

Article

# LSTM-Based Forecasting for Urban Construction Waste Generation

Li Huang <sup>1,2</sup> , Ting Cai <sup>3</sup>, Ya Zhu <sup>4</sup>, Yuliang Zhu <sup>1,5</sup>, Wei Wang <sup>1,5,\*</sup>  and Kehua Sun <sup>6</sup>

<sup>1</sup> Key Laboratory of Coastal Disaster and Defence of Ministry of Education, Hohai University, Nanjing 210098, China; lily8214@hhu.edu.cn (L.H.); zhuy1@hhu.edu.cn (Y.Z.)

<sup>2</sup> School of Public Administration, Hohai University, Nanjing 210098, China

<sup>3</sup> Business School, Sichuan University, Chengdu 610065, China; caiting@stu.scu.edu.cn

<sup>4</sup> College of Political Science, Nanjing Agricultural University, Nanjing 210095, China; zhuya@njau.edu.cn

<sup>5</sup> College of Harbour, Coastal and Offshore Engineering, Hohai University, Nanjing 210098, China

<sup>6</sup> Shanghai Communications Construction Co., Ltd., Shanghai 200136, China; 191603010032@hhu.edu.cn

\* Correspondence: 13813826667@hhu.edu.cn; Tel.: +86-25-8370-1905

Received: 3 September 2020; Accepted: 13 October 2020; Published: 16 October 2020



**Abstract:** Accurate forecasts of construction waste are important for recycling the waste and formulating relevant governmental policies. Deficiencies in reliable forecasting methods and historical data hinder the prediction of this waste in long- or short-term planning. To effectively forecast construction waste, a time-series forecasting method is proposed in this study, based on a three-layer long short-term memory (LSTM) network and univariate time-series data with limited sample points. This method involves network structure design and implementation algorithms for network training and the forecasting process. Numerical experiments were performed with statistical construction waste data for Shanghai and Hong Kong. Compared with other time-series forecasting models such as ridge regression (RR), support vector regression (SVR), and back-propagation neural networks (BPNN), this paper demonstrates that the proposed LSTM-based forecasting model is effective and accurate in predicting construction waste generation.

**Keywords:** environmental engineering; construction waste; short and long-term memory (LSTM) network; time-series forecasting; deep learning

## 1. Introduction

As urbanization and transformation of old urban districts progress, the volumes of construction waste that are generated constantly increase [1]. Statistics show that the total annual disposal volume of construction waste in China has ranged from approximately 1.55 to 2.4 billion tons in recent years, accounting for approximately 40% of the total volume of urban solid waste. Construction waste disposal in China is relatively extensive, with landfilling and dumping as the primary disposal approaches. Research indicates that with effective planning and technical measures, most construction waste can be reused as renewable resources. However, thoughtless disposal will produce serious environmental issues, including air, land, and water pollution. These issues will pose a threat to human health and generate waste of potentially recyclable resources.

Recycling of construction waste is considered an effective means to digest urban waste. Methods for quantifying construction waste are the basis for management of its recycling. Forecasting of trends and variations of construction waste has significance because it enables a government to estimate the landfill capacity requirements in advance and formulate relevant policies [2]. In 2014, Wu et al. [3] broadly reviewed the literature on construction waste management. They noted that relatively few studies had been conducted to predict construction waste generation, the lack of support information

that can be used in forecasting, and that additional efforts should be made to investigate this topic. Forecast analysis is performed in this study for the generation of construction waste.

## 2. Literature Review

Historical construction waste generation data are mostly presented as time series and can show the dynamic evolutionary process of a system. When the sample size is sufficiently large, forecasts can be made using multiple data-driven methods, such as multiple linear regression (MLR), gray models (GMs), support vector machine regression (SVR), and artificial neural networks (ANNs).

In 2014, Zuo and Fu [4], and Zhang et al. [5] in 2013 established an autoregressive integrated moving-average model and a GM based on estimated historical data, respectively. These models can satisfactorily fit and predict continuous growth and relatively regular curves. For complex and irregular time series, in 2012, Wu [6] introduced gene-expression programming into the forecasting of urban construction waste generation and combined the advantages of time-series and factor-regression models. In 2013, Antanasijevic et al. [7] established a backpropagation neural network (BPNN) and a general regression neural network (GRNN) model for forecasting urban solid-waste generation. They discovered that the GRNN model significantly outperformed the conventional BPNN model and exhibited relatively high stability even for long-term forecasting. In 2016, Abbasi and Hanandeh [8] demonstrated that the adaptive neuro-fuzzy inference system (ANFIS), k-nearest neighbor (KNN), and SVR models can be employed to establish forecasting models and provided accurate and reliable monthly waste-generation forecasts, and that the ANFIS and KNN models outperformed the SVMR model in terms of forecast accuracy.

Considering the limitations of single models, in 2009, Li et al. [9] introduced a variable-weight combination forecasting model and determined that in terms of forecast accuracy, this model outperformed various single forecasting models, such as unary linear regression, exponential smoothing, and GM. In 2014, Yuan [10], proposed a forecasting method that consists of principal component analysis (PCA) and the hidden Markov model, and used it to predict waste generation based on a small sample size. In 2017, Song et al. [11] established a PCA–SVR-combined forecasting method. In this method, modeling and forecasting are achieved after characteristic components are extracted from the original data. Better experimental results were obtained using this combined model than single models. However, in practical application, combined forecasting methods are disadvantageous due to the complexity of their construction processes and strong reliance on human involvement.

With rapid urbanization, China's data on urban construction waste generation exhibit complex, nonlinear characteristics. Due to various factors (e.g., statistical specifications and data storage) that relate to these data, the data published by the statistics bureaus of China's major cities are limited. The use of conventional time-series forecasting methods to capture the complex patterns contained in these small volumes of data is often difficult. On the other hand, making accurate predictions for such irregular time series must be solved, since many management decisions are based on these forecasts. However, accurate forecasting of construction waste generation has practical significance and is an important research problem that must be effectively addressed.

In recent years, forecasting methods based on deep learning have been gradually applied in research on time-series data. In particular, recurrent neural networks (RNNs) exhibit relatively high adaptability in time-series analysis. As a typical RNN variant, long short-term memory (LSTM) can better represent long- and short-term dependency than RNNs, and effectively address the problems of RNNs, such as vanishing and exploding gradients, and rapid long-term memory loss. It is applicable to forecasting problems with relatively small volumes of data. LSTM models have been successfully applied in research on time-series data in various fields, including machine translation, speech recognition, and traffic flow forecasting [12]. These models are expected to provide effective means to predict construction waste generation based on relatively small volumes of historical data. However, the application of LSTM models in time-series forecasting of urban construction

waste generation has not been reported, indicating that this problem needs to be carried out with advanced techniques.

This study will contribute to the literature by developing a novel and effective LSTM-based method for the forecasting of construction waste generation. A forecasting method based on a three-layer LSTM neural network structure is proposed for time-series data for construction waste generation. Network design and implementation algorithms for network training and forecasting are described. Additionally, the proposed model is compared with multiple time-series forecasting models via numerical experiments. The experimental results show that the proposed LSTM-based forecasting model exhibits exceptional performance in time-series forecasting of construction waste generation.

### 3. LSTM Network

LSTM networks are a type of RNN and an improvement over the standard RNN structure. The difference between an LSTM network and a standard RNN is that only one threshold function exists in the standard RNN cell structure, whereas three important threshold functions are introduced in the hidden-layer cell structure of an LSTM network [13], as shown in Figure 1.

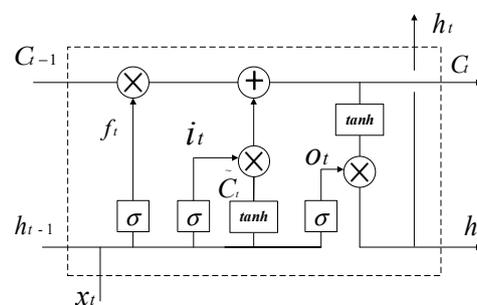


Figure 1. LSTM cell structure in hidden layer.

Compared with the other methods, the advantage of LSTM is learning long-term dependencies. For an LSTM network, information storage and interaction in the forward-propagation process are controlled by the three thresholds in the hidden-layer cell structure. The forget gate ( $f_t$ ) controls the unit state information at the previous time point that needs to be discarded. The input gate ( $i_t$ ) controls the information-input process at the current time point. The output gate ( $o_t$ ) controls the filtered output of the current unit state.

The three thresholds are calculated using the following equations [14]:

$$\begin{cases} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \end{cases} \quad (1)$$

The update equation for the cell state information for the current time point is expressed as:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2)$$

Based on the combined actions of the  $f_t$  and  $i_t$ , the cell state at the current time point ( $C_t$ ) can be represented as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (3)$$

The information returned to the hidden layer ( $h_t$ ) is:

$$h_t = o_t * \tanh(C_t) \quad (4)$$

where  $x$  is the input vector,  $h$  is the output vector,  $C$  is the cell state,  $\sigma$  and  $\tanh$  are the *sigmoid*. and *tanh*. activation functions, respectively, and  $W$  and  $b$  are the corresponding weight and deviation matrices, respectively.

The backpropagation-through-time algorithm [15] is employed to train an LSTM model in the following steps:

- Step 1: Use Equations (1)–(4) to calculate the output value of the forward propagation;
- Step 2: Inversely calculate the error term for each LSTM cell, including the longitudinal propagation between the layers and the temporal transverse propagation;
- Step 3: Calculate the gradient of each weight, based on the corresponding error term;
- Step 4: Use a gradient-based optimization algorithm to update the weights.

#### 4. LSTM Network-Based Construction Waste Generation Forecasting Model

In this study, an LSTM model was constructed based on the characteristics of univariate time-series data with limited sample points and the simple RNN design principle. Figure 2 shows the total framework based on a typical three-layer LSTM structure, which involves four main modules of data preprocessing, network training, network prediction and model evaluation. Relevant steps of LSTM time-series prediction can be described briefly as follows:

- Step 1. Preprocess historical data and split it into training set and test set;
- Step 2. Construct the model construction and train LSTM network based on training set;
- Step 3. Make predictions based on test set;
- Step 4. Evaluate the model accuracy.

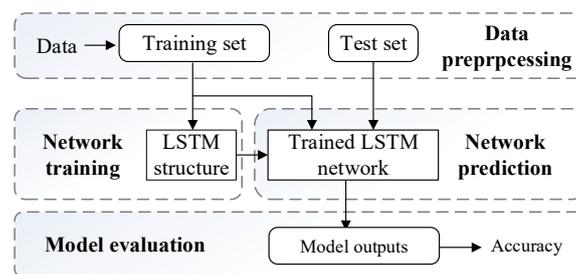


Figure 2. LSTM network-based framework for construction waste generation forecasting.

During data preprocessing, to make inputs conform to the LSTM structure, the investigated time series can be transformed to a supervised learning problem by creating inputs of delayed observations and labels of forecasts. For a given lag step  $k$  (which equals to the number of neurons in input layer of LSTM structure), the input vector  $(y_{t-k}, y_{t-k+1}, \dots, y_{t-1})$  at period  $t$  is used to calculate the forecasting value  $\hat{y}_t$ .

Network training is aimed at a trained LSTM network and its forward calculation  $f$  by optimizing loss function of mean square error (MSE). The final forecasting model can be simply expressed as follows:

$$\hat{y}_t = f(y_{t-k}, y_{t-k+1}, \dots, y_{t-1}) \quad (5)$$

Next, an iterative manner is adopted in prediction process, which means that each new forecast on test set will be taken to create new input vector to obtain the next forecast until all forecasts are collected. Lastly, all forecasts are taken to compute model accuracy with original observations on test set.

#### 5. Numerical Experiments

In this study, a forecasting model was constructed based on the historical construction waste generation data for two cities, namely, Shanghai and Hong Kong. Note that a change was made in Shanghai's statistical standards for construction waste generation. As a result, it is unreasonable to

forecast using all historical data for Shanghai. However, this study is focused on the forecasting of urban construction waste generation instead of construction waste classification.

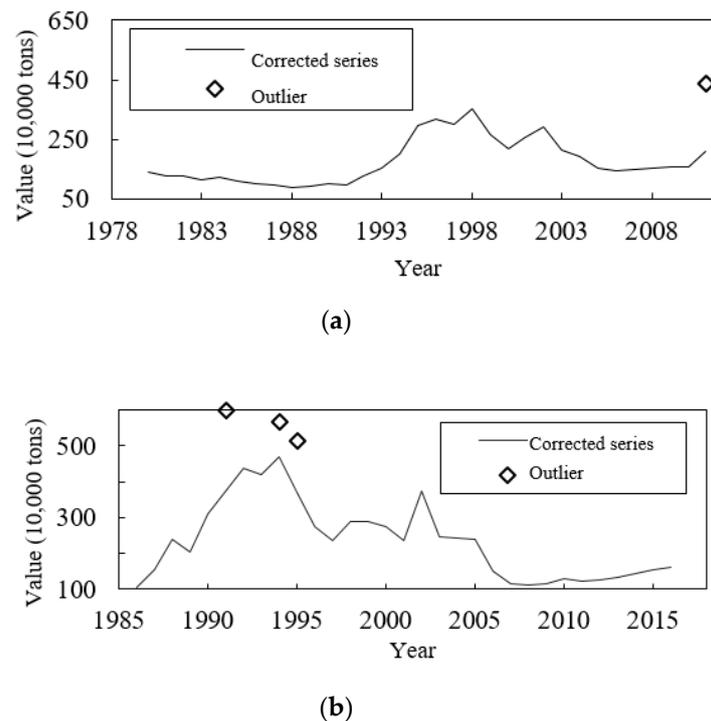
### 5.1. Experimental Data and Design

#### 5.1.1. Experimental Data and Preprocessing

##### Data Collection and Outlier Elimination

We collect two datasets of construction waste data from Shanghai and Hong Kong, respectively, for the experiments. The construction waste generation data from Shanghai statistical yearbook (<http://tjj.sh.gov.cn/tjnj/nj18.htm?d1=2018tjnj/C0618.htm>) for the pre-2012 period (a total of 32 years), which is denoted by dataset A, exhibit complex time-series features with substantial volume. The dataset B is the annual construction waste generation data for the 31-year period from 1986 to 2016, which is collected from the Waste Reduction Office in the Environmental Protection Department of Hong Kong ([https://www.wastereduction.gov.hk/en/assistancewizard/waste\\_red\\_sat.htm#top](https://www.wastereduction.gov.hk/en/assistancewizard/waste_red_sat.htm#top)).

Outliers refer to data points beyond the normal value range and generally identified by scatter or box plots. Due to the relatively small sample size of the datasets selected in this study, correcting the outliers can facilitate the use of the available data information and prevent an insufficient sample size caused by eliminating outliers. The outliers in the datasets A and B were treated by mean-value interpolation, as shown in Figure 3.



**Figure 3.** Outlier handling schematic. (a) Dataset A; (b) Dataset B.

##### Data Segmentation

In this study, each time series was divided into a training set and a test set based on the time sequence by hold-out cross-validation, and the subsets of the training set (segmented based on the time sequence) were preserved to examine the model performance, as shown in Figure 4 [16]. Due to the slight difference in sample sizes between datasets A and B, to facilitate the comparison of model performance, the sample sizes of both the test sets were set to 5 (i.e., data for the last five years).

Thus, the first 27 data points (pre-2007 data) in dataset A and the first 26 data points (pre-2016 data) in dataset B were selected to form training sets.

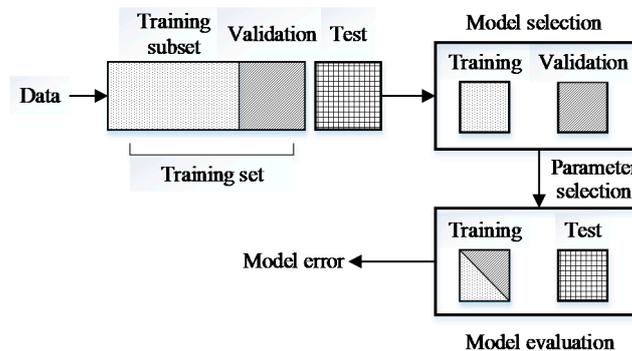


Figure 4. Hold-out cross-validation based on time series.

### Data Normalization

Data normalization can eliminate the difference in dimension among input data and increase the computational speed of the model. In this study, the input data were mapped by min-max normalization [17] to the interval of  $[0, 1]$ , which can be represented by the following equation:  $y^* = (y - y_{min}) / (y_{max} - y_{min})$ , where  $y$  is the input datum,  $y^*$  is the normalized datum, and  $y_{max}$  and  $y_{min}$  are the maximum and minimum values, respectively, of the input data. Denormalizing the output of the forecasting model is necessary to enable it to fall within the actual range and be consistent with the actual significance.

### 5.1.2. Evaluation Indices

To reduce the limitation associated with the use of a single index to evaluate the model performance, the following three indices were employed in this study to examine the forecasting models comprehensively.

Mean absolute error (MAE) was used to evaluate the closeness between the actual values and predicted values.

$$MAE = \frac{1}{T} \cdot \sum_{t=1}^T |y_t - \hat{y}_t| \quad (6)$$

Mean absolute percentage error (MAPE) was used to evaluate the relative error, which can be used to compare the forecast performance for various dataset ranges.

$$MAPE = \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times \frac{100\%}{T} \quad (7)$$

Root mean square error (RMSE) is highly sensitive to extremely large or small errors, and therefore can satisfactorily reflect the forecast accuracy.

$$RMSE = \sqrt{\frac{1}{T} \cdot \sum_{t=1}^T (y_t - \hat{y}_t)^2} \quad (8)$$

Here,  $y_t$  and  $\hat{y}_t$  are the actual and predicted values, respectively, output by the model at the time  $t$ .

### 5.1.3. Comparison Models

The following three models were selected for experimental comparative analysis.

## Ridge Regression (RR)

RR is an approach that adds L1 regularization to a general linear regression model to prevent overfitting. The RR time-series forecasting model is structurally similar to MLR [18] and can be represented by the equation:  $Y_t = a_0 + a_1Y_{t-1} + a_2Y_{t-2} + \dots + a_kY_{t-k} + e$ , where  $Y_t$  is the predicted value at time  $t$ ,  $a_1, a_2, \dots, a_k$  are the regression coefficients of  $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$ , respectively, at various time points and  $a_0$  and  $e$  are the bias and error terms, respectively. In this study, the regularization coefficient  $\alpha$  of the RR model was set to 0.5. Additionally, the step size of the input series (i.e.,  $k$ ) was selected by trial and error.

## SVR

SVR determines the nonlinear mapping relations between low-dimensional data and output indices by regression after mapping the low-dimensional data to a high-dimensional space via a nonlinear kernel function [19]. In this study, the Gaussian-radial basis function was selected as a nonlinear kernel function. The corresponding penalty coefficient  $C$  and kernel function coefficient  $\gamma$  were set to certain values, and  $k$  was also determined.

## ANN

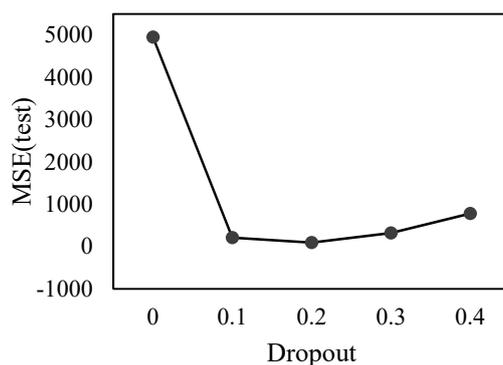
Three-layer BPNNs exhibit excellent performance in approximating nonlinear data [20,21]. Thus, a three-layer BPNN was employed in this study. The *sigmoid* activation function was applied between the input layer and the hidden layer and between the hidden layer and the output layer. The network was trained using the gradient descent with momentum algorithm. The learning rate was set to 0.1. The momentum parameters were set to the following values: momentum = 0.9 and Nesterov = true. The random seeds for network initialization, maximum number of iterations, and expected error were set to 0, 2000, and  $1 \times 10^{-6}$ , respectively. In this study, point-by-point forecasting was performed for the test-set data. Therefore, the number of neurons in the output layer was set to 1. Additionally, the number of neurons in the input layer ( $i$ ) and the number of neurons in the LSTM hidden layer ( $n$ ) were determined by trial and error ( $i = k$ ).

## 5.2. Experimental Results and Discussions

### 5.2.1. Analysis of Model Parameters

First, dataset A is used as an example. The training set of the time-series data for construction waste generation was normalized using the proposed method. Additionally, an LSTM-based forecasting model was constructed. The numbers of nodes in the input and hidden layers of a three-layer LSTM model exert an extremely significant impact on the network scale. The selected optimizer also affects the convergence rate of the network during the training process. Thus,  $n$  and  $k$  were determined by trial and error to be 64 and 2, respectively. An adaptive-gradient optimizer was selected [22,23] (default parameters were preserved). For the non-key model parameters, the random seeds for network initialization, maximum number of iterations, and expected error were set to 0, 2000, and  $1 \times 10^{-6}$ , respectively.

Due to the relatively large  $n$  value, the model is relatively complex, which may cause overfitting and relatively poor generalization performance. In this study, the dropout method [24] was employed. Specifically, a dropout layer was added after the LSTM hidden layer to randomly invalidate some neurons and prevent them from updating by forward propagation and backpropagation, which prevented overfitting to a certain extent. In this study, an optimal dropout rate was determined by comparing the MSE values on the test set with respect to various dropout rates, as shown in Figure 5. The MSE on the test set was the largest before adding a dropout layer to the model. At a dropout rate of 2, the MSE for the test-set data was the smallest, and the optimal training performance was achieved on the test set.



**Figure 5.** MSE of test set under different dropout probabilities (dataset A).

To examine its forecast effectiveness and accuracy, the LSTM-based model was compared with various time-series forecasting models. Table 1 summarizes the parameter settings for each model. Additionally, to determine the advantages of the LSTM-based model among RNNs, the hidden-layer cells in the LSTM-based model were replaced by the RNN structure, and experiments were conducted using the same parameters.

**Table 1.** Model parameter settings for datasets A and B.

Model Parameter	Dataset A	Dataset B
RR ( $k$ )	(4)	(3)
SVR ( $k, C, \gamma$ )	(4, 4, 2)	(2, 10, 4)
BPNN ( $k, n$ )	(4, 4)	(3, 4)
RNN ( $k, n$ , optimizer)	(2, 64, Adagrad)	(4, 14, RMSprop)
LSTM ( $k, n$ , optimizer)		

The experimental process for dataset B was similar to that for dataset A. The parameters of the LSTM-based model were adjusted based on experience. The  $k$  and  $n$  were set to 4 and 14, respectively. A root-mean-square propagation (RMSProp) optimizer was selected (default parameters were preserved). The non-key parameters were set to the same values selected for the non-key parameters for dataset A. For dataset B, because the  $n$  value was moderate, no dropout layer was added to the network structure.

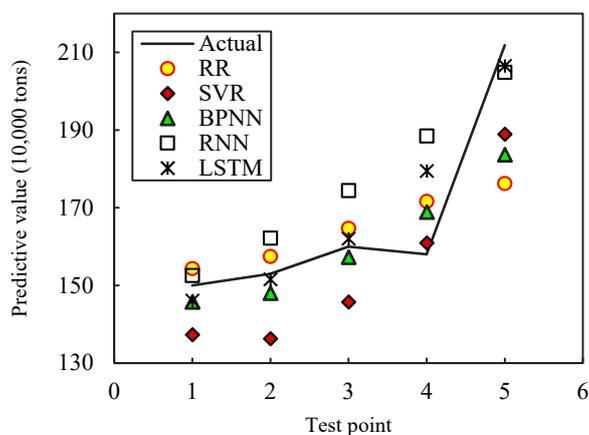
### 5.2.2. Forecasts and Comparative Analysis

First, dataset A is used as an example. The network was trained using the previously determined LSTM network structure and parameters. Values fitted to the training set were obtained. Additionally, point-by-point recursive forecasting was performed on the test set. The corresponding predicted values (for the period 2007–2011) were obtained. The fitting and forecast performance of various time-series forecasting models were comparatively analyzed using the evaluation indices described in Section 5.1.2, as shown in Table 2.

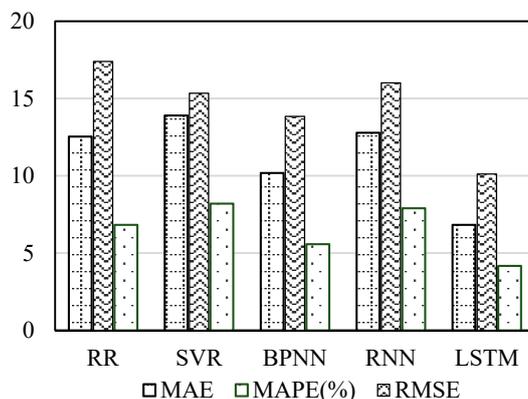
Figures 6 and 7 display the performance of the LSTM-based model and each comparison model in terms of the predicted value and error. As demonstrated in Figure 6, the forecast curves of the RR, SVR, and BPNN models tend to be stable and unable to accurately predict the data for the fifth test point, which produces relatively large prediction errors. In comparison, the LSTM-based model exhibits relatively excellent forecast performance. The prediction error of each model for the fourth data point is relatively large. However, the forecast curve of the LSTM-based model is closer to the variation trend in the actual data.

**Table 2.** Comparison of fitting and prediction performance of different models (dataset A).

		MAE	MAPE (%)	RMSE
RR	training set	34.54	19.17	40.98
	test set	12.54	6.83	17.39
SVR	training set	19.3	12.06	21.23
	test set	13.9	8.2	15.34
BPNN	training set	30.38	17.16	36.08
	test set	10.18	5.58	13.84
RNN	training set	26.65	15.26	35.14
	test set	12.78	7.9	16.01
LSTM	training set	22.15	12.02	30.00
	test set	6.79	4.17	10.14



**Figure 6.** Prediction results of different models for test sets (dataset A).



**Figure 7.** Comparison of prediction performance of different models for test sets (dataset A).

Figure 7 further compares the LSTM-based model and other forecasting models in terms of performance on the test set. The LSTM-based model is employed as an example. The values of the evaluation indices for the LSTM-based model are lower than those for other forecasting models, which suggests that the proposed LSTM-based model satisfactorily tracks the actual data in the test set and achieves relatively ideal forecast performance.

Dataset A has low timeliness. The data for the following six years (2012–2017) predicted by the LSTM-based model were converted using a conversion factor (i.e., ratio of the mean construction waste generation (excluding the construction waste soil generation) in the period 1980–2011 to the mean construction waste generation (including the construction waste soil generation) in the period

2012–2017). Table 3 show the forecasting result for the years 2012–2017, and the result was compared with the trend of construction waste generation (including the construction waste soil generation) in the period 2012–2017. The correlation coefficient ( $R$ ) between the two trendlines is 0.87. This finding indirectly demonstrates that the LSTM-based model is effective in forecasting.

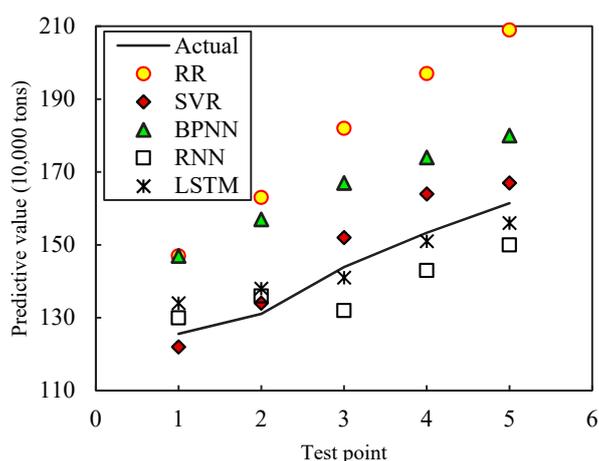
**Table 3.** Forecast for the next 6 years.

Year	2012	2013	2014	2015	2016	2017
Predictive value (10,000 tons)	281.80	312.26	287.55	260.11	248.04	246.14

The applicability of the proposed method to forecast construction waste generation in other cities was further examined based on dataset B. Table 4 and Figure 8 show the experimental results, which demonstrate that the proposed LSTM-based model outperforms the comparison models in terms of forecast accuracy.

**Table 4.** Comparison of fitting and prediction performance of different models (dataset B).

		MAE	MAPE (%)	RMSE
RR	training set	48.71	19.68	59.95
	test set	36.7	25.29	37.84
SVR	training set	41.81	17.88	51.18
	test set	6.13	4.2	6.73
BPNN	training set	50.82	21.26	60.77
	test set	27.94	15.57	22.09
RNN	training set	5.94	2.68	7.53
	test set	12.16	8.12	14.19
LSTM	training set	5.54	2.83	7.07
	test set	5.48	3.98	5.95



**Figure 8.** Prediction results of different models for test sets (dataset B).

According to the aforementioned case study of two datasets, the proposed LSTM model performs well for obtaining more accurate forecasts than four benchmarking models. Our proposed model aims to solve univariate time series forecasting problem with nonlinearity and non-stationarity due to the lack of multivariate data of influencing factors. However, in practice, our model can be used in multivariate time series prediction problem by slightly changing inputs to LSTM structure, and its effect can be discussed in the further research.

## 6. Conclusions

The forecasting of construction waste generation enables local governments to manage construction waste landfill and formulate construction waste management policies. However, a review of the available literature indicates that the methods for this type of forecasting and their accuracy require improvements.

Available data for urban construction waste generation have the characteristics of small volumes and high nonlinearities. A three-layer LSTM network is proposed in this study for forecasting based on univariate time series with limited sample points. The proposed model is compared with regression and neural-network models. The results show that the LSTM-based model is highly effective in solving univariate nonlinear forecasting problems. Additionally, a dropout layer is added to effectively address the overfitting problem of the LSTM-based model and improve its generalization performance.

The applicability of the LSTM-based model to forecasting construction waste generation is demonstrated. The scope of the deep-learning technique is expanded. This study has theoretical and practical significance in quantifying and managing construction waste. Obtaining construction waste data for cities in China is difficult. As a result, the proposed method is validated based on data for only Shanghai and Hong Kong. In the future, the proposed method can be used to predict construction waste generation in other cities and solve forecasting problems in other fields. Additionally, the application performance of various network-training optimization algorithms, such as Adam and RMSProp, can be compared.

**Author Contributions:** All authors contributed equally to this work. In particular, L.H. reviewed and revised the manuscript. T.C. drafted the manuscript. Y.Z. (Ya Zhu) performed formal analysis. Y.Z. (Yuliang Zhu) performed the investigation. W.W. developed the original idea for the study and designed the methodology. K.S. performed the case study. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Fundamental Research Funds for the Central Universities, (2019B20414), the Key Laboratory of Coastal Disaster and Defence of Ministry of Education, Hohai University, (201913), the National Natural Science Foundation of China, (71974052) and the Jiangsu Social Science Fund, (18GLB013).

**Acknowledgments:** Thanks to Guo Zhaoxia from Sichuan University, Wu Xiaonan from Shanghai Communications Construction Co., Ltd. We also acknowledge the data support from Shanghai Communications Construction Co., Ltd.

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

## References

1. Yuan, H.; Shen, L. Trend of the research on construction and demolition waste management. *Waste Manag.* **2011**, *31*, 670–679. [[CrossRef](#)] [[PubMed](#)]
2. Wu, H.; Zuo, J.; Zillante, G.; Wang, J.; Yuan, H. Status quo and future directions of construction and demolition waste research: A critical review. *J. Clean. Prod.* **2019**, *240*, 118163. [[CrossRef](#)]
3. Wu, Z.; Yu, A.T.; Shen, L.; Liu, G. Quantifying construction and demolition waste: An analytical review. *Waste Manag.* **2014**, *34*, 1683–1692. [[CrossRef](#)] [[PubMed](#)]
4. Zuo, H.K.; Fu, S.L. Output forecast of construction waste in Beijing and constructing and distributing of its disposal facilities. *Environ. Sanit. Eng.* **2011**, *19*, 63–64.
5. Zhang, H.Y.; Yang, F.H.; Li, G.X. Analysis and forecast of construction waste production based on the ARIMA model in ChaoYang district. *Environ. Eng.* **2014**, *32*, 696–699.
6. Wu, Z.Z. Research on Construction and Demolition Waste Quantification and Management Strategies. Master's Thesis, Chongqing University, Chongqing, China, 2013. (In Chinese).
7. Antanasijević, D.; Pocajt, V.; Popovic, I.G.; Redzic, N.; Ristic, M. The forecasting of municipal waste generation using artificial neural networks and sustainability indicators. *Sustain. Sci.* **2013**, *8*, 37–46. [[CrossRef](#)]
8. Abbasi, M.; El Hanandeh, A. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. *Waste Manag.* **2016**, *56*, 13–22. [[CrossRef](#)]
9. Li, Z.T.; Zhang, Y.F.; Yao, F.Z.; Shi, X.R. Combined forecasting model with variable weights for prediction of municipal solid waste. *J. Nanjing Tech. Univ. Nat. Sci. Ed.* **2009**, *31*, 97–101.
10. Yuan, Y.; Liu, X. Municipal Waste Generation Forecasting Based on Advanced Hidden Markov Model. Master's Thesis, Shanghai Jiao Tong University, Shanghai, China, 2014. (In Chinese).

11. Song, Y.; Wang, Y.; Liu, F.; Zhang, Y. Development of a hybrid model to predict construction and demolition waste: China as a case study. *Waste Manag.* **2017**, *59*, 350–361. [[CrossRef](#)]
12. Ma, X.; Tao, Z.; Wang, Y.; Yu, H.; Wang, Y. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transp. Res. Part. C Emerg. Technol.* **2015**, *54*, 187–197. [[CrossRef](#)]
13. Graves, A. *Long Short-Term Memory*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 1735–1780.
14. Greff, K.; Srivastava, R.K.; Koutnik, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A Search Space Odyssey. *IEEE Trans. Neural Networks Learn. Syst.* **2016**, *28*, 2222–2232. [[CrossRef](#)] [[PubMed](#)]
15. Werbos, P. Backpropagation through time: What it does and how to do it. *Proc. IEEE* **1990**, *78*, 1550–1560. [[CrossRef](#)]
16. Duchi, J.; Hazan, E.; Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **2011**, *12*, 257–269.
17. Wang, W.; Huang, L.; Liang, X. On the Simulation-Based Reliability of Complex Emergency Logistics Networks in Post-Accident Rescues. *Int. J. Environ. Res. Public Health* **2018**, *15*, 79. [[CrossRef](#)] [[PubMed](#)]
18. Yu, X.; Liong, S.-Y. Forecasting of hydrologic time series with ridge regression in feature space. *J. Hydrol.* **2007**, *332*, 290–302. [[CrossRef](#)]
19. Brereton, R.G.; Lloyd, G.R. Support Vector Machines for classification and regression. *Analyst* **2010**, *135*, 230–267. [[CrossRef](#)]
20. Wang, L.; Zeng, Y.; Chen, T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Syst. Appl.* **2015**, *42*, 855–863. [[CrossRef](#)]
21. Huang, L.; Huang, J.; Wang, W. The Sustainable Development Assessment of Reservoir Resettlement Based on a BP Neural Network. *Int. J. Environ. Res. Public Health* **2018**, *15*, 146. [[CrossRef](#)]
22. Wang, W.; Yang, J.; Huang, L.; Proverbs, D.; Wei, J. Intelligent Storage Location Allocation with Multiple Objectives for Flood Control Materials. *Water* **2019**, *11*, 1537. [[CrossRef](#)]
23. Wang, W.; Huang, L.; Gu, J.; Jiang, L. Green port project scheduling with comprehensive efficiency consideration. *Marit. Policy Manag.* **2019**, *46*, 967–981. [[CrossRef](#)]
24. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.

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



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).