applied sciences

MDPI

Article

Sustainable Transport in a Smart City: Prediction of Short-Term Parking Space through Improvement of LSTM Algorithm

Bowen Jin [1], Yu Zhao [2] and Jing Ni [2,*]

[1] College of Engineering, Northeastern University, Boston, MA 02115, USA
[2] College of Economic Management, Beijing Institute of Petrochemical Technology, Beijing 102617, China
* Correspondence: nijing@bipt.edu.cn; Tel.: +86-1060228013

Abstract: The carbon emission of fuel vehicles is a major consideration that affects the dual carbon goal in urban traffic. The problem of "difficult parking and disorderly parking" in static traffic can easily lead to traffic congestion, an increase in vehicle exhaust emissions, and air pollution. In particulate, when vehicles make an invalid detour and wait for parking with long hours, it often causes extra energy consumption and carbon emission. In this paper, adding a weather influence feature, a short-term parking occupancy rate prediction algorithm based on the long short-term model (LSTM) is proposed. The data used in this model is from Melbourne public data sets, and parking occupancy rates are predicted through historical parking data, weather information, and location information. At the same time, three commonly prediction models, i.e., simple LSTM model, multiple linear regression model (MLR), and support vector regression (SVR), are also used as comparison models. Taking MAE and RMSE as evaluation indexes, the parking occupancy rate during 10 min, 20 min, and 30 min are predicted. The experimental results show that the improved LSTM method achieves better accuracy and stability in the prediction of parking lots. The average MAE and RMSE of the proposed LSTM prediction during intervals of 10 min, 20 min, and 30 min with the weather influence feature algorithm is lower than that of simple LSTM, MLR and that of SVR.

Keywords: short-term parking occupancy; LSTM; weather influence factor

check for updates

Citation: Jin, B.; Zhao, Y.; Ni, J. Sustainable Transport in a Smart City: Prediction of Short-Term Parking Space through Improvement of LSTM Algorithm. *Appl. Sci.* **2022**, *12*, 11046. https://doi.org/10.3390/app122111046

Academic Editor: Xianpeng Wang

Received: 8 October 2022
Accepted: 27 October 2022
Published: 31 October 2022

1. Introduction

In the vision of the smart city, a city should be a living agent that senses, analyzes, and integrates the key information of the core system of urban operation, based on cloud computing, big data, artificial intelligence, information technology and other advanced technologies, to respond intelligently to various demands, such as people's livelihood, environmental protection, public safety, urban services, industrial and commercial activities, so as to build a sustainable urban ecosystem [1,2].

Intelligent transportation systems [3,4] (ITS) have become an important part of the construction of smart cities. They apply the new generation of information technologies such as big data analysis, artificial intelligence, and mobile internet to intelligent transportation, and deeply mine related data to optimize transportation systems. Parking guidance and information (PGI) systems are an important part of intelligent transportation systems [5,6]. The use of PGI systems can not only effectively alleviate traffic congestion and parking difficulties in the urbanization construction, but also reduce exhaust emissions and the noise pollution of detours. The utilization rate of the original urban parking lots and parking spaces has been significantly improved, and good social and economic benefits can also be achieved. From the perspective of the development of the intelligent parking guidance system, it is an inevitable trend that the system will maintain the full life cycle management of data collection, integration, analysis, and application of the big data platform.

With the improvement of the Chinese economy, vehicle ownership and utilization rates have greatly increased (see Figure 1), causing disorderly urban parking. In addition, the number of parking spaces in urban planning cannot meet the year-on-year growth of the number of vehicles, resulting in serious parking problems and frequent urban traffic congestion. According to the data of the China White Paper on Parking Industry Development (2020) [7], there is a huge gap between the number of vehicles and parking spaces in China. This problem has led to a long-term failure to solve the parking problem in many big cities, especially in crowded commercial areas.

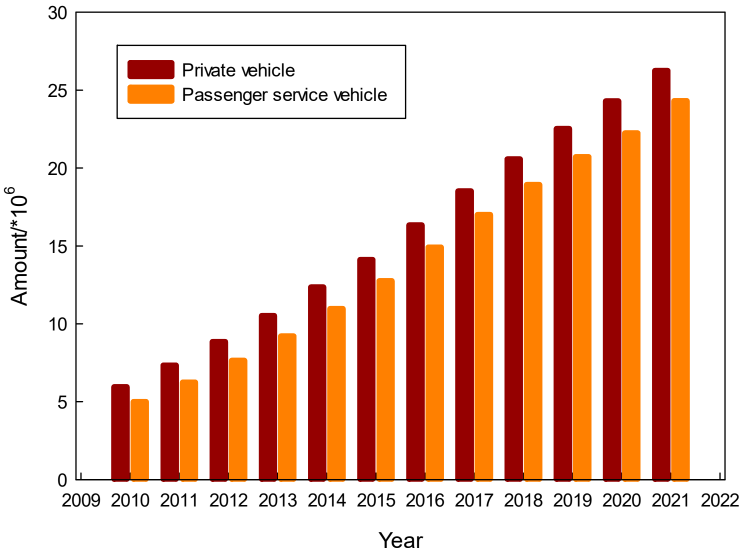

Figure 1. The statistics of private vehicle and passenger service vehicle from 2010 to 2021.

The problem of parking difficulty in large cities must be controlled from three perspectives: city governors should continue to optimize urban roads and parking space [8,9], solve the problem of high carbon emissions caused by the usage of fossil fuels by vehicles, and reduce the carbon emissions generated by unnecessary driving activities, such as car owners looking for parking spaces. From the perspective of enterprises, advanced information technology should be applied to improve service quality and operation efficiency. The innovative and practical parking guidance system [10] should also be developed to provide car owners with prediction information of both short-term and long-term parking spaces, in order to reduce carbon emissions, protect the environment, and relieve urban traffic pressure. Car owners would be able to improve their urban traffic awareness and quickly find parking lots and parking positions with such a system. Therefore, it is very necessary to develop a parking guidance information system with prediction algorithms in order to let car owners can find parking spaces purposefully, which would not only improve the efficiency of parking but also save time. In this way, carbon dioxide emissions and air pollution can be reduced, and the goals of smart city construction, carbon peak, and carbon neutralization can be better achieved.

It is challenging to use the actual parking availability data in intelligent transportation systems and parking guidance systems because parking availability is a random process. With the change of time and location, the occupation information of parking spaces will also change. Therefore, improving the accuracy of parking guidance systems and providing accurate parking availability prediction services for car owners is the key to effectively solving the problem of "parking difficulty".

The main purpose of this paper is to design, construct, and evaluate a prediction framework of parking space availability. The main contents are presented through the following contributions:

(1) In order to ensure the validity of the commercial block data about a big city, the data of eight commercial blocks with high vehicle access density, shortage of parking spaces, and urgent need for parking spaces are selected for analysis and evaluation.

(2) The improved LSTM, simple LSTM, MLR, and SVR algorithms are compared with MAE and RMSE, and the short-term prediction model is obtained to facilitate the more accurate real-time parking occupancy prediction for large-scale parking data.

The remaining sections of this paper are listed as follows. The Section 2 reviews related literature and research. The Section 3 introduces the method of parking information analysis based on the data-driven method. The Section 4 gives the discussion and the analysis results for four data-driven models based on the actual parking data, as well as the comparison results among the four model methods. Finally, the Section 5 summarizes the research and gives suggestions for future work.

2. Related Work and Research

With the arrival of the 5G era and the breakthrough of key technologies, such as sensors, intelligent transportation has become a reality. Most of the experimental data shows that in the 5G era, systems can transmit data every five minutes in cloudy conditions [11]. How to deal with the data collected by sensors in real time, and improve the accuracy of the algorithms, have become urgent problems to be solved. From parking data, the large amount of historical information recorded by infrastructure sensors provide information guarantees for parking occupancy prediction. At the same time, the existing big data processing capacity has been greatly improved, and the rapid development of parking guidance system will make it a reality. A wide range of prediction models have been developed in the literature and used to estimate current parking occupancies and parking demand in transportation networks [12]. The prediction of parking occupancy availability is closely related to traffic conditions and estimated arrival time based on predicted traffic flows and travel times.

Currently, the statistical theory used to predict the count (cnt) of parking spaces in the interval of a certain time is based on the Poisson distribution, such as the multiple regression model [13,14], Markov model [15], Kalman filter [16], and auto-regressive integrated moving average models (ARIMA) [17]. Rajabioun T and Loannou P A [14] studied a vector regressive model at the estimated arrival time of the driver with the highest probability using real-time parking data. Ye X et al. [18] combined the wavelet neural network model and wavelet transform to estimate the short-term prediction of APS. Dutta N et al. [19] deduced a general average field relationship using the mechanism of statistical physics and graph theory, and took the parking search time as a function of parking space occupancy. Bock F et al. [20] smoothed the original parking data using support vector regressions (SVR), and then trained a multidimensional SVR model, which represents the availability of parking space and is suitable for parking predictions.

Fan J et al. [21] proposed the number of vacant parking spaces in a specific period of time based on SVR with the fruit fly optimization algorithm (FOA). Xiao J et al. [22] proposed a discrete-time Markov model to present a model framework to estimate parking occupancy from actual occupancy data. Shen J et al. [23] studied a multiple linear regression model with an autoregressive moving average model including the characteristics of the parking lots and external environmental factors. Ji Y et al. [24] proposed a new multi-step forecasting model, which has a more accurate performance and better learning ability in short-term prediction. Yin C et al. [25] proposed an integrated path analysis discrete-choice model with the car ownership and travel distance as model variables using Mplus software. Amini M et al. [26] presented an auto-regressive integrated moving average (ARIMA) model for medium-term parking demand-forecasting. The simulation results show the high accuracy of the proposed method for parking occupancy prediction. From the above reviews, multivariate or multi-scale regression models can simply and quickly predict parking occupancy. However, the model parameters are affected by multi-level and multi-factor interaction, and the weight of each parameter in the model cannot be accurately measured. Moreover, environmental factors, such as the convenience of surrounding traffic and weather factors, seriously affect parking prediction, thus leading to inaccuracies.

In the last decade, models of deep neural networks have become more and more popular in the area of traffic modeling and prediction. Compared with traditional algorithms, various machine learning and deep learning algorithms have greatly improved the accuracy of short-term traffic state prediction and data processing scales [27].

Taking the street parking occupancy rate and vehicle departure probability in a specific area as the prediction performance indicators, Shao W et al. [28] used the long short-term memory (LSTM) model to predict the parking occupancy rate. Feng Y et al. [29] proposed a hybrid deep learning framework to intelligently predict the availability area of vacant parking spaces in the short term (within 30 min) and long term (more than 30 min). The hybrid deep learning framework can obtain considerably high accuracy in both short-term and long-term predictions. Zhang F et al. [30] proposed a novel periodic weather-influenced LSTM model, which successfully predicts the parking occupancy rate according to actual data, weather feature, environment condition, and weekdays. Feng N et al. [31] discussed the effect of weather features on parking behavior and found that the random forest model can make the most accurate parking behavior prediction. However, the above two papers obtained data at an interval of one hour, which failed to consider the need of real-time parking. Liu F et al. [32] constructed an optimized LSTM model based on LSTM and a bi-directional LSTM network with a better prediction precision and training speed. Samaranayake P et al. [33] presented a travel demand management system including parking demand, parking occupancy, and future changes. However, this model did not consider the demand of actual parking and limited on-street parking application. Chen H [34] proposed a forecasting model of short-term unoccupied parking space using the method of combining wavelet transformation and an extreme learning machine. The results of the prediction example show that the method shortens the training time and improves the prediction results. Rong Y et al. [35] used long short-term memory (LSTM) to model the temporal closeness and period and the current general factors. Zhang W et al. [36] proposed a semi-supervised hierarchical recurrent graph neural network for predicting city-wide parking availability within the city from the spatial and temporal domains. Lu K et al. [37] applied the fuzzy c-means clustering method to identify the traffic state of the roundabouts according to the headway changes in the inner and outer ring-road traffic flow to provide the time basis of signal control for traffic managers. Liu J et al. [38] presented a game model between a parking guidance system and users based on Stackelberg's multi-round game theory, including the profit function, game tree, and equilibrium conditions. However, the above study did not consider the influence of weather factors and surrounding condition [30].

Meanwhile, some researchers used other deep learning methods for the prediction of parking occupancy [39,40]. Zheng Y et al. [41] presented a parking occupancy prediction mechanism with feature datasets with selected parameters to illustrate these features, and analyz the comparative advantages of deep learning methods for parking occupancy prediction. Zhang C et al. [42] proposed a parking guidance model with a solution algorithm, and verified the effectiveness of the model and algorithm through simulation. Liu S [43] gave a weighted first-order local-region method to forecast unoccupied parking space based on the actual data and the chaotic time series.

Accurate and reliable short-term parking occupancy prediction are the most important factors of parking guidance and information systems. Recently, the most developed deep learning methods, such as recurrent neural network (RNN) and long-short term memory (LSTM), have shown great advantage in other fields' prediction. Zhang F et al. [30] presented a long-term car parking behavior prediction with a periodic weather-aware LSTM model with one hour, two hour, and three hour predictions. However, for a big city, the parking spaces in a parking lot are in higher demand and subject to frequent changes within short intervals, as seen in the Melbourne public dataset [44] and shown in Bourke Street Parking Lot in Figure 2. Most vehicles' arrival and departure duration are within 30 min, and the histogram shows a significant decline after 30 min. Therefore, it is particularly important to predict short-term car parking behavior.

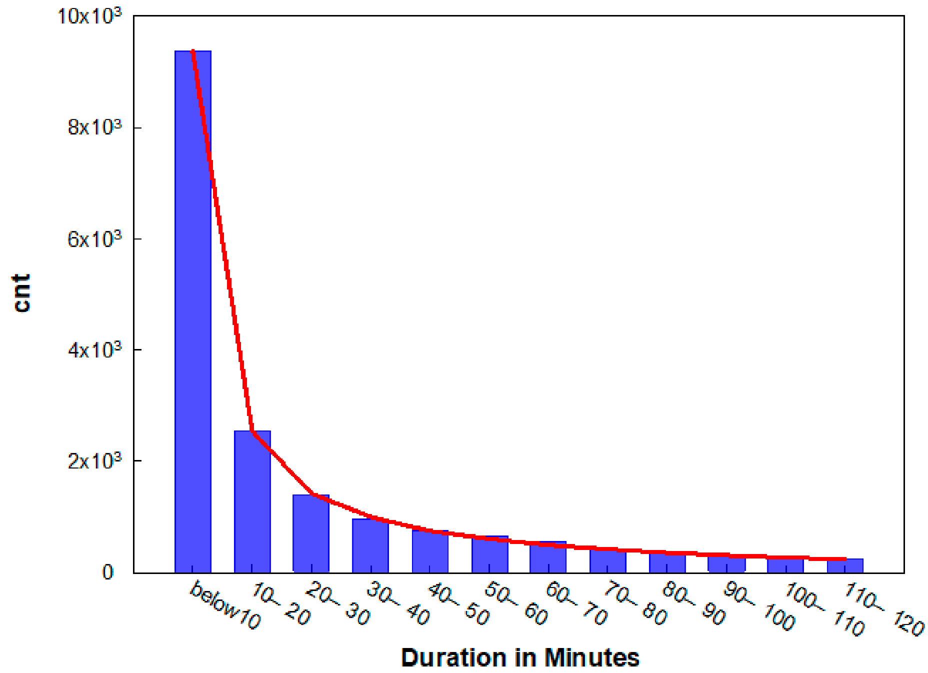

Figure 2. The number of parking with duration in minutes in a typical parking lot.

Based on the above literature review, this paper combines the advantages of LSTM (that is, it can process time series and weather features to improve training efficiency) and Melbourne's public data set [44], which includes parking history data (about 65,8900 parking records), weather information, and location information. Through location information, parking spaces in the block are divided, and specific parking lots can be simulated and predicted. The occupancy rate of parking spaces for 10 min, 20 min, and 30 min are predicted, respectively, to meet the demand for real-time parking in the business district.

3. Methodology

3.1. Research Framework

The framework of parking prediction in this paper is shown in Figure 3. The research framework consists of three parts, including data collection, methods, and evaluation.

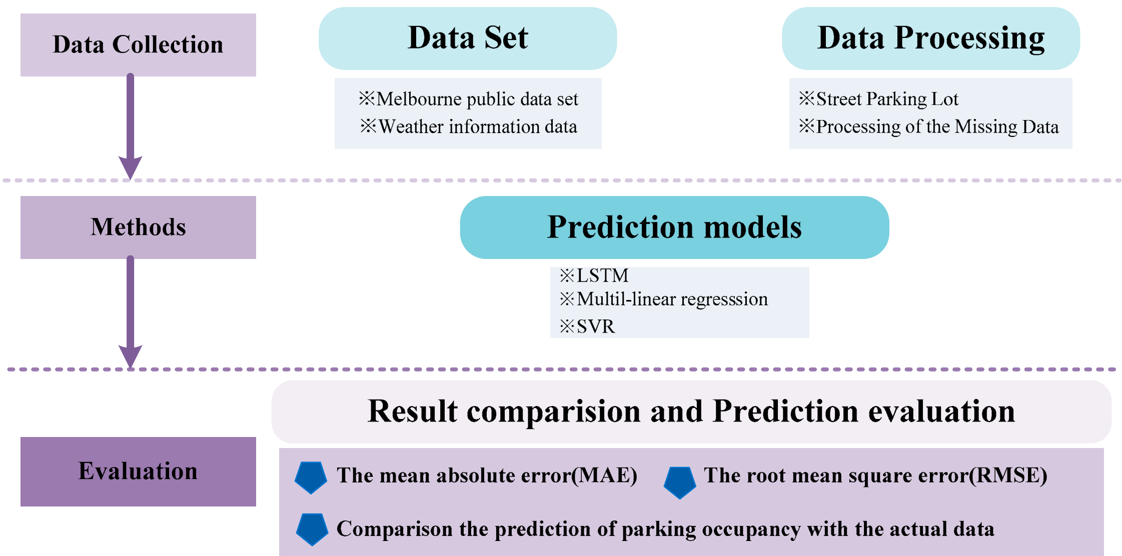

Figure 3. The research framework of short-term parking occupancy prediction.

3.2. Data Set and Data Information

The data set used in this paper is the Melbourne public dataset [44], since it is hard to obtain open parking data from official websites elsewhere. The data set is from 1 January 2017 to 31 December 2017. The minimum and maximum values of regional longitude are 144.9275761 and 144.9833373 respectively, while the minimum and maximum values of the regional latitude are −37.83126814 and −37.78907112. Zhao Z et al. [27] gave some geographical locations of placed sensors. The detailed information of placed sensors refer to reference [44]. From the data set, we selected the eight districts (streets) as parking spaces, and the details are shown in Table 1. The map of Melbourne in this region is shown in Figure 4. In Table 1, PL1-PL8 are labels of commercial parking lots. These 8 parking lots are located near the CBD of Melbourne, including Swanston Street, Collins Street, Flinders Lane, Elizabeth Street, Bourke Street, Little Bourke Street, William Street, and Queensberry Street. This area is home to retail, finance, tourism, entertainment, and educational establishments. These lots are busy every day with frequent traffic, and the parking lots are mainly of a commercial type. The detailed format of parking sensor data is shown in Table 2. The weather data set is used from Bureau of Meteorology, Commonwealth of Australia, including temperature, wind speed, pressure, and relative humidity [45]. The detailed format of weather data is shown in Table 3.

Table 1. The basic information of the eight parking lots.

Street Name	**Swanston Street**	**Collins Street**	**Flinders Lane**	**Elizabeth Street**	**Bourke Street**	**Little Bourke Street**	**William Street**	**Queensberry Street**
Label of parking lot	PL1	PL2	PL3	PL4	PL5	PL6	PL7	PL8
Area	University	Docklands/City Square	Twin Towers	China Town/ Victoria Market	Windsor	Windsor/Princes Theatre	Queensberry	Queensberry
Scale	31	111	31	61	195	48	169	60
Data sizes	12367	163334	30953	63900	197075	45372	112702	33253
Type	Commercial parking	Streetside Parking	Commercial Parking	Mixed Functional Parking	Commercial Parking	Commercial Parking	Commercial Parking	Commercial Parking

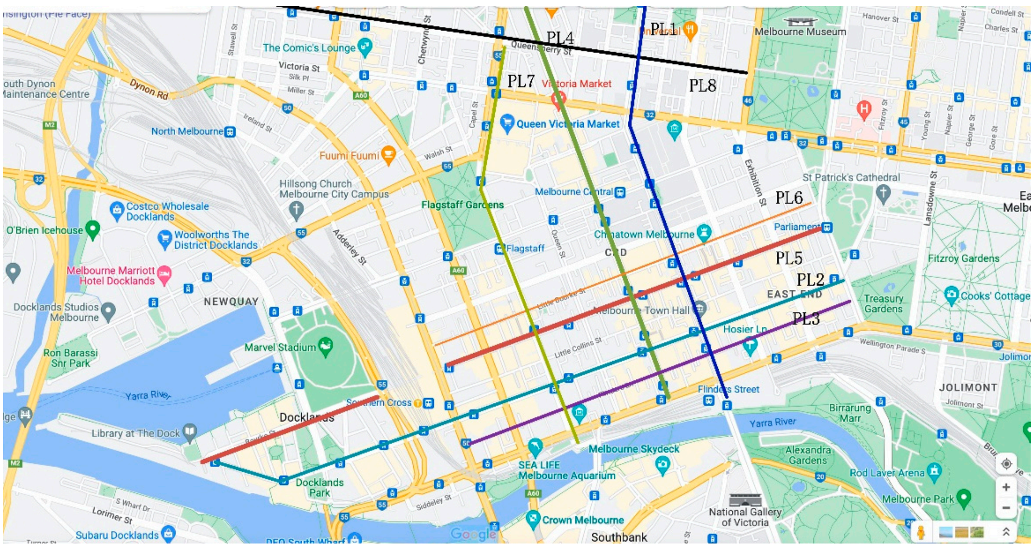

Figure 4. Map of Melbourne.

Table 2. Formats of parking data sets.

Features	Describe
Parking lots	Unique identification of parking space sensor
Arrival time	Date and time when the sensor detected that the vehicle arrives
Departure time	Date and time when the sensor detected that the vehicle departs
During time	Time difference between arrival time and departure time events, in seconds
Street name	Street signs for each group of parking spaces

Table 3. Formats of weather data sets.

Features	Describe
Temperature	°C
Wind	km/h
Pressure	mmHg
Relative humidity	Pa/Pa
Precipitation in the past hour	mm

The parking occupancy data was obtained through sensors. Due to the long working time of the sensor, some irresistible or human factors including data reading and writing errors may occur during the recording process. Therefore, it is necessary to preprocess the data, which will cause the loss of the collected data and reduce prediction accuracy. The mean filling method was used to deal with these missing and error values in this paper. The formula for calculating the average value of normal time in these days is shown in Equation (1).

$$\overline{m}(t) = \frac{1}{d}\sum\nolimits_{i=1}^{d} m_i(t) \quad (1)$$

where $\overline{m}$ (t) represents the average value in the time period of t, $m_i(t)$ represents the actual value in the time period of the previous days, and d represents the number of days in the time period.

At the same time, some parking lots have conducted time-series noise processing, such as turning on the sensors at about 7:30 a.m. and turning off the sensors at about 6.30 p.m. This means that the parking data of the parking lot changes greatly, which greatly interferes with the prediction of parking occupancy. Therefore, the first and last data of the day will be truncated.

3.3. Analysis of Experimental Data

The parking occupancy prediction, based on deep learning methods, aims to predict the future parking occupancy rate according to the historical parking data and contextual weather information. Because the time series is characterized by randomness, trend, periodicity, and time variability, it is necessary to analyze the data of the time characteristics, i.e., the parking occupancy rate per minute, to obtain the possible trend fluctuation and periodicity of the parking occupancy rate, and to analyze the influence of weather features and adjacent parking lots on the parking occupancy rate.

This paper randomly selected the parking change of the grouped parking lot in a week. The period of each data collection for an interval of 60 min in PL6 is shown in Figure 5. From Figure 5, we can see the changing trend of parking and the regularity of periodicity in the parking lot; the amount of parking on weekdays and non-weekdays conforms to the basic laws of people's activities: more parking during the day, less parking at night. Compared with the weekend, the number of parking occupancy in working days is relatively large. The possible reason maybe that the parking lot is close to the commercial center. People usually work from Monday to Friday, while they might go shopping on Saturday and Sunday; a small peak appears at 20:00 every night, indicating that people come out for dinner or get together during this period. Figure 5 also shows that the peak of the working day usually occurs at 8:00 a.m. and 6:00 p.m. The peak at 6 p.m. indicates that people are busy getting off work in the evening rush hour, driving from the company's parking lot to the roadside for dinner, shopping, or other activities. From 1:00 a.m. to 6:00 a.m., there are a few cars parked on the roadside.

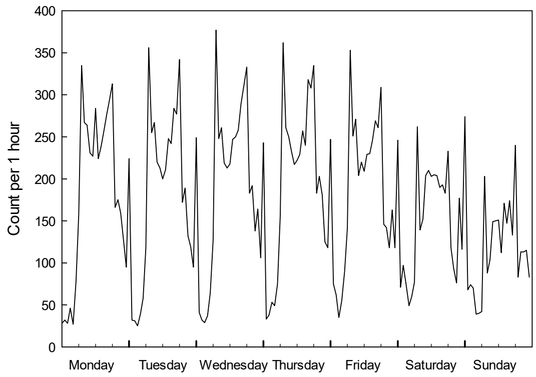

Figure 5. The periodicity trend of parking occupancy for the interval of 60 min in a week.

3.4. Weather Feature Analysis

In order to analyze the correlation of the weather data, in this part, we analyze the weather features, such as temperature, wind speed, pressure and relative humidity. First, we normalize the features using Equation (2), and then check for weather feature and Gaussian distribution. Figure 6 shows the correlation between the parking occupancy rate and the weather data. It can be seen that the temperature, wind speed, and parking occupancy rate are positively correlated, while humidity is negatively correlated with the parking occupancy rate. The air pressure has very little impact on the parking occupancy.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

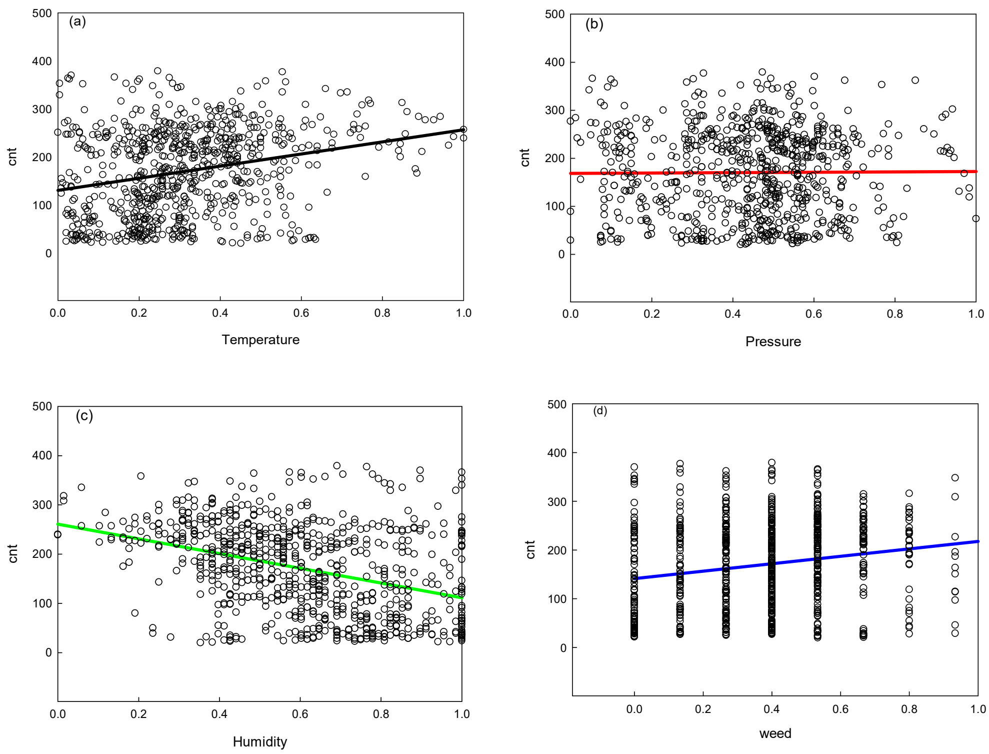

Figure 6. The regression trend for the weather features to cnt. (**a**) the correlation between the parking occu-pancy rate and the temperature; (**b**) the correlation between the parking occu-pancy rate and the pressure; (**c**) the correlation between the parking occu-pancy rate and the humidity; (**d**) the correlation between the parking occu-pancy rate and the wind speed.

Gaussian distribution, also known as normal distribution, is the mean value of all distributions' functions, approximating a normal distribution for large enough sample numbers. Figure 7 shows the distribution of these features accord with Gaussian distribution. By adding a fitting curve (density function curve) to the histogram, it was found that the normal curve is basically symmetrical and presents a "bell shaped" distribution, indicating that each attribute of the data basically meets the normal distribution. A small amount of data needs cleaning.

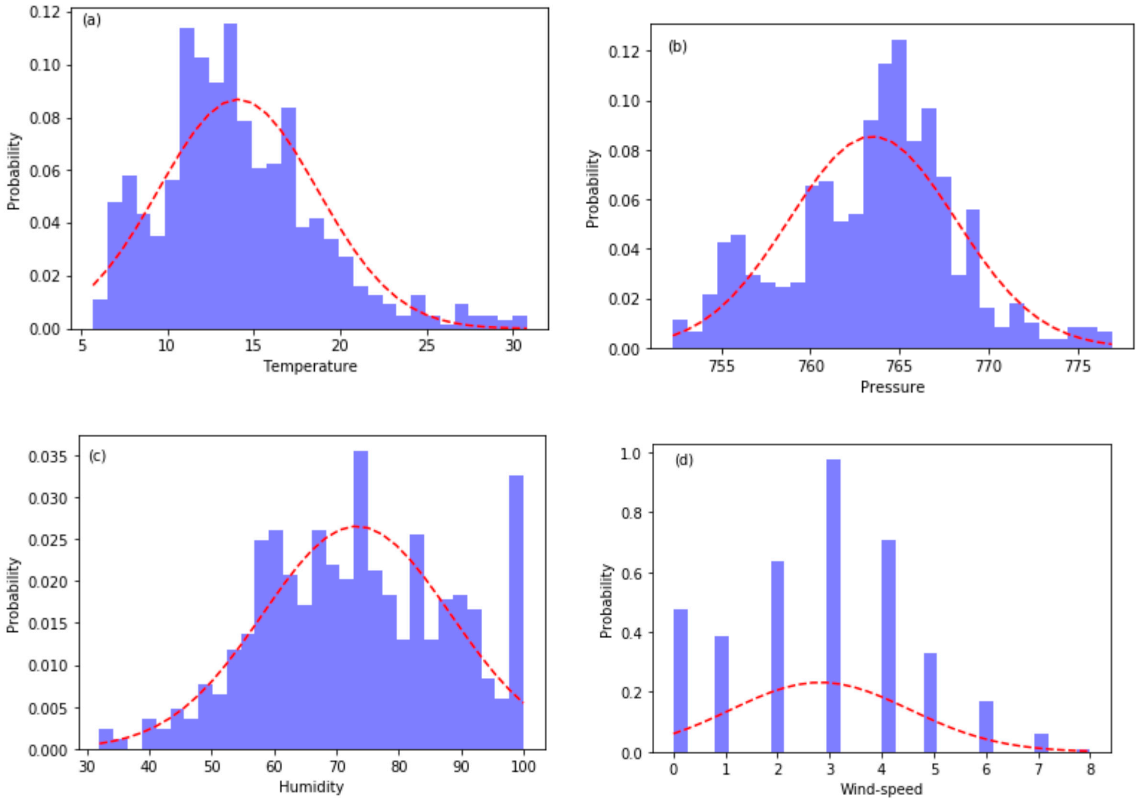

Figure 7. Features accord with Gaussian distribution. ((**a**). Temprature (**b**). Pressure (**c**). Humidity (**d**). Wind-speed).

3.5. LSTM Model Design

Traditional network neurons process data in parallel, and the input of the previous moment cannot affect the output of the next moment. The recurrent neural network (RNN) is a network structure with short-term memory for processing sequence data [46–49]. At the same time, the RNN framework can be used for sequence-to-sequence or sequence-tone learning [50,51].

LSTM is an improvement of RNN which can effectively alleviate the above problems [52]. LSTM introduces a gating mechanism to control the path of information transmission, including input gate, forgetting gate, and output gate [30]. The scheme of single-layer LSTM deployed in time sequence is shown in Figure 8. LSTM provides a direct channel from left to right for data flow, which can effectively avoid gradient disappearance and explosion. An LSTM cell is constructed with three gates, input gate i_t, forget gate f_t, and output gate o_t, and a memory cell c_t is used to store the current state of the LSTM block. The formulation of LSTM is presented in Equations (3)–(7).

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{3}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{4}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{5}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (6)$$

$$h_t = o_t \cdot tanh(c_t) \quad (7)$$

where x_t and h_t are the input and output vector of the current time point, respectively; W, U, and b are the weight and bias parameters of the block, and σ is the sigmoid function.

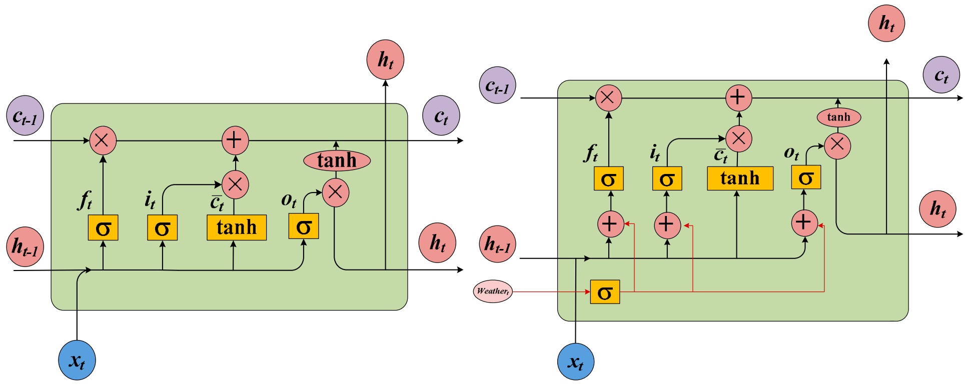

Figure 8. Internal structure of LSTM model and LSTM model with weather influence feature.

The overall structure of LSTM model design with weather influence feature is shown in Figure 8. Through considering weather factors, we compared the simple LSTM model and improved LSTM model. In fact, the parking behavior caused by travel behavior is often affected by weather factors. Therefore, the prediction of parking lot vacancies is determined by weather conditions [30], parking patterns, and previous parking records, as feature attributes. The weather vector is integrated with f_t, $i_{t,}$ and o_t. That is to say, the unit state (c_t) and hidden state (h_t) are not only affected by h_{t-1} and c_{t-1}, but also affected by weather (w_t). The LSTM algorithm has the advantage itself to extract the dependence of the time series to improve the prediction accuracy of the short-term parking occupancy rate, as following Equations (8)–(10).

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + V_f w_t + b_f) \quad (8)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + V_i w_t + b_i) \quad (9)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o w_t + b_o) \quad (10)$$

where V is the weight and bias parameters of the block for weather feature.

The experimental steps are as follows: (1) Data were extracted from eight parking lots around the business district. (2) The parking data of roadside parking occupancy were preprocessed, including missing value processing and noise processing, and then normalize all data. (3) The weights for each layer were initialized, and the preprocessed data were inputted into the LSTM model for training. (4) The characteristic data obtained from the training were predicted, including the parking occupancy rate for the intervals of 10 min, 20 min, and 30 min. (5) MAE and RMSE were calculated, the prediction of the parking occupancy rate was inputted.

3.6. The Evaluation Index

MAE is the mean absolute error, which represents the average of the absolute error between the predicted value and the observed value. MAE is a linear score. All individual differences have the same weight on average, which directly calculates the average of the residuals. RMSE is root mean square error, which represents the sample standard deviation of the residuals between the predicted value and the observed value. RMSE is used to explain the dispersion degree of the sample. For nonlinear fitting, the smaller the RMSE, the more accurate the model prediction is. Therefore, RMSE can punish more for high differences than MAE. MAE and RMSE are calculated as follows:

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}|obersered_t - predicted_t| \tag{11}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(oberved_t - predicted_t)^2} \tag{12}$$

4. Results and Discussion

4.1. Data Set Introduction and Analysis

Figure 9 shows the change curve of parking occupancy in 30 min intervals in a week for PL1 to PL8. When we superimpose the trend on the figure, it can be seen that each parking lot has the same trend for the week, except individual times are different, and the overall law is not affected; namely the change from Monday to Sunday is similar. From 0:00 a.m. to 6:00 a.m., the occupancy of parking lots can be negligible. The occupancy of parking lots is increased from 6 a.m., and there are still some differences between these parking lots. The occupancy rate of the parking lots peaked from 7:00 a.m. to 8:00 a.m. From 11:00 a.m. to 12:00 p.m., the occupancy rate of parking lots decreased slightly. The peak value is basically maintained within working hours (from 8:00 a.m. to 6:00 p.m.), generally reaching the upper limit of the parking lot, indicating that the utilization rate of parking spaces in various regions is high during this period. The occupancy rate of the parking lot reached a high value between 7:00 p.m. and 8:00 p.m., but did not exceed the rate of the working hours' occupancy. This means people need to go out for entertainment, parties, or dinner, which increases the utilization rate of street parking space during this period of time. The trend of parking occupancy on Saturdays and Sundays is similar to that on weekdays. The peak value on Saturday is slightly lower than that on weekdays, and the peak value on Sunday is lower than that on Saturday. Therefore the parking occupancy can be predicted using the deep learning method.

4.2. Comparison between Prediction Results and Actual Data for Four Models

Figure 10 shows the comparison between prediction results and actual data for the four models' evaluation results of PL5, in the intervals of 10 min, 20 min, and 30 min under the four models with MLR (in red), SVR (in green), LSTM (in dark red), and improved LSTM (in blue). Among the four prediction models, LSTM with weather influence feature performed the better for the parking data from PL5. From the comparison above, the improved LSTM model using the weather feature showed more accuracy in its prediction of parking occupation. The experimental results show that the MLR model is not fitting well in the time period when the wave peaks. The result of the SVR model indicates that it can only predict a relatively stable occupancy period, i.e., in a prediction interval of 10 and 20 min, and the model has a large error for intervals of 30 min. The simple LSTM model shows a periodic line change like a sine wave. It presents the same parking occupation trend every day, since it does not contain the weather feature.

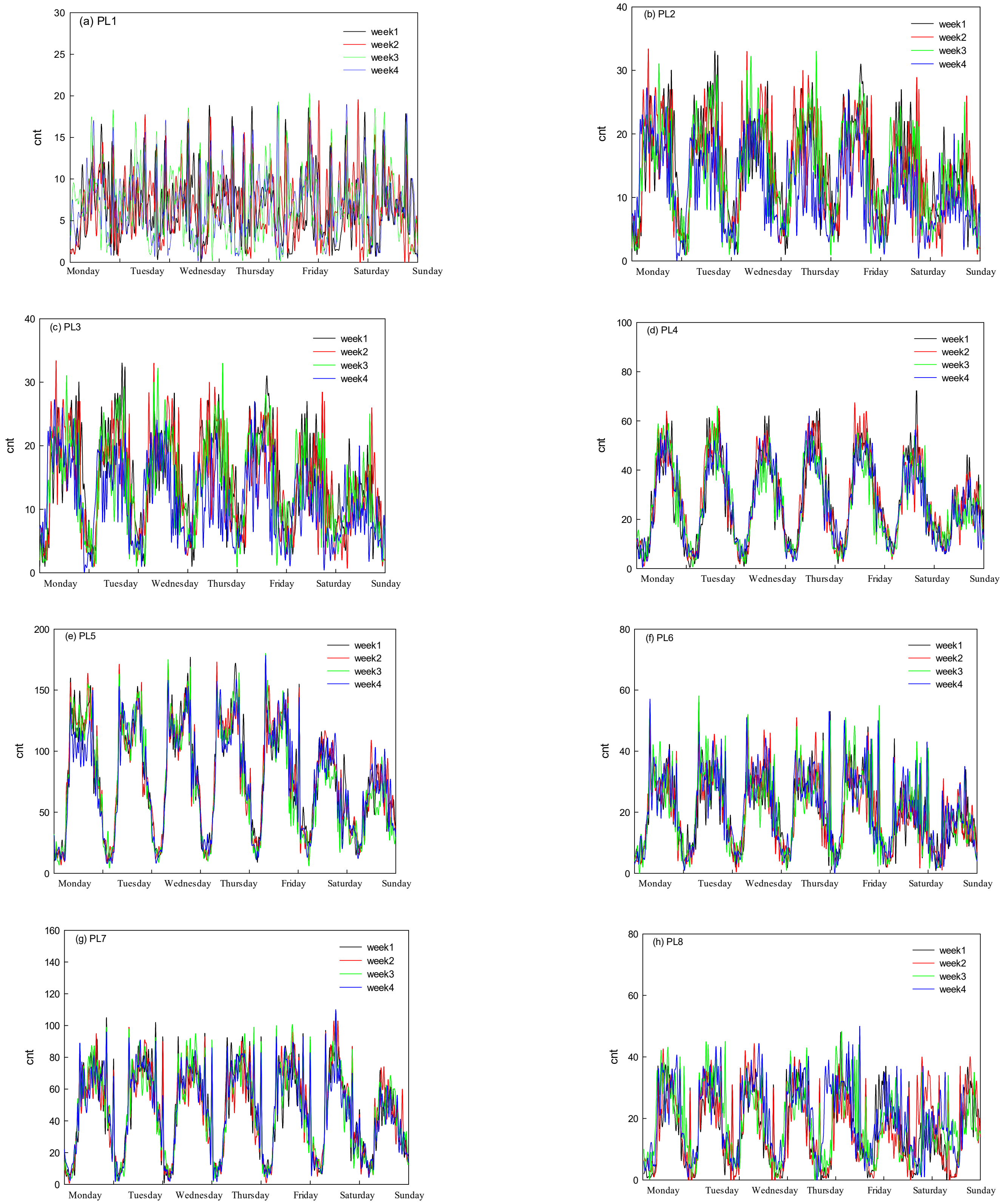

Figure 9. Relation between parking occupancy rate and time for four week.

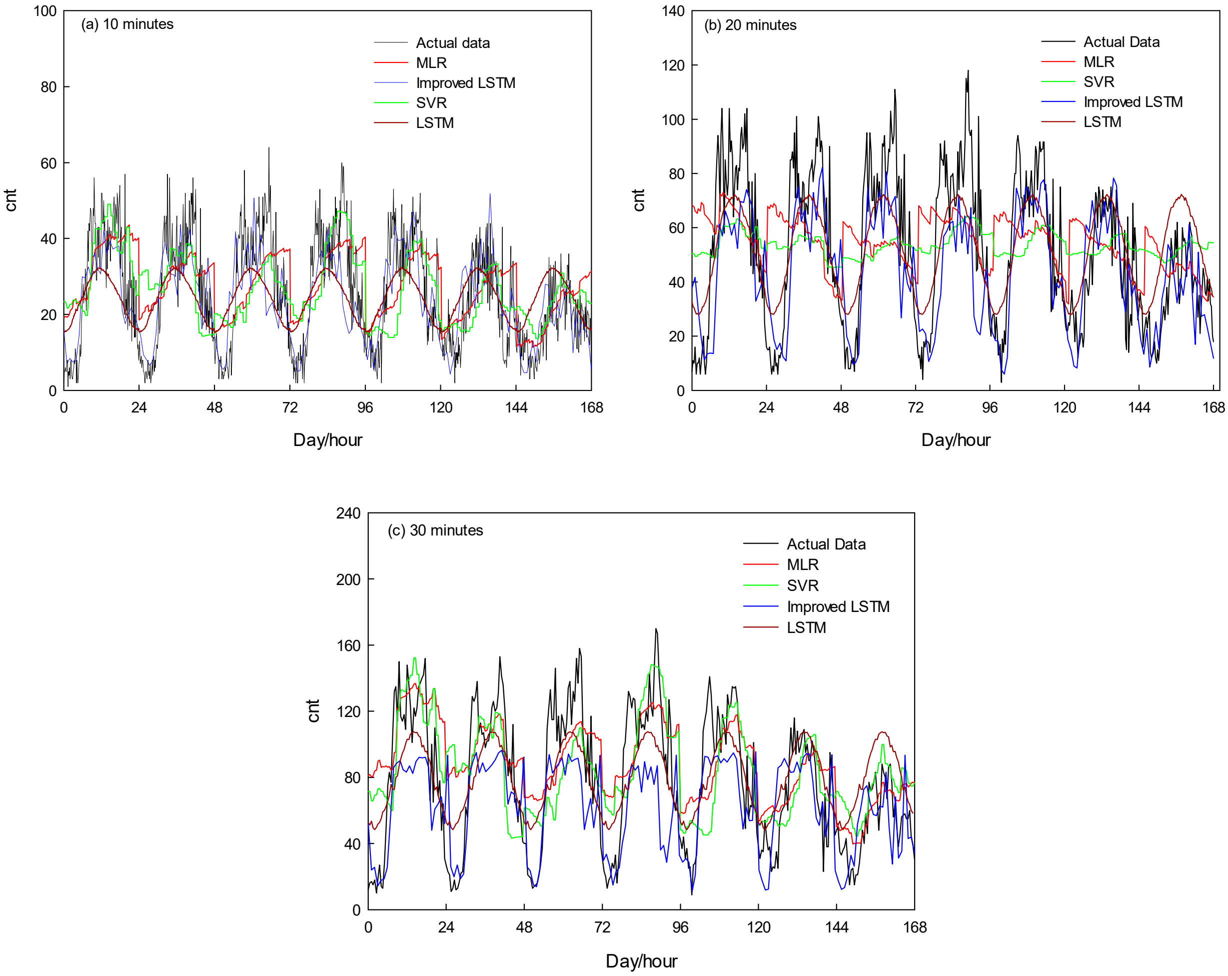

Figure 10. Evaluation results of PL5 using four models for 10 min, 20 min, and 30 min.

4.3. Comparison of Evaluation Index

Using MAE and RMSE as evaluation indicators, the improved LSTM with weather influence feature, simple LSTM, MLR, and SVR algorithms for all parking spaces are compared and analyzed. This paper predicts the effects of each parking lot, as shown in Figure 11. The eight block parking lots (PL1-PL8) in Figure 11 shows that the improved LSTM with weather influence feature algorithm has the better overall prediction effect. The comparison with simple LSTM and improved LSTM is shown in Table 4. From Table 4 we can see that the mean prediction accuracy of the improved LSTM model is higher than that of the simple LSTM model, indicating that the influence of weather on the parking prediction rate should be considered. The parking lot prediction effect is relatively better in MAE in intervals of 10 min, while the parking lot prediction effect needs to be improved in intervals of 30 min according to the MAE index; especially in PL1 and PL3, the MAE prediction error at the interval of 10 min is smaller. In the interval of 10 min, 20 min, and 30 min, the least prediction error of RMSE is still the LSTM with the weather influence feature algorithm proposed in this paper.

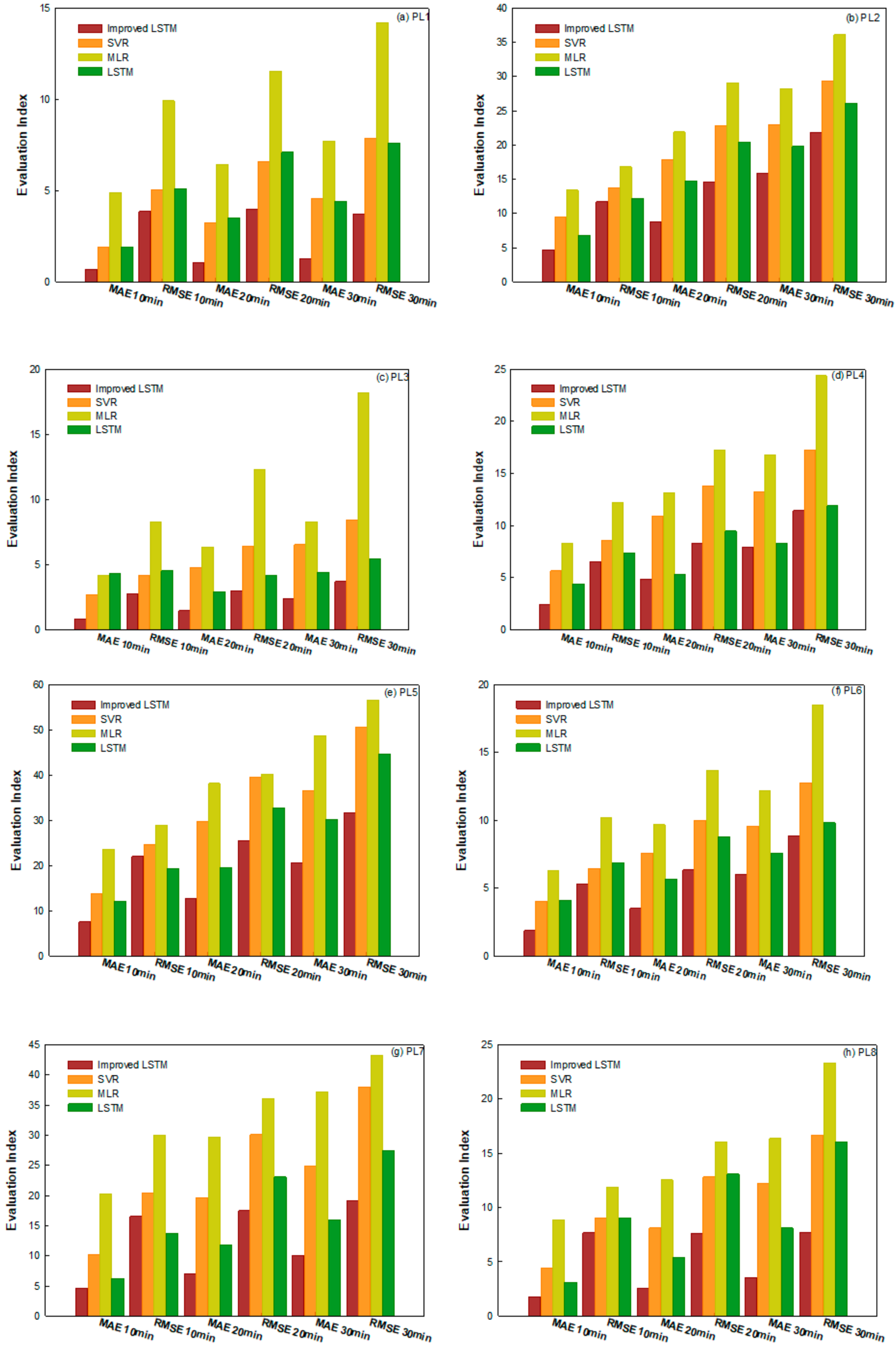

Figure 11. Comparison of evaluation index for PL1 to PL8.

Table 4. The comparison with simple LSTM and improved LSTM.

Models	Time Interval	Mean Accuracy
Simple LSTM	10 min	60.31%
	20 min	65.43%
	30 min	66.10%
Improved LSTM	10 min	77.08%
	20 min	78.28%
	30 min	77.15%

For the overall prediction effect of eight parking lots, the prediction results of parking lots in all blocks were applied in this paper, as shown in Figure 12. From Figure 12, the prediction results of MLR in the intervals of 10 min, 20 min, and 30 min are the worst. The main reason is that the MLR solving time series problem needs more stable data, while the data in this paper fluctuates relatively. Because the main feature of the prediction of parking occupancy rate is time series, LSTM performs nonlinear conversion through activation function. As it is a nonlinear model, LSTM captures the complex dependencies between sequences and is not limited to a single logic rule. It can find the expected prediction results with low requirements for data stability. In addition, the evaluation index of simple LSTM is better than MLR and MVR, but worse than that of improved LSTM. This proves that it is very important to consider the features of weather influence. The MAE of LSTM with weather influence feature in 10 min, 20 min and 30 min intervals follows 3.07, 5.27, and 8.48, and the RMSE is 9.55, 10.84, and 13.23, respectively. The average MAE of classic LSTM in 10 min, 20 min, and 30 min is 5.37, 8.61, and 12.36, and the RMSE is 9.76, 14.87, and 18.64. The average MAE of SVR prediction in 10 min, 20 min, and 30 min intervals follows 6.52, 12.75, and 16.30, and the RMSE is 11.54, 17.77, and 20.61, respectively. LSTM with the added weather influence feature model is obviously more accurate than simple LSTM and SVR model, which proves that the improved algorithm in this paper has the better effect when compared to other methods.

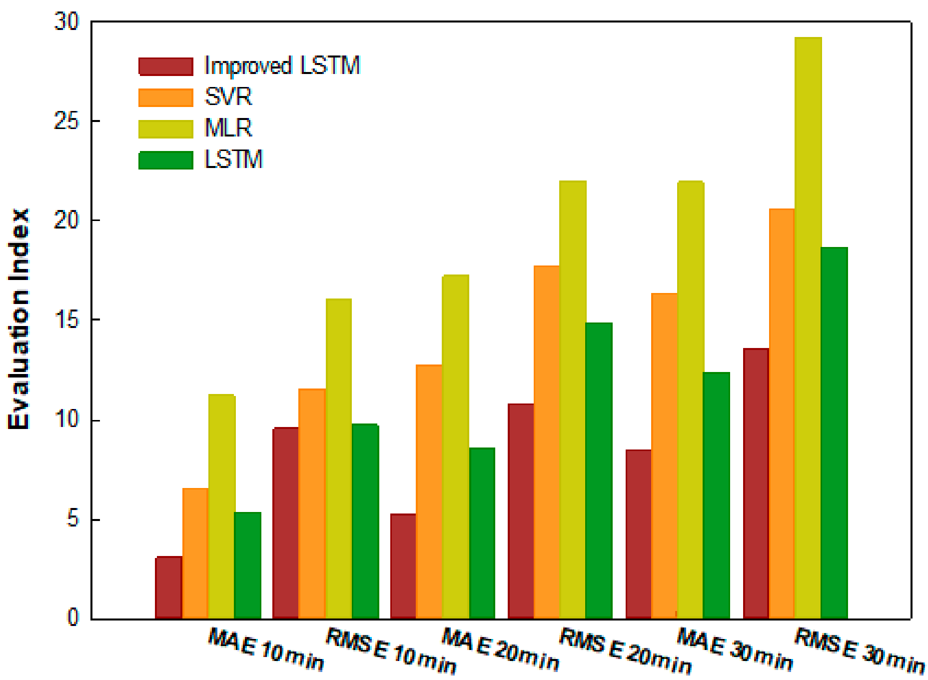

Figure 12. The average evaluation results of all parking lots for 10 min, 20 min, and 30 min.

5. Conclusions

The main purpose of this paper is to design, construct, and evaluate a prediction model of parking space availability, and estimate the parking occupancy rate of the parking lot (from about 658,900 parking records). Four machine learning and deep learning methods, i.e., improved LSTM, simple LSTM, MLR, and SVR, were compared and evaluated, and the short-term prediction model was obtained to achieve more accurate parking occupancy predictions with large-scale parking data. In addition, data processing and evaluation methods are presented, based on data from different parking lots. Considering the factors of real-time parking, the short-term prediction model for intervals of 10 min, 20 min, and 30 min is raised through improved LSTM, which reveals the spatial-temporal features of the available parking spaces around the business center in the big city. Then, using RSME and MAE as evaluation indexes, the accuracy of four models of parking occupancy rate for the intervals of 10 min, 20 min, and 30 min were evaluated. The MAE and RMSE of the improved LSTM with weather influence feature model at 10 min, 20 min, and 30 min were lower than that of the simple LSTM model, SVR algorithm, and MLR. Therefore, LSTM with an added weather influence feature model is obviously the most accurate, and has the best performance.

Short-term accurate prediction of parking demand is an important part of refined parking management. Especially in the process of parking guidance, accurate prediction of parking demand can greatly improve the effectiveness of information release and solve the dilemma that the information released by parking guidance is inconsistent with the actual information after the vehicle arrives at the parking lot. In future research, in addition to weather factors, the parking environment in different areas should also be further considered. There will be different characteristics around hospitals, schools, and residential communities. At the same time, other prediction methods in depth learning can be considered for cross-comparison analysis, and a combination prediction model can be built to improve the prediction accuracy.

Author Contributions: Writing—original draft, B.J.; methodology, formal analysis, Y.Z.; writing—review, formal analysis, methodology, J.N.; writing—review and editing, investigation, resources, formal analysis, methodology. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Beijing Natural Science Foundation (Project No.: 9123026) and Beijing Social Science Foundation, grant number (Project No.: 19XCB006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Chang, V.; Sharma, S.; Li, C.-S. Smart cities in the 21st century. *Technol. Forecast. Soc. Change* **2020**, *153*, 119447. [CrossRef]
2. Al-qudah, R.; Khamayseh, Y.; Aldwairi, M.; Khan, S. The smart in smart cities: A framework for image classification using deep learning. *Sensors* **2022**, *22*, 4390. [CrossRef] [PubMed]
3. Li, L.; Li, Y. Short-term prediction of parking demand for parking delicacy management. *J. Tongji Univ.* **2021**, *49*, 1301–1306.
4. Tudor, C.; Catalin, G.; Raluca, C.; Ciprian, N.; Ciprian, D. Chapter 1-Cloud Services for Smart City Applications. In *Intelligent Data-Centric Systems, Adaptive Mobile Computing*; Mauro, M., Alessio, M., Sherenaz, A.B., Eds.; Academic Press: Cambridge, MA, USA, 2017; pp. 3–28.
5. Zhu, L.; Yu, F.R.; Wang, Y.; Ning, B.; Tang, T. Big Data Analytics in Intelligent Transportation Systems: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 383–398. [CrossRef]
6. Paul, A.; Chilamkurti, N.; Daniel, A.; Rho, S. Chapter 2-Intelligent Transportation Systems. In *Intelligent Vehicular Networks and Communications*; Paul, A., Chilamkurti, N., Daniel, A., Rho, S., Eds.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 21–41.
7. National Bureau of Statistics. China's private car ownership over the years (1985–2020), 2022-08-01. Available online: https://m.shujujidi.com/hangye/106.html. (accessed on 1 September 2022).
8. Wang, L.; Ding, L.Q.; Wang, D. The research review of parking index in big cites. *J. Munic. Technol.* **2022**, *40*, 193–199.

9. Shu, S.-N.; Li, S.; Tu, Q.; Chen, G. Interpretation on DB11/T 1813-2020, Parking requirement standard for auto of public buildings. *China Stand.* **2022**, *2*, 2105–2112.
10. Gai, S.-X.; Zeng, X.-Q.; Yue, X.-Y.; Yuan, Z.-H. Parking guidance model based on user and system bi-level optimization algorithm. *J. Jilin Univ.* **2022**, *52*, 1344–1352.
11. Khanh, Q.V.; Hoai, N.V.; Manh, L.D.; Le, A.N.; Jeon, G. Wireless Communication Technologies for IoT in 5G: Vision, Applications, and Challenges. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 3229294. [CrossRef]
12. Inam, S.; Mahmood, A.; Khatoon, S.; Alshamari, M.; Nawaz, N. Multisource data integration and comparative analysis of machine learning models for on-street parking prediction. *Sustainability* **2022**, *14*, 7317. [CrossRef]
13. Stathopoulos, A.; Karlaftis, M.G. A multivariate state space approach for urban traffic flow modeling and prediction. *Transp. Res. C Emerg. Technol.* **2003**, *11*, 121–135. [CrossRef]
14. Rajabioun, T.; Loannou, P. On-street and off-street parking availability prediction using multivariate spatiotemporal models. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2913–2924. [CrossRef]
15. Tilahun, S.L.; Di Marzo Serugendo, G. Cooperative multiagent system for parking availability prediction based on time varying dynamic Markov chains. *J. Adv. Transp.* **2017**, *2017*, 1760842. [CrossRef]
16. Xu, B.; Wolfson, O.; Yang, J.; Stenneth, L.; Philip, S.Y.; Nelson, P.C. Real-Time Street Parking Availability Estimation. In Proceedings of the 2013 IEEE 14th International Conference on Mobile Data Management, Milan, Italy, 3–6 June 2013; Volume 1, pp. 16–22.
17. Yu, F.; Guo, J.; Zhu, X.; Shi, G. Real Time Prediction of Unoccupied Parking Space Using Time Series Model. In Proceedings of the 2015 International Conference on Transportation Information and Safety (ICTIS), Wuhan, China, 25–28 June 2015; pp. 370–374.
18. Ye, X.; Wang, J.; Wang, T.; Yan, X.; Chen, J. Short-term prediction of available parking space based on machine learning approaches. *IEEE Access* **2020**, *8*, 174530–174541. [CrossRef]
19. Dutta, N.; Charlottin, T.; Nicolas, A. Parking search in the physical world: Calculating the search time by leveraging physical and graph theoretical methods. *arXiv* **2022**, arXiv:2202.00258. [CrossRef]
20. Bock, F.; Di Martino, S.; Origlia, A. A 2-Step Approach to Improve Data-Driven Parking Availability Predictions. In Proceedings of the 10th ACM SIGSPATIAL Workshop on Computational Transportation Science, Redondo Beach, CA, USA, 7 November 2017; pp. 13–18.
21. Fan, J.; Hu, Q.; Tang, Z. Predicting vacant parking space availability: An SVR method with fruit fly optimisation. *IET Intell. Transp. Syst.* **2018**, *12*, 1414–1420. [CrossRef]
22. Xiao, J.; Lou, Y.; Frisby, J. How Likely am I to Find Parking?—Stochastic Models of Parking Process and Probabilistic Estimation of Parking Availability. In Proceedings of the Transportation Research Board 96th Annual Meeting, Washington, DC, USA, 8–12 January 2017.
23. Shen, J. Research on Application of Data Mining Technology in Urban Roadside Parking Management. Master's Thesis, Zhejiang University, Hangzhou, China, 2020.
24. Ji, Y.; Tang, D.; Guo, W.; Blythe, T.; Wang, W. Forecasting available parking space with largest Lyapunov exponents method. *J. Cent. South Univ.* **2014**, *21*, 1624–1632. [CrossRef]
25. Yin, C.; Shao, C.; Wang, X. Influence of urban built environment on car commuting considering parking availability. *J. Jilin Univ.* **2020**, *50*, 115–120.
26. Amini, M.; Karabasoglu, O.; Ilic, M.; Boroojeni, K.G.; Iyengar, S.S. ARIMA-Based Demand Forecasting Method Considering Probabilistic Model of Electric Vehicles' Parking Lots. In Proceedings of the 2015 IEEE Power & Energy Society General Meeting, Denver, CO, USA, 26–30 July 2015; pp. 1–5.
27. Zhao, Z.; Zhang, Y.; Zhang, Y. A comparative study of parking occupancy prediction methods considering parking type and parking scale. *J. Adv. Transp.* **2020**, *10*, 5624586. [CrossRef]
28. Shao, W.; Zhang, Y.; Guo, B.; Qin, K.; Chan, J.; Salim, F.D. Parking Availability Prediction with Long Short Term Memory Model. In *Lecture Notes in Computer Science, Proceedings of the Intenational Conference on Green, Pervasive, and Cloud Computing. GPC 2018, Hangzhou, China, 11–13 May 2018*; Li, S., Ed.; Springer: Cham, Switzerland, 2019; p. 11204.
29. Feng, Y.; Xu, Y.; Hu, Q.; Krishnamoorthy, S.; Tang, Z. Predicting vacant parking space availability zone-wisely: A hybrid deep learning approach. *Complex Intell. Syst.* **2022**, *8*, 4145–4161. [CrossRef]
30. Zhang, F.; Feng, N.; Liu, Y.; Yang, C.; Zhai, J.; Zhang, S.; He, B.; Lin, J.; Du, X. PewLSTM: Periodic LSTM with Weather-Aware Gating Mechanism for Parking Behavior Prediction. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence {IJCAI-PRICAI-20}, Yokohama, Japan, 7–15 January 2020; pp. 4424–4430.
31. Feng, N.; Zhang, F.; Lin, J.; Zhai, J.; Du, X. Statistical Analysis and Prediction of Parking Behavior. In *Lecture Notes in Computer Science, Proceedings of the IFIP International Conference on Network and Parallel Computing, Hohhot, China, 23–24 August 2019*; Springer: Cambridge, UK, 2019; pp. 93–104.
32. Liu, F.; Hao, F.; Hao, J.; Zhou, Y.; Xin, G. Parking prediction algorithm based on optimized LSTM model. *J. Comput. Appl.* **2019**, *39* (Suppl. 1), 65–69.
33. Samaranayake, P.; Gunawardana, U. Parking Assessment in the context of growing construction activity and infrastructure changes: Simulation of impact scenarios. *Sustainability* **2022**, *14*, 5098. [CrossRef]
34. Chen, H.; Tu, X.; Wang, Y.; Zhen, J. Short-term parking space prediction based on wavelet-ELM neural networks. *J. Jilin Univ.* **2017**, *55*, 388–392.

35. Rong, Y.; Xu, Z.; Yan, R.; Ma, X. Du-Parking: Spatio-Temporal Big Data Tells You Realtime Parking Availability. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining 2018, London, UK, 19–23 August 2018; 2018; pp. 646–654.
36. Zhang, W.; Liu, H.; Liu, Y.; Zhou, H.; Xiong, H. Semi-supervised hierarchical recurrent graph neural network for city-wide parking availability prediction. *arXiv* **2019**, arXiv:1911.10516.
37. Lu, K.; Lin, M.; Deng, X.; Xu, G.; Xu, J. A dynamic allocation and guidance model for parking spaces with minimum total parking costs. *J. South China Univ. Technol.* **2018**, *46*, 82–91+98.
38. Liu, J.; Zhu, H.; Peng, L. Real-time parking guidance model basd on game between system and user. *Comput. Simul.* **2019**, *36*, 166–170.
39. Jelen, G.; Podobnik, V.; Babic, J. Contextual prediction of parking spot availability: A step towards sustainable parking. *J. Clean. Prod.* **2021**, *5*, 127684. [CrossRef]
40. Yang, S.; Ma, W.; Pi, X.; Qian, S. A deep learning approach to real-time parking occupancy prediction in spatio-temporal networks incorporating multiple spatio-temporal data sources. *Transp. Res. Part C: Emerg. Technol.* **2019**, *107*, 248–265. [CrossRef]
41. Zheng, Y.; Rajasegarar, S.; Leckie, C. Parking Availability Prediction for Sensor-Enabled car Parks in Smart Cities. In Proceedings of the 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Singapore, 7–9 April 2015; pp. 1–6.
42. Zhang, C.; Wang, X.; Ma, Z. Free parking space prediction and reliability analysis based on big data analysis. *IEEE Access* **2020**, *8*, 66609–66614. [CrossRef]
43. Liu, S.; Guan, H.; Yan, H.; Yin, H. Unoccupied Parking Space Prediction of Chaotic Time Series. In Proceedings of the Tenth International Conference of Chinese Transportation Professionals (ICCTP), Beijing, China, 4–8 August 2010; pp. 2122–2131.
44. City of Melbourne Open Data, 2019. Available online: https://data.melbourne.vic.gov.au/ (accessed on 1 November 2019).
45. Bureau of Meteorology, Commonwealth of Australia. Available online: http://www.bom.gov.au/climate/dwo/IDCJDW3050.latest.shtml (accessed on 1 September 2022).
46. Fokker, E.S.; Koch, T.; van Leeuwen, M.; Dugundji, E.R. Short-term forecasting of off-street parking occupancy. *Transp. Res. Rec.* **2022**, *2676*, 637–654. [CrossRef]
47. Cho, K.; van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
48. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
49. Chung, J.; Gulcehre, C.; Cho, K.; Bengo, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
50. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *arXiv* **2014**, arXiv:1409.3215.
51. Gers, F.A.; Schraudolph, N.N.; Schmidhuber, J. Learning precise timing with LSTM recurrent networks. *J. Mach. Learn. Res.* **2002**, *3*, 115–143.
52. Lei, T.; Zhang, Y.; Wang, S.; Dai, H.; Artzi, Y. Simple recurrent units for highly parallelizable recurrence. *arXiv* **2017**, arXiv:1709.02755.