Transfer learning for transportation system resilience estimation using floating car data

A thesis presented in part fulfilment of the requirements of the Degree of Master of Science in Transportation Systems at the Department of Civil, TUM School of Engineering and Design, Technical University of Munich.

Supervisor M.Sc. Qinglong Lu

M.Eng. Cheng Lyu

Prof. Dr. Constantinos Antoniou

Chair of Transportation Systems Engineering

Submitted by Ningkang Yang

Submitted on Munich, 09.05.2023

Abstract

Understanding the reaction of a transportation system under extreme events is significant for evaluating the resilience of the system. However, collecting traffic data after a natural disaster is usually expensive and the data volume is always insufficient for building a robust model. In this thesis, an unsupervised pattern extractor is proposed to extract the patterns of transportation resilience from grid-wise aggregating floating car data (FCD), which improves the transferability of deep learning models and enables the prediction of target cities to learn resilience patterns from other cities. The unsupervised pattern extractor is a combination of the k-shape time series clustering method and the Bayes inference method. The k-Shape method is used to capture the similarity of pre-disaster patterns between different cities while the Bayes inference method is used to generate the posterior distribution for the traffic volume at each time step for the time series of each cluster.

To verify the effectiveness of our method, we conducted a case study under the context of the COVID-19 pandemic using the grid traffic volume resulted from aggregating floating car data of three cities, in which the source domain cites include Antwerp and Bangkok, and the target domain city is Barcelona. Results show that the extracted resilience patterns by our method can improve the prediction performance of transfer learning neural networks at both grid and macro levels with less pre-disaster information and limited data volume. Furthermore, we explored the impact of grid size through a sensitivity analysis and discussed the balance between data stability and data volume. The results show that our method achieved better performance as the grid size increased, although the data volume decreased.

Acknowledgement

First of all, I would like to express my sincere gratitude to my supervisors at the Chair of Transportation Systems Engineering (TSE) at the Technical University of Munich (TUM), who provided me with unconditional support. Thanks to Cheng Lyu's guidance on the methodologies and to Qinglong Lu's guidance on experiment design and thesis writing. During this thesis, I developed a strong interest in scientific research and significantly improved my research abilities and understanding of traffic data analysis and transportation system resilience. I also want to thank Vishal Mahajan and Prof. Dr. Constantinos Antoniou for their advice in the midterm presentation of this thesis, which helped me to recognize my shortcomings.

Additionally, I would like to express my gratitude to all of my friends and teachers at TUM for their help in both daily life and studies. I really appreciate TUM for providing me with a conducive learning environment. During this one and a half years, I have gained abundant expertise and met many like-minded friends. This study experience would be the most precious wealth of my life, which will continuously inspire me to become a better person and researcher in the rest of my life.

Last but not least, I would like to express my gratitude to my family, in particular to my parents, for their understanding and unconditional support.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Motivation

The frequent natural disasters have led to numerous socio-economic problems for the past decades. Under the impact of natural disasters, transportation infrastructure is often at significant risk [3]. Natural disasters in transportation systems can highly damage the systems' performance and impact citizens' travel behavior directly or indirectly. Thus, understanding the resilience of transportation systems under extreme events to bring the systems back to a normal state in a timely manner becomes imperative [4].

According to [1], resilience refers to the capacity of a system to avoid shocks, to absorb the shock when it happens, and to restore within a specific time after a shock event. They further proposed a "resilience triangle" as the measurement for this process. As shown in Figure 1, the performance of the system drops sharply after the occurrence of the shock events at t_0. After reaching the lowest point, the system gradually recovers to a new equilibrium at t_1.

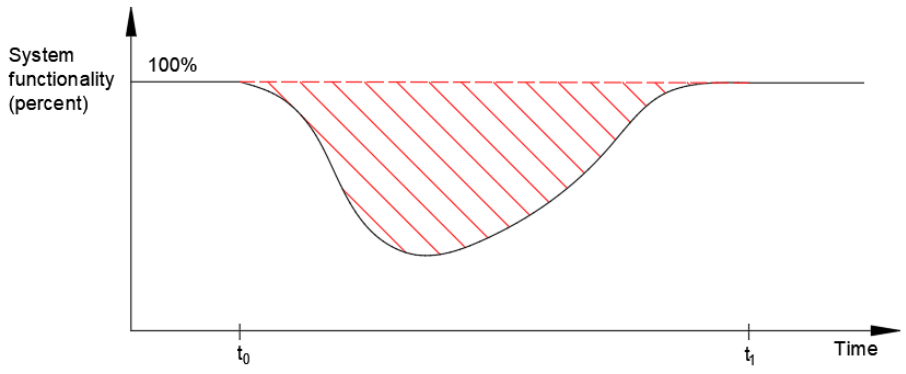


Figure 1 Measure of resilience—conceptual definition, adapted from [1].

The "resilience triangle" theory provides opportunities for researchers to investigate the resilience of transportation systems. Although several studies have applied various methods to model the resilience of transportation systems during natural disasters, the research scope of these studies is often limited, and most of their models are non-transferable. Moreover, most current studies only rely on local traffic detection data from fixed locations, which fails to reflect the resilience patterns of the entire network and can only focus on certain natural disaster events. For example, [5] applied a logistic function to model the resilience of the New York City transportation system under the Hurricanes through taxi trip data and subway ridership data; [6] used snowplow data and traffic speed detectors data to analyze the resilience of city road network after the snow storm through some performance index.

In order to build more comprehensive models, the floating car data (FCD) is receiving increasing scholarly attention in these years, which makes it possible to apply artificial neural networks to predict the resilience patterns of transportation systems. However, for a single city road network, the limited historical data and insufficient data volume always present significant challenges to traffic prediction tasks, and the temporal domain shift among cities increases the difficulties of applying transfer learning. Therefore, little research focuses on transportation resilience pattern prediction.

To address the gap in this research area, this thesis proposed a transferable model for predicting transportation system resilience using FCD. The proposed model consists of a resilience pattern extractor and artificial neural networks, which can alleviate the temporal domain shift among different cities and enable effective transfer learning for the target domain. By applying the proposed method, this study seeks to contribute to the development of more transferable and effective models for transportation system resilience prediction.

1.2. Objectives and Contributions

1.2.1. Research Objectives

The overall objective of this thesis is to predict the resilience patterns of the transportation system using FCD through the transfer learning method. Specific objectives are listed below:

1. Explore the dynamics of the resilience triangle of the transportation system under the natural disasters through the time series of FCD.
    a. Extract the resilience patterns of the networks that contain different pre-disaster features.
    b. Discuss the features of extracted resilience patterns.
2. Propose a transferable model to describe and predict transportation resilience by using pre-disaster FCD.
3. Explore and analyze the performance and transferability of different machine learning models in the resilience estimation task.

1.2.2. Contributions

The contributions of this thesis are listed below:

1. The thesis proposed a transferable method to address the temporal domain shift by aggregating the traffic volume resilience patterns from different cities. The proposed method achieved better transfer learning performance than traditional neural networks under limited data conditions.It overcomes the challenges of post-disaster data collec-

tion and limited historical data and provide timely predictions of both macro and local resilience patterns of cities.

2. This thesis proposed a model for predicting resilience patterns of large-scale city road transportation at the grid level. Through the analysis and observation of both micro and macro prediction results, the proposed model can learn stable traffic volume resilience trends at the grid level while resisting the interference of some unstable trends, showing more robust predictive performance than traditional neural networks.

3. This thesis also discusses the impact of grid size on the predictive performance of the model, and analyzes the importance of data quantity and stability to the model. The proposed model can achieve better transfer learning performance by enhancing data stability with fewer data than traditional neural networks. This further reduces the requirement for the volume of training data and alleviates the conflict between data stability and volume in traffic prediction tasks.

1.3. Thesis Structure

The overall structure of this thesis is illustrated in Figure 2, and the rest of this paper is structured as follows: Chapter 2 reviews the related literature, provides comprehensive definitions of transportation resilience, and summarizes the tasks and methods in the previous studies on transportation resilience research. Chapter 3 proposes the proposed framework of the applied method and describes the components of the method respectively. Chapter 4 introduces the background of the case study and describes the used floating car data. Chapter 5 describes the experimental setup. Chapter 6 presents and discusses the experiment results and some further experiments. Chapter 7 concludes the main findings and contributions of this thesis and discusses potential future research directions.

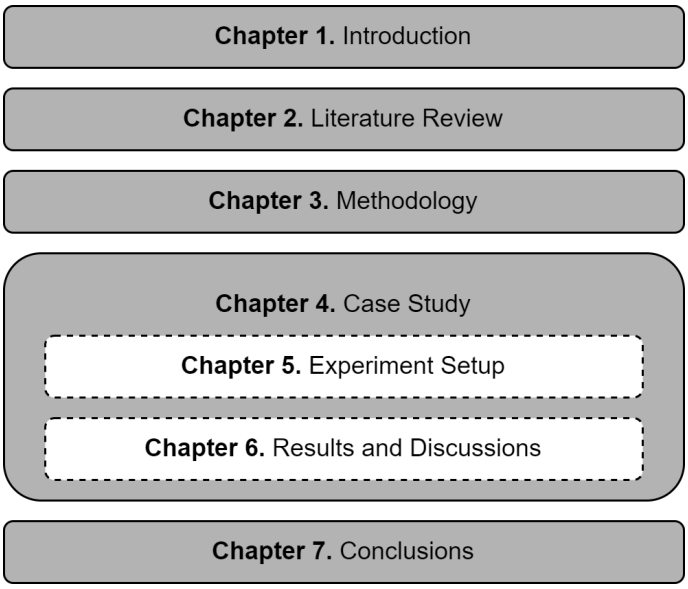


Figure 2 Structure of Thesis

2. Literature Review

In this chapter, we first clarify the basic concepts of transportation system resilience; Then we summarize the features of transportation system resilience; In the third section, we review the tasks in different research branches of transportation resilience and their methodologies; In the fourth section, we introduce the categories and strategies of transfer learning.

2.1. Definitions of Transportation System Resilience

Resilience is physically expressed as the ability of an object to recover its shape. For the past decades, this concept has been extended to various fields, such as climate analysis, psychology, economics, and engineering research. For example, [1] proposed the "4R" framework and initially presented a method to quantify system resilience. Based on their proposed concept in Section 1.1, many researchers have further proposed various definitions of resilience in the field of transportation.

[7] defined transportation resilience as the endurance and recovery capacity of transportation systems under the extreme disruptions. In order to describe the resources cost of the resilience process, transportation systems resilience is further extended to the ability of a transportation system to absorb shocks and gradually provide the needed service with its basic structure and function within an acceptable time and cost after a disturbance event [2]. Besides, [8] took the riskiness of shocks into consideration and introduced the concept of risk into the definition of system resilience. They emphasize that the degradation parameters and recovery time of the disruption that the transportation system withstands should be within an acceptable range. In the temporal dimension, the definition of transportation system resilience should also consider both their short-term and long-term adaptability, which describe the ability of the transportation system to respond in the short term and to transform in the long term [9]. Based on the aforementioned definitions, [10] have proposed a comprehensive definition of traffic resilience:" The ability of a system to resist, reduce and absorb the impacts of a disturbance (shock, interruption, or disaster), maintaining an acceptable level of service (static resilience), and restoring the regular and balanced operation within a reasonable period of time and cost (dynamic resilience)."

2.2. Features of Transportation System Resilience

The definition of system resilience contains various aspects, and these aspects have been summarized by [1] as four characteristics (4R): robustness, redundancy, resourcefulness and rapidity. Figure 3 illustrates the schema of the system resilience. Robustness describes the ability of a system or unit to withstand a certain level of stress without suffering degradation and to maintain its functionality; Redundancy describes the extent to which the system

contains replaceable components in the event of degradation and loss of functionality; Resourcefulness describes the capacity of system resilience to achieve stability through the available resources during the recovery phase; Rapidity describes the time consumption of a system to minimize losses and satisfy priorities.

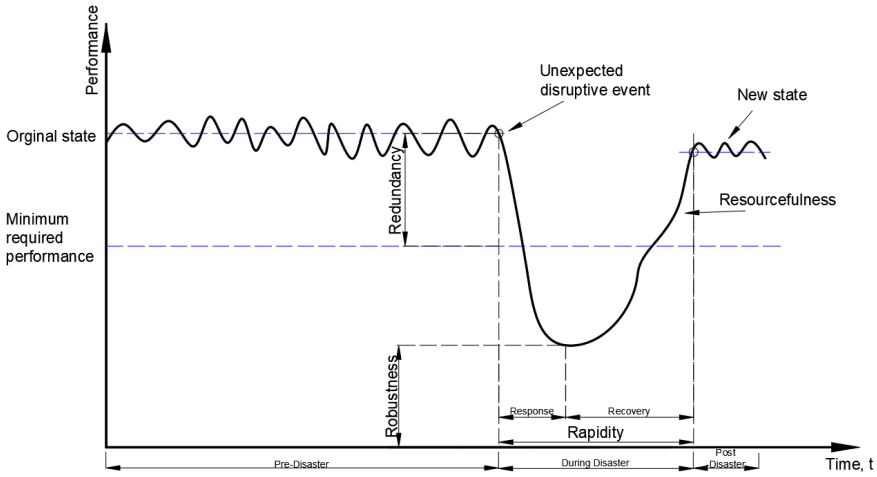


Figure 3 Schematic of performance of a resilient system, adapted from [2].

Based on the 4R, more characteristics of the resilience of the transportation system resilience were proposed, such as adaptability, preparedness, reliability, and vulnerability. Among them, adaptability emphasizes the flexibility of the system to learn from previous shock events and become more robust to better adapt to new stresses [11, 12]; Preparedness refers to the degree to which the system is prepared for a catastrophic event before being shocked, with the aim of reducing the negative impact of the shock event on the system [13]; Reliability measures the probability that the network will still be able to meet part of the demand after a disruptive event [14]. In the context of transportation systems vulnerability is defined as the sensitivity of the road system to events that could lead to a significant reduction in road service capacity [15]. Although these features are actually the further complement of the 4R, they are not comprehensive enough to describe system resilience.

Therefore, this thesis applies 4R as the features of system resilience and extends the definition of 4R to transportation systems field:

- Robustness is defined as the ability of the transportation system to satisfy primary traffic demand under a certain level of stress.
- Redundancy refers to the capacity of a transportation system to provide alternative routes or services for the traffic demand in the early stage of the disaster.
- Resourcefulness describes the ability of a transportation system to provide enough service for the potential traffic demand during the recovery phase.
- Rapidity is the ability of a transportation system to reach a new equilibrium state after a disruptive event in a short time.

2.3. Tasks and Methodologies in Different Research Branches of Transportation Resilience

The introduction of the above concepts and characteristics has inspired many scholars to analyze transportation system resilience from various aspects. The tasks of system resilience analysis can be divided into several categories: Assessment metrics exploration, resilience regression model developing, causal impact and resilience patterns prediction. Different methods usually need to be applied for different tasks, including regression models, statistical time series models, and deep learning methods.

2.3.1. Assessment Metrics Exploration

In order to quantify the resilience of systems, considering different disaster types and considering different aspects, numerous scholars have developed different indicators. For instance, [16] introduced a framework to evaluate the impact of floods on road traffic disruption. They used traffic flow betweenness centrality, overall travel time, and person-minute delay as resilience indicators. They explored the importance of intersections of major roads and networks during hazardous events and identified priority intervention locations. Meanwhile, [17] defined the resilience of a system as the ratio of the system's recovery level to the loss it suffered and used the system delivery function to quantify this resilience indicator. [18] proposed an importance-impedance-based rail transit network resilience measurement, using variables such as node betweenness centrality, passenger volume, and travel time among different origin-destination pairs under varying levels of impact to define the extent of degradation of rail transit networks during extreme events. Some scholars further used optimization algorithms to calculate system resilience metrics under optimized recovery measures for transportation networks. [19] used the expected mean of the area between the actual performance curve and the target performance curve for different stages of transportation network system performance during disasters as the resilience index. They proposed an O-D pair connectivity optimization model based on different resilience actions and calculated the resilience index for each stage of the transportation network under optimized measures. However, these methods cannot flexibly consider the time of each stage of the micro-level resilience pattern, and the complexity of calculating the indicator increases when considering complex urban networks.

2.3.2. Resilience Regression Model

It has been proven that the logistic model can simulate natural growth. [5] analyze the number of taxi trips and subways ridership in New York City before and after the impact of Hurricanes, and applied a logistic function to model the recovery rate. They divided the city into six zones according to the degree of impact and fitted the resilience models for each zone. The regression models require low computational resources and can reflect the approximate trends of the resilience patterns. However, their model failed to capture the temporal dependencies of the resilience patterns and cannot reflect the seasonality of the model. [3] developed a time-dependent recovery rate regression model based on Cox's proportional hazards regression

model. The independent variables of their model are region, type of natural disaster, and post-disaster reconstruction cost, while the dependent variable is the post-disaster reconstruction duration. They conducted a case study in New South Wales, Australia, and used the likelihood function to calculate the cumulative probability of time to the recovery of transport infrastructure. However, their model only considered the overall recovery time, i.e., the rapidity of system, and did not consider the other resilience features, which means that their model failed to describe the specific impact of disasters at different stages.

2.3.3. Causal Impact Analysis

The other type of task is to analyze the causal impact of disasters. The causal impact aims to estimate the causal effect of a particular intervention on the outcome of an event. The widely used methods in transportation resilience causal impact analysis are statistical time series models. By comparing with the prediction of the system performance under a no-disaster scenario, the effect of the disaster could be captured. The autoregressive integrated moving average (ARIMA) model has high efficiency in stable time series analysis and prediction [20]. [21] applied ARIMA model to predict the short-term GDP of the no-earthquake scenario by the pre-disaster time series. They focused on the post-disaster macroeconomic recovery ratio (RE) by comparing the simulated GDP and actual value. The Bayesian structural time series (BSTS) model is another method for inferring causal impact. The BSTS model allows the researchers to add regression components and estimate their potential contributions through Bayes methods. [7] applied BSTS to infer the non-disaster ridership of four public stations and used a regression tree to explore the relationship between the resilience of the rail transit system and possible influencing factors, such as built environment, socioeconomic disparities and the COVID-19 cases. [22] calculated dynamic time warping (DTW) distance to measure the similarity between the smooth historical data and the shocked serial data. They measured the resilience of the ecosystem by developing disturbance magnitude, recovery strength and recovery rate as the resilience metrics. However, These models usually require long-term historical data and the performance of the non-disaster scenarios prediction tasks cannot be guaranteed.

2.3.4. Resilience Patterns Prediction

The above conventional methods are usually used to estimate the resilience of certain disaster events. These models are non-transferable and can hardly be applied in large-scale scenarios. Currently, deep learning methods such as recurrent neural network (RNN), Long Short-Term Memory (LSTM), and temporal convolutional network (TCN) have achieved great success on time series prediction tasks. They provide vital support to predict the whole resilience patterns directly. [23] proposed a bidirectional diffusion graph convolutional layer to predict the transportation system resilience patterns under extreme weather. Their research also provided the performance of RNN-based methods, which shows that the RNN-based methods still have strong prediction performance. However, training a deep neural network is usually time-consuming, and the insufficiency of the data is always a challenging problem. Thus, transfer learning methods are widely used in traffic prediction tasks. Due to the limited

knowledge about the transferability of the GNNs, this thesis only adopts conventional neural network methods as one of the components for traffic prediction tasks.

2.4. Transfer Learning Strategies

The distributions of the traffic data are usually inconsistent among different cities, which is the so-called domain shift. However, most machine learning methods require the training data and test data to obey the same distribution. A challenge for traffic forecasting is insufficient data [24] and using past traffic data for a data imputation is always unreliable [25]. Hence, transfer learning methods are popular in traffic data mining tasks. According to [26], transfer learning techniques can be divided into three categories: inductive transfer learning, transductive transfer learning, and unsupervised transfer learning. The differences between these techniques depend on the availability of the labels and the similarity of the tasks between the source domain and the target domain. In the transportation system resilience prediction task, the main purpose of transfer learning is to use the prediction knowledge from the other cities to improve the learning of the target city, which corresponds to the definition of the inductive transfer learning task. [27] summarized the characteristics of three main transfer learning strategies: transfer without fine-tuning, fine-tuning with freezing transferred layers and fine-tuning without freezing transferred layers. Their results show that the second strategy requires fewer transferred layers and can achieve better performance in the prediction tasks. Although under the third strategy the model converges more rapidly, the overfitting problem significantly affects its performance.

2.5. Literature Gaps

Based on the above literature review, this thesis aims to address the following literature gaps.

1. Few works have used transfer learning for transportation resilience prediction and proposed solutions for temporal domain shift.
2. Few works have performed large-scale resilience patterns prediction at the grid level.
3. Few works have developed general approaches for transportation resilience prediction and focused on the predictive performance of the models.

3. Methodology

In this chapter, we first review the concepts and characteristics of FCD in Section 3.1. In the rest sections, we introduce the components of the proposed method sequentially. Section 3.2 presents the overall framework of the proposed method and basic assumptions; Specifically, Section 3.3 introduces the time series clustering methods; Section 3.4 describes the application of the Bayes method; Section 3.5 introduces several conventional neural networks; Finally, Section 3.6 introduces the transfer learning process.

3.1. Floating Car Data

3.1.1. Characteristics of FCD

To build a robust model that can describe the resilience patterns for the entire city road network, the input traffic data of the proposed model should be able to cover as wide as an urban area and provide specific details to reflect the citizen's travel patterns and behavior. Nowadays, floating car data (FCD), which is collected by Global Position System (GPS)-equipped vehicles, play a vital role in traffic data mining [28]. Compared to traditional methods of obtaining traffic data, the FCD can be provided by various types of vehicles in the whole city road network in real time. Therefore, the floating car data has a broader coverage and is more flexible as road sensors and traffic cameras are always fixed [29], which overcomes some technical and terrain limitations of certain areas or sections. In addition, the FCD includes dimensions, such as positions, speed, time and traffic volume, which strongly support the research on intelligent transportation systems (ITS) [30]. Therefore, we applied grid FCD as the input to the proposed model.

3.1.2. Data Availablility

For the input grid FCD, the grid size should be appropriately selected to provide sufficient and stable data for the machine learning models. The grid FCD usually contains traffic volume and average speed. For traffic volume data, which is sensitive to the number of vehicles connected, a relatively large grid size could increase the order of magnitude and alleviate the instability of the data. However, the city road network is sparse, and applying a large grid size can further reduce data volume after data cleaning. For average speed data, which is not affected by the number of floating vehicles, a small grid size can be applied to guarantee a sufficient data volume. However, the speed data fail to reflect the resilience patterns for some types of disasters. In this thesis, we applied traffic volume to represent the system's performance and paid more effort into data cleaning to maintain as much data as possible for the proposed model.

3.2. Overview and Basic Assumption

The proposed method of this thesis combines time series clustering, Bayes methods, and artificial neural networks (ANN). This method aims to perform traffic volume resilience patterns prediction on the test set and target domain data set using only pre-disaster data. The framework of the method is presented in Figure 4, which outlines the steps involved in the method.

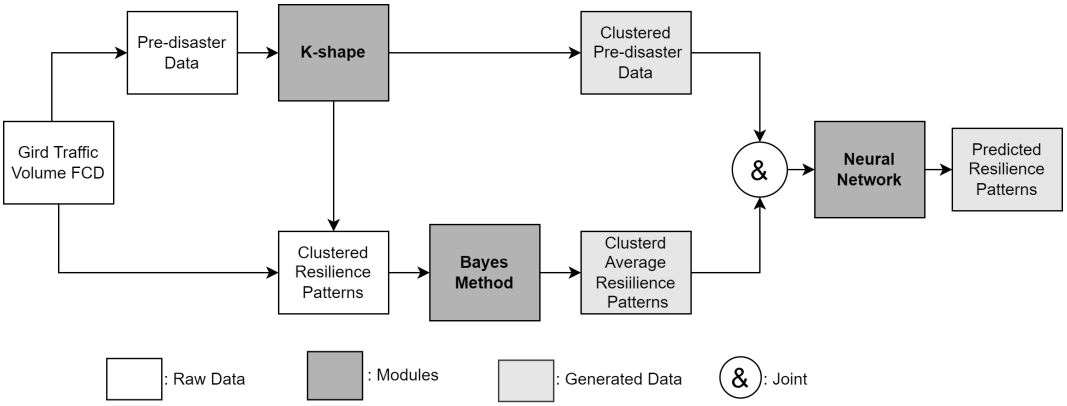


Figure 4 The framework of the proposed method.

In order to predict traffic volume resilience patterns using only pre-disaster data, some prompt features of the resilience patterns are necessary, which can only be extracted from the training data, as both the pre-disaster sequence and the during-disaster sequence of the training set are given. To match the extracted features to the pre-disaster data set, the approach for resilience pattern extraction should depend on the pre-disaster data. Although the distribution of overall traffic volume time series varies from city to city, there are still similarities among the pre-disaster patterns at the grid level. Therefore, we presented the basic assumption that grids with similar pre-disaster characters tend to have similar resilience patterns. To capture the similarities of the pre-disaster data, we applied the k-shape time series clustering method to cluster the pre-disaster time series of the training set. In order to extract general resilience patterns, we mixed the pre-disaster part of the raw grid traffic volume FCD from different cities as the input to the k-shape method. Then, the clusters, each containing similar pre-disaster time series from different cities, are generated, and the corresponding during-disaster time series of the elements in each cluster could be labeled based on the clustering outcomes of their pre-disaster parts. In this way, the clusters of resilience patterns can be further generated.

Based on the “resilience triangle” model, for each cluster of resilience patterns, an overall resilience pattern can be observed. To extract the overall resilience pattern, we applied the Bayes method for each cluster to provide inference for the posterior distribution of during-disaster traffic volume time series at each time step. Thus, for the grids in the same cluster, traffic volume distributions of each time step can be obtained. In order to extract unbiased

features, the mean value of the distributions is denoted as the extracted prompt features. The sequence of extracted prompt features represents the average resilience pattern for each cluster.

Then, the clustered pre-disaster time series and their corresponding average resilience patterns are joined together and fed into different neural networks to predict the actual resilience patterns. In this way, the pre-trained models can be obtained. For the test set, only pre-disaster data is given. The pre-disaster data of the test set is first matched to the corresponding cluster directly and then joined together with the corresponding average resilience patterns extracted from the training set as the input of the neural network.

3.3. Time Series Clustering

In order to find the homogeneity of grids in a city road network, we applied time series clustering to identify the girds with similar pre-disaster patterns. Here we define a data set of traffic volume time series in a n-grid road network as $G = \{g_1, g_2, ..., g_n\}$. The pre-disaster part of the data set is $P = \{p_1, p_2, ..., p_n\}$ and the resilience patterns of the data set is $R = \{r_1, r_2, ..., r_n\}$. The pre-disaster time series P is divided into k clusters $C = \{c_1, c_2, ..., c_k\}$, where $c_i = \{p_j | j \in [1, n]\}$. Based on the aforementioned basic assumption in Section 3.2, the cluster set of resilience patterns $D = \{d_1, d_2, ..., d_k\}$ can be generated, in which each cluster d_i should contain the resilience patterns r from the same grids as the cluster c_i in C, i.e., $d_i = \{r_j | j \in [1, n]\}$. On the one hand, for the time series in the same cluster, their patterns, trends and other characteristics are similar. On the other hand, the data from different clusters are heterogenous in some features. Thus, the measurement of similarity between time series directly highly determines the performance of clustering. The similarity between two pre-disaster time series can be also the described as distance [31].

3.3.1. Distance Measures of Time Series

Euclidean distance (ED)[32] and time dynamic warping (DTW) [33] are two popular distance measures for time series clustering. The main difference between them is evident in the alignment of the time series and the time complexity, as shown in Figure 5. ED can be directly calculated through the sum of squared differences between two sequences $P_S = \{s_1, s_2, ..., s_m\}$ and $P_Q = \{q_1, q_2, ..., q_m\}$ in a point-to-point way as follows:

$$ED\left(P_S, P_Q\right) \equiv \sqrt{\sum_{i=1}^{m} \left(s_i - q_i\right)^2} \tag{3.1}$$

ED is a simple and competitive measure for time series clustering, with a time complexity of only $\mathcal{O}(m)$. However, when the lengths of the time series are different or when the time series are not aligned at the same time points, ED is not robust enough to reflect the real distance between two time series, i.e., ED fails to achieve shift-invariance. [34].

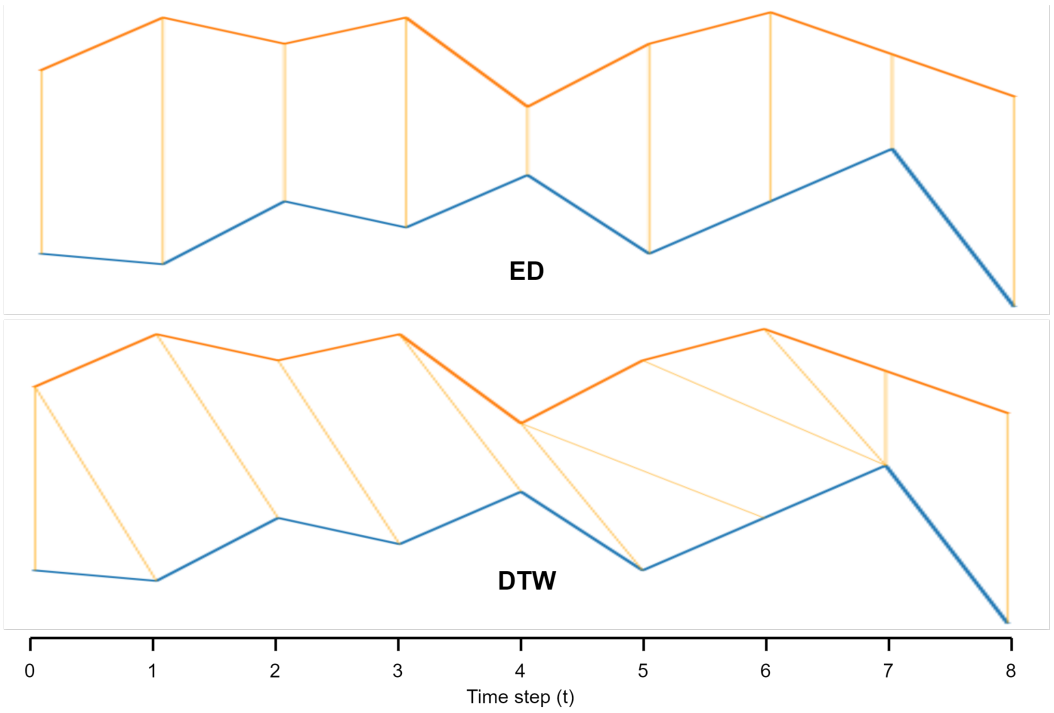


Figure 5 Time series alignment under ED and DTW

In contrast, DTW can handle the unequal length and unaligned time series. As shown in Figure 6, in the DTW algorithm, an m-by-m cumulative distance matrix M needs to be calculated and a warping path $W = \{w_1, w_2, ..., w_n\}, n \geq m$ through the matrix represents the distance between two time series, which is given by the following formula:

$$DTW\left(P_S, P_Q\right) = \min \sqrt{\sum_{i=1}^{n} w_i} \tag{3.2}$$

The warping path is selected from M by a dynamic programming process. Given a distance measure D between two time points, the elements of the cumulative distance Matrix M are given as follows:

$$M(i,j) = \begin{cases} D\left(s_i, q_0\right) + M(i-1, 0), & j = 0 \\ D\left(s_0, q_j\right) + M(0, j-1), & i = 0 \\ D\left(s_i, q_i\right) + \min\{M(i-1, j), M(i, j-1), M(i-1, j-1)\}, & \text{others} \end{cases} \tag{3.3}$$

Although the DTW algorithm has achieved the best performance in many kinds of research, the time complexities for DTW and the dynamic programming process of calculating minimum time warping path are both $\mathcal{O}(m^2)$. Therefore, applying DTW to a time series clustering algorithm requires a long calculation time for a large time series data set and leads to low efficiency. In this thesis, due to the large data volume, DTW is applied for evaluating the

performance of the model rather than time series clustering.

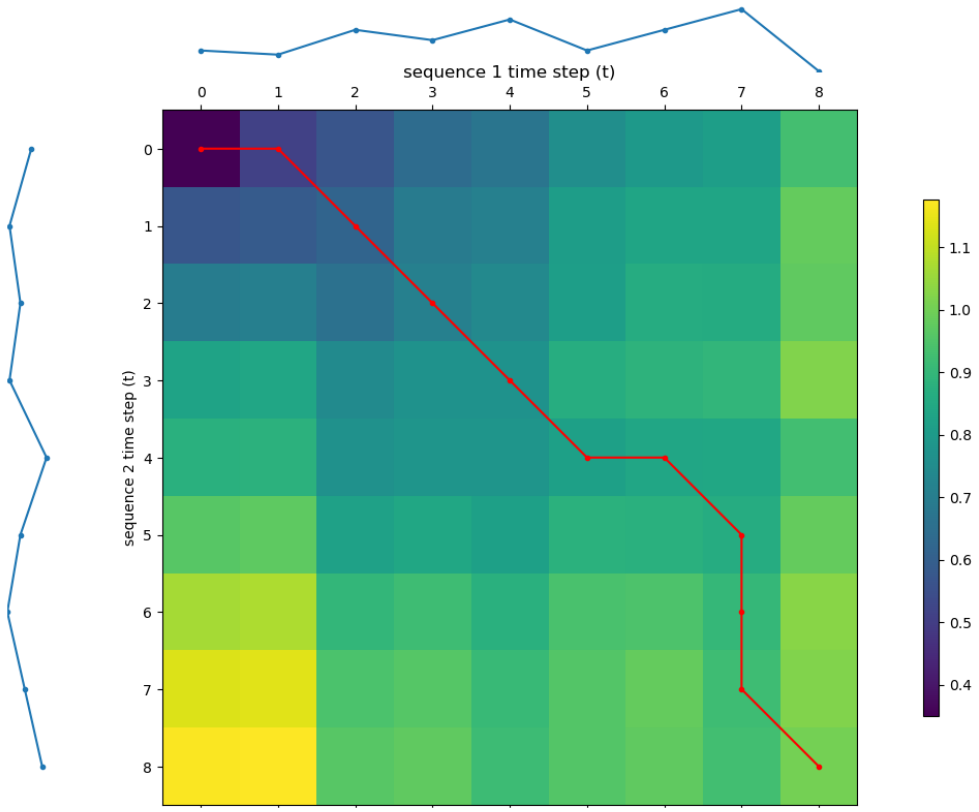


Figure 6 Heatmap of cumulative distance matrix and warping path

To tradeoff the time complexity and the robustness of the model, in this thesis, we applied k-Shape algorithm for time series clustering [35]. The similarity measure of k-Shape algorithm is the Shape-based distance (SBD). The SBD is calculated through coefficient normalized cross-correlation (NCC_c) between two time series. According to [35], when computing the NCC_c, one of the sequences $P_Q = q_1, q_2, ..., q_m$ keeps static and the other sequence $P_S = s_1, s_2, ..., s_m$ slides over Q to calculate the inner product of their overlapping part for each shift of P_S. Since P_S can slide both forward and backward, a total of 2m-1 slides can be generated, which can result in 2m-1 NCC_c. The cross-correlation $CC_k(P_S, P_Q)$ can be computed as:

$$CC_k\left(P_S, P_Q\right) = \begin{cases} \sum_{l=1}^{m-k} P_{S_{l+k}} \cdot P_{Q_l}, & k \geq 0 \\ CC_{-k}\left(P_Q, P_S\right), & k < 0 \end{cases} \tag{3.4}$$

Where k is the sliding length of sequence P_S. When $k \geq 0$, the direction of sliding is forward and when $k \leq 0$, the backward sliding of P_S can be seen as the forward sliding of P_Q. Then, NCC_c is defined as follow equation:

$$NCC_c = \frac{CC_k\left(P_S, P_Q\right)}{\sqrt{R_0\left(P_S, P_S\right) \cdot R_0\left(P_Q, P_Q\right)}} \tag{3.5}$$

And the SBD is derived as:

$$\mathrm{SBD}(P_S, P_Q) = 1 - \max_k \left(\frac{CC_k\left(P_S, P_Q\right)}{\sqrt{R_0\left(P_S, P_S\right) \cdot R_0\left(P_Q, P_Q\right)}} \right) \tag{3.6}$$

The time complexity of SBD is $\mathcal{O}(m^2)$ due to the computation of $CC_k(P_S, P_Q)$. In order to increase the efficiency of the calculation, [35] optimizes the computation of SBD by using a Fast Fourier Transform, which reduces the time complexity of SBD to only $\mathcal{O}(mlog(m))$. The characteristics of the above distance measures are summarized in Table 1:

Table 1 Characteristics of the distance measures

Distance measure	Time complexity	Shift-invariance
ED	$\mathcal{O}(m)$	False
DTW	$\mathcal{O}\left(m^2\right)$	True
SBD	$\mathcal{O}(mlog(m))$	True

3.3.2. K-Shape Time Series Clustering

K-means algorithm are one of the renowned methods for time series clustering [36]. According to [37], the basic steps of the k-means clustering procedure are as follows:

Step 1: Specify the number of K clusters to assign.

Step 2: Randomly initialize K centroids within the feature space.

Step 3: Assign each data point to its closest centroid based on a distance measure.

Step 4: Compute the new average centroid of each cluster.

Step 5: Repeat step 3 and step 4 until the centroids no longer change or the maximum iterations is reached.

K-Shape algorithm is a k-means-based partitional clustering method with SBD as the distance measure, and the centroids are calculated by solving a maximization problem [38]. Based on Equation 3.4, the centroids of k-Shape can be computed as follow:

$$\mu_k{}^* = \underset{\mu_k}{\operatorname{argmax}} \sum_{P_i \in c_k} \left(\sum_{l \in [1,m]} P_{i_l} \cdot \mu_{kl} \right)^2 \tag{3.7}$$

Where $\mu_k{}^* = \mu_{k+1}$ is the new centroid based on the optimization and P_i is the ith time series from P, i.e., $P_i = p_i$.

By applying k-Shape time series clustering method, the pre-disaster time series P is divided into k clusters $C = \{c_1, c_2, ..., c_k\}$, where $c_i = \{p_j | j \in [1, n]\}$. For the pre-disaster time

series in each cluster c_i, the label i is assigned to their corresponding resilience patterns from G and the new cluster of resilience patterns $d_i = \{r_j | j \in [1, n]\}$ is generated. Finally, as presented in Fig 7, k clusters of resilience patterns are extracted from G. In the next section, we introduce how to extract the shape of resilience patterns for each cluster.

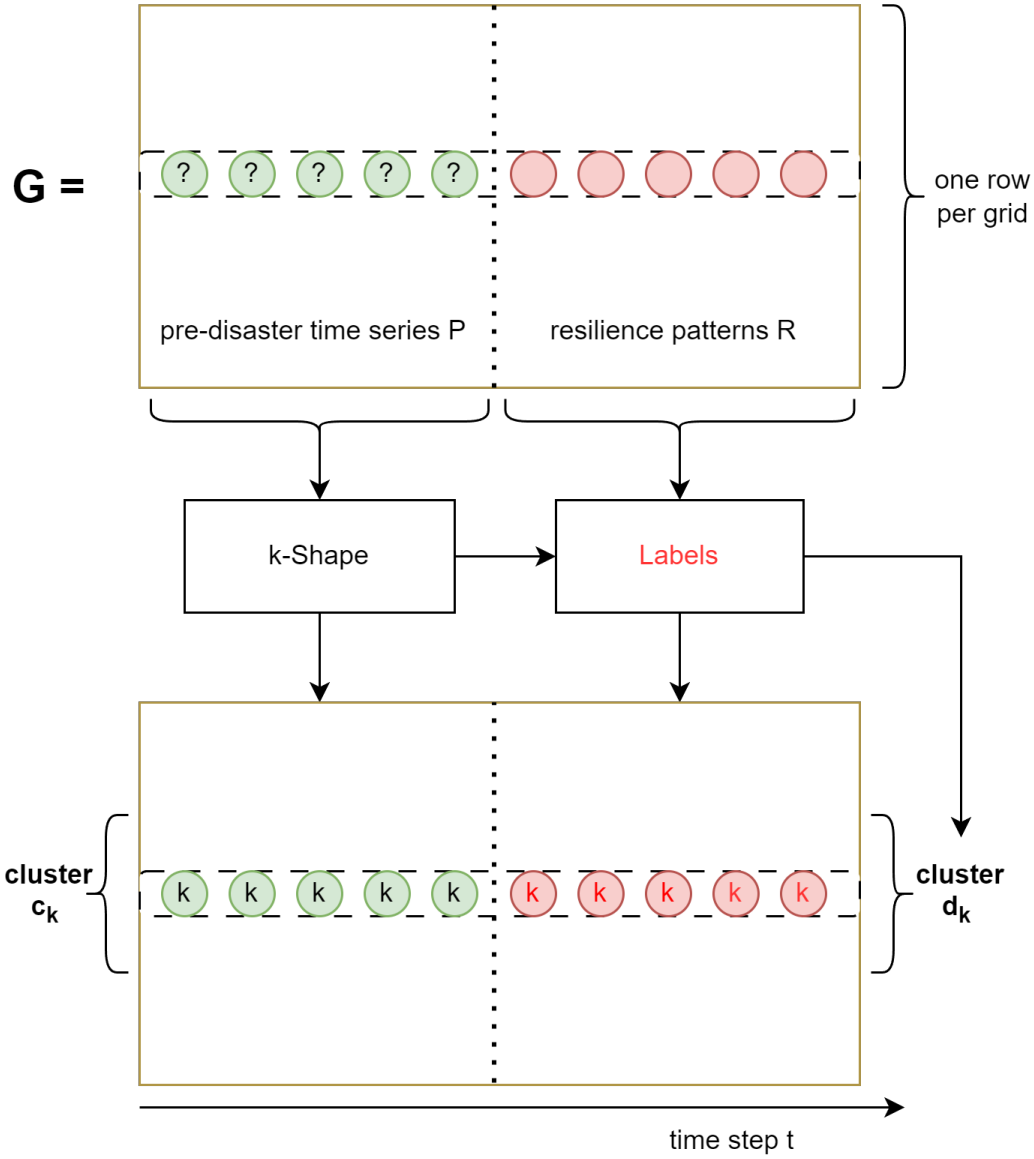


Figure 7 K-shape time series cluster for grid floating car data

3.4. Bayes-Based Shape Extractor

Assuming that for each resilience patterns cluster d_i, the number of elements is N and the duration of the resilience phase is N_d days. Therefore, for each time step d_{it} in d_i, a maximum of N values can be sampled. Therefore, for each cluster, N_d data sets are generated, which represent the days of the resilience patterns. In order to represent the resilience patterns for each cluster, a Bayes inference method is implemented to generate the posterior distribution

for each data set. The parameter field of each data set is denoted as θ, and the calculation of Bayes inference is given as follows:

$$\overbrace{p\left(\theta \mid d_{it}\right)}^{posterior} = \frac{\overbrace{p\left(d_{it} \mid \theta\right)}^{likelihood} \cdot \overbrace{p(\theta)}^{prior}}{\underbrace{p\left(d_{it}\right)}_{evidence}} \propto p\left(d_{it} \mid \theta\right) \cdot p(\theta) \tag{3.8}$$

where

$$\begin{aligned} P\left(d_{it} \mid \theta\right) &= P\left(\left\{d_{it}^{(1)}, \ldots, {d_{it}}^{(N_i)}\right\} \mid \theta\right) \\ &= \prod_{j=1}^{N} P\left({d_{it}}^{(j)} \mid \theta\right) \end{aligned} \tag{3.9}$$

Here, for each time step d_{it} in cluster d_i, $p(d_{it}|\theta)$ represents the likelihood function and $p(d_{it})$ is a normalizing constant called evidence. $p(\theta)$ is the prior distribution for the parameter θ. In our research, we applied Gaussian distribution as posterior distribution and uniform priors for both μ and σ. For k clusters, $k \cdot N_d$ posterior distributions are computed. An example of Bayes resilience patterns extraction is presented in Figure 8 and for each posterior, we applied the mean value mu as the representing resilience patterns.

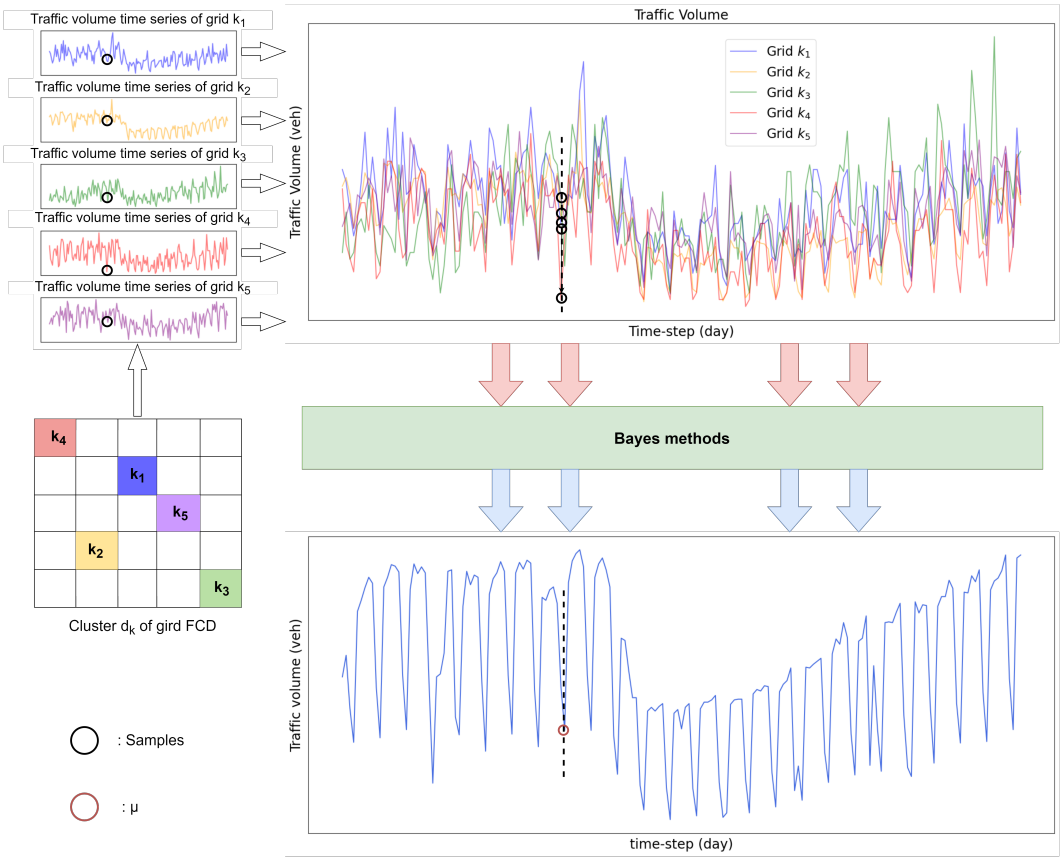


Figure 8 Example of Bayes resilience patterns extraction. Here we marked an example of resilience pattern extraction with a single time step. The samples of each time step form a data set. For each cluster di, the resilience pattern extraction process is repeated N_d times.

In the next section, we summarize the architecture of different neural networks and clarify how these networks are implemented for the prediction tasks.

3.5. Neural Networks

By applying the k-Shape clustering and Bayes method, prior knowledge about resilience patterns is extracted, which enriches the features for the pre-disaster data of the test set and transfer learning data set. In order to acquire deep embedding of the features and accomplish prediction task, deep learning models with the prediction heads are applied to automatically learn the prediction parameters. In this thesis, two kinds of deep learning models were applied, namely multi-layer perceptron (MLP) and recurrent neural networks (RNNs). For RNNs, we used a conventional recurrent neural network (RNN) and long short-term memory (LSTM), which are variants of RNN.

3.5.1. Multi-Layer Perceptron (MLP)

MLP is a basic type of feed-forward neural network (FNN). As presented in Fig 9, the nodes of FNN are connected in a directed graph without a circular structure. The inputs of FNN only flow from the input layer through hidden layers to the output layer in one direction and the length of inputs and outputs is fixed. The computation of each hidden layer h_l is given as follows:

$$h_l = f\left(\mathrm{w}_l^T x_l + b_l\right) \tag{3.10}$$

Where x_l is the input of the current layer, w_l represents the weight matrix of the lth layer, b_l denotes the bias vector and f is the non-linear activation function.

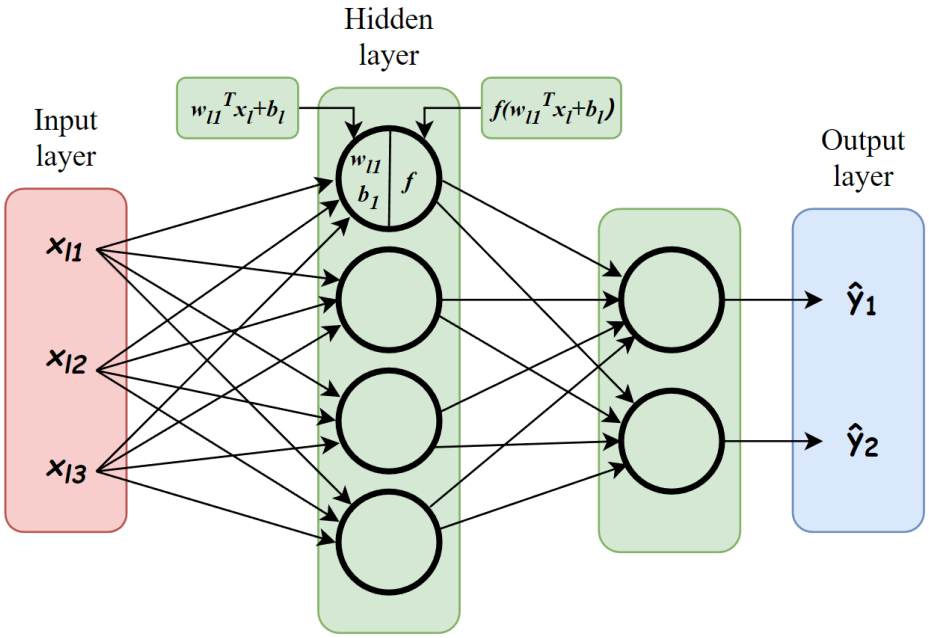


Figure 9 Sample of a feed-forward neural network

In this thesis, the inputs of the MLP are the joined sequences comprising the pre-disaster time series and the corresponding Bayes resilience patterns. The outputs are the predicted resilience patterns and the labels are the actual resilience patterns.

3.5.2. Recurrent Neural Network (RNN)

In contrast to FNN, in RNNs, the parameters are shared between different time steps to allow the model to handle the variable-length sequence. As shown in Figure 10, the inputs are processed one by one through time in the RNN model, and for each layer, the inputs are not only the features of the current time, but also the hidden features from the last time step, so that the RNN can capture the temporal dependency of the time series. The computation of the hidden feature h_t and output $\hat{y}_t$ is given as follows:

$$\begin{aligned} h_t &= f_h\left(x_t W_{hx}^T + h_{t-1} W_{hh}^T + b_h\right) \\ \hat{y}_t &= f_y\left(h_t W_h^T + b_y\right) \end{aligned} \tag{3.11}$$

Where W_{hx}^T is the weight matrix of the input x_t, W_{hh}^T denotes the weight matrix of the hidden features h_{t-1} from the last time step, W_h^T is the weight matrix of the h_t, $\hat{y}_t$ is the output and b_h and b_y represent the bias. f_h and f_y are non-linear activation functions.

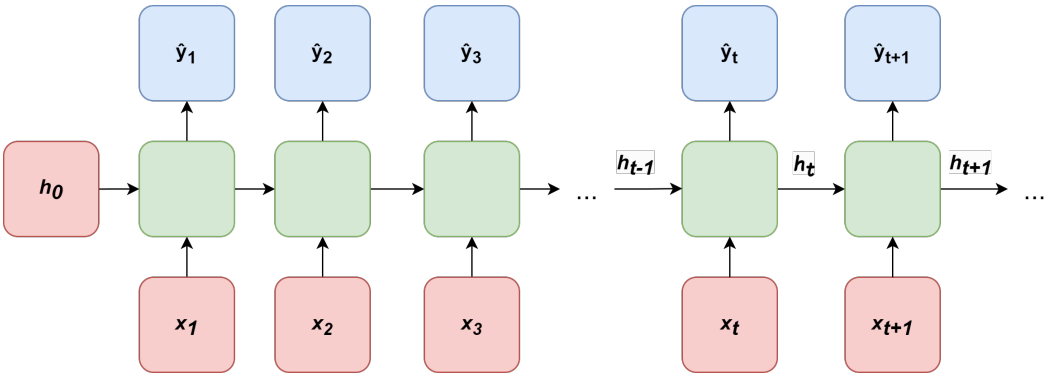


Figure 10 Sample of a recurrent neural network

In this thesis, the inputs of the RNN are also the joined sequences comprising the pre-disaster time series and the corresponding Bayes resilience patterns. Only the outputs from the last few layers are optimized to predict the resilience patterns.

3.5.3. Long Short-Term Memory (LSTM)

Long short-term memory (LSTM) is a variant of RNN, which is designed to solve the gradient problem in order to remember more information from the past by alleviating the gradient problem. In RNN, the outputs can only be optimized through hidden features. Once the weights are smaller than zero or larger than one, based on the backward propagation through time (BPTT) and chain rule, the successive derivations of the latest outputs can result in their gradients to the previous inputs converging either zero or infinity when predicting for long sequence. Therefore, the previous information is challenging to propagate to distant future

units. As shown in Figure 11, an LSTM unit contains a memory cell, an input gate, a forget gate and an output gate. The memory cell is introduced to aggregate the past and current information, the flow of which is adjusted by the gate units so that the information can be selectively transmitted to the following LSTM units to keep a long-term temporal dependency. The equations for an LSTM cell are as follows:

$$
\begin{aligned}
i_t &= \sigma\left(W_{ix}x_t + W_{ih}h_{t-1} + b_i\right) \\
f_t &= \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \\
c_t &= f_t \odot c_{t-1} + i_t \odot \tanh\left(W_{cx}x_t + W_{ch}h_{t-1} + b_c\right) \\
o_t &= \sigma\left(W_{ox}x_t + W_{oh}h_{t-1} + b_o\right) \\
h_t &= o_t \odot \tanh\left(c_t\right) \\
y_t &= W_{yh}h_t + b_y
\end{aligned}
\tag{3.12}
$$

Where i_t, f_t, o_t denote the input gate, forget gate and output gate respectively, c_t is the memory cell, h_t represents the hidden state, x_t is the input of current time step and y_t is the output of current time step. For the learning parameters, W denotes the weight matrix and b is the bias. $\odot$ denotes the Hadamard product. For activation functions, σ is the sigmoid function and $tanh$ is the hyperbolic tangent function.

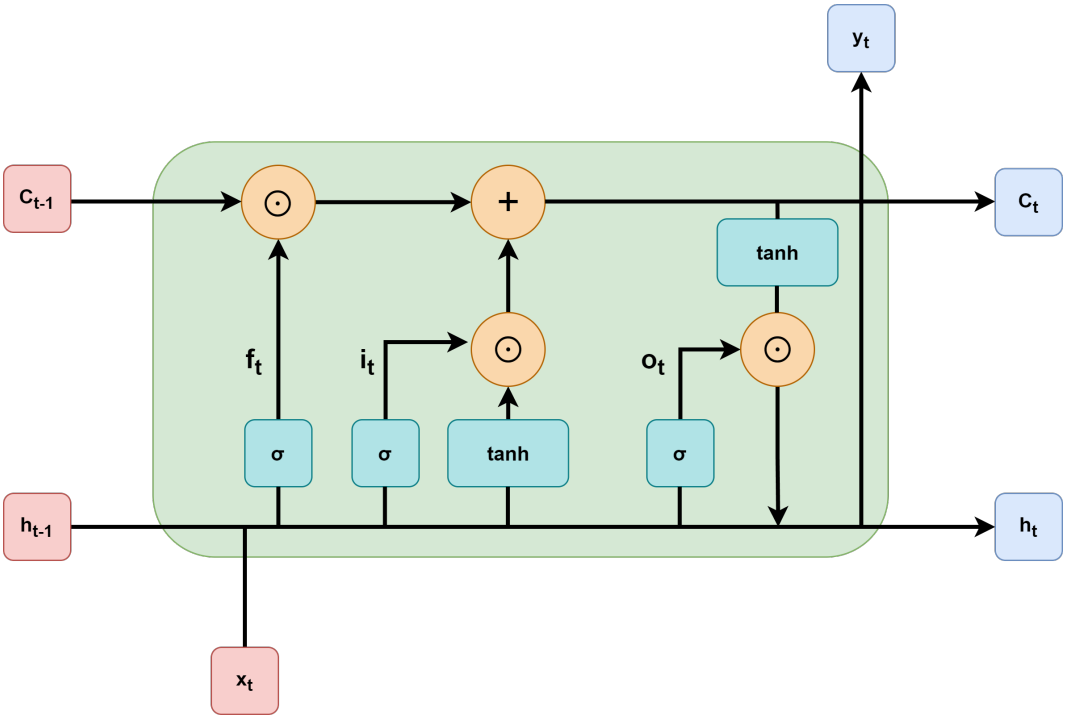


Figure 11 Sample of a long short-term memory network.

3.6. Transfer Learning

In the transfer learning experiment, the input data is only the pre-disaster grid floating car data from the other city. As the framework presented in Figure 12, the input traffic volume time series is first assigned to the corresponding cluster by measuring their SBD to the cluster centers of each group. The corresponding resilience patterns are then matched to the input data, and the joined sequences comprising the pre-disaster time series and the resilience patterns extracted from the cities of the source domain model are fed into the pre-trained neural networks with frozen parameters. Each pre-trained neural network is stacked with a MLP for parameter learning. Through this transfer learning process, the transfer learning model can predict the resilience patterns only with the pre-disaster data from the other city. from the other city.

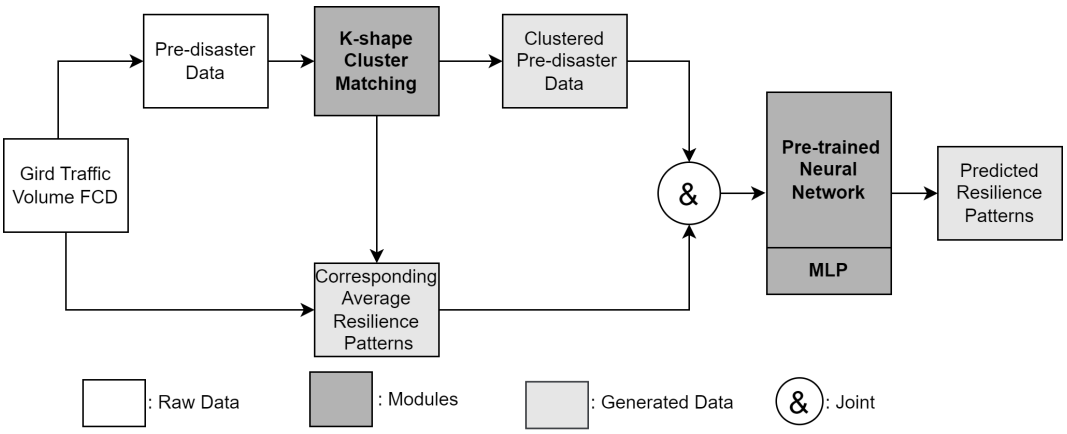


Figure 12 The transfer learning process.

In this thesis, transfer learning was implemented on all of the pre-trained neural networks from the source domain, aiming to explore and analyze how the extracted average resilience patterns affect the performance of both the source and target domain models. As the transfer learning data set is only the FCD from one city, the data volume is far less than the source domain data set. Besides, the input pre-disaster time series of the transfer learning experiment is different from the source domain models. Specifically, the sequence length of the inputs is shorter than in the source domain model, which means that they contain fewer pre-disaster patterns. Therefore, the transfer learning model should overcome both insufficient data volume and less historical data. As mentioned before, the average resilience patterns are extracted from the mixed FCD, which means that they balanced the different temporal distributions of the grid traffic volume and are the embeddings of them in a common feature space so that the model could capture the target city traffic volume distribution from the pre-disaster input and the resilience trend from the average resilience patterns. In this way, we seek to alleviate the domain shift by the unsupervised machine learning methods before training a neural network and further improve the robustness of the transfer learning models.

4. Case Study

In this chapter, we introduce the case study for our experiment, which is conducted in the context of the COVID-19 pandemic. Section 4.1 presents the background and related literature of the case study. Section 4.2 describes the situation of the study areas and Section 4.3 provides a description of the floating car data used in the study.

4.1. Background

COVID-19 virus emerged in late 2019 and spread globally in 2020, posing enormous challenges to public health and economies around the world. As of December 2022, approximately 650 million people worldwide have been infected with COVID-19 and over 6 million died. Although modern cities have ever experienced pandemics before, the knowledge of urban resilience under the impact of pandemics is still limited. The blind adoption of travel control and the urban lockdown has still led to unpredictable consequences, greatly stalling economic development and reducing the quality of life of citizens [39]. In the future, transportation systems are still likely to face pandemic challenges. Therefore, it is imminent to understand the resilience of urban transportation systems under the impact of the pandemic, which can help governments to react more quickly to the pandemic and to take more effective pandemic intervention measures in order to ensure the satisfaction of travel demands and logistics trade under the pandemic.

Unlike natural disasters, the COVID-19 pandemic did not destroy the transportation infrastructure directly but affected travel behaviors and limited the travel opportunities of citizens indirectly. In order to protect the health of citizens and reduce the economic losses caused by the pandemic, governments of different countries and regions have taken a series of emergency measures. Among them, the lockdown of event venues, short-term travel control of citizens, and quarantine policies are the most commonly used measures. According to [40], the mobility of the population is sensitive to the government's stay-at-home announcement, which changed the travel behavior and limited the travel opportunities of citizens, significantly reducing traffic demand for a period. Citizens may or have to choose to stay at home and reduce the possibilities of travel because of the government-imposed lockdown policies and the possibilities offered by businesses to work from home.

Under the pandemic control policies, the traffic volumes of many cities showed a sharp decline and then gradually recovered as the control measures were relaxed. Although "resilience triangles" showed in most of the city road networks, the impact of COVID-19 varied greatly in different cities due to their differences in topology and response policies [41]. Therefore, some researchers analyze the resilience patterns for certain cities. [42] applied the speed reduction index (SRI) as an indicator to investigate the traffic on a small road network

in South Africa during the COVID-19 pandemic. Their study shows that under the strictest lockdown policy, i.e., with an alert level (AL) equal to AL5, the traffic system experienced little congestion. In their research, as the alert level decreased to AL1, the performance of the transportation system recovered gradually but still failed to return to its pre-pandemic state [42]. However, they only implemented the experiment in a small test zone and did not exclude the impact of citizen relocation. The advances in communication and transportation technologies allow people to remote work and obtain necessary household around their residence area, so some citizens could choose to move to the outskirts of the city. However, with the reopening of urban amenities, some people were inspired to move to the urban core. [43]. Since the coverage of COVID-19 is broader than natural disasters, the resilience of urban transportation systems should be considered on a large scale. [44] explored the changes of citizens' travel behavior before and during the pandemic by traffic volume data at a city level. Their results show that traffic volumes at urban intersections during the pandemic were generally lower than before the pandemic, and the distributions of morning and evening peak hours of traffic volumes were flatter during the pandemic. In addition, there were fixed morning and evening peak hours at city intersections in specific months before the outbreak, while the peak hours became blurred in specific months during the pandemic. Their results illustrate that the COVID-19 pandemic radically changes citizens' travel behavior and only considering the situation for a certain city is inadequate to understand the changed scenario.

Therefore, learning from the experience of other cities and studying how to transfer the knowledge of resilience patterns from them is essential. In our research, we applied the proposed method on the grid traffic volume FCD from three cities: Antwerp, Bangkok and Barcelona. We used the data of Antwerp and Bangkok for source domain model training and the FCD from Barcelona for transfer learning. In the next section, we introduce the basic situation and the pandemic intervention measures of these cities.

4.2. Study Area

During the COVID-19 pandemic, the transportation systems of these three cities showed different resilience patterns, which depend on their pandemic intervention policies and the rate of disease transmission. The cities where cases first appear tended to implement prevention measures earlier to slow the spread of the disease and the cities that took measures later might be more severely affected. Overall, all these three cities implemented strict prevention measures in March 2020 and gradually eased these measures in mid-2020. Now, we introduce the basic situation and pandemic intervention policies of these three cities: Antwerp, Bangkok and Barcelona, respectively.

Antwerp is the largest city in Belgium, which is located in the Antwerp Province in the Flemish region with a geographical coordinate of 51°15'N and 4°24'E. As presented in Figure 13, the road network of Antwerp has a radial structure with a relatively dense road network in the city

center and a sparse road network in the outskirts.

Figure 13 ANTWERP road network

In order to combat the spread of the COVID-19 virus, Belgium implemented lockdown polices on March 18, 2020. The lockdown measures affected schools, restaurants, and workplaces across the entire nation. The traffic volume of the Antwerp road network decreased sharply under the pandemic intervention measures, and until May 4, 2020, as the lockdown measures were gradually eased, some urban amenities were allowed to reopen and the traffic volume started to recover.

The second city is Bangkok, the capital of Thailand, which is located in the center of the country with a geographical coordinate of 13°45'N and 100° 29'E. As presented in Figure 14, the road network of Bangkok has a ring structure and the roads within the ring have a mix of

grid and radial layouts.

Figure 14 BANGKOK road network

The pandemic intervention measures in Bangkok were implemented since January 3, 2020, when the Thai Ministry of Public Health started to screen the temperature and issue health declaration cards to travelers. On March 3, 2020, the Thai government started to prohibit large gatherings and close schools and entertainment venues. Shortly after the closure, the government imposed a curfew since March 26 and until May 17, 2020, the entertainment places were allowed to reopen. Because of the timely implementation of pandemic intervention policies, the transportation system of the Bangkok road network was relatively less affected by the pandemic and showed more resilient patterns.

Barcelona is the city for the transfer learning experiment in this research, which is located on

the northeast coast of Spain with a geographical coordinate of 41°23'N and 2° 11'E. As presented in Figure 15, Barcelona has a comprehensive network structure that consists mainly of ring and radial roads, and some blocks in the city center have a grid structure.

Figure 15 BARCELONA road network

The government of Spain implemented lockdown measures since March 14, 2020 and extended the measures until April 26. After that, the prevention measures began to ease and on May 11, 2020, citizens were gradually allowed to resume social activities. During the pandemic, the traffic volume of Barcelona experienced a sharp decline and then gradually recovered at an unstable rate.

The governments of these cities have all implemented strict prevention policies and achieved remarkable results. In the next section, we introduce the grid traffic volume FCD used in this

thesis and present the resilience patterns of these cities through data.

4.3. Data Description

The grid traffic volume floating car data is provided by HERE [45] and was used for NeurIPS Traffic4cast competitions [46]. Half of the data of each city is from February 02, 2019, to June 30, 2019, before the appearance of the COVID pandemic and the other half is from February 02, 2020, to June 30, 2020, during the first outbreak of the pandemic. Therefore, the data contains 180 days of pre-COVID patterns and 181 days during and after the first outbreak patterns. As shown in Figure 16, the data for our experiment is a (288,495,436,4) tensor for one day. The first three-dimension encode the number of 5-min time intervals and the number of 100m×100m grids for each city and the four channels encode the traffic volume of four directions of each grid. In our experiment, we merged the time interval into one day to avoid multiple seasonality and then applied two different grid sizes, i.e., 100m square and 200m square, to evaluate the stability of the traffic volume FCD.

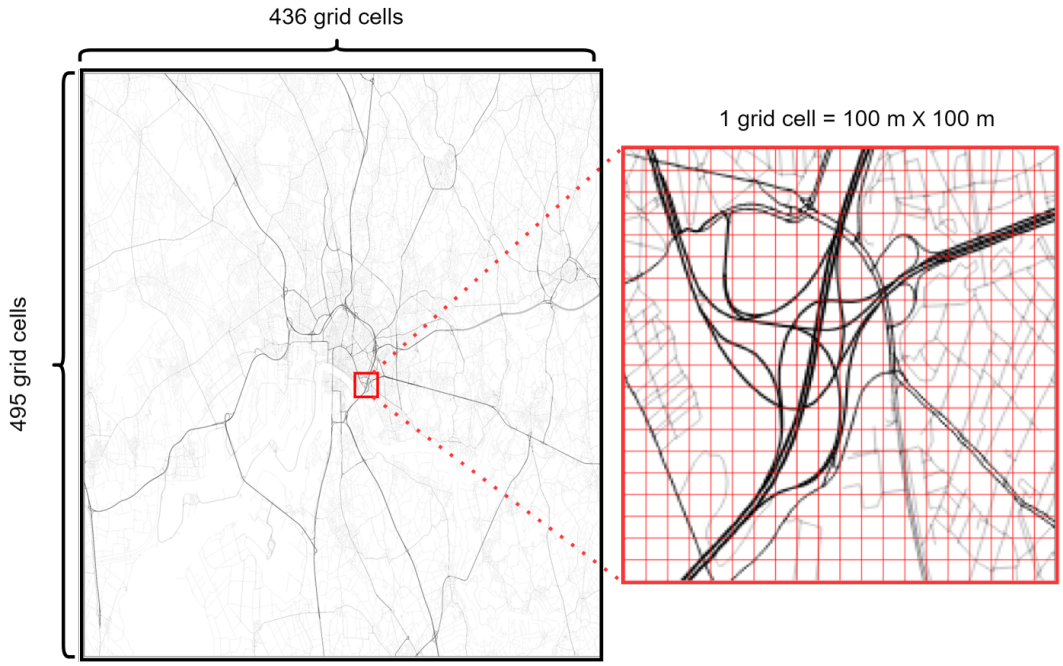


Figure 16 Grid floating car data

4.3.1. Macroscopic Traffic Volume Resilience Patterns

The traffic volume time series of the whole cities are presented in Figure 17, 18 and 19 respectively. Compared with their pandemic intervention timeline, it can be observed that they have different patterns before and during the pandemic.

As shown in Figure 17, the trend of traffic volume time series of Antwerp was relatively stable before the COVID pandemic. In the first half year of 2019, the values showed a slightly

increasing trend and remained at the same level in the first three months of 2020. After that, the government of Antwerp implemented pandemic preventive measures in March 2020, and then the traffic flow experienced a sharp decline and then reached its minimum point, which was only half of the pre-disaster level, by the end of March and maintained a low level for a period. In early May, as the preventive policies were eased, the traffic volume gradually recovered with a slightly accelerating trend. By the end of June, the overall traffic volume in Antwerp had recovered to and even surpassed its pre-pandemic level.

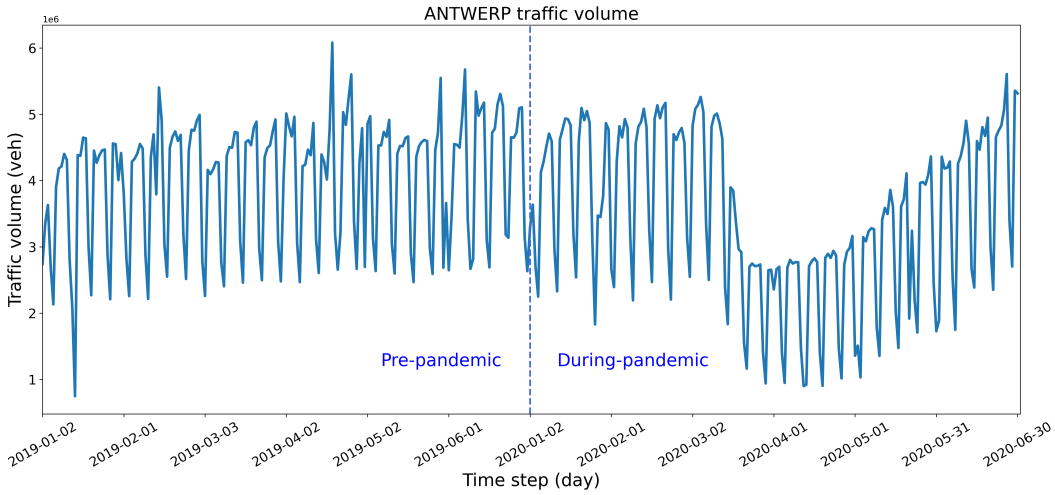


Figure 17 ANTWERP traffic volume time series

As presented in Figure 18, the trend of traffic volume of Bangkok was relatively stable before April 2019 and then declined slightly. In early 2020, because of the timely prevention measures, the traffic volume had a slightly decreasing trend and declined to the minimum point at the beginning of April. On March 2020, the curve recovered gradually but failed to reach its original state. The traffic volume of Bangkok did not show a sharply decreased during the pandemic and has an overall decreasing trend.

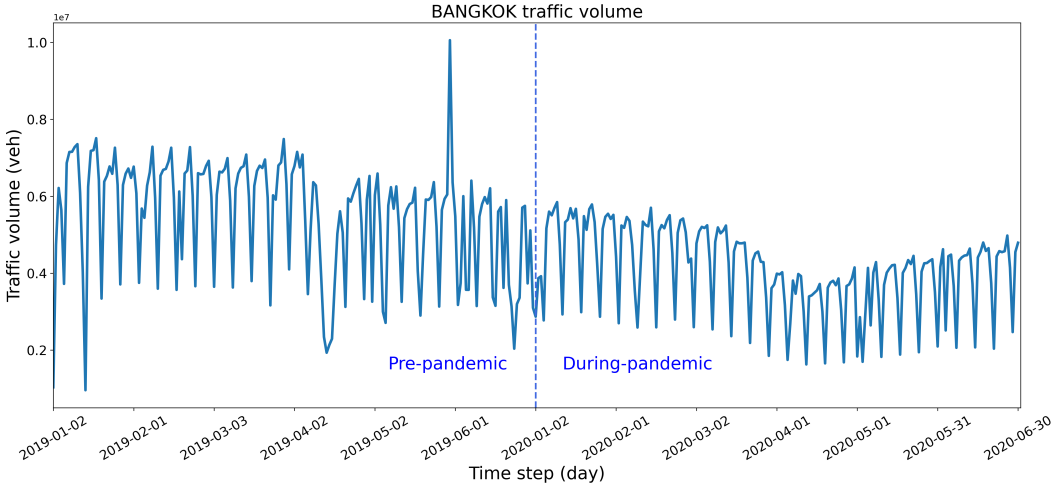


Figure 18 BANGKOK traffic volume time series

Figure 19 shows the traffic volume patterns in Barcelona. Unlike the previous two cities, the curve in 2019 fluctuated with an overall ascent trend and the traffic volume was relatively stable in early 2020. Then a plunge showed in the curve since the implementation of the lockdown policies in March and remained at a low level. After the end of the lockdown, the traffic volume showed an erratic upward trend and recovered to the pre-pandemic level in the last week.

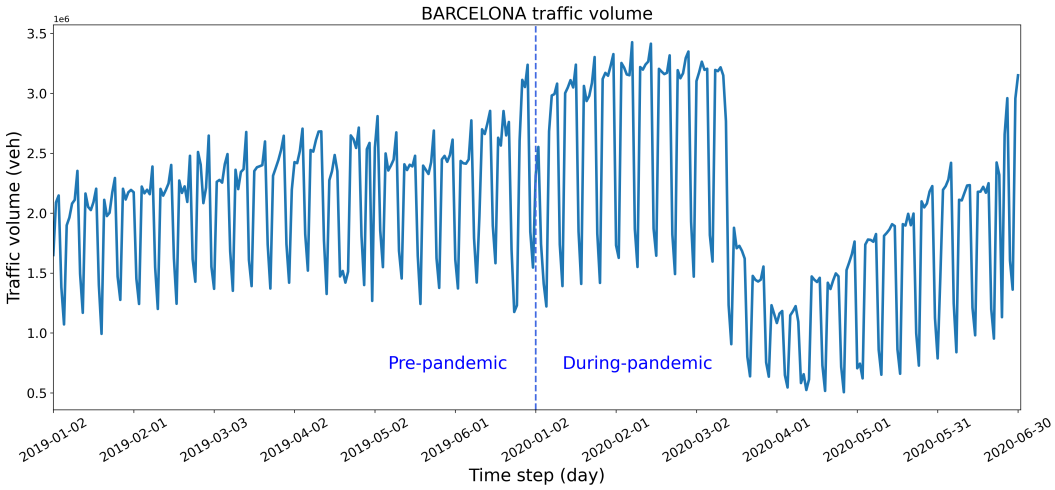


Figure 19 BARCELONA traffic volume time series

From the macro level, the overall traffic volume in each of the three cities has different trends and resilience patterns. However, from a microscopic view, the grids of each city contain various patterns, and for different cities, some grids of them could contain the same patterns. The smaller the grid size, the more the patterns. In other words, when we evaluate the patterns for the whole city, we consider the maximum grid size. However, the small grid size could increase the instability and complexity of data. Therefore, it is necessary to evaluate the transportation system resilience patterns in different grid sizes.

4.3.2. Microscopic Traffic Volume Resilience Patterns

In order to precisely show the changes in microscopic traffic volumes for each city at different resilience stages, for each city, four heatmaps are presented below for each stage. As shown in Figure 20, 21, 22, and 23, the pre-disaster heatmap is generated using the data in 2019, while the minimum point heatmap reflects the average traffic volume patterns during the lockdown policies. The post-disaster heatmap is extracted from the last week of the data set.

For a 100×100m grid size, the traffic patterns for the city can be presented more detailedly, which is beneficial for capturing the changes in traffic volume as well as traffic congestion. As shown in Figure 20, the overall traffic volume of Antwerp is mainly concentrated on the central circular road network and extended outwards. During the pre-pandemic period, there were traffic hot spots with high traffic volume in the city center and circular road network.

Further research is required to determine whether these areas are at high congestion risk. During the lockdown period, the traffic volume of the main arterials decreased significantly, while the traffic volume of some dispersed grids in the periphery remained stable. In the post-pandemic period, traffic volume on the main arterials and in the city center recovered and even exceeded the pre-pandemic level. The average traffic volume patterns in Antwerp remained at pre-pandemic levels.

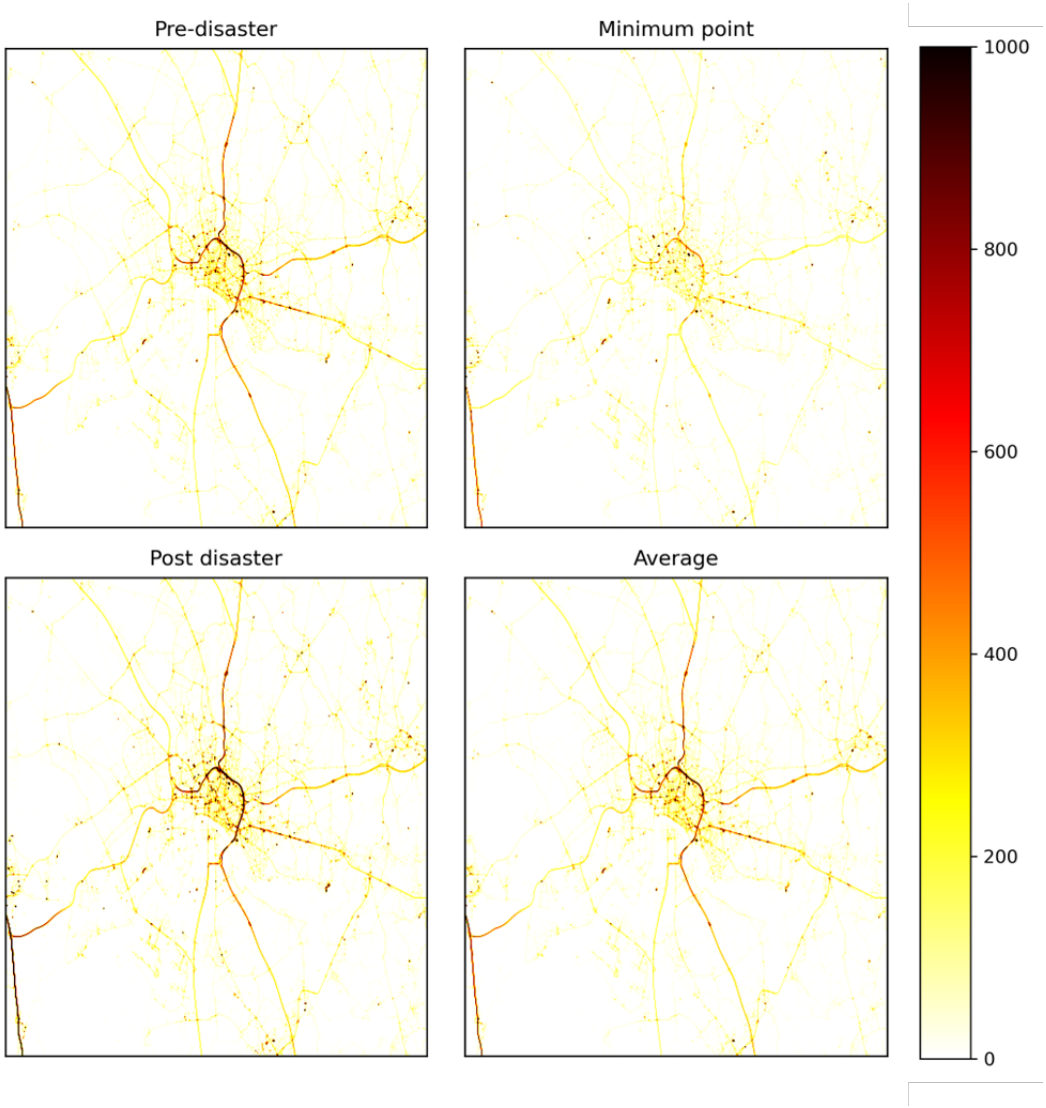


Figure 20 ANTWERP traffic volume heatmap (grid size: $100m \times 100m$ unit: veh/day)

Figure 21 illustrates the traffic volume distribution of Bangkok, which was mainly distributed on the peripheral circular and radial road networks of the city. During the pre-pandemic period, the traffic volume was primarily concentrated on the main roads close to the intersections and intersections. During the lockdown period, the traffic volume of the city center decreased significantly, and the traffic volume of the circular road network also experienced a certain degree of reduction. However, on some road sections near the intersections, the traffic volume still remained at a high level. In the post-pandemic period, the traffic volume on the circular

road network showed a trend of recovery to the pre-pandemic level, but the traffic volume of the city center still had a gap compared to the pre-pandemic level. Compared to Antwerp, the impact of the lockdown measures on the traffic volume of Bangkok was relatively mild, and the average traffic volume was close to its normal level.

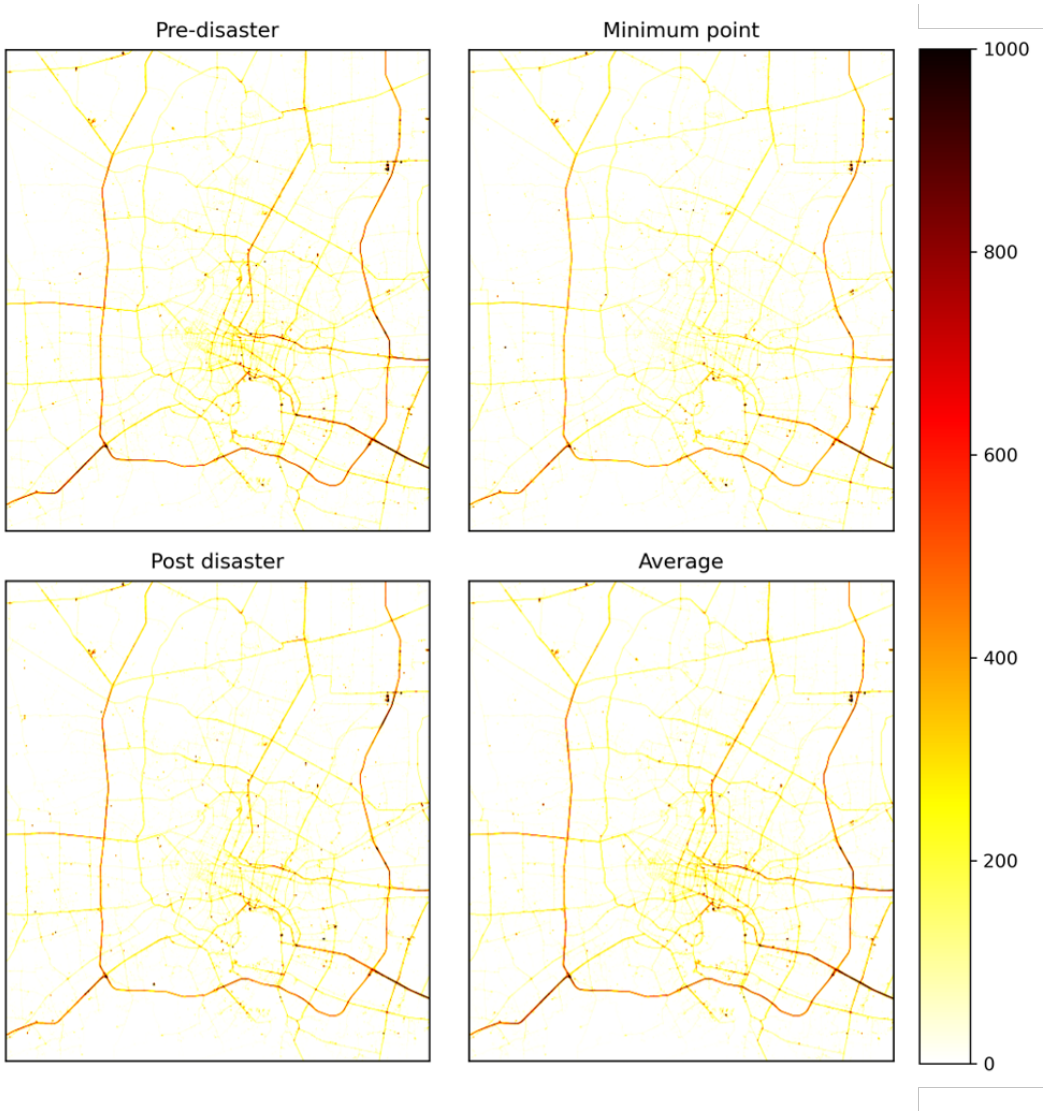


Figure 21 BANGKOK traffic volume heatmap (grid size: $100m \times 100m$ unit: veh/day)

As presented in Figure 22, the traffic volume on the main roads of Barcelona was relatively evenly distributed before the pandemic, and there were some grid clusters with high traffic volume around the main roads. During the lockdown period, the traffic volume on the main roads decreased significantly, but the traffic volume in some grid clusters did not show too much change. In addition, more clusters with relatively high traffic volume appeared in the periphery of the city, which reflects a trend of citizens moving to the outskirts of the city. In the post-pandemic period, the traffic volume of the main road network has basically returned to pre-pandemic levels, and the number of traffic-dense clusters in the periphery increased compared to before the pandemic.

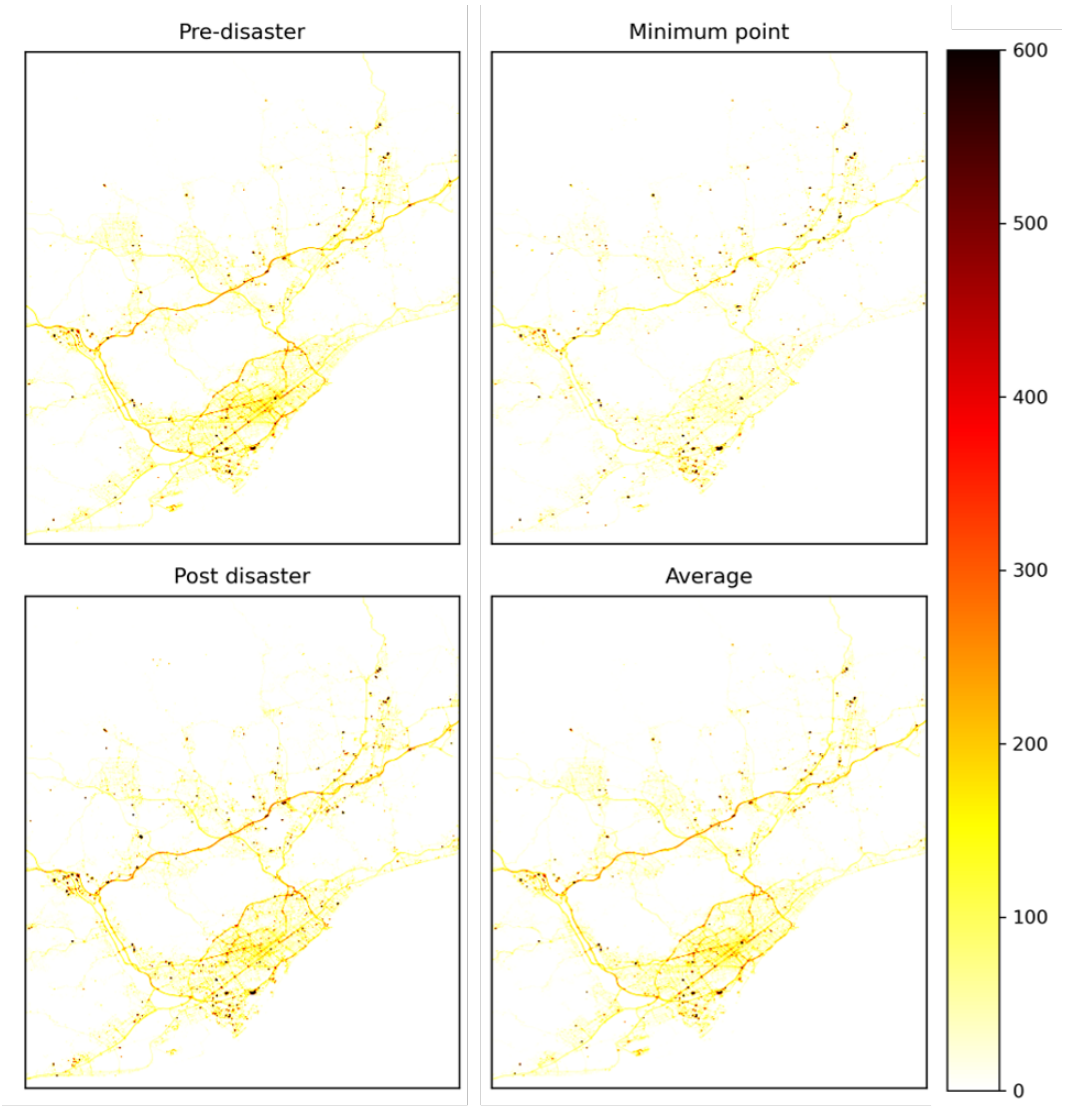
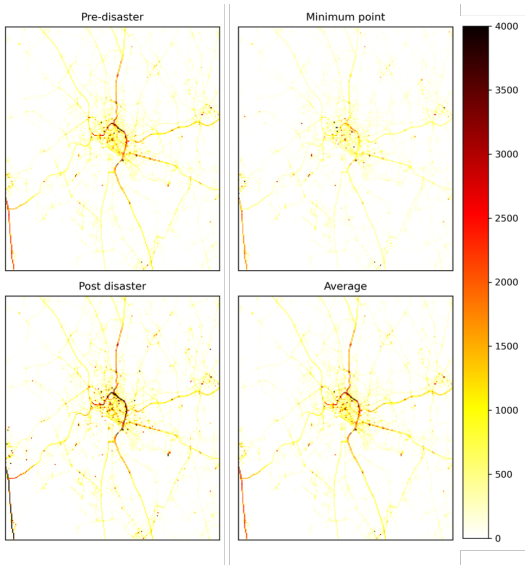
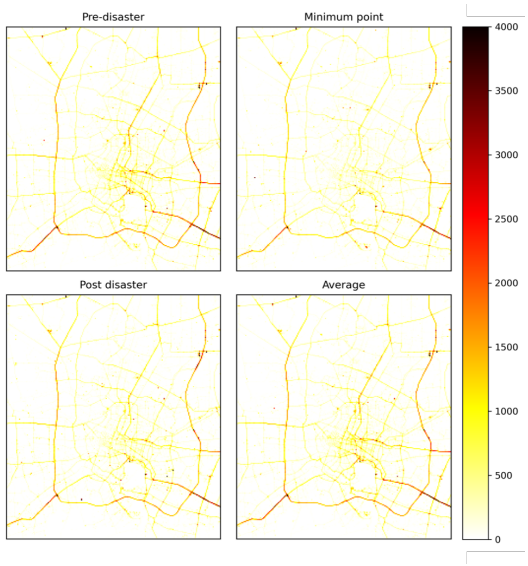


Figure 22 BARCELONA traffic volume heatmap (grid size: $100m \times 100m$ unit: veh/day)

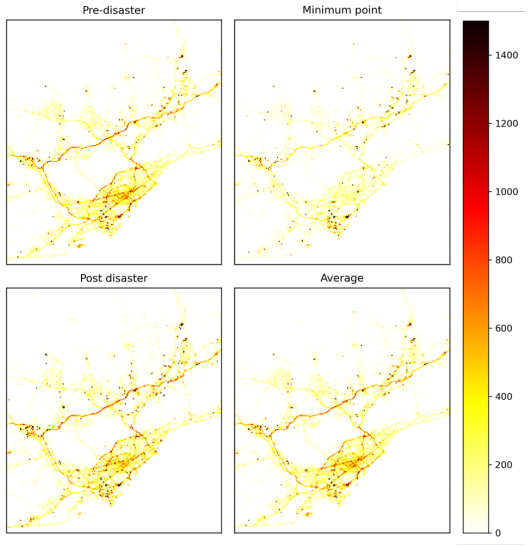
From the heatmaps with $100m \times 100m$ grid size, some mobility trends of citizens as well as the level of traffic density can be observed. In addition, some grids from different cities had similar resilience patterns, which can facilitate inter-city knowledge transfer. The heatmaps also reveal that the road network of each city is sparse, which results in a large amount of invalid grid data for each city. Moreover, for smaller grid sizes, the trends of some grids with low traffic volume are usually unstable and require more data-cleaning efforts.



(a) ANTWERP traffic volume heatmap



(b) BANGKOK traffic volume heatmap



(c) Barvelona traffic volume heatmap

Figure 23 Traffic volume heatmap (grid size: $200m \times 200m$ unit: veh/day)

Figure 23 illustrates the distribution of traffic volume in our study cities using a $200m \times 200m$ grid size. On the one hand, the traffic volume is averaged to larger grid cells, which results in a reduction of the traffic volume density in some grids and also decreases the amount of valid data for our model. On the other hand, a larger grid size can help to reduce the number of grids with unstable trends and also the noise of the grid FCD, which can better reflect the overall transportation system patterns. Although the number of available data is decreased when using a larger grid size, the percentage of valid data is higher compared to smaller grid sizes. Both large grid size and small grid size have their own advantages and disadvantages. In the next chapter, we will present our experimental process and results and discuss the performance of our model under different grid size conditions.

5. Experiment Setup

5.1. Data Preprocessing

Data preprocessing involves several steps: (1) aggregate the traffic volume into one-hour intervals; (2) pad and join the data based on the grid size (2) reshape the data to three dimensions (cities, samples, time steps); (3) set a threshold to eliminate part of abnormal data; (4) run a k-shape clustering for each city and delete the abnormal cluster; (5) normalize the rest data and reshape the data for source domain cities to the shape (samples, time steps) and shuffle the data. (6) shuffle the data of the target domain city. (7) For both the source domain and target domain, 60% of the data are used for training, 20% are used for testing and the rest 20% for validation.

After the data preprocessing, the data splitting is shown in Table 2. Although the grid area has been increased fourfold, we are still able to extract a sufficient amount of data from the FCD dataset, which represents 60% of the $100m \times 100m$ dataset. For the source domain cities, we used 180-day pre-disaster traffic volume FCD to predict the 181-day resilience patterns. For the target domain task, we used only 50 days of pre-disaster data to predict 181-day resilience patterns. Therefore, the total input sequence length of the source domain time series is 361 and 231 of the target domain.

Table 2 Data volume

Data	Antwerp+Bangkok		Barcelona	
Grid size	$100m \times 100m$	$200m \times 200m$	$100m \times 100m$	$200m \times 200m$
Training set	(48881,361)	(30634,361)	(14370,231)	(8805,231)
Test set	(16295,361)	(10213,361)	(4791,231)	(2935,231)
Validation set	(16295,361)	(10213,361)	(4791,231)	(2935,231)
Total	(81471,361)	(50700,361)	(23952,231)	(214675,231)

For $100m \times 100m$ grid size, each city contains 495×436 grid cells and each grid contains the traffic volume time series from four different directions, which means that for each city, there is a maximum of 863280 traffic volume time series can be extracted. For $200m \times 200m$ grid size, each city contains 248×218 grid cells after padding and each grid also contains the traffic volume time series from four different directions. Therefore, the dataset of each city contains a maximum of 216256 time series. Based on the data cleaning results, the valid rate of the datasets can be derived as Table 3:

Table 3 Data valid rate

Data	Antwerp+Bangkok		Barcelona	
Grid size	$100m \times 100m$	$200m \times 200m$	$100m \times 100m$	$200m \times 200m$
Total data volume	1726500	432512	863280	216256
Valid data volume	81471	50700	23952	14675
Valid rate	4.71%	11.72%	2.77%	6.78%

For $100m \times 100m$ grid size, the valid rate of the Antwerp+Bangkok dataset is 4.71%, and the valid rate of the Barcelona dataset is only 2.77%. However, when the grid size is extended to $200m \times 200m$, the valid rate of the datasets is 11.72% and 6.79% respectively, which is 2.5 times higher than the 100 m * 100 m grid size.

5.2. Experiment Settings

The architectures of the neural networks are shown in Table 4. Our experiment is deployed in a Windows 11 System with 32 GB RAM and NVIDIA RTX 3080 Laptop Graphics Cards (16 GB). The data processing and modeling are performed using Python 3.8 with Tslearn package for k-Shape time series clustering, Pymc package for Bayes method, Pandas, Numpy, and Pytorch package for deep learning. For more effective learning, we employed the ReLU activation function, Adam optimizer, Mean Square Error (MSE) loss function, EarlyStoping technique and learning rate decay technique. For learning rate decay, the initial learning rate is set to 0.001, and we applied the step decay schedule with the decay rate of 0.75 per 50 steps. The batch size for all models is taken as 64.

For our proposed model, we first applied a Bayes patterns extractor to generate the average resilience patterns for each cluster, the sequence length of which is 181. Therefore, the input sequence length of each neural network is the sum of the pre-disaster sequence length and the average resilience patterns length, which is 361. And for the ablation experiment, we directly feed the pre-disaster sequence into the neural network with the same architecture as our proposed model. For each RNN, we added a feedforward network as the prediction head to generate the output. For transfer learning, the parameters of all models were frozen, and their outputs were fed into a feedforward neural network for fine-tuning.

Table 4 Model architecture

Model	Number of Layers	Hidden Size	RNN Predition Head	Input Sequence Length	Output Sequence Length
Bayes+MLP	3	Layer 1 = 256 Layer 2 = 128 Layer 3 = 181		361	181
Bayes+RNN Bayes+LSTM Bayes+BiLSTM	3	Layer 1 = 16 Layer 2 = 16 Layer 3 = 16	Layer 1 = 256 Layer 2 = 128 Layer 3 = 1	361	181
Bayes+MLP	3	Layer 1 = 256 Layer 2 = 128 Layer 3 = 181		180	181
RNN LSTM BiLSTM	3	Layer 1 = 16 Layer 2 = 16 Layer 3 = 16	Layer 1 = 256 Layer 2 = 128 Layer 3 = 1	180	181

For model performance evaluation, we employed Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Time Dynamic Warping Distance (DTW) as metrics, which are defined as:

$$MAE = \frac{1}{n \times t} \sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right| \tag{5.1}$$

$$RMSE = \frac{1}{t} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2} \tag{5.2}$$

$$DTW = \frac{1}{n} \sum_{i=1}^{n} |W_i| \tag{5.3}$$

Where Y_i refers to the ground truth, $\hat{Y}_i$ represents the predictions, n denotes the number of samples, t is the total prediction length, and W_i is the warping path length between the ground truth and predictions (see Section 3.3.1).

6. Results and Discussions

This chapter summarizes the performance and characteristics of our models. The first two sections present the results and performance of the models using $100m \times 100m$ grid size datasets in both the source and target domains. In section 6.3, we conduct a sensitivity analysis of the models using $200m \times 200m$ grid size datasets. Finally, in section 6.4, we present the results of the ablation experiment.

6.1. Results of Source Domain Model

6.1.1. Average Resilience Patterns

The number of clusters in this thesis was defined by the Elbow method. The Elbow method involves plotting the within-cluster sum of squares (WCSS) against the number of clusters. WCSS is the sum of the squared distance between each point and the centroid in a cluster, and the "elbow" of the WCSS curve is selected as the number of clusters to use. As shown in Figure 24, we picked four clusters rather than three clusters as the number of clusters for the normalized source domain training pre-disaster data to generate more average resilience patterns.

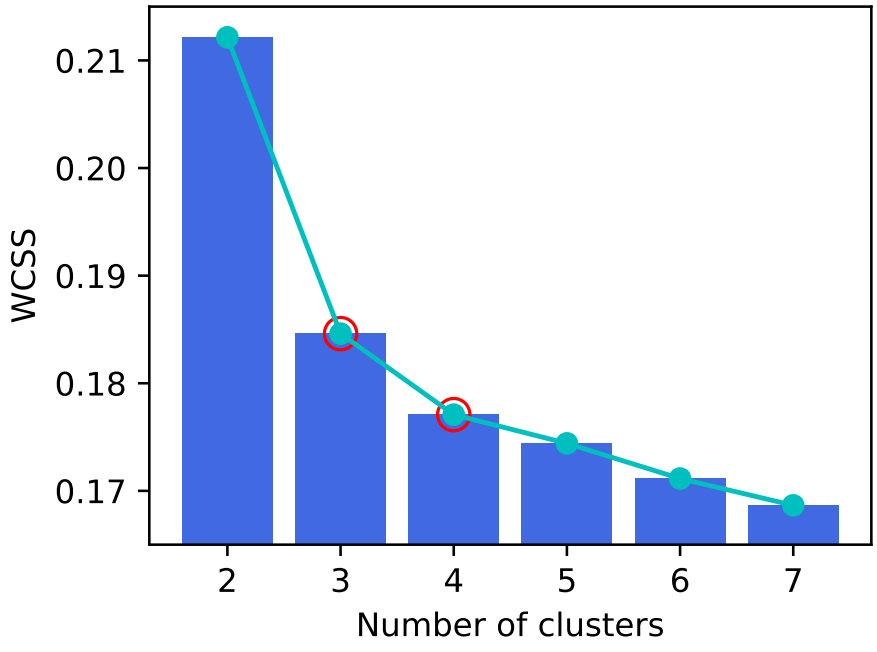
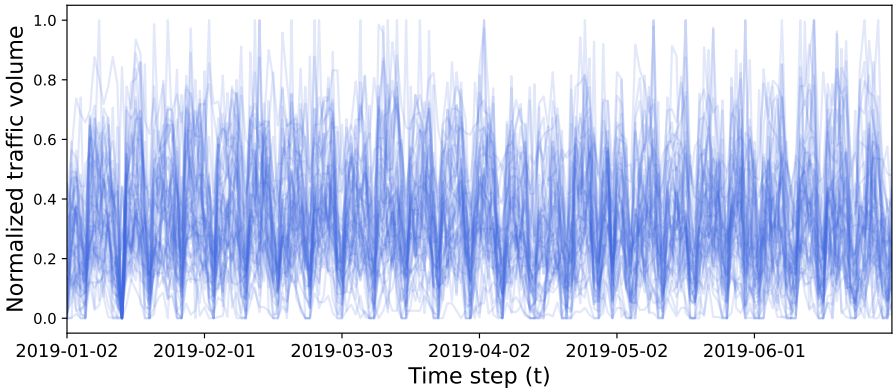
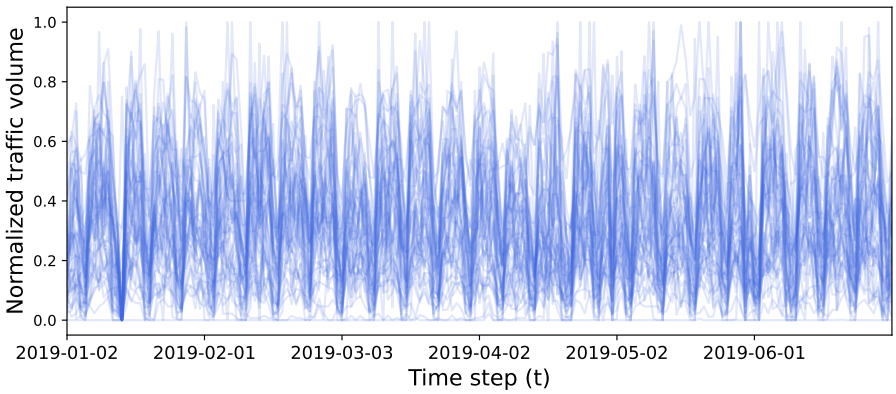


Figure 24 Within-Cluster Sum of Square of different number of clusters (Grid size: $100m \times 100m$)

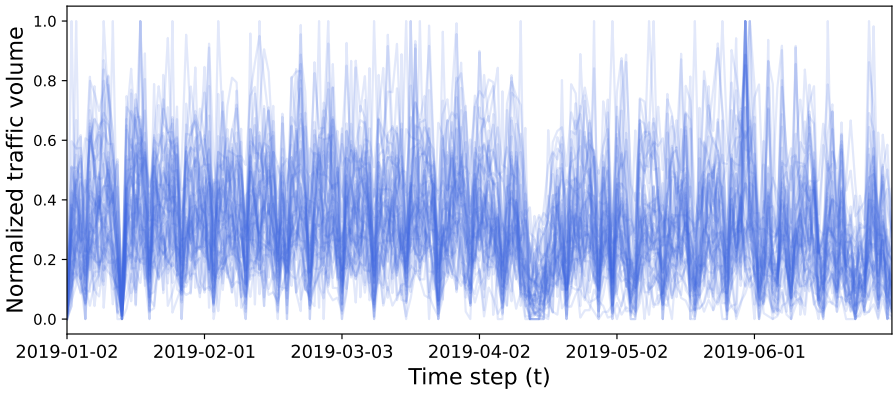
Based on the WCSS curve, the k-shape method was applied to cluster the pre-disaster time series into four clusters. Figure 25 presents 50 samples of pre-disaster time series in each cluster. As shown in Figure 25(a) and 25(b), the first two clusters contained the normalized traffic volume pre-disaster time series with overall stable trends, while the elements of the last two clusters (see Figure 25(c) and 25(d)) contained some unstable trends. In addition, an apparent bolded trend curve can be observed among the samples of each cluster, which means that some time series in the same cluster have high similarity, and the shades of color reflect the number of similar time series in these samples.



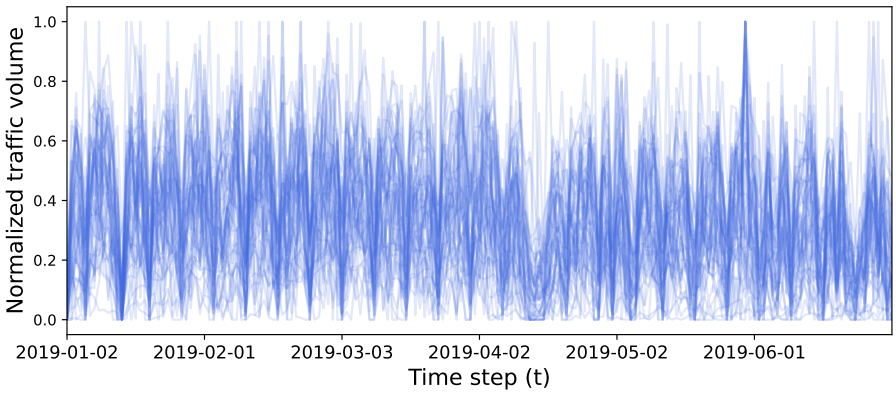
(a) Cluster 0 (pre-disaster)



(b) Cluster 1 (pre-disaster)



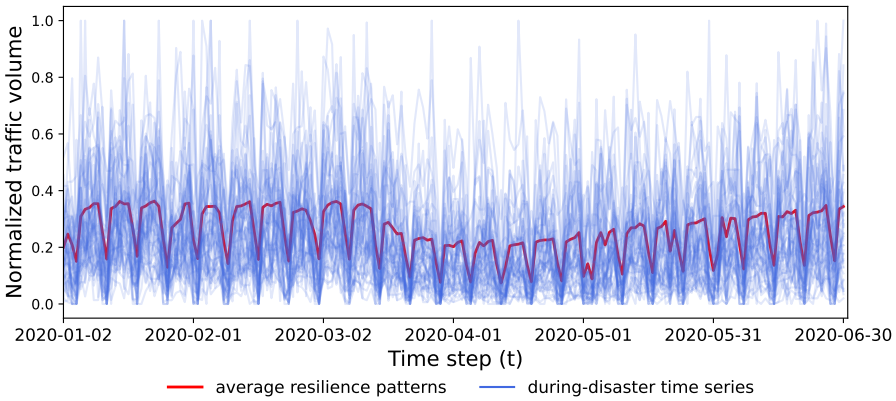
(c) Cluster 2 (pre-disaster)



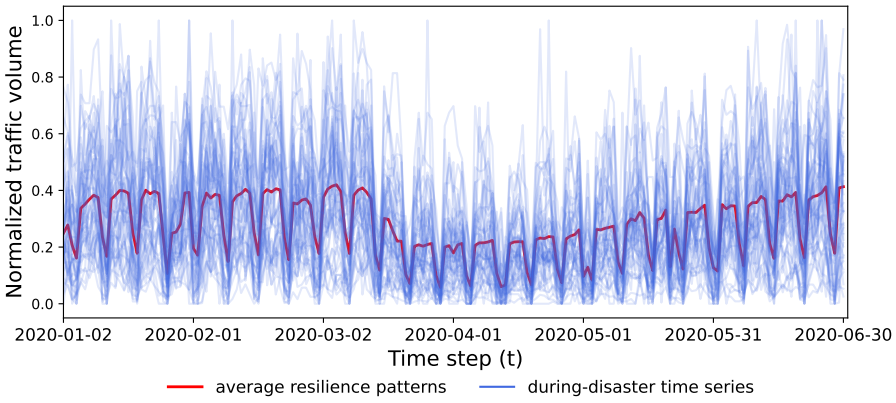
(d) Cluster 3 (pre-disaster)

Figure 25 Samples of pre-disaster clusters (Grid size: $100m \times 100m$)

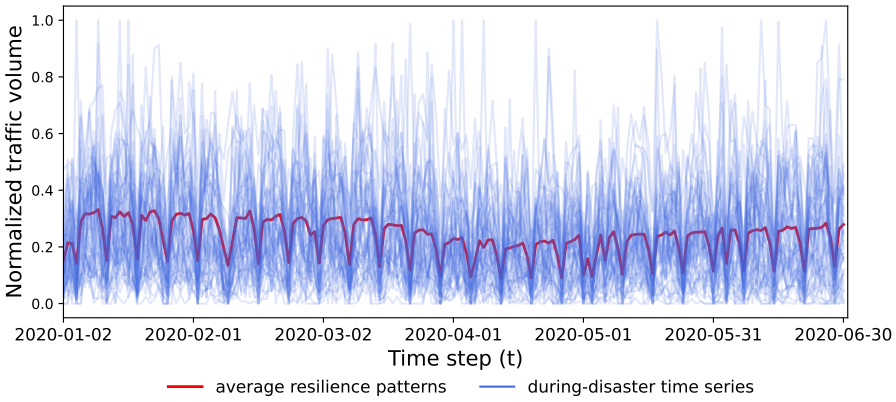
According to the clustering results of the pre-disaster time series, the resilience patterns can also be correspondingly divided into four clusters. Figure 26 presents the samples of during-disaster time series and the extracted average resilience patterns from all elements in each cluster. Consistent with pre-disaster clusters, an apparent bolded trend curve can also be observed among the samples of each cluster. The average resilience patterns reflect the possible normalized traffic volume with the highest degree of confidence in each cluster. As shown in Figure 26(a), the average resilience patterns of cluster 0 showed a slightly downward trend in mid-March 2020 and a slow upward trend since April 2020. In contrast, the average resilience patterns of cluster 1 (see Figure 26(b)) presented a more distinct "resilience triangle" shape. The extracted time series in Figure 26(c) and 26(d) had an overall downward trend and slight resilience patterns.



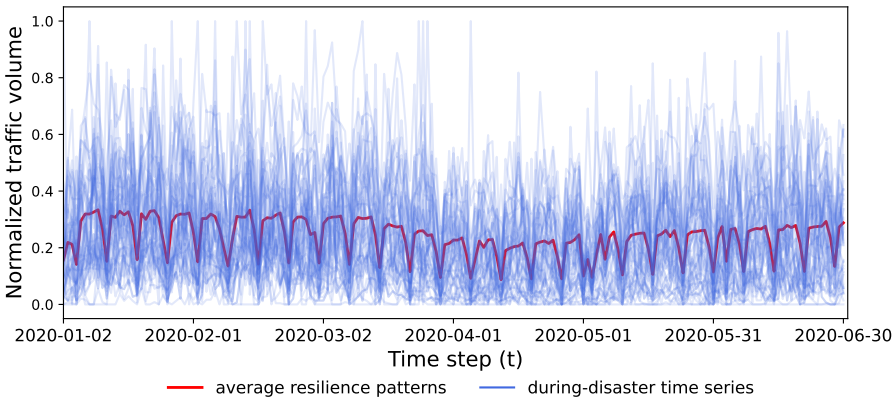
(a) Resilience patterns of cluster 0



(b) Resilience patterns of cluster 1



(c) Resilience patterns of cluster 2



(d) Resilience patterns of cluster 3

Figure 26 During-disaster clusters and average resilience patterns (Grid size: $100m \times 100m$)

The above-presented average resilience patterns of each cluster are the output of the unsupervised resilience patterns extractor and also part of the input to the neural networks. In the next subsection, we summarize and discuss the results of the source domain neural networks.

6.1.2. Results of Neural Networks

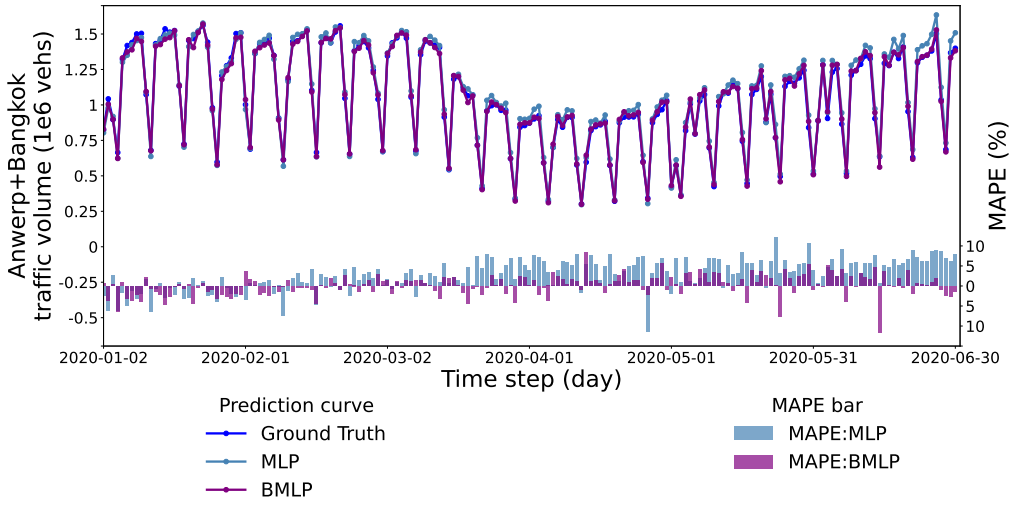
Table 5 summarizes the performance of different models under different metrics when predicting traffic volume resilience patterns at the grid level. DTW measures the similarity between time series, MAE measures the difference between prediction and actual value of each time step, and RMSE emphasizes the impacts of the significant errors.

Table 5 Source domain loss (Grid size: $100m \times 100m$)

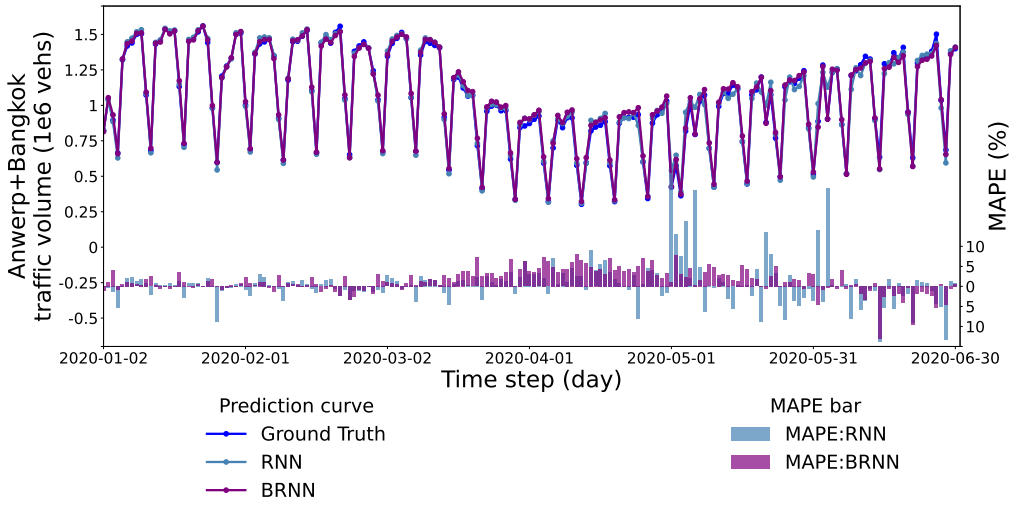
Model & Input features	DTW	MAE	RMSE
BMLP (180+181)	260.607	17.223	40.0486
BRNN (180+181)	273.541	19.134	44.745
BLSTM (180+181)	262.951	17.498	40.887
BBiLSTM (180+181)	261.848	17.747	41.504

In the source domain, it is evident from the results that the Bayes patterns extractor-based MLP model (BMLP) achieved the best performance under all three metrics. In contrast, the Bayes patterns extractor-based RNN (BRNN) model performed worst. The Bayes patterns extractor-based LSTM model (BLSTM) and Bayes patterns extractor-based BiLTSM model (BBiLSTM) achieved similar performance close to BMLP. For DTW, the performance of the BMLP, BLSTM, and BBiLSTM were almost identical, with a difference of less than 1%, while the performance of BRNN was worse. It appears that the BRNN model failed to capture the similarity of the traffic volume time series. For MAE, the BLSTM and BBiLSTM models achieved similar performance as the BMLP model, while the MAE of the BRNN model was significantly lower than other models, which means that the BRNN model had a relatively poor prediction accuracy. For RMSE, the BMLP and BLSTM models had relatively low values of around 40, significantly lower than the value of the BRNN model. The BBiLSTM model showed an RMSE value of 41.504 which is slightly larger than the BLSTM model. The RMSE values show that the prediction of the BMLP and BLSTM model were less influenced by significant errors than BiLSTM and BRNN models.

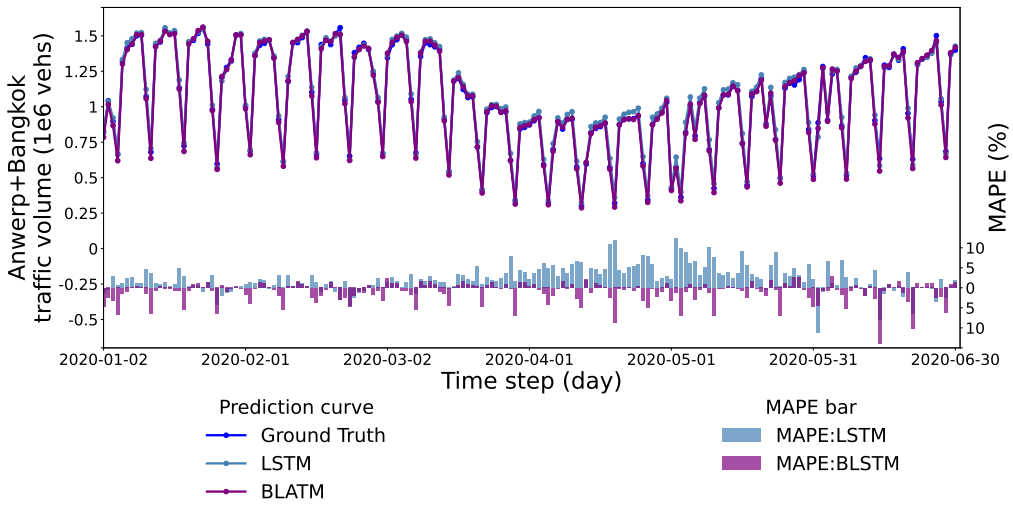
It can be noticed that the BMLP model exhibited more robust prediction performance in our experiment than RNN-based models in the source domain. The possible reason could be the cumulative error caused by the long prediction length. In RNN-based models, the BLSTM and BiLSTM performed better than BRNN because the gate mechanism can enhance the ability of the LTSM-based model to capture the long-term temporal dependencies and achieve similar performance as the BMLP model.



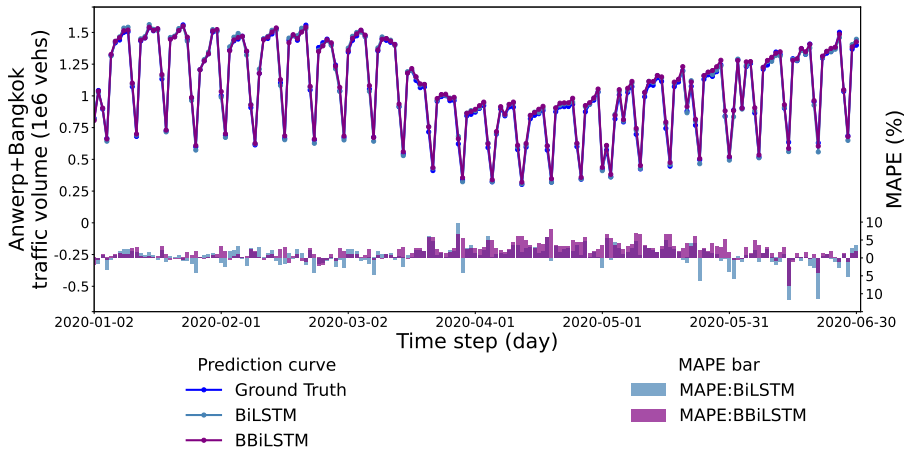
(a) MLP model



(b) RNN model



(c) LSTM model



(d) BiLSTM model

Figure 27 The macro level prediction and true values in source domain (Grid size: $100m \times 100m$)

To visualize the overall absolute error of the models, the line charts of Figure 27 present the prediction curves of our models and the Bayes ablated models (refer to Section 6.4 for more details). It can be observed from the line chart that the predictions of the Bayes-based model are almost consistent with the ground truth. As shown in Figure 27(a), the BMLP exhibited strong robustness over the entire forecasting interval but visible bias in predicting the peak value for the first month. Additionally, it had relatively low precision at the time steps with significant changes in trends, and the occurrence times of error were relatively random, which indicates that the BMLP can not capture the dependency between adjacent time steps but have a strong ability in learning the overall shape of the resilience patterns. As shown in Figure 27(b), the BRNN performed well in the prediction from January to March. However, the model showed relatively poor performance when predicting the period with large fluctuations, i.e., April and May. This result suggests that the BRNN has limitations in capturing the long-term dependencies of the sequence and is not sensitive enough to the changes in the trends of the time series. The BLSTM model was relatively robust on the overall prediction interval, as displayed in Figure 27(c). Compared to BMLP, BLSTM performed better in peak value prediction. In contrast to BRNN, BLSTM can capture significant changes in trends. However, the BLSTM was unstable in predicting the valley values of each period in our experiment and focused more on the trends with large traffic volumes. As shown in Figure 27(d), the BBiLSTM performed better in predicting peak-to-peak values in the first few months. However, for the prediction from mid-March to the end of May, BiLSTM continuously overestimated the traffic volume. Although BiLSTM can capture future temporal dependencies, the upward trend during the recovery phase might lead to the overestimation of the redundancy and robustness of the transportation systems.

In order to quantify and compare the performance of different models in different experiments for the macro-level prediction, we further introduced the mean absolute percentage error (MAPE) as the metric. The bar charts of Figure 27 present the MAPE of different neural

networks for each time step. For the proposed models, BMLP and BLSTM achieved better overall performance than BRNN and BBiLSTM in MAPE. For RNN-based models, BRNN and BBiLSTM showed better predictive performance in the first three months and relatively large MAPE in April and May, while BLSTM had a stable performance in the whole prediction length. For BMLP, although some significant bias in the line chart can be observed in the first month, the large absolute traffic volume keeps the MAPE in a small range.

In general, our models showed strong robustness in predicting the overall traffic volume of the city in the source domain, and each model has its advantages and disadvantages. Therefore, the models should be selected under the consideration of specific tasks and goals.

6.2. Results of Target Domain Model

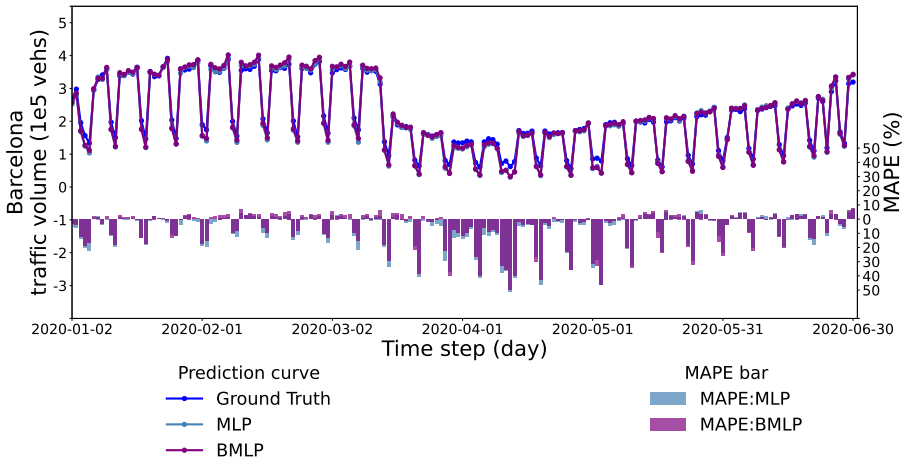
In the target domain, the input was only 50 days of the Barcelona pre-disaster traffic volume time series, and a feedforward neural network was stacked to each pre-trained model for fine-tuning. Table 6 lists the prediction performance comparison of different models under different metrics at the grid level.

Table 6 Target domain loss (Grid size: $100m \times 100m$)

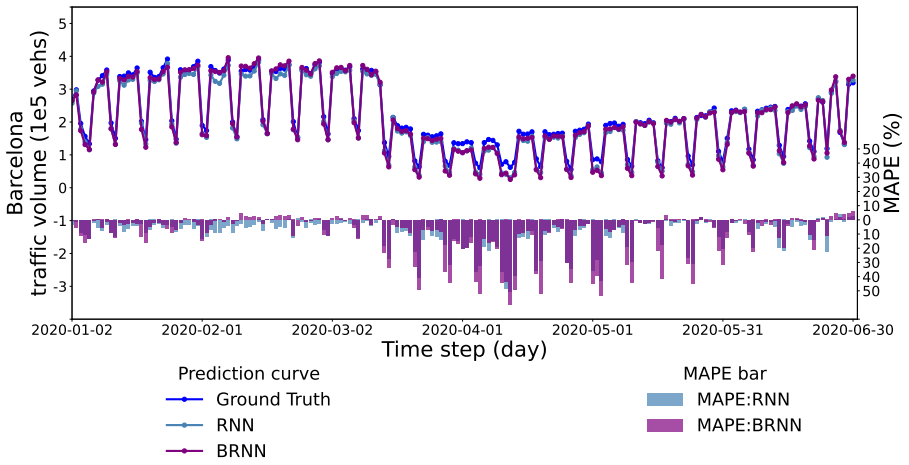
Model & Input features	DTW	MAE	RMSE
BMLP (50+181)	309.585	20.458	82.508
BRNN (50+181)	317.416	21.580	84.702
BLSTM (50+181)	311.488	20.830	85.225
BBiLSTM (50+181)	311.215	21.501	84.558

In general, similar to the results of the source domain models, the BMLP showed the strongest robustness across all three metrics, while the BRNN performed the worst, and the BLSTM and BBiLSTM exhibited similar performance. The DTW results were consistent between the target and source domains. The BMLP-based transfer learning model still exhibited the most prominent performance in capturing the similarity between traffic volume time series. BLSTM and BBiLSTM exhibited similar DTW performances, which were close to BMLP. However, the BRNN still performed poorly in learning time series similarity. The MAE results show that both BLSTM and BMLP achieved high accuracy in our traffic volume prediction task, while BBiLSTM and BRNN had approximately 4% lower prediction accuracy. In terms of RMSE, BMLP achieved significantly lower errors than other models. Although BLSTM exhibited the highest accuracy in MAE error, BBiLSTM and BRNN models were less affected by significant errors. Compared to the source domain tasks, the performance of the BRNN was closer to other models. In the source domain, the performance of BRNN and other models differs by approximately 5% in terms of the DTW metric. However, in the transfer learning experiment, the gap had narrowed to approximately 2-2.5%. Regarding MAE, the performance of BRENN was even closer to BBiLSTM and exhibited better performance than in the source domain.

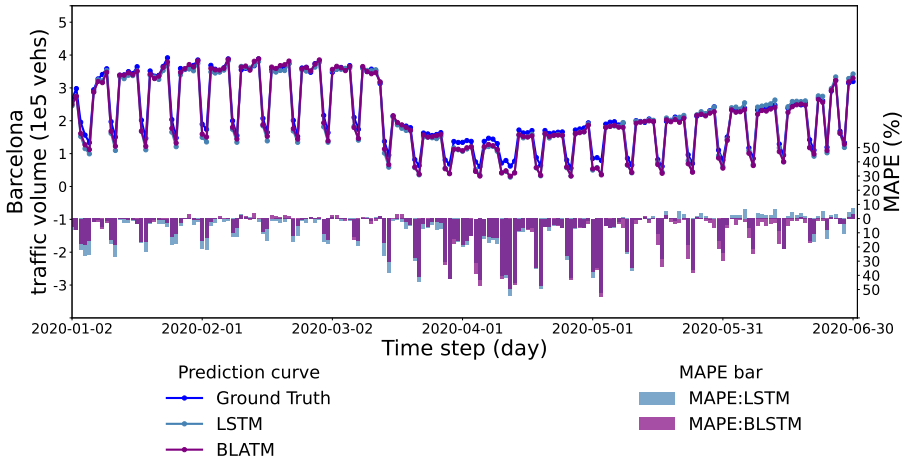
For RMSE, BRNN's performance surpassed BLSTM and was close to BBiLSTM. This result suggests that although RNN models have a relatively poor ability to capture the similarity of traffic volume time series at the micro level, they exhibited good generalization ability in predictive accuracy.



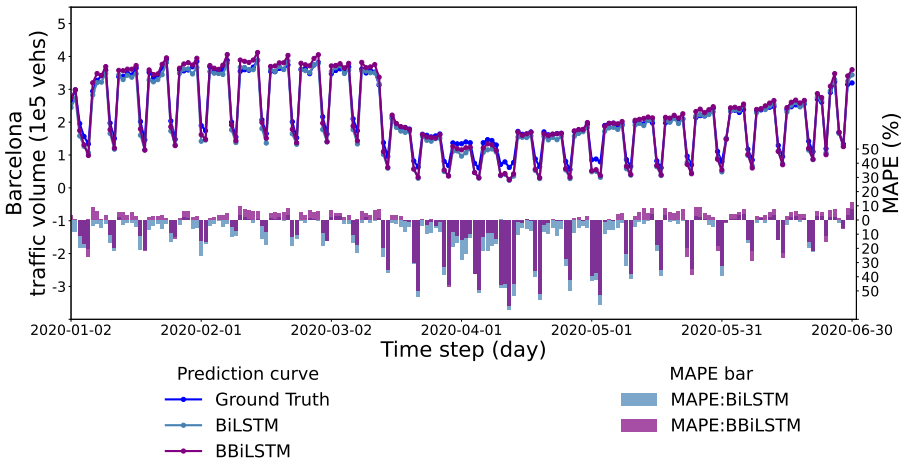
(a) MLP model



(b) RNN model



(c) LSTM model



(d) BiLSTM model

Figure 28 The macro level prediction and true values in target domain (Grid size: $100m \times 100m$)

From a macro perspective, all four models achieved robust prediction performance for relatively stable trends but exhibited particular bias in predicting the valleys of the overall resilience pattern, as shown in the line charts of Figure 28. Figure 28(a) illustrates the prediction results of the BMLP model in the target domain. It can be observed that the BMLP model had generally robust performance and better prediction for the periods with significant trend changes. Figure 28(b) shows that although the BRNN model could capture the overall trend of the time series, its predictive ability for the valleys of the resilience patterns is poor. Compared to the BRNN model, as shown in Figure 28(c), the BLSTM model captured the downward trend of the resilience patterns better but overestimated the robustness and resourcefulness of the transportation systems. Figure 28(d) shows the prediction results of the BBiLSTM, which showed a more robust performance in predicting the traffic volume of the recovery phase of the time series than the BLSTM.

However, as shown in the bar charts, all of the target domain models experienced large MAPE when the traffic volume was relatively low. Due to the inconsistency of traffic volume distribution between the source domain cities and Barcelona and the instability of data due to the small grid size, the models may show poor prediction performance for the grids and periods with small traffic volume. The influence of low data stability on the RNN-based models may be more significant than MLP because the predictions of RNN-based models highly depend on the correlation between traffic volumes at different time steps. On the other hand, MLP does not consider the dependency between the prediction results. Additionally, as the amount of the input pre-disaster information decreased in the transfer learning, the stability of the grids could significantly determine the quality of the model's input, which is particularly important for the prediction performance of the models. Therefore, to analyze the impact of grid size on models' prediction performance, we conducted a sensitivity analysis for the input data with a grid size of $200m \times 200m$.

6.3. Sensitivity Analysis

6.3.1. Average resilience patterns

As shown in Figure 29, the elbow points of the WCSS curve appeared when the number of clusters was 3 and 4. Similar to the previous experiment, we picked four rather than three as the number of clusters for the normalized source domain training pre-disaster data to generate more average resilience patterns.

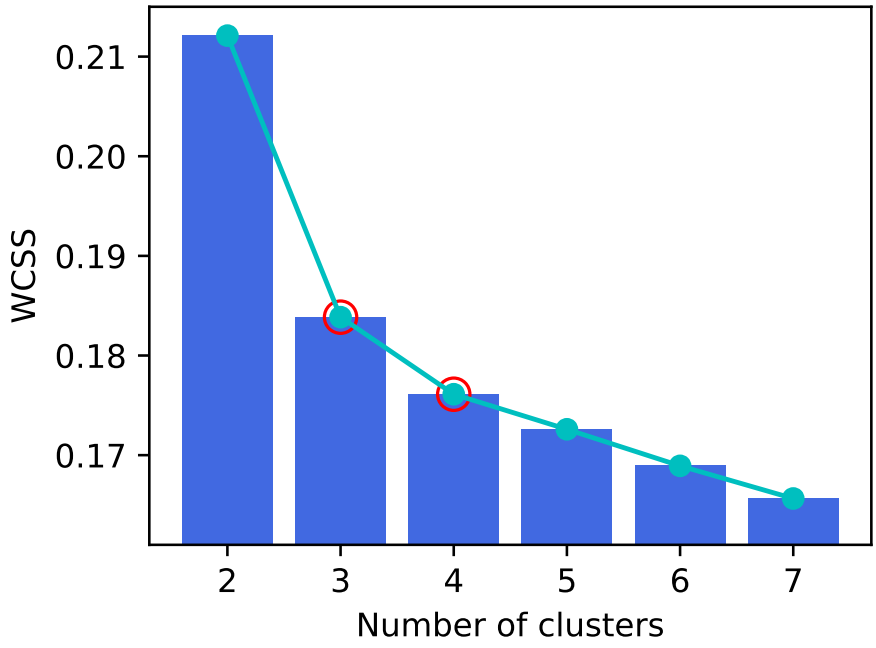
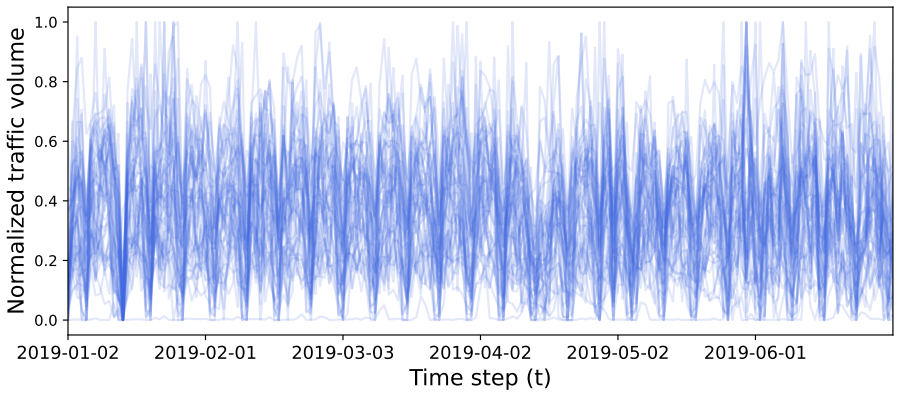
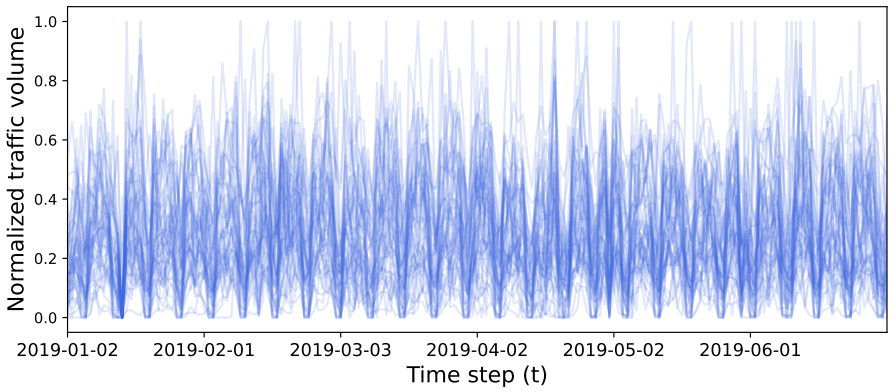


Figure 29 Within-Cluster Sum of Square of different number of clusters (Grid size: $200m \times 200m$)

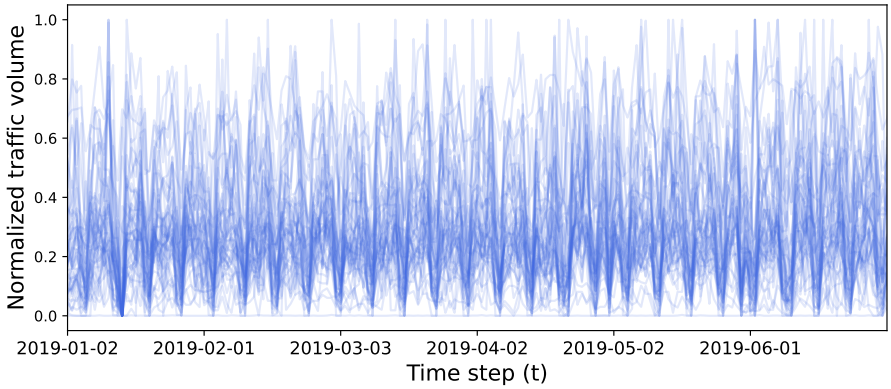
Figure 30 presents 50 samples of pre-disaster time series in each cluster, which is generated through the k-shape method. As shown in Figure 30(a), 30(b), and 30(c), the first three clusters contained the normalized traffic volume pre-disaster time series with overall stable trends, while the elements of the last cluster (Figure 30(d)) contained some unstable trends.



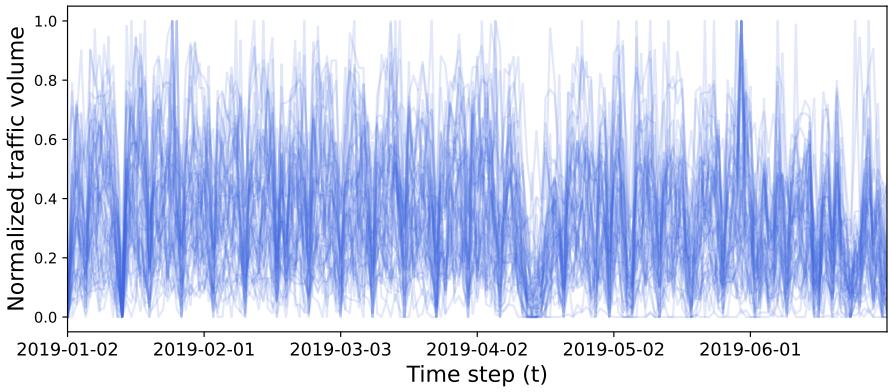
(a) Cluster 0 (pre-disaster)



(b) Cluster 1 (pre-disaster)



(c) Cluster 2 (pre-disaster)

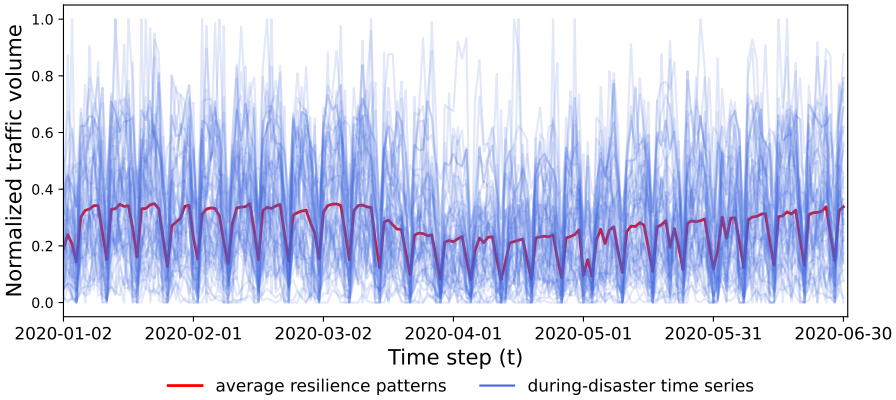


(d) Cluster 3 (pre-disaster)

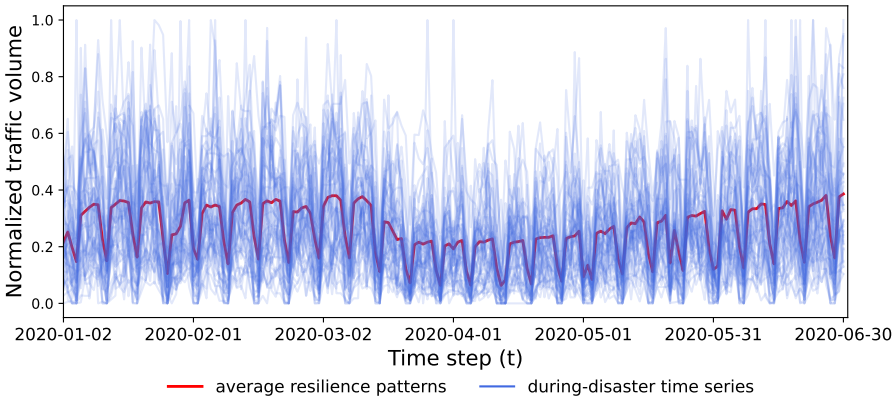
Figure 30 Samples of pre-disaster clusters (Grid size: $200m \times 200m$)

Figure 31 presents the samples of during-disaster time series and the extracted average resilience patterns from all elements in each cluster. As shown in Figure 31(a), the average resilience patterns of cluster 0 showed a slightly downward trend in mid-March 2020 and a slow upward trend since April 2020. In comparison, the average resilience patterns of cluster 1 and cluster 2 (see Figure 31(b) and 31(c)) had a slightly upward trend since January 2020 and presented a more distinct “resilience triangle” trend. Consistent with the small grid size

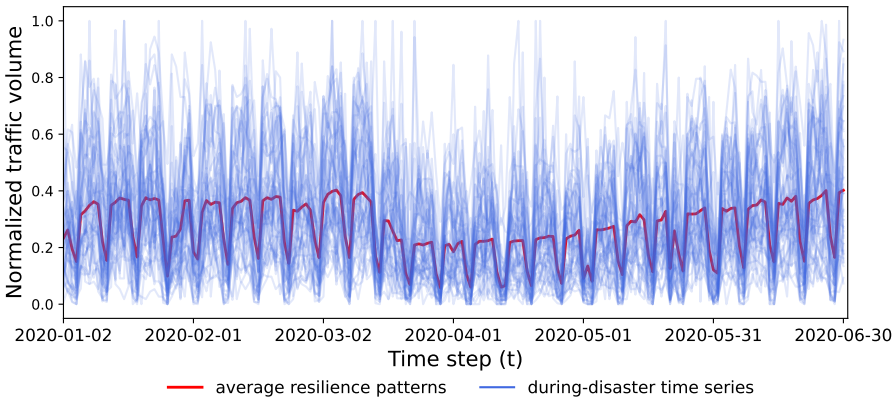
experiment, the extracted time series in Figure 31(d) had an overall downward trend and slight resilience patterns.



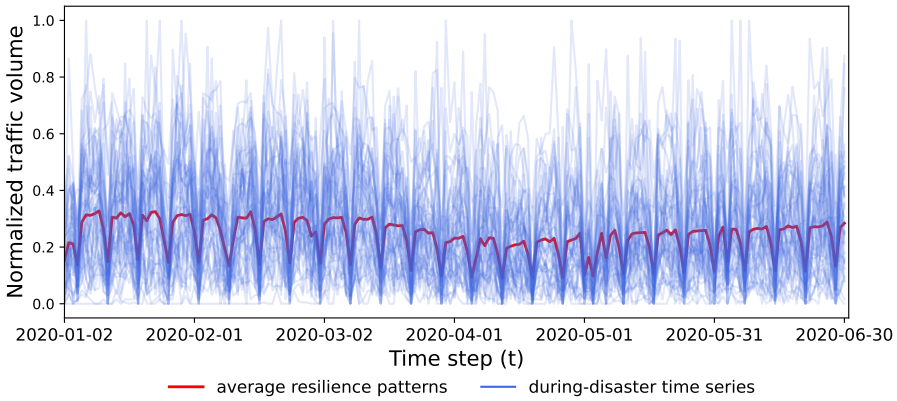
(a) Resilience patterns of cluster 0



(b) Resilience patterns of cluster 1



(c) Resilience patterns of cluster 2



(d) Resilience patterns of cluster 3

Figure 31 During-disaster clusters and average resilience patterns (Grid size: $200m \times 200m$)

6.3.2. Source Domain Sensitivity Analysis

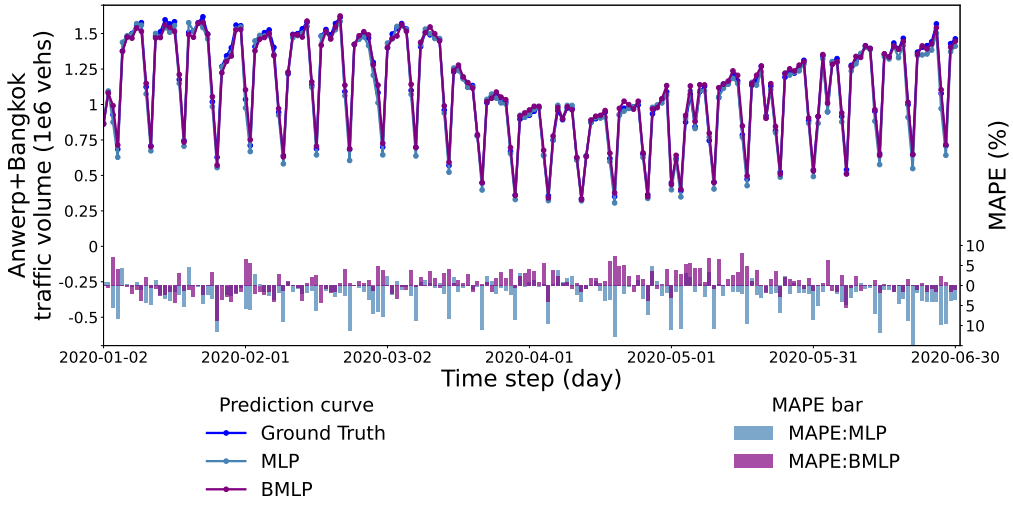
Table 7 presents the performance of proposed methods in the source domain under different metrics for the input data with a grid size of 200m x 200m. Overall, the BMLP model still outperformed the other three models in all three metrics, and the performance gaps among the models became more significant (see Table 5).

Table 7 Source domain loss (Grid size: $200m \times 200m$)

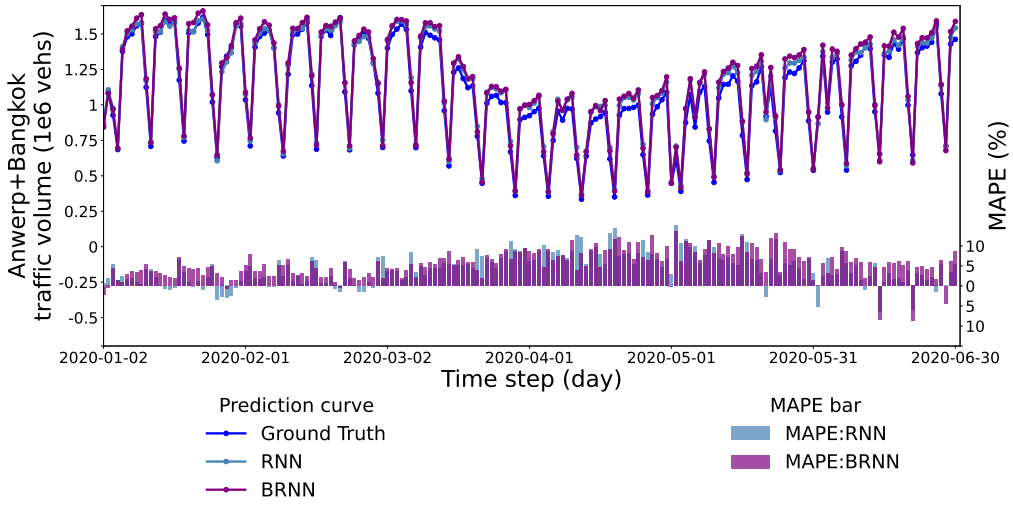
Model & Input features	DTW	MAE	RMSE
BMLP (180+181)	422.888	27.910	72.265
BRNN (180+181)	457.916	33.763	91.761
BLSTM (180+181)	441.835	30.104	80.205
BBiLSTM (180+181)	433.922	29.419	84.595

For DTW, the BRNN performed the worst, with the error approximately 8% larger than BMLP, 5.5% larger than BBiLSTM, and 3.6% larger than BLSTM. However, in our previous experiment, the performance of BRNN was only about 2.5% lower than BMLP and about 2% lower than BLSTM and BBiLSTM. This finding indicates that increasing the grid size and reducing the data volume can improve the differences in the neural networks' performance to learn time series similarity. For MAE, the accuracy of BRNN was 22% lower than BMLP, while the performances of BLSTM and BBiLSTM were similar, which were between BRNN and BMLP. Compared to the previous experiments, the performance gaps of MAE between the networks became more significant, but the accuracy of BBiLSTM exceeded BLSTM. For RMSE, BMLP achieved the best performance. Unlike the previous experiments, BRNN performed worse than BLSTM. Additionly, the performance of BLSTM was better than BBiLSTM, which indicates that BLSTM was less affected by significant errors in data with a larger grid size than other models.

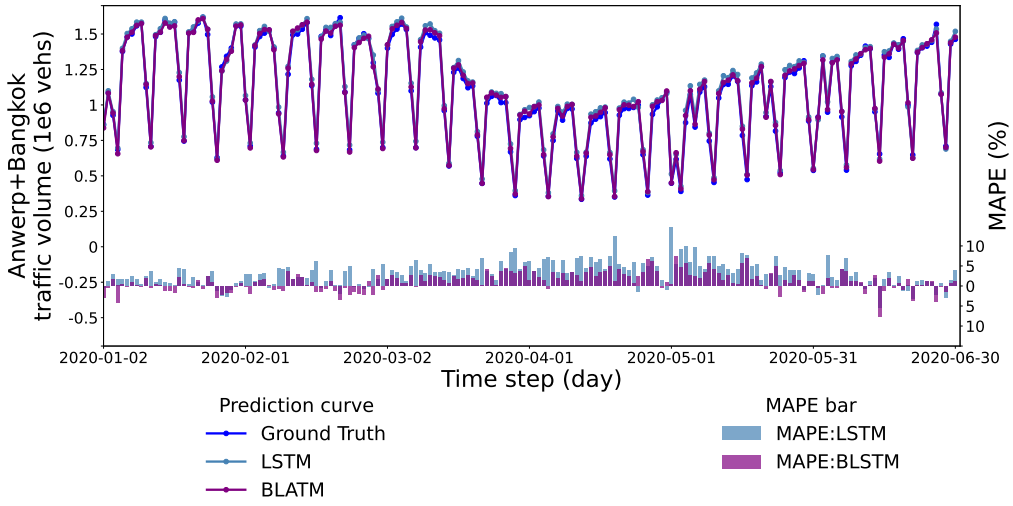
At the macro level, the prediction performance of the neural networks is shown in Figure 32. Compared to the results of the previous experiments with a small grid size (see Figure 27), the performance of BRNN on the resilience patterns prediction became worse, while other models still achieved robust prediction performance. As shown in Figure 32(a), BMLP performed relatively well throughout the entire prediction interval. However, a visible bias occurred when predicting the peak of the first month, as in the previous experiment. Figure 32(c) shows that the ability of BLSTM to predict the valley values of resilience patterns had decreased compared to the previous experiment. Compared to the bar charts of Figure 32(d) and Figure 27(d), it can be clearly observed that the robustness of BBiLSTM in predicting valley values and time steps with significant changes in trend had improved.



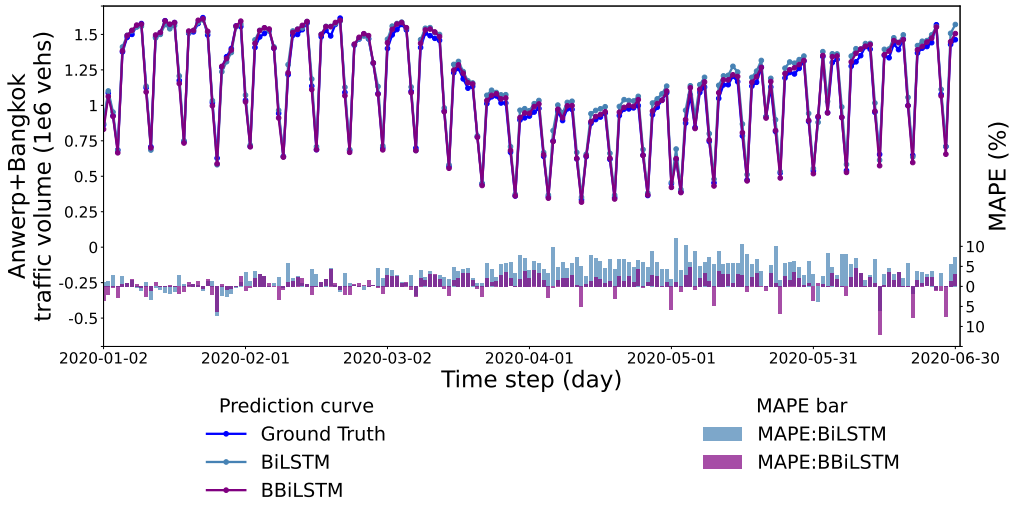
(a) MLP model



(b) RNN model



(c) LSTM model



(d) BiLSTM model

Figure 32 The macro level prediction and true values in source domain (Grid size: $200m \times 200m$)

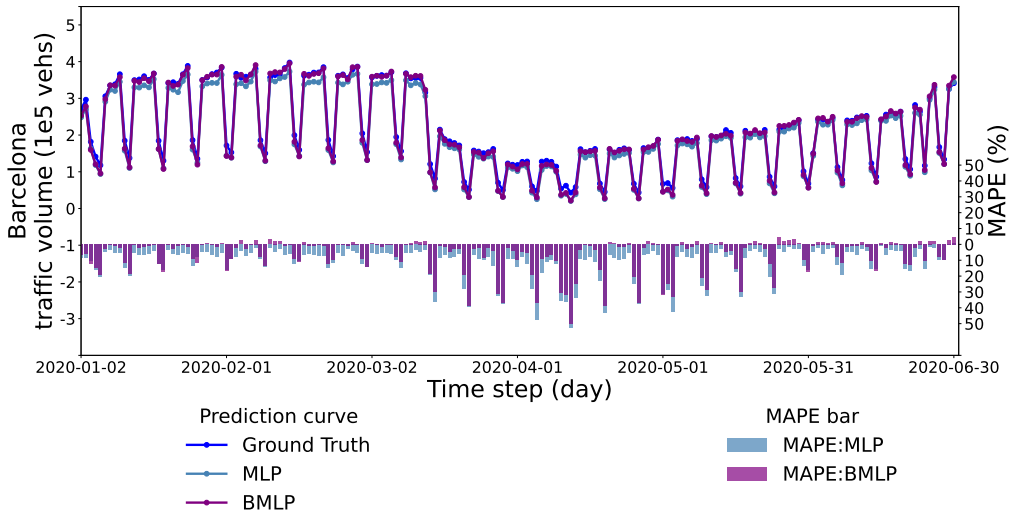
6.3.3. Target Domain Sensitivity Analysis

Table 8 presents the performance of the proposed methods under different metrics for the target domain experiment when the grid size was 200m*200m. Overall, BMLP showed the most robust performance among all models.

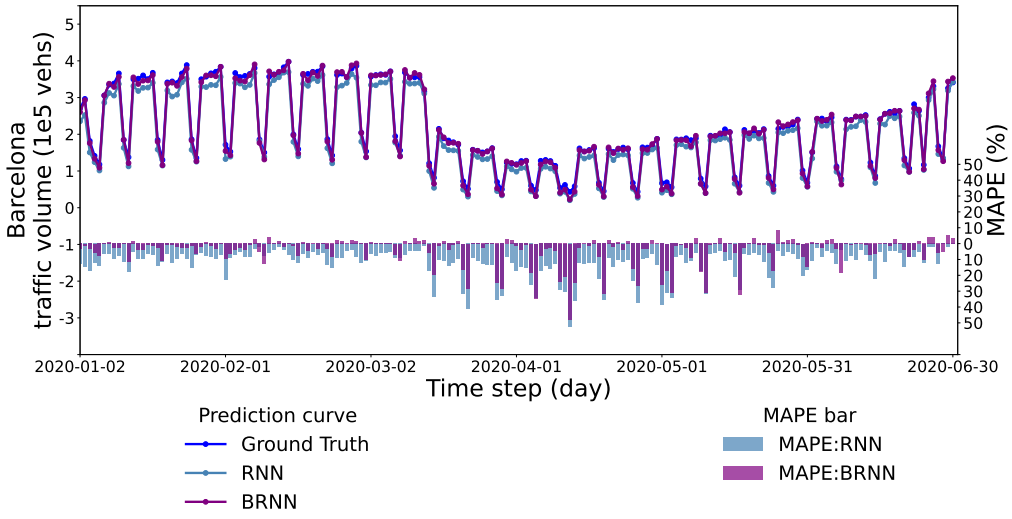
Table 8 Target domain loss (Grid size: $200m \times 200m$)

Model & Input features	DTW	MAE	RMSE
BMLP (50+181)	440.767	29.049	75.771
BRNN (50+181)	453.752	31.33	80.998
BLSTM (50+181)	450.478	30.651	80.478
BBiLSTM (50+181)	445.896	30.384	77.429

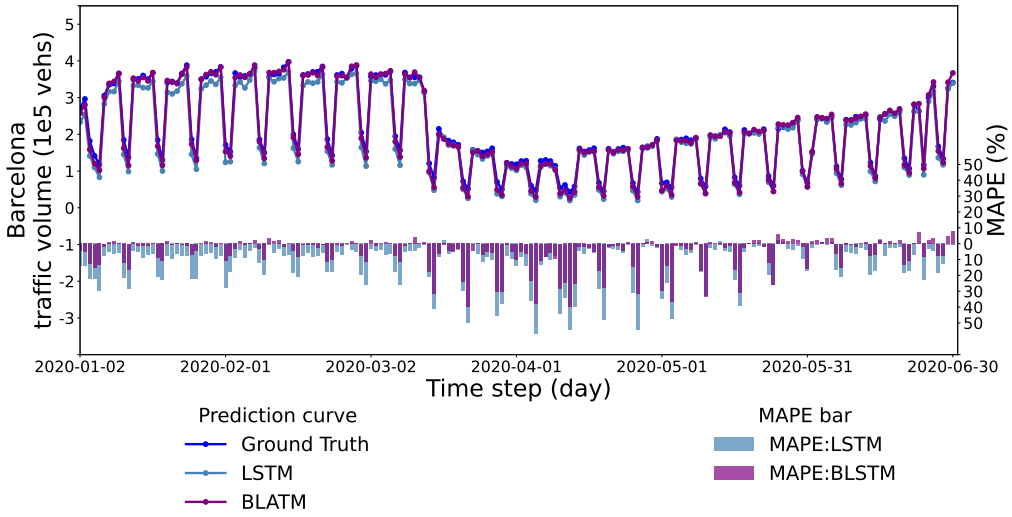
For DTW, although the gap between each neural network and BMLP became more significant compared to the previous experiment (see Table 6), the difference between BRNN and BLSTM, and between BBiLSTM and BLSTM, had decreased. For MAE, compared to the experiment with a small grid size, the performance of BBiLSTM had improved relative to BRNN and BLSTM. For RMSE, BRNN and BLSTM show similar performance, while BBiLSTM outperformed them by approximately 4%. Since a large grid size can contain a large traffic volume and result in a decrease in the input data volume, which can impact the neural network's performance, therefore, in the sensitivity analysis of the source domain, the values DTW, MAE, and RMSE of the models were higher than those in the previous source domain experiment. In the target domain, the model's DTW and MAE were also higher than in the previous experiment. However, it can be noticed that the RMSE of the models is lower than the previous target domain experiment. Although the data volume for training in the target domain of the sensitivity experiment is only 8805, the number of significant errors is lower than the transfer learning of the previous transfer learning experiment. This finding indicates that increasing the grid size improves the stability of grid traffic volume time series.



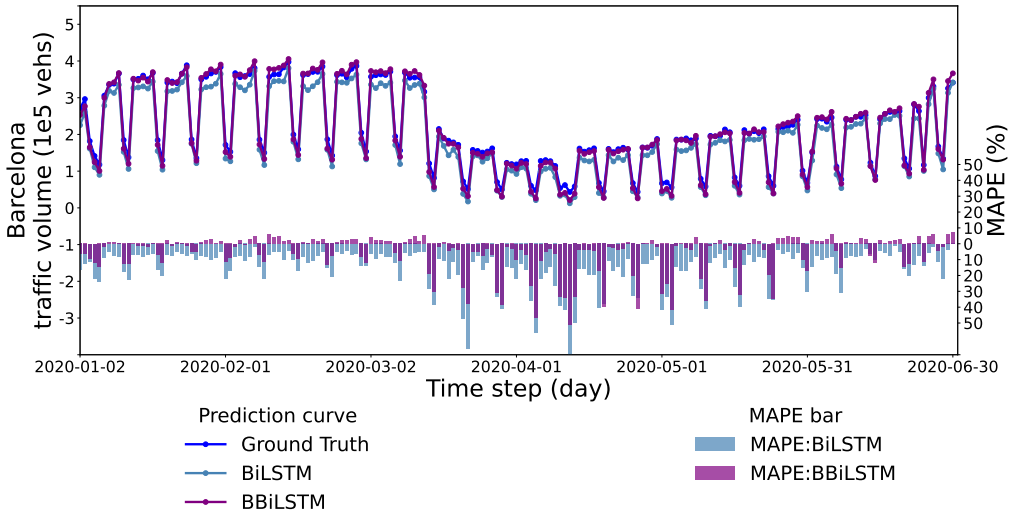
(a) MLP model



(b) RNN model



(c) LSTM model



(d) BiLSTM model

Figure 33 The macro level prediction and true values in target domain (Grid size: $200m \times 200m$)

Figure 33 shows the prediction results of transfer learning of various neural networks in the sensitivity analysis experiment. Compared to Figure 28, although the predictive performance of BMLP was almost similar to the previous experiment, the performance of RNN-based models had improved compared to the previous transfer learning experiment. The RNN-based models could capture more information and temporal dependencies on resilience patterns and became more robust in predicting the months with significant changes in trends. Among RNN-based models, BRNN achieved the best prediction performance at the macro level, which was inconsistent with the previous experimental results. The possible reason is that BRNN had a better generalization ability for grids with unstable trends, while BiLSTM and BLSTM focused more on these grids. Therefore, BRNN could better express stable trends at the macro level.

Table 9 displays the average MAPE loss of the during disaster time series for the experiments with different grid sizes in predicting traffic volume resilience patterns at the macro level. Comparing the different Bayes patterns extractor-based models, the evidence suggests that BMLP achieved the best prediction performance in both the source and target domains for the 100m x 100m grid size experiment, and the BRNN performed the worst. However, in the sensitivity experiment, BLSTM and BBiLSTm performed better in the source domain than BMLP. BLSTM and BRNN outperformed BMLP in the target domain. This evidence indicates that increasing the grid size can enhance the temporal correlation of resilience patterns and facilitate RNN-based models to capture temporal dependencies more accurately. Compared with the micro-level results, it can be inferred that BMLP had a better explanation ability for unstable time series of grid data, while RNN-based models failed to handle the unstable grid information. However, the ability of the RNN-based models to capture the temporal dependence made them a better prediction than BMLP on a macro-level when the proportion of stable time series data increases.

Table 9 MAPE of the models in macro level traffic volume resilience patterns prediction

Model	MAPE			
	$100m \times 100m$		$200m \times 200m$	
	Source domain	Target domain	Source domain	Target domain
MLP	3.73%	11.91%	2.67%	14.39%
BMLP	1.91%	11.33%	2.22%	9.66%
RNN	2.99%	12.85%	4.59%	15.26%
BRNN	2.39%	15.74%	5.43%	7.12%
LSTM	2.89%	15.29%	3.37%	17.05%
BLSTM	1.89%	14.18%	1.93%	8.72%
BiLSTM	1.98%	18.08%	3.67%	22.04%
BBiLSTM	2.18%	15.47%	1.97%	10.61%

On the other hand, increasing the grid size can reduce the input data volume of the models. By comparing the prediction results of the source domain experiments, we found that when the data volume was relatively sufficient, the model with a larger data volume (100 m x 100 m grid size) performed better generally. However, in cases of insufficient data volume, the transfer learning performance of Bayes patterns extractor-based models in the sensitivity experiment far exceeded that of the original experiment. This result suggests that increasing the grid size can help stabilize the traffic volume time series and significantly impacted the performance of our Bayes patterns extractor-based models.

6.4. Ablation Experiment Analysis

In the previous section, we discussed the transfer learning performance of the proposed Bayes patterns extractor-based models. To evaluate the contribution of the Bayes patterns extractor, we further conducted an ablation experiment by removing the Bayes patterns extractor from the proposed models and directly training different types of neural networks (MLP, RNN, LSTM, and BiLSTM) to predict the traffic volume resilience patterns.

6.4.1. $100m \times 100m$ Grid Size

Table 10 lists the prediction results of the proposed methods and the ablation experiment models in the source domain with a $100m \times 100m$ grid size. It can be observed that the resilience patterns extracted by the Bayes patterns extractor could improve the overall per-

formance of RNN-based models while did not improve the performance of MLP. Although in RNN-based models, the improvement of BRNN compared to RNN was insignificant, BLSTM and BBiLSTM outperformed LSTM and BiLSTM in all metrics. The possible reason is that the memory capability of RNN for pre-disaster information is inferior to LSTM and BiLSTM when the prediction length is long.

Table 10 Source domain ablation experiment (Grid size: $100m \times 100m$)

Model & Input features	DTW	MAE	RMSE
MLP (180)	258.456	17.290	39.888
BMLP (180+181)	260.607	17.223	40.0486
RNN (180)	275.614	19.371	44.916
BRNN (180+181)	273.541	19.134	44.745
LSTM (180)	273.553	18.884	44.033
BLSTM (180+181)	262.951	17.498	40.882
BiLSTM (180)	266.648	17.963	42.597
BBiLSTM (180+181)	261.848	17.747	41.504

Table 11 shows the prediction results of the proposed methods and the ablation experiment model in the target domain under different metrics. As evident from the table, our proposed models showed stronger robustness in transfer learning. For DTW, compared to the Bayes component ablated models, BMLP, BRNN, BLSTM, and BBiLSTM models performed better on capturing time series similarity by 4.6%, 7.0%, 6.2%, and 3.3%, respectively. The accuracy of proposed models also improved by 6.0%, 3.8%, 9.2%, and 4.9% in MAE. Moreover, for RMSE, the accuracy of our models increased by 1-3%.

Table 11 Target domain ablation experiment (Grid size: $100m \times 100m$)

Model & Input features	DTW	MAE	RMSE
MLP (50)	319.442	21.679	83.732
BMLP (50+181)	309.585	20.458	82.508
RNN (50)	332.087	22.401	82.213
BRNN (50+181)	317.416	21.580	84.702
LSTM (50)	333.436	22.745	87.104
BLSTM (50+181)	311.488	20.830	85.225
BiLSTM (50)	330.359	22.547	87.788
BBiLSTM (50+181)	311.215	21.501	84.558

From the results of both Table 10 and Table 11, we can conclude that the proposed methods enhanced the performance of RNN-based models in the source domain and improved the transfer learning performance of all models.

6.4.2. $200m \times 200m$ Grid Size

Table 12 lists the prediction results of the proposed methods and the ablation experiment models in the source domain with a 200m * 200m grid size. It is observed that the Bayes patterns extractor improved the prediction performance of MLP, LSTM, and BiLSTM, while it did not improve the overall prediction performance of RNN.

Table 12 Source domain ablation experiment (Grid size: $200m \times 200m$)

Model & Input features	DTW	MAE	RMSE
MLP (180)	425.397	28.499	73.701
BMLP (180+181)	422.888	27.910	72.265
RNN (180)	463.907	33.755	87.998
BRNN (180+181)	457.916	33.763	91.761
LSTM (180)	447.590	31.138	84.824
BLSTM (180+181)	441.835	30.104	80.205
BiLSTM (180)	448.240	31.191	84.537
BBiLSTM (180+181)	433.922	29.419	84.595

Table 13 displays the results of the proposed models and the ablation experiment models in the target domain with a 200m*200m grid size. According to the table, the proposed models exhibited a more robust ability to capture the time series similarity and more precise prediction accuracy. Compared to the Bayes component ablated models, BMLP, BRNN, BLSTM, and BBiLSTM models had better prediction performance by 5.4%, 7.7%, 5.7%, and 7.7%, respectively, in terms of DTW. The accuracy of the proposed models also improved by 5.9%, 7.5%, 6.9%, and 7.4% in MAE and by 7.0%, 8.8%, and 11.1%, respectively, in RMSE.

Table 13 Target domain ablation experiment (Grid size: $200m \times 200m$)

Model & Input features	DTW	MAE	RMSE
MLP (50)	466.852	31.298	83.757
BMLP (50+181)	440.767	29.049	75.771
RNN (50)	487.631	33.054	86.694
BRNN (50+181)	453.752	31.334	80.998
LSTM (50)	483.807	33.015	87.595
BLSTM (50+181)	450.478	30.651	80.478
BiLSTM (50)	476.700	32.103	86.023
BBiLSTM (50+181)	445.896	30.384	77.429

The results of Table 12 and Table 13 illustrate that the proposed Bayes patterns extractor-based models can significantly improve the prediction ability of traffic volume resilience patterns when the data volume is insufficient and pre-disaster information is limited.

From the macro perspective, the proposed models achieved superior performance in predicting overall traffic volume than the Bayes component ablated models in almost all metrics (see Table 9). In the experiment with a 100m x 100m grid size, the BMLP, BRNN, and BLSTM models in the source domain achieved lower MAPE losses than MLP, RNN, and LSTM models by 1.8%, 0.6%, and 1%, respectively. In the target domain, BMLP, BLSTM, and BBiLSTM models showed lower MAPE losses than MLP, LSTM, and BiLSTM models by 0.6%, 1.1%, and 2.6%, respectively. In the experiment with a $200m \times 200m$ grid size, the BMLP, BLSTM, and BBiLSTM had lower losses than MLP, LSTM, and BiLSTM models by 0.5%, 1.4%, and 1.7% in the source domain, respectively. In the target domain, the MAE losses of the BMLP, BRNN, BLSTM, and BBiLSTM models were reduced by 4.73%, 8.14%, 8.33%, and 11.43%, respectively. It can be noticed that MLP, RNN, LSTM, and BiLSTM models showed a significant decline in performance in the target domain when the grid size increased. However, the Bayes pattern extractor-based models demonstrated an increasing prediction performance in such conditions. This result suggests that our Bayes pattern extractor-based models exhibited better robustness for more stable data and enhanced the transferability of the neural networks when predicting traffic volume resilience patterns. Additionally, by comparing Figures 27, 28, 32, and 33, we can infer that Bayes pattern extractor-based models primarily contributed to the model's robustness in predicting peak and valley values. Although the Bayes component ablated models (MLP, RNN, LSTM, BiLSTM) could learn temporal dependencies between different patterns, they all continuously overestimated or underestimated the traffic volume in predicting the overall city resilience patterns. The reason could be the over-capturing of the local unstable traffic trends, which made them fail to express the global traffic volume trends. In long-term prediction tasks, the accumulative error also was a limitation for RNN-based models (RNN, LSTM, and BiLSTM). However, introducing the Bayes pattern extractor can provide global information to the neural networks and makes the input grid time series more stable, which could enhance the models' robustness in predicting the overall traffic volume at the macro level. On the other hand, as the traffic volume floating car data highly depends on the number of vehicles connected, which can increase the instability of data with a small grid size. However, the extracted resilience patterns might prevent the models from learning such unstable trends. Hence, some models' performance did not show significant improvements at the grid level.

7. Conclusion

This thesis aims to build a transferable model for capturing and predicting transportation systems resilience patterns using floating car data. The method proposed in this thesis includes an unsupervised resilience patterns extractor and different kinds of neural networks. By conducting a case study under the context of the COVID-19 Pandemic, we have proven the effectiveness of our proposed model and achieved satisfactory results.

7.1. Main Contributions and Findings

In this thesis, we applied grid traffic volume FCD as the model inputs and sought to capture the transportation resilience patterns from the FCD by a series of machine learning methods. The proposed framework in Section 3.2 combines both unsupervised machine learning methods and supervised deep learning methods. The unsupervised machine learning methods include the time series clustering method and the Bayes inference method. Through these methods, we created prompt features for the grids with homogeneous resilience patterns and formed the average resilience patterns as the additional input of the neural networks. In the experiments, we evaluated the performance of different neural networks and further discussed the pros and cons of them through ablation analysis and sensitivity analysis experiments.

Additionally, the average resilience patterns extracted from the mixed FCD enable the models to learn the experience from other cities. In the existing literature, the inputs to the models for the resilience pattern prediction tasks are mostly only the traffic data at the local level or only the macro level traffic data. In our experiment, the extracted resilience patterns from different cities were proven helpful in enriching the features of the input data and could provide macro information for the deep learning models. Therefore, our proposed methods also achieved robust performance at the macro level.

Moreover, we conducted a sensitivity analysis (see Section 6.3) to explore the importance of data stability. In our experiment, although the values of a larger grid size dataset were larger and the input data volume was decreased, the models achieved less RMSE than the experiment with a small grid size in the target domain.

More importantly, we explored and analyzed the performance and transferability of the models with different neural network components in the transportation resilience prediction task. By conducting an ablation experiment (see Section 6.4), we found that our models achieved better performance in almost all situations of the transfer learning experiments. When increased the grid size, our models exhibited a more robust performance with only thousands of input data at the macro level, while the traditional neural networks experienced a drop in

performance.

7.2. Limitations

Although the proposed method achieved satisfactory performance in our case study, there are still limitations require further research.

1. The extracted resilience patterns allowed the proposed method to learn experiences from different grids and cities, and we successfully applied it in our case study. However, the COVID-19 pandemic lasted for a relatively consistent duration globally. Therefore, the proposed method has not been proven effective between the time series with different disaster duration. In reality, most disasters have different duration, and collecting data with the same disaster duration requires considerable effort and is not always practical.

2. This thesis assumed that cities with similar pre-disaster traffic volume patterns would exhibit similar resilience patterns during the pandemic. Therefore, we applied time series clustering to capture the similarity of the grids. However, the performance of time series clustering depends on how similar are the pre-disaster patterns between different grids and cities. In this thesis, we only collected floating car data from two cities for resilience patterns extraction, which cannot guarantee that each grid from the target city could have the right category to correspond. Therefore, collecting resilience patterns from more cities to enrich pre-disaster pattern categories can potentially enhance the model's performance.

3. This paper only used traffic volume to measure the resilience patterns of the transportation system but did not consider other features of the floating car data. However, traffic volume floating car data is probably unstable at the grid level and highly depends on the number of GPS-equipped vehicles. Therefore, features like average speed should also be considered. Furthermore, this paper only considered the similarity of the data itself but neglected the spatial similarity between grids. Therefore, introducing spatial features may help classify grids more precisely and extract resilience patterns more accurately.

7.3. Outlook

Considering the limitations of the proposed method, future efforts should be paid in several potential research directions.

1. **Learn a deep embedding for the extracted resilience patterns**. In order to address the problem of potential differences in the duration of resilience patterns between different cities it is necessary to learn deep embeddings of the extracted resilience patterns.

A feasible solution is to first perform the Bayes-based feature extraction on the grids from the same city after time series clustering between different cities. Then, using a representation learning method to learn the embeddings of resilience patterns of the same cluster from different cities.

2. **Increase the sources of Floating Car Data.** In our case study, we only extracted the resilience patterns of the FCD data from two cities, but some grids in the target city did not belong to any cluster. Therefore, collecting FCD data from more cities to generate more clusters and extract more types of resilience patterns is necessary.

3. **Enrich the input features for the model.** For pandemics, demand loss is the reason for the performance decline of the transportation system. However, for other disasters, the average speed is also an important feature that can reflect the resilience patterns of urban transportation systems. Moreover, even though some grids have the same pre-disaster patterns, their difference in spatial structures may result in entirely different resilience patterns. Thus, clustering of node features (node degree, node centrality, graphlets) should also be considered as a future research direction.

Bibliography

[1] M. Bruneau, S. E. Chang, R. T. Eguchi, G. C. Lee, T. D. O'Rourke, A. M. Reinhorn, M. Shinozuka, K. Tierney, W. A. Wallace, and D. von Winterfeldt, "A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities," *Earthquake Spectra*, vol. 19, no. 4, pp. 733–752, 2003.

[2] C. Wan, Z. Yang, D. Zhang, X. Yan, and S. Fan, "Resilience in transportation systems: A systematic review and future directions," *Transport Reviews*, vol. 38, no. 4, pp. 479–498, 2018.

[3] M. Mojtahedi, S. Newton, and J. Von Meding, "Predicting the resilience of transport infrastructure to a natural disaster using Cox's proportional hazards regression model," *Natural Hazards*, vol. 85, no. 2, pp. 1119–1133, 2017.

[4] B. Rouhanizadeh and S. Kermanshachi, "Post-disaster reconstruction of transportation infrastructures: Lessons learned," *Sustainable Cities and Society*, vol. 63, p. 102505, 2020.

[5] Y. Zhu, K. Ozbay, K. Xie, and H. Yang, "Using Big Data to Study Resilience of Taxi and Subway Trips for Hurricanes Sandy and Irene," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2599, no. 1, pp. 70–80, 2016.

[6] R. Mirjalili, H. Barati, and A. Yazici, "Resilience Analysis of New York City Transportation Network After Snow Storms," *Transportation Research Record: Journal of the Transportation Research Board*, p. 036119812211010, 2022.

[7] W. Xiao, Y. D. Wei, and Y. Wu, "Neighborhood, built environment and resilience in transportation during the COVID-19 pandemic," *Transportation Research Part D: Transport and Environment*, vol. 110, p. 103428, 2022.

[8] A. Deloukas and E. Apostolopoulou, "Static and dynamic resilience of transport infrastructure and demand: The case of the Athens metro," *Transportation Research Procedia*, vol. 24, pp. 459–466, 2017.

[9] S. Amoaning-Yankson and A. Amekudzi-Kennedy, "Transportation System Resilience: Opportunities to Expand from Principally Technical to Sociotechnical Approaches," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2604, no. 1, pp. 28–36, 2017.

[10] L. Gonçalves and P. Ribeiro, “Resilience of urban transportation systems. Concept, characteristics, and methods.,” *Journal of Transport Geography*, vol. 85, p. 102727, 2020.

[11] D. R. Godschalk, “Urban Hazard Mitigation: Creating Resilient Cities,” *Natural Hazards Review*, vol. 4, no. 3, pp. 136–143, 2003.

[12] P. Murray-tuite, “A Comparison of Transportation Network Resilience under Simulated System Optimum and User Equilibrium Conditions,” in *Proceedings of the 2006 Winter Simulation Conference*, (Monterey, CA, USA), pp. 1398–1405, IEEE, 2006.

[13] X. Zhang and E. Miller-Hooks, “Scheduling Short-Term Recovery Activities to Maximize Transportation Network Resilience,” *Journal of Computing in Civil Engineering*, vol. 29, no. 6, p. 04014087, 2015.

[14] W. Szeto, “Cooperative game approaches to measuring network reliability considering paradoxes,” *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 2, pp. 229–241, 2011.

[15] N. Khademi, B. Balaei, M. Shahri, M. Mirzaei, B. Sarrafi, M. Zahabiun, and A. S. Mohaymany, “Transportation network vulnerability analysis for the case of a catastrophic earthquake,” *International Journal of Disaster Risk Reduction*, vol. 12, pp. 234–254, 2015.

[16] M. Pregnolato, A. Ford, C. Robson, V. Glenis, S. Barr, and R. Dawson, “Assessing urban strategies for reducing the impacts of extreme weather on infrastructure networks,” *Royal Society Open Science*, vol. 3, no. 5, p. 160023, 2016.

[17] D. Henry and J. Emmanuel Ramirez-Marquez, “Generic metrics and quantitative approaches for system resilience as a function of time,” *Reliability Engineering & System Safety*, vol. 99, pp. 114–122, 2012.

[18] Q.-C. Lu, “Modeling network resilience of rail transit under operational incidents,” *Transportation Research Part A: Policy and Practice*, vol. 117, pp. 227–237, 2018.

[19] T.-Y. Liao, T.-Y. Hu, and Y.-N. Ko, “A resilience optimization model for transportation networks under disasters,” *Natural Hazards*, vol. 93, no. 1, pp. 469–489, 2018.

[20] Z. Malki, E.-S. Atlam, A. Ewis, G. Dagnew, A. R. Alzighaibi, G. ELmarhomy, M. A. Elhosseini, A. E. Hassanien, and I. Gad, “ARIMA models for predicting the end of COVID-19 pandemic and the risk of second rebound,” *Neural Computing and Applications*, vol. 33, no. 7, pp. 2929–2948, 2021.

[21] Y. Zhu, Y. Wang, T. Liu, and Q. Sui, “Assessing macroeconomic recovery after a natural hazard based on ARIMA—a case study of the 2008 Wenchuan earthquake in China,” *Natural Hazards*, vol. 91, no. 3, pp. 1025–1038, 2018.

[22] Y. Meng, X. Liu, C. Ding, B. Xu, G. Zhou, and L. Zhu, “Analysis of ecological resilience to evaluate the inherent maintenance capacity of a forest ecosystem using a dense Landsat time series,” *Ecological Informatics*, vol. 57, p. 101064, 2020.

[23] H.-W. Wang, Z.-R. Peng, D. Wang, Y. Meng, T. Wu, W. Sun, and Q.-C. Lu, “Evaluation and prediction of transportation resilience under extreme weather events: A diffusion graph convolutional approach,” *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102619, 2020.

[24] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, “Short-term traffic forecasting: Where we are and where we’re going,” *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.

[25] A. Abadi, T. Rajabioun, and P. A. Ioannou, “Traffic flow prediction for road transportation networks with limited traffic data,” *IEEE transactions on intelligent transportation systems*, vol. 16, no. 2, pp. 653–662, 2014.

[26] S. J. Pan and Q. Yang, “A Survey on Transfer Learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[27] J. Li, F. Guo, A. Sivakumar, Y. Dong, and R. Krishnan, “Transferability improvement in short-term traffic prediction using stacked LSTM network,” *Transportation Research Part C: Emerging Technologies*, vol. 124, p. 102977, 2021.

[28] V. Cerqueira, L. Moreira-Matias, J. Khiari, and H. van Lint, “On Evaluating Floating Car Data Quality for Knowledge Discovery,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 11, pp. 3749–3760, 2018.

[29] Y. Guo, B. Li, Z. Lu, and J. Zhou, “A novel method for road network mining from floating car data,” *Geo-spatial Information Science*, vol. 25, no. 2, pp. 197–211, 2022.

[30] R. Hu, Y. Xia, C.-Y. Hsu, H. Chen, and W. Xu, “Traffic Intersection Detection Using Floating Car Data,” in *2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, (Xiamen, China), pp. 116–120, IEEE, 2020.

[31] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah, “Time-series clustering–a decade review,” *Information systems*, vol. 53, pp. 16–38, 2015.

[32] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, “Fast subsequence matching in time-series databases,” *ACM Sigmod Record*, vol. 23, no. 2, pp. 419–429, 1994.

[33] H. Sakoe and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” *IEEE transactions on acoustics, speech, and signal processing*, vol. 26, no. 1, pp. 43–49, 1978.

[34] G. E. Batista, E. J. Keogh, O. M. Tataw, and V. M. De Souza, “Cid: an efficient complexity-invariant distance for time series,” *Data Mining and Knowledge Discovery*, vol. 28, pp. 634–669, 2014.

[35] J. Paparrizos and L. Gravano, “K-Shape: Efficient and Accurate Clustering of Time Series,” in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, (Melbourne Victoria Australia), pp. 1855–1870, ACM, 2015.

[36] T. W. Liao, “Clustering of time series data—a survey,” *Pattern recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.

[37] J. MacQueen, “Classification and analysis of multivariate observations,” in *5th Berkeley Symp. Math. Statist. Probability*, pp. 281–297, University of California Los Angeles LA USA, 1967.

[38] J. Paparrizos and L. Gravano, “Fast and Accurate Time-Series Clustering,” *ACM Transactions on Database Systems*, vol. 42, no. 2, pp. 1–49, 2017.

[39] M. Acuto, “COVID-19: Lessons for an Urban(izing) World,” *One Earth*, vol. 2, no. 4, pp. 317–319, 2020.

[40] S. Engle, J. Stromme, and A. Zhou, “Staying at Home: Mobility Effects of COVID-19,” *SSRN Electronic Journal*, 2020.

[41] Z. Chu, M. Cheng, and M. Song, “What determines urban resilience against COVID-19: City size or governance capacity?,” *Sustainable Cities and Society*, vol. 75, p. 103304, 2021.

[42] M. M. Bruwer and S. J. Andersen, “Exploiting COVID-19 related traffic changes to evaluate flow dependency of an FCD-defined congestion measure,” *Environment and Planning B: Urban Analytics and City Science*, p. 239980832210815, 2022.

[43] R. Florida, A. Rodríguez-Pose, and M. Storper, “Cities in a post-COVID world,” *Urban Studies*, p. 004209802110180, 2021.

[44] E. Macioszek and A. Kurek, “Extracting Road Traffic Volume in the City before and during covid-19 through Video Remote Sensing,” *Remote Sensing*, vol. 13, no. 12, p. 2329, 2021.

[45] HERE Developer, “Sample map data for students.” `https://developer.here.com/sample-data#download`, 2021.

[46] C. Eichenberger, M. Neun, H. Martin, P. Herruzo, M. Spanring, Y. Lu, S. Choi, V. Konyakhin, N. Lukashina, A. Shpilman, N. Wiedemann, M. Raubal, B. Wang, H. L. Vu, R. Mohajerpoor, C. Cai, I. Kim, L. Hermes, A. Melnik, R. Velioglu, M. Vieth, M. Schilling, A. Bojesomo, H. A. Marzouqi, P. Liatsis, J. Santokhi, D. Hillier, Y. Yang, J. Sarwar, A. Jordan, E. Hewage, D. Jonietz, F. Tang, A. Gruca, M. Kopp, D. Kreil, and S. Hochreiter, “Traffic4cast at neurips 2021 - temporal and spatial few-shot transfer learning in gridded geo-spatial processes,” in *Proceedings of the NeurIPS 2021 Competitions and Demonstrations Track* (D. Kiela, M. Ciccone, and B. Caputo, eds.), vol. 176 of *Proceedings of Machine Learning Research*, pp. 97–112, PMLR, 06–14 Dec 2022.

Declaration

I hereby confirm that the presented thesis work has been done independently and using only the sources and resources as are listed. This thesis has not previously been submitted elsewhere for purposes of assessment.

Munich, 09.05.2023. Ningkang Yang

München, 09.05.2023, Signature