sciendo

Transport and Telecommunication, 2024, volume 25, no. 4, 409-426
Transport and Telecommunication Institute, Lauvas 2, Riga, LV-1019, Latvia
DOI 10.2478/ttj-2024-0030

INFERRING TRAFFIC PATTERNS OF DHAKA CITY: A SPATIO-TEMPORAL ANALYSIS OVER A YEAR

***Md. Moshiur Rahman*[1], *Naushin Nower*[2]**

[1]*Department of Software Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University*
Gazipur 1750, Bangladesh
[2]*Institute of Information Technology, University of Dhaka*
Dhaka, 1200, Dhaka, Bangladesh
naushin@iit.du.ac.bd

Numerous large cities' sustainability and livability are severely hampered by traffic congestion. To create efficient measures to relieve this congestion, it is vital to comprehend the spatiotemporal patterns of this congestion. Our study addresses the scarcity of long-term spatiotemporal analysis for traffic congestion in developing megacities, focusing on Dhaka, the fifth most congested city. Utilizing big data analytics with a sample of 350,400 records from Google Maps over a year, we identify temporal patterns, spatial distribution, and recurrent congestion patterns. Through extensive image processing, peak hours, congestion variations, inter-zone relationships, and causes of extreme congestion are analyzed. Dhaka is divided into ten zones, revealing distinct congestion patterns with variations between weekends and weekdays. These findings offer crucial insights for urban planning, traffic control, and infrastructure development in rapidly expanding megacities, contributing to the alleviation of congestion and enhancement of sustainability and livability.

Keywords: Congestion analysis, Dhaka city, traffic patterns, spatial-temporal analysis

1. Introduction

In transportation engineering, traffic analysis is the study of how drivers and their vehicles interact with infrastructure, such as highways, signage, and traffic control devices, with the goal of understanding and creating the best possible transportation system (Mannering and Washburn, 2019). It is crucial for dwellers and city planners since traffic analysis is the foundation for effective urban mobility and transportation planning. This analysis research also provides valuable insights that help communities optimize infrastructure, alleviate traffic congestion, and improve overall road safety by methodically evaluating the patterns and dynamics of vehicular movement (Neilson *et al.*, 2019). Recently, traffic volume has been increasing every year in almost every mega city which further causes traffic congestion due to over-supplied traffic on limited roadways (Sheikh & Peng, 2022). As a result, the significance of traffic pattern analysis research is gaining attraction (Kulshrestha, 2022). In addition, the pattern of traffic volume varies spatially and temporally and has recurrent and non-recurrent nature (Zhang *et al*., 2016). Thus, a data-driven pattern analysis strategy helps to make strategic decisions, such as identifying bottlenecks, implementing public transportation hubs, deploying intelligent traffic management systems, and many more. Furthermore, this study shows the peak traffic hours and bottlenecks, which aids in the development of proactive measures such as congestion prediction to reduce gridlock and travel time (Nama *et al.*, 2021). Besides these, understanding traffic patterns gives benefits to all levels of private and government agencies to promote a smoother, more sustainable urban environment, and ensure a quality of life.

To investigate geographic patterns of traffic is the first step in developing efficient policies to address the traffic-related problem, which has been extensively studied in major western cities. For example, Turner (1992) calculated the levels of congestion in fifty large and medium-sized American cities and examined the relationships between various indicators and congestion levels. In the paper (Zhao and Hu, 2019), the authors have analysed long-term geographical and temporal patterns of traffic congestion in Beijing City using a big data analytic technique for a six-month period. Their analysis revealed four types of congestion patterns in Beijing each with unique spatial and temporal features. Another research (Bachechi *et al.*, 2022) focuses on the two European cities (Modena, in Italy, and Santiago de Compostela, in Spain) identifies traffic trends, abnormal events, and seasonal events, and analyses the effects of traffic volume on the environment of the mentioned cities. Extensive traffic analysis study has already been done on the many mega cities in the developed countries, nevertheless, developing cities still require greater focus. For example, Dhaka city, the capital of Bangladesh is the fifth most congested city in the world, and the sixth most densely populated city in the world according to Numbeo's Traffic Index by City 2023 (Tolaini, 2020). The city's ever-increasing population and economic activities have resulted in a huge increase in automobile traffic, resulting in a chronic problem of traffic congestion. This problem has not only hampered the free flow of people and commodities, but it has also had an impact on the city's overall quality of life.

Various statistics and studies demonstrate the severity of traffic congestion in Dhaka. According to Bangladesh Institute of Planners research, the number of registered automobiles in Dhaka has been steadily increasing over the last decade (Hossain and Nower, 2022*)*. The city had roughly 1.5 million registered vehicles as of 2021, including individual automobiles, motorcycles, buses, and lorries (TBS Report, 2022). In addition, Dhaka is the sixth world's most densely populated city, with an estimated 17 million people living in an area of 1,528 square kilometres (Dhaka Tribune, 2018). This great population density places enormous strain on the city's infrastructure, notably its road networks. During peak hours, average speeds in some areas have decreased to as low as 7-10 kilometres per hour (km/h) according to the Asian Development Bank (ADB) (Rahman and Nower, 2024). Usually, traffic data is collected from fixed detectors such as sensors, loop detectors, speed cams, etc on the road. However, there is no such reliable architecture exists in Bangladesh to collect traffic data as a result, research on traffic pattern analysis on Dhaka city is not so extensive. For example, Momin *et al.* (2023) collect traffic data only for three hours from their recorded video and apply Kalman filtering on the collected data. They count the number of vehicles from their recorded footage.

Rahman *et al.* (2018) provides a comprehensive analysis of traffic intensity patterns in Dhaka using GPS data collected over 15 days on 11,769 road segments of Dhaka city. They collect GPS data from a private company named Gobd.co. With this short period of data, they examine the effects on traffic intensity of marketplaces, the quantity of road intersections, and the impacts of rickshaw on roadways. They also provide zone clustering based on similarity. However, the short data collection period, absence of real-time traffic attributes, lack of intra-zone traffic pattern analysis, and consequent lack of analysis on long-term patterns, seasonal, and periodic pattern analysis, are the limitations of this study.

Following that, Jaman and Amin (2023) collected speed data from a single road segment between Technical and Shyamoli, focusing exclusively on cars. Their research is confined to time series prediction and does not analyze influencing factors or spatial relationships. Due to data limitations and insufficient infrastructure, their prediction accuracy is around 70%. The study overlooks spatial relationships, temporal trends, and other critical traffic issues.

Because of lack of traffic data, Amin Al Noor and Mehanaz (2022) explored travel time dynamics in Dhaka, using a questionnaire survey that gathered responses from 721 diverse road users. The study identified peak travel times in the morning, afternoon, and evening, with moderate peaks in the late evening. However, it did not consider spatial and temporal correlations, real traffic attributes, or the suitability of index values in its analysis. Details comparison among the existing pattern on Dhaka city is presented on Table 1.

Table 1. Comparative analysis of existing traffic patterns based work on Dhaka city

Methods	**Duration of data collection**	**Method used for data collection**	**Area Coverage**	**Identifying peak hour**	**Spatial analysis**		**Temporal Analysis**				**Zone clustering**
					Intra zone	**Inter zone**	**Hour**	**Week**	**Month**	**Year**	
Fahim *et al***.** (Jaman, 2023)	15 days	GPS data provided by private company	13 zones of Dhaka city	Yes	No	No	Yes	No	No	No	Yes
(Rahman *et al*., 2018)	3 hours	Manual	2.5 Km	No	No	No	No	No	No	No	No
(Amin *et al*., 2022)	-	Survey	721 people	Yes	No	No	No	No	No	No	No
Our Proposed	365 days	Web scraping	95.475 km^2 Entire Dhaka city	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes

Most of the developed countries have the facilities to gather traffic data every second from many sources, including sensors, speed cameras, motion detectors, in-vehicle GPS, etc., and can benefit from big data analysis (Arooj *et al*., 2022). Big data has a tremendous impact on traffic congestion analyses, transforming urban transportation management. Big data provides sophisticated knowledge of traffic patterns, congestion triggers, and mobility behaviour by collecting extensive traffic data at several time scales hourly, daily, weekly, and yearly. However, infrastructure-based data collection facilities in developed countries are not available in developing countries, for example, Bangladesh. As far as we know, in Bangladesh, there is very limited infrastructure to collect real-time spatial-temporal traffic data though it suffers severe traffic congestion, and most of the traffic data is collected from manual counting which is not a feasible approach to cover an area/city (Momin *et al*., 2023). However, regular traffic data is of paramount importance as it serves as a vital tool for urban planning and transportation management.

To solve the traffic data collection problem for an infrastructure-less environment, we have developed a data collection tool that processes Google maps traffic images and extracts traffic information without using its paid API. A preliminary version of the tool is proposed in Hossain and Nower (2022) and Rahman & Nower (2023). The automated and detailed version of the of tool is presented in this paper which can produce traffic data sets for any time interval that can be used directly for data analysis, modelling, prediction, etc. The tool has undergone additional enhancements, enabling it to generate traffic data for specific road segments or areas. The aim of this paper is to provide the details of the proposed traffic data collection tool named 'Traffic Data Drive'. The designed tool can automatically extract the traffic data for any given duration using extensive image processing. Using the developed tool, we can explore and visualize temporal and spatial traffic patterns using big data analytics to reduce traffic congestion in expanding megacities in the developing world. As a case study, we have selected Dhaka city, Bangladesh's capital. Dhaka, with a population of over 20 million people and a vehicle count of almost 6 million as of the end of 2017, symbolizes the prototypical quickly expanding congested megacity in the developing world.

For traffic congestion pattern analysis throughout Dhaka city, we divide the entire city into ten zones. For each zone, we have collected data from 1st January 2022 to 31st December 2022 in 15-minute intervals. Thus, the total data instances for a zone are 35040, and for ten zones it is 350400 instances. We have applied a variety of analytical techniques to investigate the spatial-temporal patterns of traffic congestion in this data. To identify trends, seasonal events, and unusual events, we analyse traffic time series on an hourly, daily, weekly, and monthly basis. Additionally, we have used dynamic time wrapping (DTW) and spatial autocorrelation to assess the spatial dependency of congestion in ten different zones.

Finally, we applied hierarchical clustering to group areas more effectively with similar traffic patterns. Throughout the analysis of Dhaka city over the year 2022, we have some following interesting findings.

- Approximately 31% of the study area experiences severe traffic congestion persistently over the year, resulting in a maximum speed of around 15 kilometres per hour. Additionally, around 28% of the study area grapples with moderate congestion, allowing for a maximum speed of approximately 25 kilometres per hour. Finally, roughly 39% of the study area enjoys minimal congestion levels year-round in Dhaka city, enabling a maximum speed of 40 kilometres per hour.
- The most congested area in Dhaka city is Old Dhaka which contains Lalbagh, Chak Bazar, Hazaribagh, Kamrangir Char, and Kotwali because of the business hub and narrow road areas. In addition, the least congested zone is Uttra which is far away from the centre of Dhaka city.
- The traffic congestion patterns in various zones of Dhaka city exhibit notable differences, accentuated by variations between weekends and weekdays.
- Exit points of Dhaka city, such as Jatrabari, Saydabad, and Gabtoli, experience their highest congestion levels on weekends and the day before weekends such as Thursday through Saturday compared to weekdays.

2. Data and methodology

In this section, we discuss the details process of our data collection tool 'Traffic Data Drive' and the details of our study zones and provide a solid foundation for our research methodology.

2.1. Data Collection Tool: Traffic Data Drive

For traffic congestion analysis we need historical data that contains spatiotemporal information that refers to information that combines both location and temporal aspects related to traffic flow (Amato *et al*., 2020; Yuan and Li, 2021). It includes data on the location, speed, direction, and volume of vehicles, as well as the time and duration of their movements. Spatiotemporal traffic data can be collected using various technologies, such as GPS, sensors, cameras, loop detectors, webcams, and other expensive technology.

This information can be used to analyse traffic patterns, identify congestion hotspots, and develop strategies for managing and improving traffic flow. To facilitate traffic congestion analysis research in developing countries, we developed a data collection tool that processes Google Maps traffic images and extracts traffic information without using its paid API. The developed tool is used to produce traffic data set for any time interval that can be used directly for data analysis, modelling, prediction, etc.

Traffic data collection tools are valuable in developing countries where no infrastructure is available to collect traffic data. Some advantages of the traffic data collection tool are given below:

- Provides a cost-effective solution for collecting and analysing traffic data for research and planning purposes.
- Access to comprehensive traffic data allows for more in-depth analyses and informed decision-making regarding transportation planning and management.
- Traffic data sets with different time intervals help identify traffic patterns, peak hours, congestion hotpots, and accident-prone areas.
- The insights from the data can inform the design and implementation of transportation policies and interventions.
- Traffic management strategies such as adjusting signal timings, implementing tolls or congestion charges, and constructing new roads can be guided by the analysis of traffic data.

To collect traffic data, we have designed a tool that extracts traffic information from Google map images without using paid API. The developed tool continuously captures Google map images with traffic information and provides information about traffic congestion.

The created tool continuously takes screenshots of Google Maps that represent traffic data and then we can calculate traffic intensity from these images.

The initial concept of the proposed tool was first described in the paper (Hossain & Nower, 2022) and applied on a single route. However, traffic data for the overall city is required for a city or area-wise traffic prediction. We have extended and automated the previous tool thus it can extract area-wise data at a time. The overall process of data collection is shown in Figure 1. The following steps are used to develop the data collection tool:

- Capture image using selenium web driver.
- Cropping image
- Image Masking and Colour Detection
- Semantic cross-sectional segmentation
- CSV generation

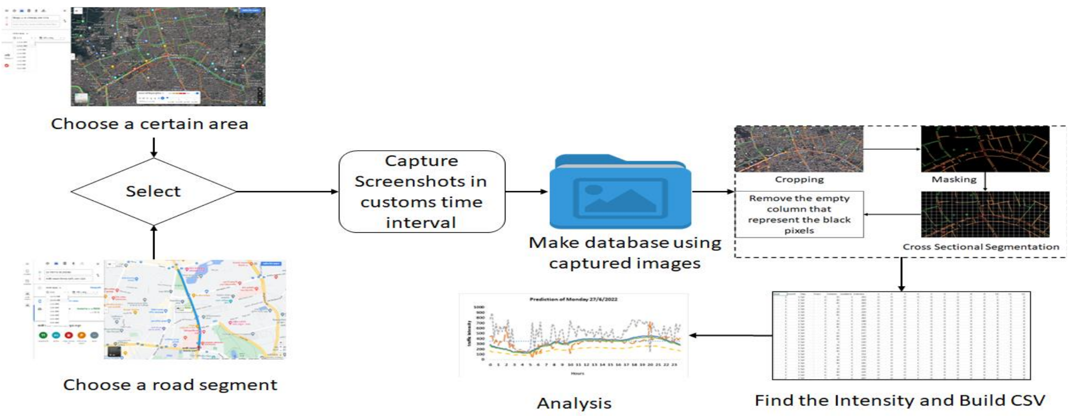

Figure 1. The process of obtaining the traffic congestion intensity data from Google Maps

2.1.1. Capture image using selenium

Our proposed traffic data collection tool utilizes the Selenium web driver to take screenshots from Google Maps. To use the tool, the user is prompted to input the URL of the desired Google Map view, as well as the date and time interval for which traffic data is desired. Sample tool then launches a Firefox browser instance using the web driver and navigates to the specified URL. The input of our proposed tool is shown in Table 2. Once the Google Map view has loaded, the tool uses the Selenium web driver to interact with the page and capture a screenshot of the traffic data. The captured screenshot is then processed using image-processing techniques.

Table 2. Input of our proposed traffic data collection tool

URL	Start Date	End Date	Time Interval (minutes)
https://www.google.com/maps/dir/Mirpur+10+Roundabout,+Dhaka+1216@23.8069243,90.359943,3191m/data=!3m2!1e3!4b1! 4m13!4m12!1m5!1m1!1s0x3755c0d6f6b8c2ff:0x3b138861ee9c8c30!2m2!1d90.3686978!2d23.8069245!1m0!2m3!6e0!7e2!8j1641035100!3e0	01/01/2022	31/12/2022	15

We used the Selenium web driver to take screenshots at fifteen-minute intervals. Figure 2 shows a sample screenshot of the Mirpur 10 area that is captured by our proposed tool. Figure 3 represents a screenshot of a road segment from Banglamotor to Shahbag.

Figure 2. Capturing screenshot using selenium web driver at Mirpur 10 zone

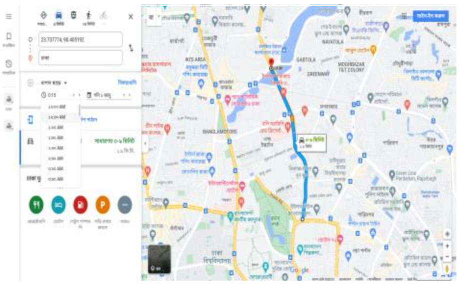

Figure 3. Capturing screenshot using selenium web driver of a road segment from Kakrail circle to Hatirjheel Lake road

2.1.2. Cropping image

After capturing the screenshot of the Google Map view using the Selenium web driver and Firefox browser, the next step in our proposed traffic data collection tool is to remove any unnecessary parts of the image. This includes removing the source and destination text box area, direction button, and weekly legend at Google Maps, which can be considered noise.

To remove this noise, we crop the image using image processing techniques. This involves selecting the maximum area of the image that contains the relevant traffic data while excluding the unnecessary parts. Once the noise has been removed, we resize the image to a standard size of 1250x690 pixels using image resizing techniques. By removing the noise and resizing the image, we can ensure that the captured traffic data is accurate and can be processed efficiently. This can lead to more accurate predictions and analyses of traffic flow patterns. Figure 4 shows a sample picture after cropping in area selection and Figure 5 shows a sample picture after cropping in a single road segment.

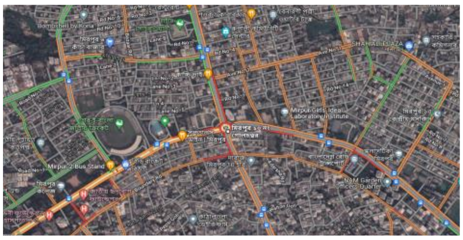

Figure 4. Removing the unnecessary part from the image using cropping

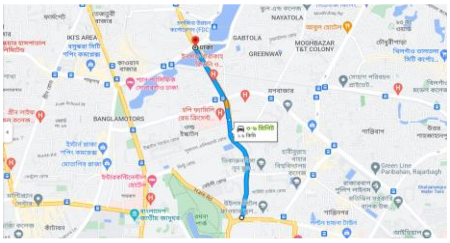

Figure 5. Removing the unnecessary part from the image using cropping

2.1.3. Image Masking and Color Detection

In our proposed traffic data collection tool, we use different map types (Default and Satellite) for collecting traffic data based on the type of analysis required. For collecting traffic data for a single road segment, we use the default map type provided by Google Maps, while for area-wise data collection, we use the satellite map type. Google Maps uses different colors to indicate the level of traffic congestion in the default and satellite map types. In the default map type, the blue color is used to indicate no traffic congestion, the orange color indicates mild traffic congestion, and the red color indicates extreme traffic congestion. On the other hand, in the satellite map type, the green color indicates light traffic congestion, the orange color indicates mild traffic congestion, and the red color indicates extreme traffic congestion. To extract the traffic data accurately from the captured images, we use different masking HSV (Hue, Saturation, Value) values for the default and satellite map types. These masking values help to identify the relevant color regions that indicate the level of traffic congestion and extract the required information from the images. Using color masking road network traffic congestion can be identified. In the proposed tool, the HSV (Hue Saturation Value) color model is used. To identify the single road segment, we use [0, 150, 150] to [10, 255, 255] for red color, [10, 100, 100] to [15, 255, 255] for orange color, and [60, 110, 100] to [70, 255, 255] for green color. On the other hand, identify the area-wise road traffic intensity [0, 150, 150] to [10, 255, 255], [10, 100, 100] to [15, 255, 255], and [60, 110, 100] to [70, 255, 255] color range is used to represent red, orange, and green respectively. Figure 6 and 7 shows the image after masking, it perfectly identifies the road and shows the traffic intensity using color.

Figure 6. Identify the roads and vehicle intensity using color masking

Figure 7. Identify the roads and vehicle intensity using color masking

2.1.4. Semantic cross-sectional segmentation

The final stage of preparing the data set is the cross-sectional segmentation of the masked images. We mainly divide the height of the image by 23 divide the width by 25 and find a total of 575 cross-sectional rectangle boxes in every image. Figure 8 represents the cross-segmental concept properly. From the cross-sectional image, a CSV file is created that has a total of 575 columns for each image. Each column represents boxes with value 3 if boxes have more red value than others, 2 if boxes have more orange value than others, 1 if boxes have more green value than others, and 0 if boxes have all black pixel's value. Algorithm 1 represents the overall process of making CSV from the mask image. The traffic data collection tool extracts three features date, time, and intensity, as part of its feature extraction process. The date feature includes the day, month, and year on which the traffic data was collected. The time feature includes the hour and minute at which the data was collected. Users can set the date and time interval for which they want to collect data. This allows for flexibility in data collection and enables users to monitor traffic patterns and trends over specific periods. The intensity feature is calculated by summing up the traffic volume across all cross-sectional units. This feature provides an estimate of the traffic flow on the monitored route and can be used to analyse traffic congestion levels and develop strategies for improving traffic flow.

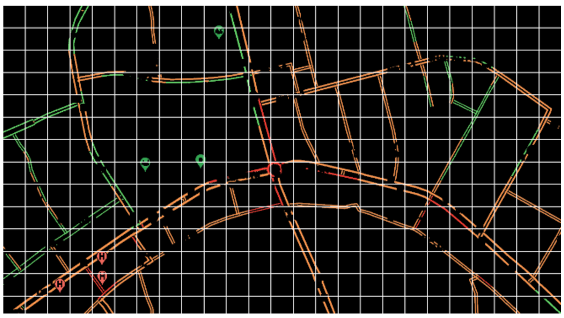

Figure 8. Semantic cross-sectional segmentation process

2.2. Data: Details of Dhaka City Context

Our primary objective is to analyse the traffic congestion patterns across the entire city of Dhaka. Initially, we adopted the thirteen Detailed Planning Zone (DPZ) divisions proposed by RAJUK (Rahman *et al*., 2018). However, we encountered issues when using our data collection tool for traffic data collection.

Our tool captures data within rectangular areas, whereas the RAJUK-designated zones are not consistently rectangular in shape. This resulted in overlapping data, which may provide redundant information for overlapping zones. To overcome this inconvenience, we decided to redefine the zones to ensure they do not overlap. We determined that a total of ten non-overlapping zones would sufficiently cover the entire city of Dhaka. Table 3 provides an overview of these ten zones and their respective coverage areas. Dividing the city into ten zones allows for a relatively even distribution of areas, which can help in capturing a representative sample of the city's congestion patterns. The city map with ten zones is shown in Figure 9.

Table 3. Study zone details - area and coordinates

Zone Name	Total area(km2)	Start	End
zone 1 (Dhanmondi, New Market, Shahbagh)	9.475	23.750076, 90.375271	23.729648, 90.410933
zone 2 (Uttara, Biman Bandar)	10.78	23.872592, 90.379920	23.849459, 90.414745
zone 3 (Darus Salam, Adabor)	8.86	23.796541, 90.333045	23.774942, 90.369008
zone 4 (Gulshan, Cantonment)	9.19	23.819657, 90.395429	23.799869, 90.433666
zone 5 (Jaabari, Sutrapur, Gendaria)	7.99	23.721895, 90.407536	23.701581, 90.444400
zone 6 (Mohammadpur, Sher-e-Bangla Nagar)	9.26	23.776920, 90.347229	23.756221, 90.384952
zone 7 (Mirpur, Kafrul, Pallabi)	9.08	23.818966, 90.353248	23.797724, 90.389168
zone 8 (Motijheel, Paltan, Ramna)	8.63	23.752629, 90.408530	23.731102, 90.444407
zone 9 (Lalbagh, Chak Bazar, Hazaribagh, Kamrangir Char, Kotwali)	12.69	23.737092, 90.369978	23.711297, 90.421185
zone 10 (Tejgaon)	9.52	23.773063, 90.380415	23.752089, 90.418395

Figure 9. Total study area and the ten zones selected for traffic flow analysis

2.3. Traffic Congestion index

The severity of traffic flow is represented by the traffic congestion index. The Congestion Index, represented in equation 1, offers a comprehensive framework to discern congestion types within various zones and at different times (Zhao & Hu, 2019). This index divides the traffic time series data into three main components: trends, seasonal fluctuations, and random fluctuations. A trend is defined as a continuous and sustained rising or downward trajectory in data over a given period. Seasonal fluctuations appear as repeated patterns of trends across time. Random fluctuations, on the other hand, signify sporadic changes in data caused by unanticipated and non-repetitive sources. Higher index values indicate more severe congestion.

$$Congestion\ Index = T \times R \times S. \quad (1)$$

By analysing trends T, which capture long-term changes, seasonal fluctuations S, which account for recurring patterns, and random fluctuations R, representing irregular variations, this index effectively characterizes congestion dynamics. Figures 10, 11 and 12 represent the seasonal fluctuations, random fluctuations, and trends of zone 1 for the year 2022, respectively. This approach enables an accurate understanding of congestion origins, aiding urban planners and policymakers in tailoring targeted solutions for diverse congestion scenarios across different zones.

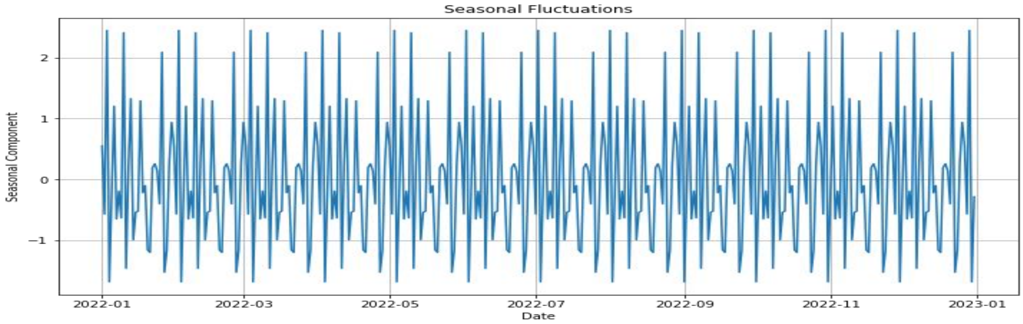

Figure 10. Seasonal (S) Fluctuations of the Zone 1 over the year 2022

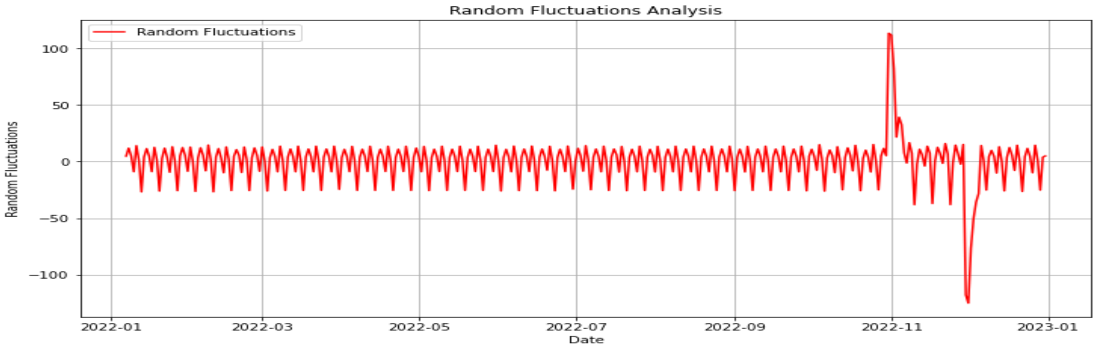

Figure 11. Random (R) Fluctuations of the Zone 1 over the year 2022

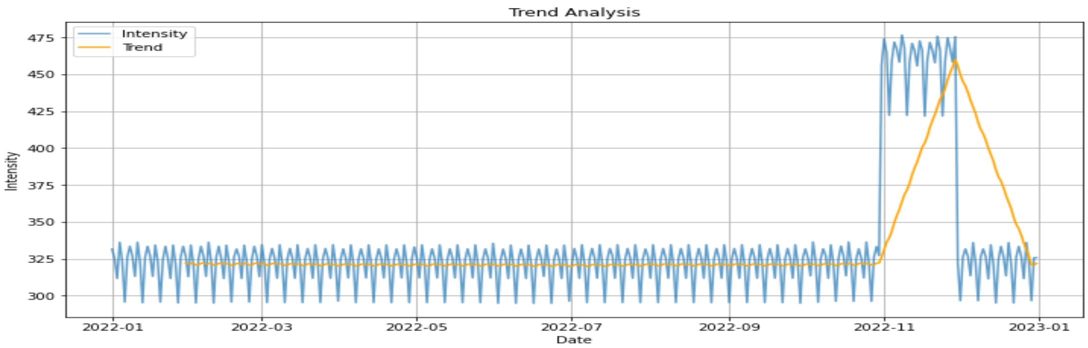

Figure 12. Trend (T) of the Zone 1 over the year 2022

Table 4. Traffic congestion index(Zhao & Hu, 2019b)

Traffic Congestion Index	Traffic Situation
0 – 1.9	No Congestion
2.0 – 3.9	Less Congestion
4.0 – 5.9	Congestion
6.0 – 7.9	Medium Congestion
8.0 – 10	Serious Congestion

Based on the congestion index we can understand the severity of traffic situation. Table 4 displays the range of congestion indices along with corresponding traffic congestion level (Zhao & Hu, 2019b). In Table 5 we can observe the congestion indices for various zones.

Zone 2 exhibits the lowest congestion index value, measuring at 0.4209, signifying the least congestion within this zone. Conversely, Zone 9 encounters the highest congestion index, indicating it to be the most congested area in the data set.

Table 5. Zone-wise traffic congestion index value and level

Zone Name	Congestion Index	Level
zone 1 (Dhanmondi, New Market, Shahbagh)	4.979	Congestion
zone 2 (Uttara, Biman Bandar)	0.4209	No congestion
zone 3 (Darus Salam, Adabor)	2.008	Less congestion
zone 4 (Gulshan, Cantonment)	2.962	Less congestion
zone 5 (Jatrabari, Sutrapur, Gendaria)	7.689	Medium congestion
zone 6 (Mohammadpur, Sher-e-Bangla Nagar)	4.935	Congestion
zone 7 (Mirpur, Kafrul, Pallabi)	3.030	Less congestion
zone 8 (Motijheel, Paltan, Ramna)	4.743	Congestion
zone 9 (Lalbagh, Chak Bazar, Hazaribagh, Kamrangir Char, Kotwali)	9.83	Serious Congestion
zone 10 (Tejgaon)	8.532	Serious Congestion

3. Temporal traffic patterns

Temporal congestion patterns refer to the variations in traffic congestion levels over different time periods within an hour day, week, month, or year (Stathopoulos and Karlaftis, 2001). Analysing these patterns is crucial for understanding how traffic flows and congestion evolve over time, which aids in effective traffic management and urban planning. Temporal congestion patterns reveal peak hours when traffic is heaviest, off-peak periods with lighter traffic, the busiest day of the week, traffic fluctuations, potential recurring congestion events, etc. These insights allow transportation authorities to implement targeted measures such as adjusting traffic signal timings, promoting alternate travel modes during peak times, or optimizing public transportation schedules.

3.1. Hourly traffic pattern

Usually, traffic flow varies from hour to hour, day to day, month to month, and obviously from year to year. Because of these variations, traffic flow needs to be analysed very crucially from different time perspectives. From hour-to-hour traffic analysis peak hours such as morning and evening peaks can be determined. For analysing daily traffic patterns, we consider three types of days such as weekdays, weekends, and national holidays. These three types have significant differences in congestion intensity and pattern.

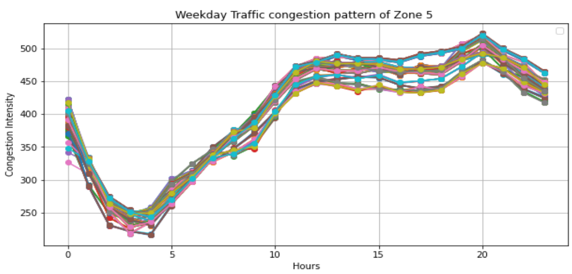

Figure 13. The hourly traffic pattern of Weekdays in Zone 5 over the year 2022

Figure 14. The hourly traffic pattern of weekends in Zone 5 over the year 2022

For instance, Figure 13 depicts the daily traffic patterns of 260 weekdays (Sunday to Thursday) over the year 2022 in zone 5, with a focus on identifying morning and evening traffic peaks. The first peak of the day appears around 12:00 AM, and the evening peak sustains from 5:00 PM to 9:00 PM. The most pronounced traffic congestion peak within this zone occurs at 8:00 PM.

Figure 14 represents the hourly pattern of weekends (Friday and Saturday) for zone 5 in 2022. All the weekends do not have the same congestion pattern in zone 5. Weekends have lower traffic congestion than weekdays. The highest traffic congestion peak of zone 5 on weekends is at 8:00 PM and the morning peak sustains from around 10:30 AM to 1:30 PM.

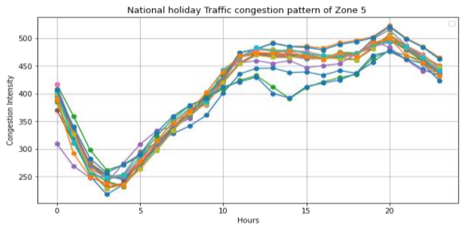

Figure 15. The hourly traffic pattern of national holidays in Zone 5 over the year 2022

Traffic congestion trends on national holidays bear some resemblance to those observed on weekends. However, during extended holidays such as Eid-ul-Fitr and Eid-ul-Azha, traffic congestion reaches its minimum levels. Figure 15 portrays the daily traffic congestion pattern on national holidays in 2022.

3.2. Daily traffic pattern

Daily traffic pattern represents the variations of traffic over the seven days in a week. Figure 16 displays the daily traffic patterns in zone 1 of Dhaka City for the year 2022 where the mean is almost the same with other days except Friday. Across weekdays, the median traffic flow remains relatively consistent, except Tuesday when the local market remains closed, resulting in a notable deviation. Both Friday and Tuesday exhibit lower median values compared to the other days. The Interquartile Range (IQR) shows a similar distribution on weekdays, decreasing on Tuesday and Friday, while experiencing a slight increase on Wednesday. Notably, during both weekends and weekdays, the maximum traffic values tend to cluster around the median. These observations underscore the importance of studying each day of the week individually. Upon conducting a one-way ANOVA, we identified a significant difference in traffic patterns between Friday, Tuesday, and the remaining days.

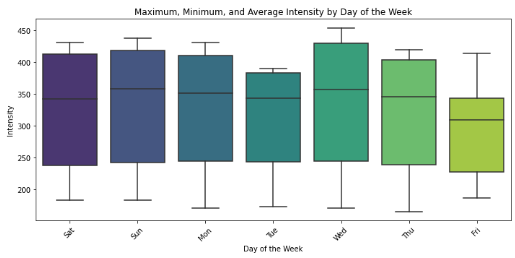

Figure 16. Daily traffic patterns of the week in the Zone 1 of the Dhaka city of year 2022

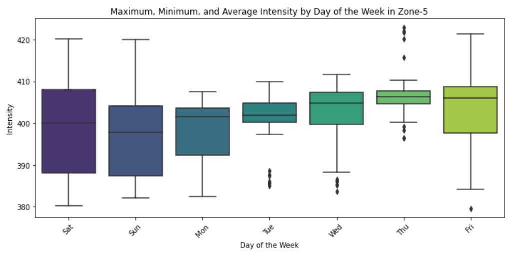

Figure 17. Daily traffic pattern of the week in the Zone 5 of the Dhaka city of the year 2022

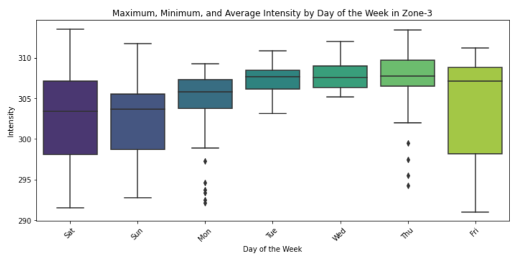

Figure 18. Daily traffic pattern of the week in the Zone 3 of the Dhaka city of the year 2022

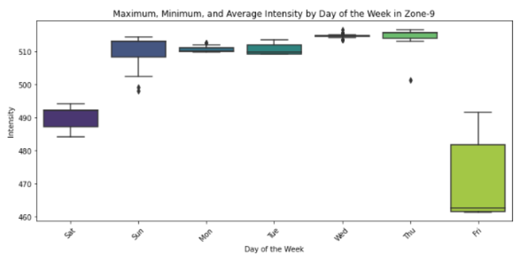

Figure 19. Daily traffic patterns of the week in the Zone 9 of the Dhaka city of the year 2022

Figure 17 delineates the daily traffic congestion patterns in Zone 5 throughout the year 2022, specifying each day of the week. Remarkably, Thursday stands out as the most consistent and busiest day for traffic congestion in this zone since its IQR is comparatively small with occasionally experiencing abrupt increases. The IQR for Thursday is only three, reflecting the minimal variation with consistent congestion. This zone exhibits the highest congestion levels from Wednesday to Friday, while Saturday through Tuesday sees traffic values closely huddled around the median. Notably, Tuesday and Thursday consistently maintain traffic patterns. In addition, it's worth noting that this zone does not exhibit a significant difference between weekdays and weekends, as indicated by the ANOVA test. However, during the weekend, Friday experiences congestion while Saturday tends to be less congested.

Figure 18 presents the daily traffic congestion patterns in Zone 3 throughout the entirety of 2022, showcasing the patterns for each day of the week. Like Zone 5, this zone demonstrates a traffic congestion pattern that is alike, albeit with a lower level of congestion. It's noteworthy that both zones serve as crucial exit points from Dhaka city.

Different daily patterns are observed in Zone 9 as shown in Figure 19. Zone 9 stands out as the most congested zone in Dhaka city, experiencing consistent congestion throughout the entire week. The minimal IQR values across all days without prominent outliers signify a remarkable consistency in congestion levels. Specifically, Wednesday exhibits the lowest IQR, while Friday records the highest. Friday and Saturday, although congested, witness slightly lower congestion levels compared to the other days. Notably, the ANOVA test underscores a significant difference between weekdays and weekends in this zone. As a result, different zones exhibit varied traffic patterns, and due to space restrictions, the daily patterns of other zones are not presented here.

3.3. Weekly traffic pattern

Weekly traffic patterns represent variations of traffic on the road that take place over the month of June 2022. We present the weekly traffic patterns of three zones: one experiencing severe congestion, one with moderate congestion, and one exhibiting lower congestion levels.

Figure 20 shows the weekly traffic pattern of seriously congested zone 9. We plot the four weeks of the entire month of June 2022. In our analysis, we observed consistent traffic in this area throughout the week, except for Fridays and Saturdays. On Fridays, this zone experiences the least traffic compared with other days. It appears that the residents of this zone tend to stay at home during the evening hours on both Friday and Saturday. Traffic pattern remains consistent on almost all weekdays and Thursday is the busiest day on the weekdays.

In Figure 21 we illustrate the weekly traffic pattern for zone 2, which experiences the lowest traffic intensity levels. This zone also exhibits similar week patterns to zone 9 except the traffic volume is lower in this zone compared to zone 9. From the analysis, we can see that traffic volume is lowest on Friday and highest on Thursday every week. Furthermore, it's evident that on Fridays, night-time congestion reaches its peak, surpassing congestion levels on other days.

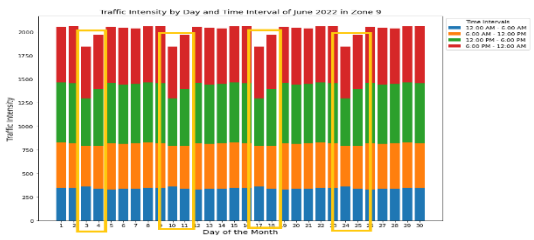

Figure 20. Weekly traffic patterns of the month of June 2022 in Zone 9

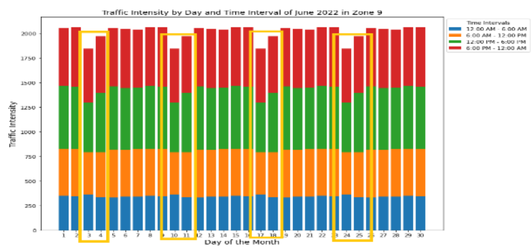

Figure 21. Weekly traffic patterns of the month of June 2022 in Zone 2

3.4. Monthly traffic pattern

Monthly traffic patterns refer to the systematic analysis of traffic flow and congestion levels over a calendar month. Unlike daily and weekly traffic patterns, monthly data can reveal long-term trends and seasonal fluctuations over the year. In this research, traffic data for ten different zones of the year 2022 have been analysed. From the analysis, it is observed that Zones 3, 4, and 7 exhibited consistent traffic congestion patterns throughout the year, with no significant fluctuations tied to specific events. These zones experienced congestion consistently throughout the year, with Zone 7 emerging as the most congested, while Zone 3 and Zone 4 demonstrated similar levels of congestion.

Zones 2, 6, and 8 share similar traffic congestion patterns characterized by frequent sudden incidents. These areas are highly susceptible to the influence of special events, causing traffic congestion to either surge or diminish rapidly in response to the prevailing circumstances.

Zone 5 and Zone 10 have great seasonal fluctuations. These zones' traffic congestion pattern was changed monthly. Zone 5 has different traffic congestion level from January to May. Zone 10 exhibits a consistent traffic pattern in January and February, followed by a shift to a different pattern in March and April. From May to August, and again in September, October, November, and December, the traffic pattern undergoes additional variations.

Due to space constraints, the monthly traffic patterns for ten zones are provided in the appendix A.

4. Spatial traffic patterns

The variations of traffic patterns throughout a geographic area are called spatial patterns. Analysing these patterns can help better urban planning, infrastructure development, and traffic management techniques by providing significant insights into where traffic congestion occurs and where it propagates. To achieve this, spatial traffic patterns of the ten zones of Dhaka city are analysed using Moran's I index, DTW, and annual average daily traffic (AADT) to find the correlated zone, zones with similar patterns, etc.

4.1. Spatial Auto-correlation Analysis

Spatial autocorrelation determines the correlations among different geographical areas. It can demonstrate whether Dhaka's congestion regions were dispersed agglomerative or randomly and sparsely. Spatial auto-correlation analysis is applied to traffic congestion data to understand and analyse patterns in congestion across different geographic areas (Jindal *et al.*, 2013). Moran's I is a widely used statistic for calculating spatial autocorrelation.

Global Moran's I index

The Global Moran's *I* index is a statistical measure used to quantify spatial autocorrelation in a data set. It assesses whether similar values of a variable tend to cluster together or whether dissimilar values are dispersed across geographic space. The formula for calculating the Global Moran's *I* index involves several components. Let n be the number of locations, S_0 be the sum of spatial weights, x_i, and x_j represent values at locations i and j respectively, and $\bar{x}$ be the mean of the values. The equation is as follows:

$$I = \frac{n}{S_0} \frac{\sum i \sum j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum i (x_i - \bar{x})^2} . \tag{2}$$

In this equation, w_{ij} denotes the spatial weight between locations i and j, reflecting the spatial relationship between them. The numerator of the equation involves computing the product of weighted deviations between values at different locations. The denominator consists of the sum of squared deviations of each location's value from the mean. The resulting index value *I* provides insights into the degree of spatial autocorrelation.

Table 6. Zone wise Moran's I index and traffic congestion level

Area	Global Moran's I index	Level
zone 1 (Dhanmondi, New Market, Shahbagh)	-1	Low
zone 2 (Uttara, Biman Bandar)	0.002666 (near to zero)	No correlation
zone 3 (Darus Salam, Adabor)	-0.17432	Low
zone 4 (Gulshan, Cantonment)	-0.17003	Low
zone 5 (Jatrabari, Sutrapur, Gendaria)	0.437471	Moderate
zone 6 (Mohammadpur, Sher-e-bangla Nagar)	-0.02245	Low
zone 7 (Mirpur, Kafrul, Pallabi)	0.401989	Moderate
zone 8 (Motijheel, Paltan, Ramna)	0.416381	Moderate
zone 9 (Lalbagh, Chak Bazar, Hazaribagh,	1	High
zone 10 (Tejgaon)	0.500295	Moderate

In traffic flow analysis, the Global Moran's I index is used to assess the spatial distribution and auto-correlation of traffic patterns across different geographic locations. The index can take on values within a range of -1 to +1, with different interpretations for negative, positive, and zero values:

- If Moran's I index is positive, it suggests that traffic congestion values tend to be positively correlated in space. In other words, areas with high congestion tend to be located near other areas with high congestion, and areas with low congestion tend to be located near other areas with low congestion. From Table 6 we can see that zone 9 is highly correlated with its adjacent zones since its I value is 1. In addition, zones 5, 7, 8, and 9 are moderately correlated with their neighbouring zones.

- If Moran's I index is negative, it indicates that traffic congestion values tend to be negatively correlated in space. In this case, areas with high congestion are surrounded by areas with low congestion, and vice versa. This phenomenon is observed in zone 1, 3, 4, and 6. Among these zones, zone 1's I value is -1 indicating that this area is congested when its neighbouring zones are free.
- If Moran's I index is close to zero, it implies that there is little or no spatial autocorrelation with its surrounding area. In other words, congestion values are randomly distributed in space, and there is no discernible spatial pattern. This situation is observed in zone 2 where the I value is almost close to 0.

Dynamic Time Warping (DTW)

DTW is a sophisticated algorithm for determining the similarity of two-time series sequences that may differ in length or display temporal aberrations (Zhang *et al*., 2022). It's especially useful for comparing sequences with differences in speed or phase. Finding zones with similar patterns is critical for targeted traffic management and urban planning, allowing for the same solutions to relieve congestion for the same patterns, optimize resources, and improve overall mobility. Moran's, I index cannot capture this because spatial correlation is calculated among adjacent zones. Non-adjacent zones also have similar patterns which can be captured by DTW. DTW can play an important role in identifying similar temporal traffic patterns in Dhaka city. DTW works by bending the time axis to align two-time series sequences, providing for the best possible match between matching points (Kafritsas, 2021). This warping allows DTW to detect similarities that might otherwise be missed when using typical Euclidean distance metrics. Traffic congestion patterns in Dhaka may vary in intensity and timing across different zones. By quantifying the temporal pattern's similarity, DTW can assist in identifying comparable traffic patterns.

$$DTW(X_{t+1}, Y_{t+1}) = min(\sqrt{(x_i - y_i)^2} + DTW(X_t, Y_t)). \tag{3}$$

In the equation (3), $X= \{x_1, x_2, \ldots, x_n\}$ is the first time series sequence of length n, $Y=\{y_1, y_2, \ldots, y_m\}$ is the second time series sequence of length m.

DTW (X, Y) is the *DTW* distance between the two temporal sequences. The cost function $(x_i y_j)^2 (x_i y_j)^2$represents the local distance between corresponding points x_i and y_j. The recursive nature of the equation aligns the sequences by considering the optimal previous alignments. In Table 7, we have shown the DTW value for all zones in Dhaka city. From this table, Zone 3 and Zone 4 have the lowest value of DTW, which means those zones have a very similar temporal traffic pattern. The most dissimilar zones are Zone 2 and zone 9. Zone 2 has the minimum traffic congestion among the other zones and zone 9 is the most congested zone in Dhaka city.

Table 7. Dynamic Time Warping (DTW) scores between zones

DTW	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7	Zone 8	Zone 9	Zone 10
Zone 1	0									
Zone 2	57220.36	0								
Zone 3	10084.89	45247.58	0							
Zone 4	9364.52	47365.17	**861.87**	0						
Zone 5	23854.3	82033.95	34913.61	34668.77	0					
Zone 6	26762.92	78967.11	35105.56	35768.3	15005.22	0				
Zone 7	10321.86	55786.95	12361.51	11301.63	22945.22	20071.62	0			
Zone 8	22851.37	80830.36	34627.08	34260.3	8639.56	12458	13136.84	0		
Zone 9	57516.08	**114136.5**	67016.13	66771.3	33710.34	20716.08	45433.19	21527.93	0	
Zone 10	19786.17	82533.86	37906.27	36931.03	13017.36	14097	24156.72	12880.18	25031.9	0

4.2. Annual Average Daily Traffic (AADT)

Annual Average Daily Traffic is defined as the average 24-hour traffic volume at a certain zone throughout a full 365 days/year. Using AADT, we can compare the amount of traffic experienced in various zones in various years and determine how it has evolved over time. In this research, we have calculated AADT for ten zones of Dhaka city over the year 2022 using the following formula.

$$AADT_{zone} = \frac{\sum_{i=0}^{365} \sum_{j=0}^{24} TrafficIntensity_{x,j,i}}{365}, \tag{4}$$

where *Traffic Intensity*$_{zone,i,j}$ is the observed traffic intensity at zone x for the j_{th} hour of the i_{th} day of the year. Table 8 represents the zone wise annual average daily traffic. From the AADT, the most congested zone is zone 9 and the least congested zone is zone 2. Overall AADT of the entire Dhaka city was 353.885 in 2022.

Table 8. Zone wise annual average daily traffic

Zone	AADT
Zone 1	332.5379
Zone 2	175.7698
Zone 3	304.8666
Zone 4	305.5374
Zone 5	400.5204
Zone 6	394.9726
Zone 7	328.6108
Zone 8	397.2229
Zone 9	**488.4725**
Zone 10	410.3429
Overall	353.8854

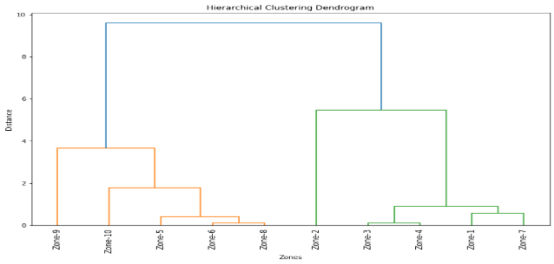

Figure 22. Generated dendrogram from hierarchical clustering analysis of various zones in Dhaka City

5. Cluster analysis

Clustering, an unsupervised learning technique refers to a grouping applied to a collection of zones to subgroup them in a way where the zones within a cluster are closely related compared to zones in different clusters (Popat and Emmanuel, 2014). Although clustering studies have been employed for a very long time in other fields, their application in the field of transportation engineering has been very limited. However, in recent years increasing availability of traffic data and the importance of clustering in analysis, modelling, and simulation have been identified. There has been an increasing interest in cluster analysis for the identification of similar traffic patterns to provide support for many traffic management systems (Sekar *et al.*, 2018).

In this research, we have used hierarchical clustering where the clusters at each level of the hierarchy are created by merging clusters at the next lower level. At the lowest level, each cluster contains a single observation while at the highest level, there is only one cluster containing all observations. Figure 22 illustrates the dendrogram resulting from hierarchical clustering, which was used to group various zones according to traffic intensity. From the dendrogram of hierarchical clustering, we identified three distinct clusters. Table 9 provides a comprehensive overview of the spatial cluster details, along with their corresponding congestion levels. **Cluster 1** groups the highly congested zones as shown in Figure 23. This cluster contains Zone 5, zone 6, zone 8, zone 9 and zone 10. The zones in this cluster exhibit multiple distinct peaks or periods of intense congestion throughout the day. Among all, zone 9 is the most congested zone which has the highest traffic intensity 600. These zones had the lowest congestion on Friday since it was the weekend.

Table 9. Spatial cluster details and their congestion level

Zones	Congestion Level	Cluster number
Zone 5, Zone 6, Zone 8, Zone 9, Zone 10	High congestion	1
Zone 1, Zone 3, Zone 4, Zone 7	Medium congestion	2
Zone 2	No congestion	3

Cluster 2 consists of zones that exhibit medium congestion. Figure 24 shows the traffic congestion patterns observed in medium congested Cluster 2, comprising Zone 1, Zone 3, Zone 4, and Zone 7. Zones in this cluster exhibit two daily congestion peaks, typically occurring in the morning and evening. It's noteworthy that congestion on Fridays lacks a morning peak, setting it apart from the other days.

Cluster 3 contains only one member, which is Zone 2, and it stands out as having minimal traffic congestion. It is notably distinct from the other clusters, featuring varying daily congestion patterns without fixed peak numbers. Additionally, its congestion pattern during the weekend differs from that observed on other weekdays. Figure 25 represents the weekly congestion pattern of Cluster 3.

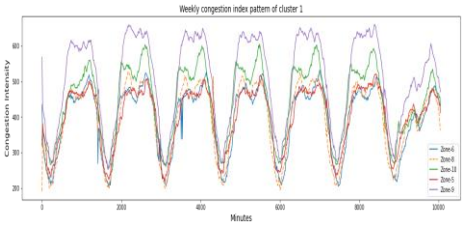

Figure 23. Weekly congestion index pattern of cluster 1

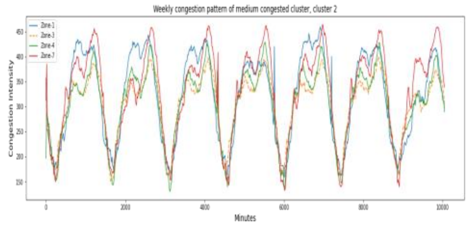

Figure 24. Weekly congestion index pattern of cluster 2

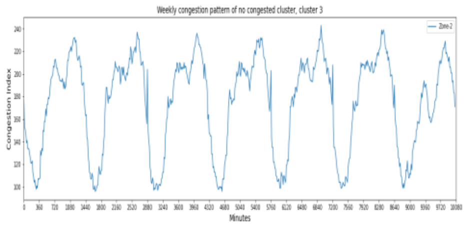

Figure 25. Weekly congestion index pattern of cluster 3

5.1. Identifying Peak Hours

Our study meticulously identifies the peak hours when traffic congestion is most pronounced. From the above analysis, we have shown that each zone has its own unique traffic congestion pattern. Although the traffic pattern of each zone is different there is a similarity based on peak hours. Within **Cluster 1**, all zones experience high levels of congestion, characterized by multiple prominent peaks. While the morning peak may not be highly pronounced, congestion begins to escalate between 7:00 AM and 9:00 AM. Congestion persists for an extended duration, spanning from 12:00 PM to 9:00 PM, with more than two distinct peaks. Notably, traffic congestion patterns on weekdays and weekends exhibit significant disparities. On Fridays, congestion peaks from 8:00 PM to 11:00 PM, while on Saturdays, the peak occurs from 6:00 PM to 9:00 PM, except for Zone 9, which maintains a consistent pattern on Saturdays as on other days, except for Friday.

All zones within **Cluster 2** exhibit two prominent traffic peaks. The first peak commences at 12:00 PM and 3:00 PM, while the second peak spans from 6:00 PM to 9:00 PM. The evening peak experiences higher traffic congestion compared to the morning peak, with zone 7 displaying the highest congestion during these peak hours, and zone 3 the lowest. On weekends, the most congested period typically falls between 8:00 PM and 10:30 PM, albeit with relatively lower congestion levels overall. Notably, zone 7 shows no significant difference between weekends and weekdays concerning peak hour congestion.

Zone 2 is the only member of **Cluster 3** and is characterized by its absence of a fixed traffic congestion pattern. This zone initially witnessed minor traffic from 7:00 AM to 9:00 AM in the morning. Notably, Fridays experience the least congestion compared to other days. Typically, this zone exhibits morning peak congestion from 10:00 AM to 3:00 PM, with evening peak congestion setting in from 6:00 PM to 9:00 PM.

6. Conclusions

Our research on traffic congestion pattern analysis in Dhaka City has revealed significant insights into the city's complex traffic dynamics. By thoroughly examining data over an extended period, we have identified key trends and patterns that can inform urban planning and transportation management strategies. Our findings highlight that congestion in Dhaka City is a pervasive issue, with high levels of traffic congestion persisting throughout the week, except for Fridays when it experiences some relief. This consistent congestion reflects the city's rapid urbanization and population growth, coupled with limited transportation infrastructure.

Notably, our analysis pinpointed specific zones that are particularly congested, shedding light on the areas that demand immediate attention for congestion alleviation measures. We observed that zones with historical significance, such as Lalbagh, Chak Bazar, Hazaribagh, Kamrangir Char, and Kotwali, have remained heavily congested due to their commercial and cultural importance. Moreover, Zone 5 emerged as a crucial exit point for Dhaka City, experiencing higher congestion on Thursdays, Fridays, and Saturdays, driven by the movement of people from various parts of Bangladesh coming to Dhaka.

The differentiation between weekday and weekend congestion patterns is noteworthy. While weekdays witness peak congestion during traditional rush hours, weekends exhibit a different pattern, with traffic intensity peaking later in the afternoon, possibly due to recreational activities and leisure travel. We also observed that Zone 2 consistently has lower congestion levels compared to other zones, making it a potential model for effective traffic management strategies.

Reducing traffic congestion in Dhaka City demands a nuanced, zone-specific approach to address the diverse factors contributing to gridlock. Zone 9 characterized by commercial and cultural significance, such as Lalbagh, Chak Bazar, Hazaribagh, Kamrangir Char, and Kotwali, is the most congested zone in Dhaka city. Factors such as non-motorized vehicles, road digging, illegal footpath encroachment, unplanned car parking, inadequate road infrastructure, and inefficient traffic management have led to significant congestion in Zone 9. As a result, the deployment of more traffic police, the improvement of road infrastructure, and the improvement of public transport system could be necessary measures to reduce traffic congestion in the area. In addition, according to our study this zone contaminates congestion to its surrounding zones since its Moran’s I index is 1. As a result, traffic issues in this zone should be seriously considered. After that zone 10 is the second most congested zone with no dedicated parking space in the large area, and offices having little parking space, and random parking make severe traffic congestion in this area. In addition, huge roads in this region are under constructed with full of construction materials scattered here and there, thus interrupting vehicle movement. Introducing a smart parking system and monitoring the development works and making the developers accountable could be a solution for this area to reduce traffic conditions.

Zone 5 is classified as a medium congested area according to our analysis. This is because it contains one of the exit points of the city (bus terminal) and it is mostly congested on the day before weekends. Optimizing traffic management and upgrading road infrastructure are imperative to streamline traffic flow in this area. Conversely, zones with lower congestion levels, such as Zone 2, offer valuable insights for replication of successful traffic management strategies, including mixed land-use development and incentives for alternative modes of transportation. Residential and industrial zones necessitate tailored interventions like traffic calming measures, flexible work schedules, and last-mile connectivity solutions to mitigate congestion effectively. Furthermore, to reduce traffic, it is imperative to develop intelligent traffic management systems, data-driven real-time analytics tools, and artificial intelligence algorithms for the purpose of predicting traffic and optimizing it.

This study marks the first attempt at traffic analysis in Dhaka city on a large scale. We comprehensively examined the entire city using a year's worth of data and facilitated data collection by dividing Dhaka into ten zones. While we successfully analysed traffic congestion on a zone basis, we were unable to conduct a road-level analysis to identify the most congested roads. Our future work aims to delve into this finer level of detail and pinpoint the most congested roadways.

Acknowledgements

We express our gratitude to the University of Dhaka for providing resources and support for this research.

References

1. Amato, F., Guignard, F., Robert, S., & Kanevski, M. (2020) A novel framework for spatio-temporal prediction of environmental data using deep learning. *Scientific Reports 2020,* 10(1), 1–11. https://doi.org/10.1038/s41598-020-79148-7.
2. Amin Al Noor, A. & Mehanaz, S. (2022) Analysis of travel time in urban area of a developing country, context: Dhaka, Bangladesh. In: *Proceedings of 5th Annual Paper Meet and 2nd Civil Engineering Congress*, Dhaka, Bangladesh.
3. Arooj, A., Farooq, M. S., Akram, A., Iqbal, R., Sharma, A., & Dhiman, G. (2022) Big Data Processing and Analysis in Internet of Vehicles: Architecture, Taxonomy, and Open Research Challenges.

Archives of Computational Methods in Engineering, 29(2), 793–829. https://doi.org/10.1007/S11831-021-09590-X/TABLES/13.
4. Bachechi, C., Po, L., & Rollo, F. (2022) Big Data Analytics and Visualization in Traffic Monitoring. *Big Data Research*, 27, 100292. https://doi.org/10.1016/J.BDR.2021.100292.
5. Dhaka Tribune. (2018) *Dhaka remains the world's most densely populated city*. (n.d.). Accessed February 16, 2024, https://www.dhakatribune.com/bangladesh/dhaka/158221/dhaka-remains-the-world%E2%80%99s-most-densely-populated.
6. Hossain, I., & Nower, N. (2022) Traffic data collection and visualization tool for knowledge discovery using google maps. *International Journal of Software Innovation (IJSI)*, 10(1), 1-12. https://doi.org/10.4018/IJSI.293270.
7. Jaman, F., Amin, M. B. (2023) Traffic speed study on technical to Shyamoli road of Dhaka City, Bangladesh. *International Journal of Research in Civil Engineering and Technology*, 4(1), 34-43. https://www.researchgate.net/publication/372914718.
8. Jindal, T., Giridhar, P., Tang, L. A., Li, J., & Han, J. (2013) Spatiotemporal periodical pattern mining in traffic data. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2505821.2505837.
9. Kafritsas. N. (2021) Time series classification using Dynamic Time Warping. *Towards Data Science*. Accessed March 19, 2024, https://towardsdatascience.com/time-series-classification-using-dynamic-time-warping-61dcd9e143f6.
10. Kulshrestha, S. K. (2022) Dynamics of Megacity Regional Development. In: *Advances in 21st Century Human Settlements*, 65–105. https://doi.org/10.1007/978-981-16-5469-5_3/COVER.
11. Mannering, F. L., Washburn, S. S. (2019) *Principles of Highway Engineering and Traffic Analysis*, 7th edn.: John Wiley & Sons.
12. Momin, K. Al, Barua, S., Jamil, M. S., & Hamim, O. F. (2023a) Short duration traffic flow prediction using Kalman filtering. In: *AIP Conference Proceedings*, 2713(1). https://doi.org/10.1063/5.0129721.
13. Nama, M., Nath, A., Bechra, N., Bhatia, J., Tanwar, S., Chaturvedi, M., &Sadoun, B. (2021) Machine learning-based traffic scheduling techniques for intelligent transportation system: Opportunities and challenges. *International Journal of Communication Systems*, 34(9), e4814. https://doi.org/10.1002/DAC.4814.
14. Neilson, A., Indratmo, Daniel, B., & Tjandra, S. (2019) Systematic Review of the Literature on Big Data in the Transportation Domain: Concepts and Applications. *Big Data Research*, 17, 35–44. https://doi.org/10.1016/J.BDR.2019.03.001.
15. Popat, S. M., Emmanuel, M. (2014) Review and Comparative Study of Clustering Techniques. *International Journal of Computer Science and Information Technologies*, 5 (1), 805-812.
16. Rahman, M. M., & Nower, N. (2023) Attention based Deep Hybrid Networks for Traffic Flow Prediction using Google Maps Data. *ACM International Conference Proceeding Series*, 74–81. https://doi.org/10.1145/3589883.3589894.
17. Rahman, M. M., Ariful Hoque, S. M., & Zaber, M. I. (2018) Understanding real time traffic characteristics of urban zones using GPS data: A computational study on Dhaka City. In: *2018 Joint 7th International Conference on Informatics, Electronics and Vision and 2nd International Conference on Imaging, Vision and Pattern Recognition, ICIEV-IVPR 2018*, 514–519. https://doi.org/10.1109/ICIEV.2018.8640961.
18. Rahman, M. M., Shuvo, M. M. M., Zaber, M. I., & Ali, A. A. (2018) Traffic Pattern Analysis from GPS Data: A Case Study of Dhaka City. In: *2018 IEEE International Conference on Electronics, Computing and Communication Technologies, CONECCT 2018*. DOI:10.1109/CONECCT.2018.8482371.
19. Rahman, Md. M., Nower, N. (2024) *Inferring Traffic Patterns of Dhaka City: A Spatio-temporal Analysis over a Year. Research Square.* https://doi.org/10.21203/rs.3.rs-3827938/v1.
20. Sekar, E. V., Anuradha, J., Arya, A., Balusamy, B., Chang, V. (2018) A framework for smart traffic management using hybrid clustering techniques. *Cluster Computing*, 21(1), 347–362. https://doi.org/10.1007/S10586-017-0855-Y/METRICS.
21. Sheikh, M. S., & Peng, Y. (2022) A Comprehensive Review on Traffic Control Modeling for Obtaining Sustainable Objectives in a Freeway Traffic Environment. *Journal of Advanced Transportation*, 2022. https://doi.org/10.1155/2022/1012206.
22. Stathopoulos, A., & Karlaftis, M. (2001) Temporal and Spatial Variations of Real-Time Traffic Data in Urban Areas. *Https://Doi.Org/10.3141/1768-16*, *1768*, 135–140. https://doi.org/10.3141/1768-16.
23. TBS Report. (2022) Which vehicle dominates Dhaka city roads? The Business Standard. Accessed February 16, 2024, https://www.tbsnews.net/infograph/numbers/which-vehicle-dominates-dhaka-city-roads-396846.

24. Tolaini, M. (2020) *Megacities and globalization of the luxury world*.
25. Turner, Sh. M. (1992) Examination of Indicators of Congestion Level. Traffic Operations. *Transportation Research Record*, 1360, 150-157: Transportation Research Board ISSN: 0361-1981.
26. Yuan, H., & Li, G. (2021) A Survey of Traffic Prediction: from Spatio-Temporal Data to Intelligent Transportation. *Data Science and Engineering*, 6(1), 63–85. https://doi.org/10.1007/S41019-020-00151-Z/TABLES/3.
27. Zhang, H., Dong, Y., Li, J., & Xu, D. (2022) Dynamic Time Warping under Product Quantization, with Applications to Time-Series Data Similarity Search. *IEEE Internet of Things Journal*, 9(14), 11814–11826. https://doi.org/10.1109/JIOT.2021.3132017.
28. Zhang, Z., He, Q., Tong, H., Gou, J., & Li, X. (2016) Spatial-temporal traffic flow pattern identification and anomaly detection with dictionary-based compression theory in a large-scale urban network. *Transportation Research Part C: Emerging Technologies*, 71, 284–302. https://doi.org/10.1016/J.TRC.2016.08.006.
29. Zhao, P., & Hu, H. (2019) Geographical patterns of traffic congestion in growing megacities: Big data analytics from Beijing. *Cities*, 92, 164–174. https://doi.org/10.1016/J.CITIES.2019.03.022.