applied sciences

MDPI

Article

Hourly Long-Term Traffic Volume Prediction with Meteorological Information Using Graph Convolutional Networks

Sangung Park [1], Mugeun Kim [2] and Jooyoung Kim [2,*]

[1] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA; park1202@purdue.edu
[2] Department of Transportation Planning & Management, Korea National University of Transportation, Chungju 27469, Republic of Korea; swq0105@naver.com
* Correspondence: jykim@ut.ac.kr; Tel.: +82-31-460-0614

Abstract: Hourly traffic volume prediction is now emerging to mitigate and respond to hourly-level traffic congestion augmented by deep learning techniques. Incorporating meteorological data into the forecasting of hourly traffic volumes substantively improves the precision of long-term traffic forecasts. Nonetheless, integrating weather data into traffic prediction models is challenging due to the complex interplay between traffic flow, time-based patterns, and meteorological conditions. This paper proposes a graph convolutional network to predict long-term traffic volume with meteorological information. This study utilized a four-year traffic volume and meteorological information dataset in Chung-ju si to train and validate the models. The proposed model performed better than the other baseline scenarios with conventional and state-of-the-art deep learning techniques. Furthermore, the counterfactual scenarios analysis revealed the potential negative impacts of meteorological conditions on traffic volume. These findings will enable transportation planners predict hourly traffic volumes for different scenarios, such as harsh weather conditions or holidays. Furthermore, predicting the microscopic traffic simulation for different scenarios of weather conditions or holidays is useful.

Keywords: traffic volume prediction; graph convolutional networks; meteorological information

Citation: Park, S.; Kim, M.; Kim, J. Hourly Long-Term Traffic Volume Prediction with Meteorological Information Using Graph Convolutional Networks. *Appl. Sci.* **2024**, *14*, 2285. https://doi.org/10.3390/app14062285

Academic Editor: Tomasz Figlus

Received: 30 January 2024
Revised: 3 March 2024
Accepted: 4 March 2024
Published: 8 March 2024

1. Introduction

Traffic volume prediction plays a pivotal role in managing traffic conditions. Previous studies have explored short-term traffic volume prediction [1–3]. The high complexity of traffic volume prediction has hindered researchers from analyzing daily or hourly volume traffic volume predictions. Nonetheless, short-term traffic volume prediction cannot fundamentally improve hourly or daily specific travel demand management since the hourly traffic volume is highly dependent on time-dependent information, the weather, and holidays. In the era of big data and real-time information technology, high-resolution traffic volume prediction can potentially mitigate hourly traffic congestion. That is because previous researchers have analyzed short-term traffic flow prediction [1–3].

Long-term traffic volume prediction is emerging for weekly and monthly traffic demand management and precise long-term traffic volume prediction. Long-term traffic volume prediction provides hourly traffic volumes for several weeks or months. Compared to short-term traffic prediction, long-term traffic volume prediction is affected by various factors, such as traffic accidents, holiday indicators, and weather information. This is because long-term traffic prediction accumulates larger errors than short-term traffic prediction.

Long-term traffic volume prediction considering various factors has been studied. Previous studies have used statistical and state–space models to predict hourly traffic volumes. The autoregressive integrated moving average algorithm (ARIMA) is a famous statistical model for predicting overall time series data [4–6]. Recent progress in neural

network methodologies has opened new avenues for traffic volume prediction. These contemporary techniques, especially those driven by neural networks, have shown promising results in predicting traffic speeds across road networks [7]. Neural network models can incorporate diverse data types, including meteorological conditions, such as precipitation, temperature, and humidity, to improve predictions. Integrating weather data into traffic volume prediction models could significantly enhance their precision and reliability. Recent works have used the multi-layer perception network (MLP) [8], long short-term memory (LSTM) [9], graph convolution network (GCN) [10,11], and graph attention network (GAN) [7] to predict traffic volumes. Despite the important role of and progress in neural network applications, the intensive study of long-term traffic volume prediction incorporating meteorological data remains limited. This gap is attributable to the intricate interactions among traffic flows, temporal patterns, and weather conditions.

Exploring long-term traffic volume prediction has been limited, primarily due to challenges associated with data acquisition. Precise long-term forecasting requires datasets spanning over a year to capture weekly, monthly, seasonal, and holiday impacts adequately. Previous studies have utilized up to 2 years of datasets for traffic volume or speed predictions [12] to explain seasonal effects carefully.

The main aim of this study was to determine whether meteorological information significantly advances long-term traffic volume prediction. The contributions of this study are three-fold. First, this paper proposes a temporal GCN-based approach to one-year traffic volume prediction by the one-week traffic volume prediction format to analyze the impacts of temporal dependencies and meteorological information. Second, we propose a neural network structure for a prediction performance that is better than that of other types in previous studies. Third, this study used meteorological information to predict the impact of weather information.

The outline of this paper is as follows. Section 2 presents some related studies on traffic volume prediction. Section 3 describes datasets on traffic volume and meteorological information. Section 4 presents the proposed model's framework, describing the graph's structure and layers. Section 5 shows the training and prediction results. Section 6 summarizes the findings of this study and proposes directions for the next steps.

2. Literature Review

The ARIMA is a widely used approach to forecasting time series data [1,13,14]. It is acceptable for forecasting the stationary time series data. Despite its widespread use, the ARIMA's applicability to traffic volume forecasting is limited due to the non-stationary nature of traffic data. The non-stationary nature can exhibit unpredictable variations influenced by external traffic flows. With the help of machine learning techniques, researchers can explore faster and finer traffic predictions with a variety of ranges [15]. Recent deep learning techniques for traffic volume prediction first considered the MPN [8] and the recurrent neural network [16]. This deep learning technique contains several hidden, fully connected layers to predict nonlinear and nonstationary relationships. The LSTM is a recurrent neural network model to strengthen the long-term and short-term memories in a neural network [9]. The LSTM is powerful in predicting periodic time-series data. The GCN considers a graph structure to identify the interdependence between data [10]. It is powerful to consider the interdependence between data. It is widely used in time-series data prediction or spatial dependence [17]. The GAN uses an attention mechanism in the GCN [7]. The attention mechanism focuses on the most related features from the input data. The attention mechanism has been utilized in various areas due to its flexibility and high predictability with dependencies. For the systematic reviews on traffic volume prediction with artificial intelligence, please refer to this article [4].

This paper intensively looks into GCN approaches most widely used to predict long-term traffic volume. The GCN is one of the emerging techniques to account for the interdependence between data using a graph structure. Previous studies related to the GCN can be classified into two types: temporal and spatial interdependence. First, temporal

interdependence is considered as temporal blocks such as temporal self-attention [11], an LSTM module [18], or a gated recurrent unit (GRU) module [19].

On the other hand, previous works have used the GCN to identify the spatial dependence between traffic volume or speed data. Table 1 represents the literature reviews on the short-term traffic volume predictions with the GCN. First, Qin et al. (2020) utilized the GCN with meteorological information to predict bike trips [12]. Qin et al. (2020) focused on spatial dependence and covered more than one year. Second, Bai et al. (2019) also described a similar model but with multiple parallel layers in the model to predict the bike-sharing volumes [20]. They considered three different cities to predict the traffic volume. Three to six months have been extensively utilized to train the model. Li et al. (2020) utilized the Manhattan area to predict the traffic volume [2]. They formed the main and minor areas and subjected the information to deep learning in advance [2]. The last study utilized the GCN and attention mechanisms unlike the GAN. The model performed better with spatial dependence.

Table 1. Literature Review.

Author (Year)	Methodology	Spatial Scope	Weather	Spatial Dependence	Temporal Scope
Qin et al. (2020) [12]	GCN	Boston in the U.S.	O	O	406 days
Bai et al. (2019) [20]	GCN, STG2Seq	Shenyang and Beijing in China and NY in the U.S.	X	X	3 to 6 months
Li et al. (2020) [2]	GCN	Manhattan in the U.S.	X	X	1 year
Jiang et al. (2023) [10]	PDFormer (GCN + attention)	Beijing in China	X	O	2 months to 1 year
This study	GCN	Chung-ju si in South Korea	O	X	4 years

From the literature review section, we find out that there is no work on long-term traffic volume prediction and no work considering both meteorological information and temporal dependence. Furthermore, no study has used data spanning more than two years to identify the yearly impacts and the other impacts. Therefore, this study aimed to develop long-term traffic volume prediction techniques to utilize meteorological information and data spanning more than 2 years to obtain stable results.

3. Data Descriptions

The dataset used in this study encompasses a comprehensive collection of information relevant to traffic volume and weather conditions within the spatial scope of the Korea University of Transportation, Chung-ju campus, from 2019 to 2022. Each row in the dataset represents an hourly observation, providing a detailed view of various attributes, including the week day, traffic volume, temperature, rainfall, wind speed, wind direction, humidity, air pressure, hour, week, and holiday indicators. We trimmed the dataset to align all row observations to be the same as the other datasets. Therefore, the trimmed dataset covered a period from 7 January 2019 to 29 December 2019, a period from 6 January 2020 to 27 December 2020, a period from 4 January 2021 to 26 December 2021, and a period from 3 January 2022 to 25 December 2022. Note that this dataset contained the COVID-19 period, such as 2020 to 2021. We used the 2019, 2020, and 2021 datasets to train the model, and we used the 2022 dataset to validate and test the model.

We collected this publicly available meteorological information from the Open MET Data Portal (https://data.kma.go.kr/ (accessed on 1 March 2024)) using the Automated Synoptic Observing System (ASOS) and publicly available hourly traffic volume from Traffic Monitoring Systems (https://road.re.kr (accessed on 1 March 2024)). The spatial scope of these data is limited to the Korea University of Transportation, Chung-ju campus. This dataset offers a comprehensive view of year-round traffic and weather patterns. The temporal granularity of this dataset allows for an in-depth examination of how these patterns change over different hours and days.

Figure 1 represents the traffic information. Figure 1a,c illustrate the weekly average traffic volumes at the upper and lower bounds, respectively, from 2019 to 2022. The datasets imply a seasonal trend, with heightened traffic volumes during the summer months (from week 22 to week 40) compared to those in winter (from week 0 to week 10 and from week 49 to week 52), suggesting a correlation between traffic patterns and seasonal variations.

Figure 1b,d depict the hourly average traffic volumes at the upper and lower bounds for the same period. These figures elucidate a consistent hourly pattern over 2019 to 2022 marked by a morning peak at around 7 a.m. at the upper bound (Figure 1b) and an evening peak at around 5 p.m. at the lower bound (Figure 1d). This bimodal distribution indicates the conventional rush hours associated with morning and evening commutes.

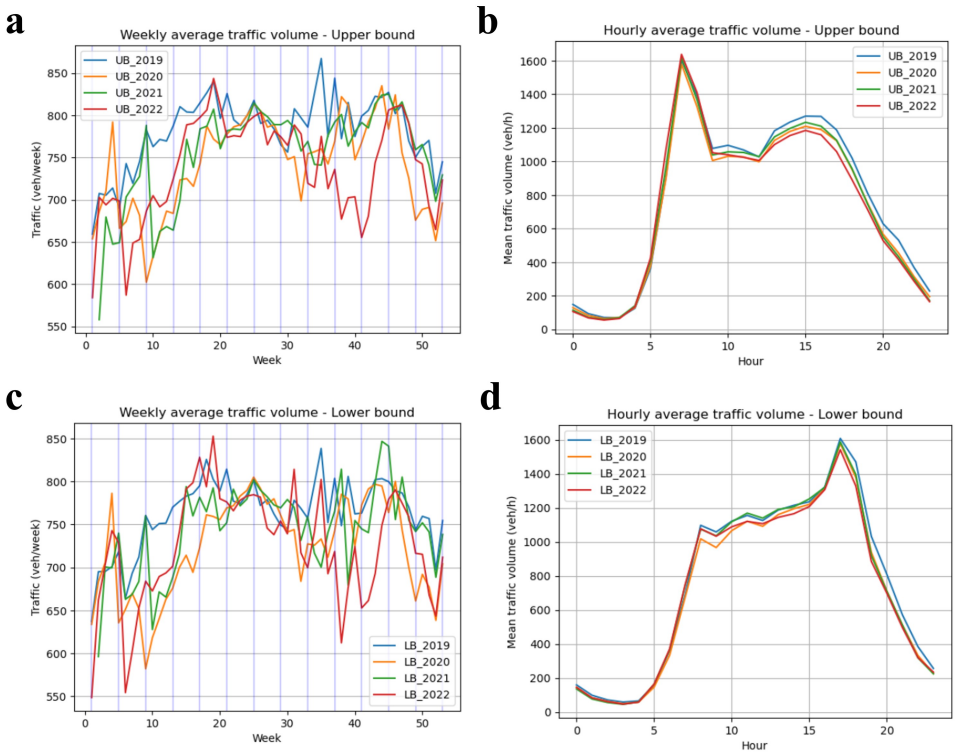

Figure 1. Temporal patterns of traffic information from 2019 to 2022.

Figure 2a presents the average weekly air temperature from 2019 to 2022, representing a consistent annual pattern with a pronounced peak around week 31, indicative of the climatic zenith typically observed during the midsummer period. Figure 2b illustrates the average weekly rainfall within the same period. The data reveal a notable peak in precipitation around week 33, implying a temporal concentration of rainfall corresponding to regional and seasonal weather patterns. Figure 2c depicts the weekly wind speed averages over the same time period. The graphical representation indicates a relatively stable wind speed pattern irrespective of the annual variations. Figure 2d visualizes the average weekly air pressure from 2019 to 2022. A peak of weekly air pressure is observed at around week 31, aligning with atmospheric pressure changes, often identifying the transition between summer and autumn.

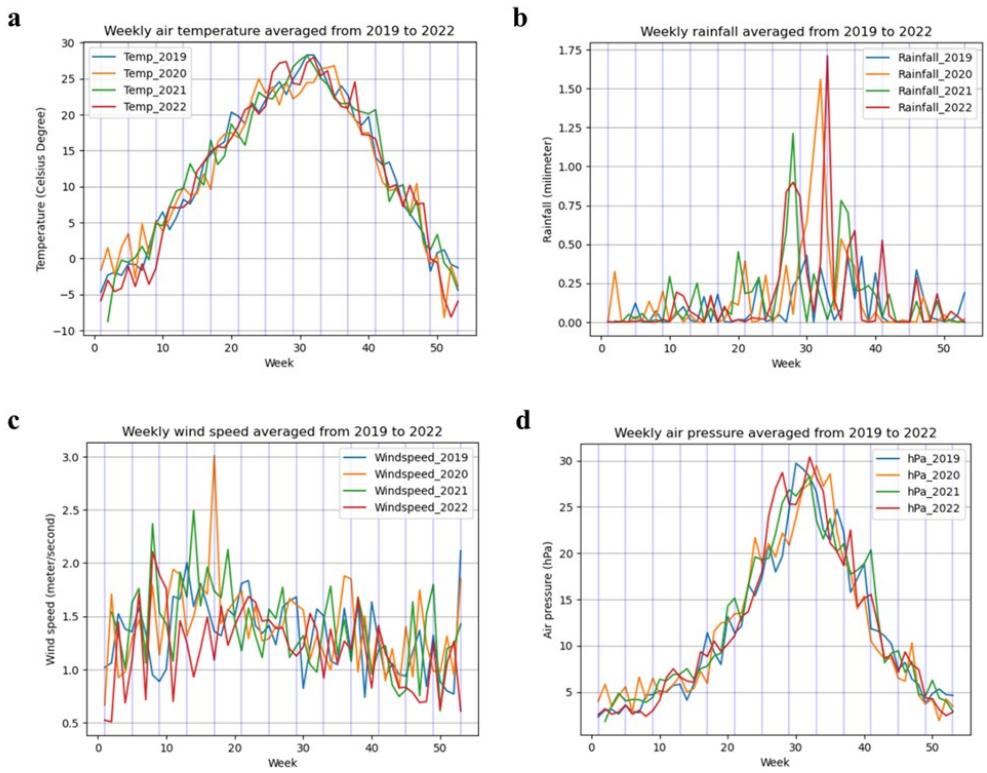

Figure 2. Weekly meteorological information from 2019 to 2022.

Table 2 represents the correlation coefficients between the variables. Traffic volumes in both directions have a strong correlation. Specifically, the lower-bound traffic volume is positively associated with wind speed and direction and negatively associated with humidity. Note that associations between the upper-bound traffic volume and the other variables are not stronger than those of the lower bound traffic volume.

Table 2. A correlation matrix.

Pearson Correlations	Traffic—UB	Traffic—LB	Temp.	Rainfall.	Wind Speed	Wind Direction	Humidity	hPa
Traffic—UB	1.00	0.80	0.20	−0.02	0.19	0.23	−0.28	0.05
Traffic—LB	0.80	1.00	0.28	−0.02	0.33	0.37	−0.48	0.04
Temp.	0.20	0.28	1.00	0.07	0.18	0.13	0.00	0.86
Rainfall	−0.02	−0.02	0.07	1.00	0.05	−0.00	0.13	0.13
Wind speed	0.19	0.33	0.18	0.05	1.00	0.50	−0.49	−0.03
Wind direction	0.23	0.37	0.13	−0.00	0.50	1.00	−0.48	−0.07
Humidity	−0.28	−0.48	0.00	0.13	−0.49	−0.48	1.00	0.40
hPa	0.05	0.04	0.86	0.13	−0.03	−0.07	0.40	1.00

4. Methods

4.1. GCN

The GCN is a type of neural network with a graph structure in each node so that the graph can utilize the graph structure's information. Compared to the original neural network, the GCN interacts with other nodes as a graph structure. It has been widely used in social networks and natural language processing. In this study, the p-th layer and $p+1$-th layer are propagated through the following equation:

$$\mathbf{X}^{p+1} = \sigma(\hat{\mathbf{D}}^{-\frac{1}{2}}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}^{p}\mathbf{W}^{p}), \tag{1}$$

where $\mathbf{X}^p$ is the p-th $|\mathcal{N}| \times F^p$ feature matrix; p is the number of hidden layers; $|\mathcal{N}|$ is the cardinality of the number of nodes $\mathcal{N}$; F^p is the number of features of each node of the

p-th layer; $\sigma(\cdot)$ is a rectified linear unit (ReLU) function, one of the widely-used activation functions in the neural network techniques; $\hat{\mathbf{D}}$ is a diagonal node degree matrix of $\hat{\mathbf{A}}$; $\hat{\mathbf{A}}$ is a sum of adjacency matrix $\mathbf{A}$; $\mathbf{W}^p$ is the p-th $F^p \times F^{p+1}$ weight matrix with trainable parameters; and identity matrix $\mathbf{I}$ is to consider self-features of each node. Note that $\hat{\mathbf{D}}^{-\frac{1}{2}}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}^p$ represents the symmetric approximation of the matrix $\hat{\mathbf{A}}\mathbf{X}^p$.

The overall data structure of this study is visualized in Figure 3. Each node of a deep learning approach has a graph structure containing 168 nodes. Each node of the graph represents 1 h. For example, 168 nodes are interconnected in two ways. The first way is the consideration of hourly interdependence. For instance, a node from 0 to 1 a.m. and another node from 1 to 2 a.m. are connected. Furthermore, a node from 0 to 1 a.m. of day 0 and a node from 0 to 1 a.m. of day 1 are connected. Therefore, the graph structure contains 311 nodes representing a 1-week graph structure.

The overall graph structure is different from that in previous studies. Instead of focusing on the spatial dependencies, this study focused on the temporal dependencies between the nodes representing every hour. We found strong temporal dependencies, as shown in Figure 1. Holidays were the exception to the strong dependencies, so we considered holidays as a node attribute. Spatial dependencies were not considered since we only targeted one detector point with two directions.

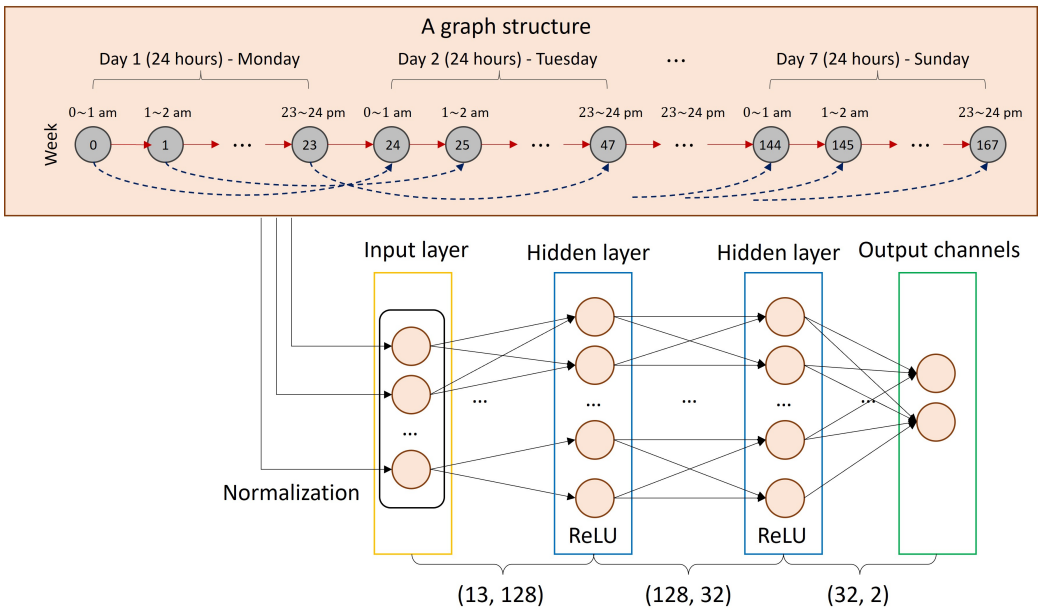

Figure 3. Schematic descriptions of the graph structure and the graph neural network.

4.2. Scenario Setup

In this study, we set up a scenario to quantify the meteorological information. A baseline scenario is the real dataset in 2022. We add the 10-millimeter rains for Friday and Saturday of the first week of 2022.

5. Results

We utilized the Pyg [21] and Pytorch package in the Python environment to develop and train the deep learning model [22].

5.1. Performance Metrics

We utilized mean squared error (MSE) and mean absolute percent error (MAPE) to evaluate the proposed model. First, MSE calculates a sum of the squared differences between the actual and predicted values.

$$MSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n} |\hat{Q}_i - Q_i|}, \tag{2}$$

where $\hat{Q}_i$ is the i-th predicted traffic volume, and Q_i is the i-th actual value.

Second, MAPE denotes a sum of the absolute percentages of the differences between the actual and predicted values. A low MAPE means that the actual values are close to the predicted values.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{|\hat{Q}_i - Q_i|}{Q_i}, \tag{3}$$

5.2. Hyper Parameters

The dropout rates were set to 30% to avoid the overfitting of the model. We tried a grid search, dropout rates (0.2, 0.3, and 0.4), learning rate (0.01, 0.001, 0.0001), and training epoch (2000, 4000, 6000, 8000) and chose the best combination. Finally, we used a learning rate of 0.01 and a training epoch of 6000 with varying batch sizes. The number of layers was 4, including the input and output layers. All baselines and proposed models had the same graph structure. The sequence of nodes in each layer was 13, 128, 32, and 2. The number of input nodes was 13, and the number of output nodes was 2.

5.3. Baselines

The MLP and GAN were used as baselines. The MLP is a simple deep learning network with multiple fully connected layers. We used three layers with the same structure as that of the GCN. The GAN is an advanced version of the GCN with attention mechanisms. The structure of the GAN is the same as that of the GCN.

5.4. Training Results

Table 3 shows the model training results. We tested six different types of models with or without weather information. The training results indicate that the GCN with meteorological information performed better in the MSE of UB, and the GCN in the MSE of LB. When we compared two models, models with meteorological information and models without meteorological information, models without meteorological information performed better. On the other hand, the GCN performed worst in the dataset without meteorological information. This is mainly because the model did not recognize the difference between upper and lower bounds in this case. Figure 4 visualizes the model prediction results. Figure 4a,b represent the first week of 2022 with the upper and lower bounds, respectively. Green lines, representing the proposed model, perform similarly to red and yellow lines, representing the MLP and GAN.

Table 3. Model training results.

No.	Method	Dropout Probability	Meteorological Information	MSE-UB	MAPE-UB	MSE-LB	MAPE-LB
1	MLP	-	O	18,690.69	0.228	14,901.66	0.264
2	MLP	0.3	O	16,731.72	0.236	20,039.50	0.266
3	GCN	0.3	O	16,001.57	0.234	16,278.24	0.286
4	GAN	0.3	O	44,036.84	0.33	38,494.36	0.398
5	MLP	-	X	14,744.30	0.167	13,110.84	0.172
6	MLP	0.3	X	16,558.00	0.279	17,039.96	0.302
7	GCN	0.3	X	24,410.24	0.319	130,680.7	0.673
8	GAN	0.3	X	59,472.62	0.956	38,152.25	0.735

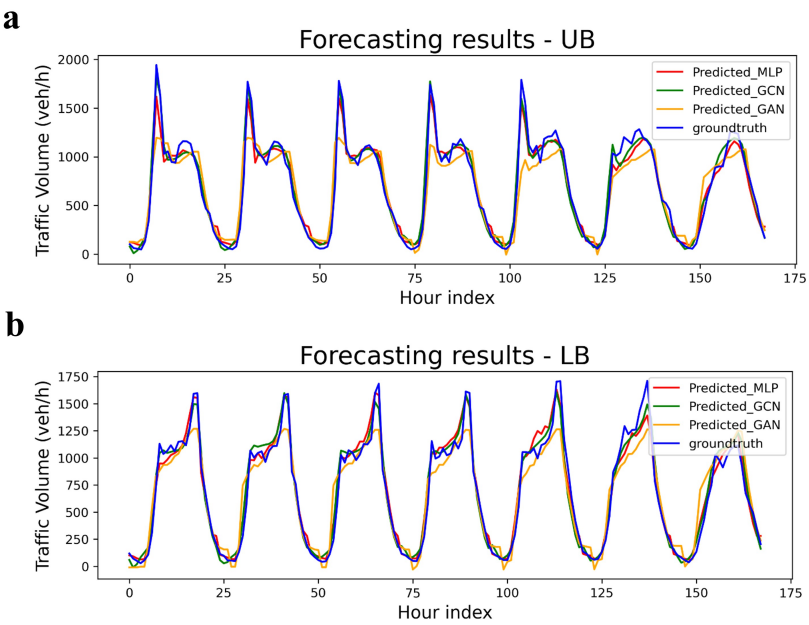

Figure 4. Model prediction results.

Heatmaps of differences between observed and predicted results over 60 days are expressed in Figure 5. Our proposed model performed the best over 60 days, except for a period from days 28 to 34. These dates contained three consecutive holidays (Lunar New Year) in South Korea. Note that morning peaks of the period accounted for 80.4% of the predicted MSE.

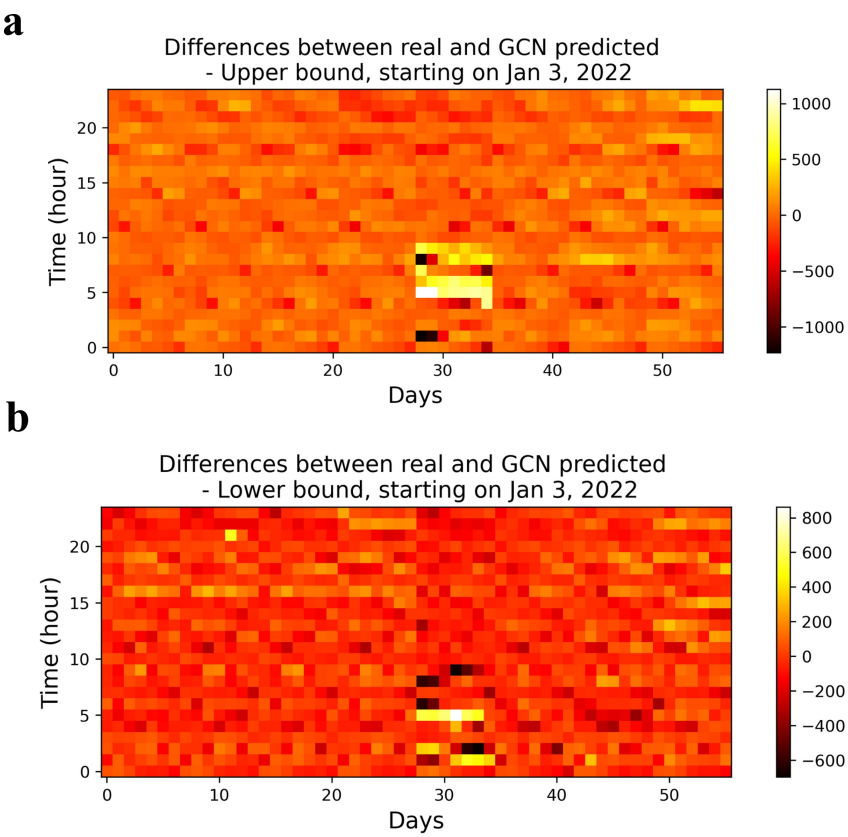

Figure 5. Heatmaps of differences between observed and GCN-predicted results on upper bound (**a**) and lower bound (**b**) starting on 3 January 2022.

5.5. Scenario Results

We changed the meteorological information manually to quantify the impact of meteorological information. In the treatment scenario, we added 10 millimeters of rainfall to every hour on days 5 and 6 (from hour index 120 to 167). Figure 6 represents the scenario result and the base scenario result. The MSE between the treatment and the base scenario was 1834.19 on the upper bound and 2254.38 on the lower bound. Specifically, in Figure 6, the predicted traffic volume of the treatment results is lower on the higher and lower bounds on Saturday than the base scenario results. On the other hand, the predicted traffic volume of the treatment results is higher on Sunday than the base scenario results. We can interpret this to mean that the impacts of small changes in meteorological information are not larger than the externalities, such as travel behavior changes.

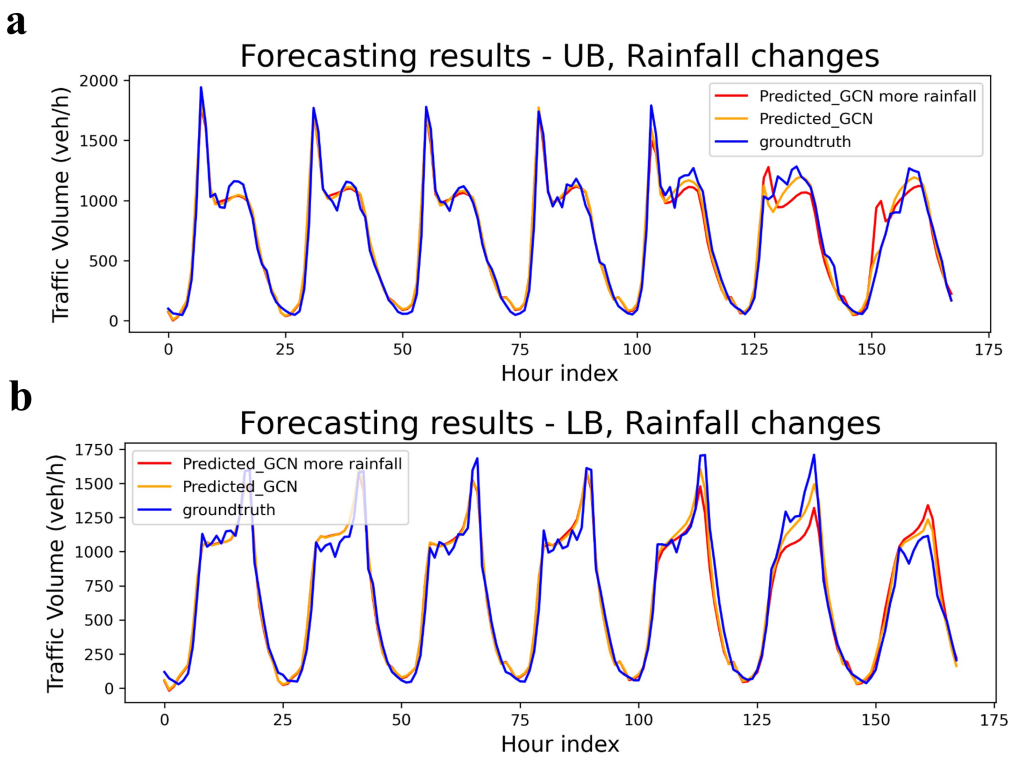

Figure 6. Scenario results.

6. Discussion

Long-term traffic volume prediction is an emerging area that can efficiently manage long-term travel demand management with meteorological prediction. This study analyzed long-term traffic volume prediction with meteorological information using the graph neural network. We tested the proposed model with eight different baselines to check the impact of the MLP, the GAN, dropouts, and meteorological information. The results showed that the proposed model (GCN) performed better than the MLP case. The counterfactual scenario results indicated that this model admits the traffic volume changes by changing meteorological information in both directions.

This study contains two main findings: (1) outperformance of GCN and (2) variability of meteorological information.

6.1. Outperformance of GCN

The findings in Table 2 indicate that the GCN with meteorological information and dropout strategies has a better prediction than the other models. The table explains that the accuracy of the GAN and MLP was less than that of the GCN. The MLP without dropout rates performed better than the GCN, but the MLP without dropouts had overfitting issues, which cannot be generalizable to the other detectors. When we consider the future direction

of the models, the GCN emerges as a promising alternative. The GCN successfully predicted stable traffic volumes with upper and lower bounds. However, the study also identified the limitation of the attention mechanism. The large weights of the neighbors of a given node make it challenging for the attention mechanism to capture temporal dependencies accurately. This suggests that a strong reliance on the interconnections between neighboring nodes may detrimentally impact the model's overall predictive capability.

The GCNs in this study successfully captured the differences between weekdays and weekends, as shown in Figure 5. However, the models did not identify the impact of consecutive holidays compared to a single holiday. In South Korea, consecutive holidays usually affect traffic patterns due to cultural differences and travel behaviors. If we identified these holidays as an indicator, the model could perform better than the proposed model.

6.2. Disparate Impacts of Meteorological Information

In this study, we quantified the traffic volume prediction with meteorological information. We expected that when we add the meteorological information, the models, especially the GCN, would perform better. Contrary to our expectations, the study showed that the models integrating meteorological data yielded results similar to those that did not integrate the meteorological data. These results imply that the influence of meteorological information spanning four years on traffic volume prediction is not as significant as expected compared to other external factors affecting the model. On the other hand, the counterfactual scenarios imply that when rainfall increases, traffic volume insignificantly drops. It implies a negative associations between rainfall and traffic volume. This observation can be supported by a previous study [23].

6.3. Future Research Directions

We identified the outperformance of the GCN and the disparate impacts of meteorological information in long-term traffic volume prediction with meteorological information. We utilize these two outcomes to suggest future research directions. The first research direction is to manage the travel demand. When we have special events, we can assess the hourly traffic volume impact of one specific location with weather prediction. This can be a pinpoint travel demand strategy, such as proactive ramp metering or hourly heterogeneous signal control strategies.

The other research direction is that this model can help transportation planners create a synthetic traffic volume dataset for analyzing microscopic traffic simulations. Common microscopic traffic simulations necessitate the use of path-specific traffic volumes. However, these methods cannot provide precise estimates without a path-based traffic assignment. We can extend the hourly link-based traffic volume prediction to the hourly-level path-based traffic volume due to the methodological similarity. Furthermore, we can predict the traffic volumes during harsh weather conditions or holidays. This can help transportation planners set up plans to mitigate the impact of harsh weather or holidays.

6.4. Limitations

There are three limitations of this study. First, we did not account for the consecutive holidays that South Korea and East Asia usually have. During consecutive holidays, such as the Lunar New Year Day (3 days, without weekends, in January or February) and Korean Thanksgiving Day (3 days, without weekends, in September or October), the proposed model did not account for the drastic patterns. We can add the consecutive holiday indicators independently and check the impact of this indicator. Second, this model did not address the spatial interdependence between neighboring detectors due to the lack of a high-quality dataset. We could not gather data from an hourly-level detector for four years. If we could use the complete dataset, we could also consider the spatial dependencies of this model. Third, we encountered the inherent prediction error in traffic volume predictions. Compared to the O/D-based traffic volume prediction, this model did

not utilize vehicle routes. Vehicle routes can be more extensively used to clearly prove the accuracy of long-term traffic volume prediction.

7. Conclusions

This study proposes a graph neural network framework to predict long-term traffic volume With meteorological information. This study predicted the hourly long-term traffic volume for one week with a GCN trained on data spanning three years, from 2019 to 2021, and validated with the 2022 dataset. The results of the proposed model showed that it maintained precision and robustness regardless of the weekend and holiday indicators by high dropout rates. Furthermore, the counterfactual scenarios of the high rainfall during two days showed the negative impact of rainfall on the hourly traffic volume. These findings show that the model can help transportation planners predict hourly traffic volumes for different scenarios, such as harsh weather conditions or holidays. Furthermore, the model is useful for predictions in microscopic traffic simulations of different scenarios of weather conditions or holidays. This study sheds light on using meteorological information for long-term traffic volume prediction. It will help us simulate various types of traffic volume prediction scenarios with harsh weather conditions or holidays.

Author Contributions: Conceptualization, S.P. and M.K.; methodology, S.P.; software, S.P.; validation, S.P. and J.K.; resources, M.K.; data curation, S.P.; writing—original draft preparation, S.P.; writing—review and editing, S.P.; visualization, S.P.; supervision, J.K.; project administration, J.K.; funding acquisition, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 23AMDP-C161756-02).

Data Availability Statement: We used two types of time-series datasets. The first, the traffic volume dataset, can be found at Traffic Monitoring Systems (https://road.re.kr (accessed on 1 March 2024)). The second, the meteorological dataset, is available at the Open MET Data Portal (https://data.kma.go.kr/(accessed on 1 March 2024)) using the Automated Synoptic Observing System (ASOS).

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

References

1. Shang, Q.; Lin, C.; Yang, Z.; Bing, Q.; Zhou, X. A hybrid short-term traffic flow prediction model based on singular spectrum analysis and kernel extreme learning machine. *PLoS ONE* **2016**, *11*, e0161259. [CrossRef] [PubMed]
2. Li, A.; Axhausen, K.W. Short-term traffic demand prediction using graph convolutional neural networks. *AGILE GIScience Ser.* **2020**, *7*, 12. [CrossRef]
3. Wang, Y.; Zhao, L.; Li, S.; Wen, X.; Xiong, Y. Short-term traffic flow prediction of the urban road using time-varying filtering based empirical mode decomposition. *Appl. Sci.* **2020**, *10*, 2038. [CrossRef]
4. Shaygan, M.; Meese, C.; Li, W.; Zhao, X.G.; Nejad, M. Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities. *Transp. Res. Part Emerg. Technol.* **2022**, *145*, 103921. [CrossRef]
5. Song, Z.; Guo, Y.; Wu, Y.; Ma, J. Short-term traffic speed prediction under different data collection time intervals using a SARIMA-SDGM hybrid prediction model. *PLoS ONE* **2019**, *6*, e0218626. [CrossRef] [PubMed]
6. Wang, Y.; Jia, R.; Dai, F.; Ye, Y. Traffic flow prediction method based on seasonal characteristics and SARIMA-NAR model. *Appl. Sci.* **2022** *12*, 2190. [CrossRef]
7. Chughtai, J.U.R.; Haq, I.U.; Muneeb, M. An attention-based recurrent learning model for short-term travel time prediction. *PLoS ONE* **2022**, *12*, e0278064. [CrossRef]
8. Qu, L.; Li, W.; Li, W.; Ma, D.; Wang, Y. Daily long-term traffic flow forecasting based on a deep neural network. *Expert Syst. Appl.* **2019**, *121*, 304–312. [CrossRef]
9. Yao, H.; Tang, X.; Wei, H.; Zheng, G.; Li, Z. Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction. *Aaai Conf. Artif. Intell.* **2019**, *33*, 5668–5675. [CrossRef]
10. Jiang, J.; Han, C.; Zhao, W.X.; Wang, J. PDFormer: Propagation Delay-aware Dynamic Long-range Transformer for Traffic Flow Prediction. *arXiv* **2023**, arXiv:2301.07945.
11. Chang, Z.; Liu, C.; Jia, J. STA-GCN: Spatial-Temporal Self-Attention Graph Convolutional Networks for Traffic-Flow Prediction. *Appl. Sci.* **2023**, *13*, 6796. [CrossRef]
12. Qin, T.; Liu, T.; Wu, H.; Tong, W.; Zhao, S. Resgcn: Residual graph convolutional network based free dock prediction in bike sharing system. In Proceedings of the 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 30 June–3 July 2020.

13. Lepage, S.; Morency, C. Impact of weather, activities, and service disruptions on transportation demand. *Transp. Res. Rec.* **2021**, *2675*, 294–304. [CrossRef]
14. Chand, S. Modeling predictability of traffic counts at signalised intersections using Hurst exponent. *Entropy* **2021**, *23*, 188. [CrossRef] [PubMed]
15. Sekula, P.; Marković, N.; V.; er Laan, Z.; Sadabadi, K.F. Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transp. Res. Part Emerg. Technol.* **2018**, *97*, 147–158. [CrossRef]
16. Khan, Z.; Khan, S.M.; Dey, K.; Chowdhury, M. Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction. *Transp. Res. Rec.* **2019**, *2673*, 489-503. [CrossRef]
17. Yi, Z.; Liu, X.C.; Markovic, N.; Phillips, J. Inferencing hourly traffic volume using data-driven machine learning and graph theory. *Comput. Environ. Urban Syst.* **2021**, *85*, 101548. [CrossRef]
18. Fu, R.; Zhang, Z.; Li, L. Using LSTM and GRU neural network methods for traffic flow prediction. In Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 11–13 November 2016.
19. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; Li, H. T-gcn: A temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 3848–3858. [CrossRef]
20. Bai, L.; Yao, L.; Kanhere, S.; Wang, X.; Sheng, Q. Stg2seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting. *arXiv* **2019**, *5*, arXiv:1905.10069.
21. Fey, M.; Lenssen, J.E. Fast graph representation learning with PyTorch Geometric. *arXiv* **2019**, arXiv:1903.02428.
22. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*.
23. Lam, W.H.; Tam, M.L.; Cao, X.; Li, X. Modeling the effects of rainfall intensity on traffic speed, flow, and density relationships for urban roads. *J. Transp. Eng.* **2013**, *139*, 758–770. [CrossRef]