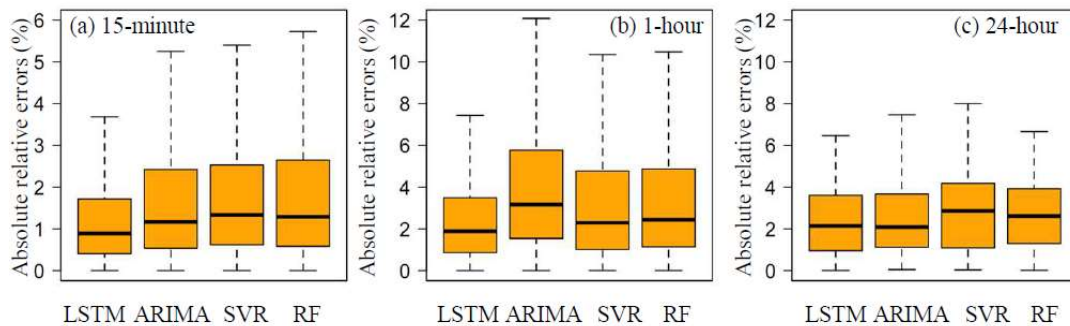
617 **resolution**

618

619

Figure 4: Predictions versus observations for the four models applied to the total water demands (TD)

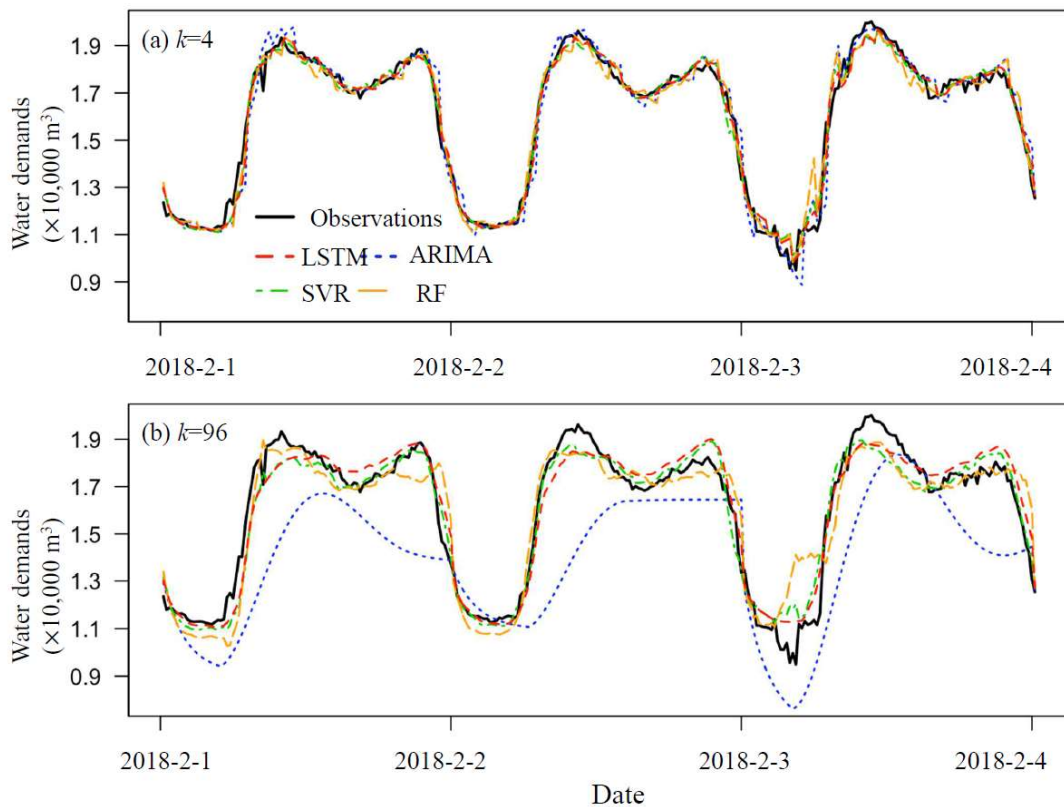**Figure 5: Absolute relative errors of the model predictions for the total water demands**

630

631
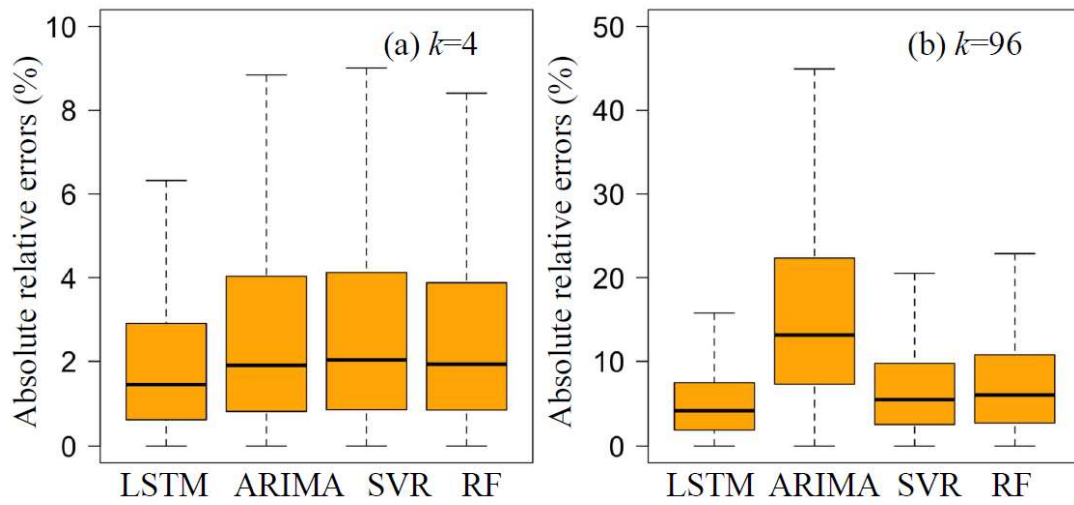
Figure 6: Absolute relative errors for models used to predict multiple successive data points, where *k* is the number of multiple successive data points

634

635
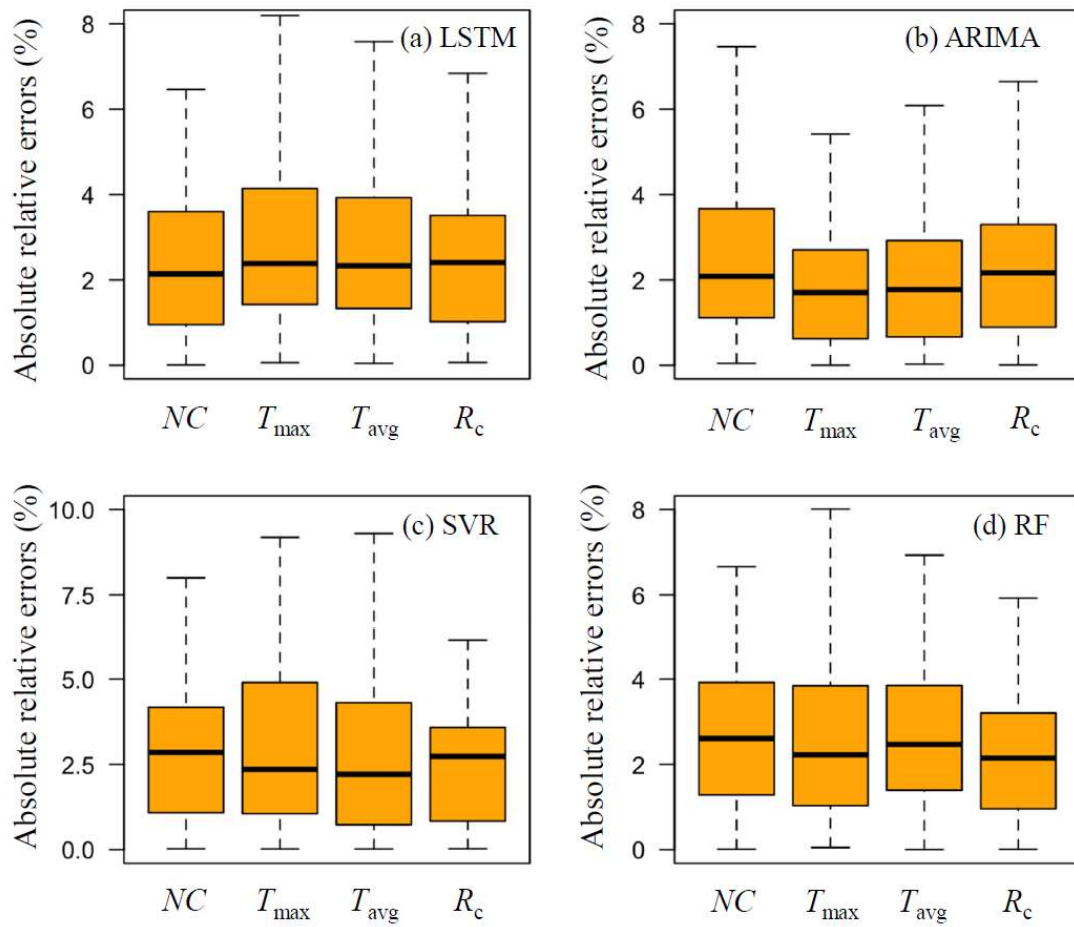

636

Figure 7: Predictions versus observations for the four models used to generate multiple successive data points

639

640

641

642   **Figure 8: Absolute relative errors of the four models with different external**

643   **parameters applied to the total water demands with the 24-hour resolution**