SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Preliminary Study: A Principal Component Analysis-Based Approach for Link Selection in Traffic Monitoring

Alya Batrisha Azhar[1], Ruzanna Mat Jusoh[1,*], Fatin Amirah Ahmad Shukri[2], Fazilatulaili Ali[1], Nurhana Rafiuddin[3], Sharifah Aishah Syed Ali[1], Arie Restu Wardhani[4]

[1] Department of Defence Science, Faculty of Defence Science and Technology, Universiti Pertahanan Nasional Malaysia, Kem Perdana Sungai Besi, 57000 Kuala Lumpur, Malaysia
[2] Department of Mathematics, Centre for Defence Foundation Studies, Universiti Pertahanan Nasional Malaysia, Kem Perdana Sungai Besi, 57000 Kuala Lumpur, Malaysia
[3] Department of Logistics Management and Business Administration, Faculty of Defence Studies and Management, Universiti Pertahanan Nasional Malaysia, Kem Perdana Sungai Besi, 57000 Kuala Lumpur, Malaysia
[4] Fakultas Teknik, Universitas Widyagama Malang, Kota Malang, Jawa Timur 65142, Indonesia

ABSTRACT

The complex and densely populated large urban network forces the transportation networks to deal with a variety of problems, including traffic congestion. Some strategy that can help to optimize traffic flow and reduce traffic congestion is through traffic management systems by monitoring the traffic conditions. The Network Fundamental Diagram (NFD) is a tool that traffic engineers can use to monitor traffic conditions and manage urban traffic networks because it provides a visual representation of the underlying network dynamics. In estimating NFD, identifying influential links (with measurements points i.e. detectors) is a very important part, which has been a key issue in preserving as much information as possible. To select a subset of links that can accurately represent the behaviour of the entire network, a principal component analysis (PCA) is adopted. PCA can be used to remove the variables (links) that contribute minimal information and retained just the variables that contribute the most. The experimental flow-occupancy dataset from 58 inductive loop detectors at urban city was tested and compared with an experiment of four selection approaches. The findings revealed that the PCA-based strategy of Experiment 2 gave result which was more favourable than the results obtained by other experiments. It was able to preserve the shape of the full measurements NFD, retain the information needed (network capacity and critical occupancy within allowable limit) as well as proven to have lowest total Root Mean Square Error (RMSE) value, in comparison with other experiments.

Keywords:

Network fundamental diagram; Principal component analysis; Sensor placement

1. Introduction

In the era of globalization, surely every person has a vehicle to facilitate their movement. Moreover, with the increase of these vehicles, the current traffic networks have to deal with many

* *Corresponding author.*
E-mail address: ruzanna@upnm.edu.my

https://doi.org/10.37934/araset.59.1.108119

issues that arise considering the increasing number of vehicles on the road, causing traffic congestion and excessive delays that may cause increased in travel time, fuel consumption [21], air pollution [5] while also decreased road safety [8] and above all, result in economic losses. For example, in Malaysia, it is reported that Klang Valley employees spend an average 44 hours a month in their vehicles as a result of traffic jams. The productivity loss, on the road is expected to be RM 308 in a month if the wage is RM 7 per hour [6]. During peak hours, KL drivers individually experienced about 159 hours per year, lost to traffic jam [12]. While that time could be better spent doing other beneficial things such as reading books, road users also paid RM 1023 on petrol for these peak hour trips. Reducing traffic congestion is a complex problem but it is an urgency to consider. Among of the strategy is through traffic management systems; by implementing traffic monitoring which can help to optimize traffic flow and reduce congestion.

In order to address this issue, the Network Fundamental Diagram (NFD) which has been introduced by [4] is a useful tool that offers managing traffic at the urban networks, describing and developing the network-level strategies and mitigating the urban traffic congestion. It has been shown to be useful in metropolitan networks as a perimeter control [11], regional route guidance [3], demand management [23], calculate travel times [15], and control of city- scale ride-sourcing systems [19]. Moreover, this concept being used by [2] to purpose control rule that maximizes the network outflow. By using this promising tool, the traffic networks can be monitored hence improve the network mobility. Figure 1 illustrates a graphical representation of NFD in a transportation network which shows the relation- ship between traffic flow (measured in vehicles per unit time) at the vertical axis and occupancy or density (measured in vehicles per unit length) at the horizontal axis. The NFD exhibits traffic states into four stages, A, B C and D. At A (initial stage), traffic volume is relatively low, and traffic is flowing freely without much interference. In this stage, traffic density is low, and vehicles are able to maintain a high-speed close to the maximum speed limit for the road. As traffic density increases, flow begins to reach its maximum capacity, resulting in congested flow. In this B stage, the flow rate decreases while density increases, causing reduced speeds and slower travel times. Noting that the capacity flow in urban road networks may be seen over a range of vehicle counts, a horizontal line is used to illustrate it [1]. The breakdown stage is reached when the speed drops below a critical threshold. At stage C, traffic density continues to increase, and speed drops even further. The congested stage is reached when the density reaches a critical value where traffic flow is no longer stable and breakdowns are frequent. In this stage, traffic queues become longer and more widespread, and drivers may experience stop-and-go traffic. Final stage D, traffic density has reached the maximum possible value, and speed is zero or close to zero. Vehicles are stuck in a gridlock, unable to move due to the high density and lack of available space to manoeuvre. Understanding these four stages of NFD is essential for predicting and managing traffic congestion. By monitoring traffic volume, density, and speed, transportation planners and engineers can implement strategies to prevent or alleviate congestion, such as improving public transportation, optimizing signal timing, or introducing variable speed limits.

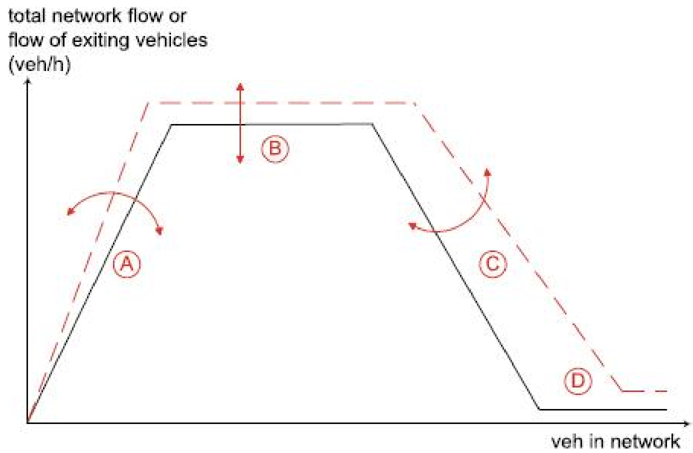


Fig. 1. Fundamental Diagram for Urban Road Networks [1]

However, there is no solid way to produce NFD. Some common factors that significantly influence the curve and shape of NFD is the location of the loop detectors in transport networks. If detectors (sensors) are placed at strategic locations (not only focus at central of business district where areas with heavy congestion) along the road network, they can capture real-time average traffic data, which can be used to dynamically adjust traffic signal timing and manage congestion. This can help to reduce delays and improve the overall flow of traffic, resulting in a good approximation of fundamental diagram. The placement of detectors can also influence the accuracy of traffic flow measurements, which can impact the shape of the NFD. For example, if detectors are placed too close together, they may not accurately capture the flow of whole traffic, resulting in a bad approximation of NFD. Hence, the first motivation in this study is to determine the strategic location of detectors for reliable NFD estimation as a basis input for the Intelligent Transportation Systems (ITS). The quality of traffic data collected by detectors has a direct influence on the reliability of NFD performance.

Moreover, starting in year 2013, [11] was the first who proposed a subset NFD concept, a simplified version of the full NFD that is estimated from a subset of links equipped with detectors in whole networks, i.e. using a sub-amount of information. The urban network in Greece was used for the investigation in their study. Based on the study, the results showed that subset NFDs, which are created using significantly fewer measurements than for a complete NFD, have a critical range of traffic states that is nearly comparable to its equivalent in the whole NFD. The results also shown that using appropriately smaller measurement amounts produces nearly identical NFD approximation results as the original. Hence the second motivation of this paper is to figure out (by selecting or discarding), the links to place the detector, in order to estimate the subset NFD. It is believed with subset NFD, a simpler model than the full NFD, which means that it requires fewer variables and calculations is easier and quicker to use for simple traffic engineering tasks. The subset NFD also offers a more feasible option since detailed data on traffic flow may not always be available. Furthermore, it is usually less expensive to monitor a small subset of links than to monitor the entire links in network. For transportation authorities with limited resources who want to utilize a more focused and specific approach to traffic monitoring, this will be very helpful.

The quantity and locations of the NFD subset have only been the subject of a few studies. Another related studies of traffic sensor location problem (TSLP) can be located in [17]. In 2015, [16] conducted a study where the aim of the study was to evaluate the information needed for a potential application of a control system in the urban area. This study expanded the [11] work and the information from ZuriTraffic links was used to investigate into the correlation between network coverage (i.e., the percentage of streets monitored) and the macroscopic fundamental diagram's accuracy. The results demonstrated that regardless of the strategy of selecting links, 25 percent is a minimum network coverage. In a recent study, [24] developed a mathematical model to optimize the fixed measurement points (i.e., loop detector locations) and sampling of probe vehicle trajectories while taking into account the limited budget, heterogeneity, and asymmetry in Origin-Demand (OD) matrices in real-world networks. They aimed to reduce the estimation error between the ground-truth MFD and the MFD obtained from the scarce resources. Nevertheless, this model also needs the ground-truth MFD as an input, which is unavailable in real-world networks. Furthermore, [10,13] studied about the development of operational or sparse-measurement NFD by selecting fewer sensors and the associated measurement by using information theoretic approach. The results demonstrated that the obtained Sparse-Measurement diagrams from the chosen models successfully preserve the shape and key characteristics of a complete measurement diagram. In recent study, [20] estimating MFD using flow and density measurements from critical link which are suggested by Principal Component Analysis (PCA). According to the results obtained, applying PCA helps to find the key traffic features and identify the critical links that are connected to these critical links. It is supported by [9] to further research subset NFD as its accuracy estimation is the foundation of accuracy in further NFDs' applications and network-level traffic models. Inspired by [20], this research employed a PCA-based approach to selectively choose sensors for constructing subset NFDs. This theoretically reduces costs as it requires fewer measurement resources.

2. Methodology

In this section, the data used for this study, PCA and some comparative studies and estimation of the subset NFD based of selected links will be discussed.

2.1 Traffic Detector Data

The traffic data used for analysis purposes are collected using loop detectors at 58 locations around the urban road network of the Chania, Greece, as illustrated in Figure 2. Real data of flow and occupancy are stored every 90-sec signalization cycle and aggregated at time intervals of 1.5-minute duration. The dataset includes measurements of data on Friday spanning 24 hours. The occupancies of lane were aggregated to 6-min all-lane and averaging, and flow were aggregated to 6-min all-lane and summing.

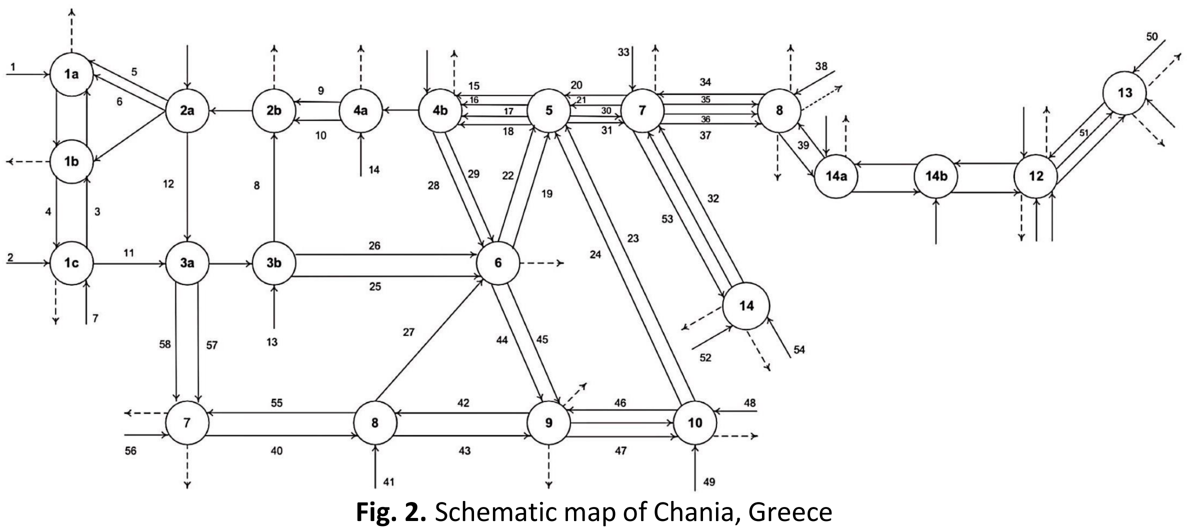


Fig. 2. Schematic map of Chania, Greece

2.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a helpful tool in data processing and dimensional reduction technique [22] founded by [18] and [7]. It serves a simple and non-parametric method that is used to reduce the dimensionality and extract relevant information from confusing data sets [22]. PCA works in reducing number of variables in data set by maintaining the important information. By using the PCA, the data set can be reduced while retaining as much of the dataset's information as possible, to a smaller set of uncorrelated variables [20]. Basically, in PCA, there are several steps that need to be considered as depicted in Figure 3.

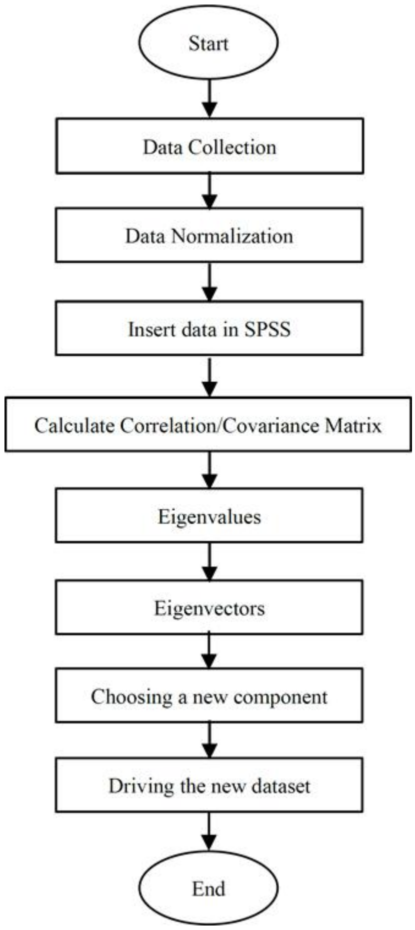


Fig. 3. General Steps in PCA [14]

PCA aims to convert the initial variables into a small number of linear combinations, referred to as principal components (PCs). PCA reduces the dimensions of a dataset with d variables by projecting it onto a subspace with k dimensions, ensuring that the maximum variance of the original data is preserved (where $k < d$). In order to study the impact of original variables on each PC, loading matrix needs to be calculated. Specifically, the loading matrix indicates the correlation between each original variable and each PC. In this work, it is attempted to identify if variable (links) from first PC with high and low loading values signifies an importance in networks, means produced a good subset NFD. For that, important links was chosen based on first PC, either to have high loading values or low loading values. Detail explanation can be found on next section.

2.3 Construction of Subset NFD

Few experiments were conducted with k links (out of 58 links) chosen. The idea is to determine which most informative k links that can be represent average traffic condition from whole network. The $k = 14$ links was used in this study to indicate a low percentage of data used, but still manage to preserve the true information of all links. Experiment 1 and 2 were based on PCA approach, to detect the representative links. In the initial stages of the work, both experiments utilized the first principal component, as it was deemed to account for the greatest variance in comparison to the other principal components. In Experiment 1, links with first 14 highest loading value were selected. In contrast, links from the first 14 with the lowest loading values were chosen in Experiment 2. The purpose was to investigate how two extreme groups (min and max) in loading value affected the NFD estimation.

Two alternative detector selection strategies are taken into consideration to evaluate the efficacy of the PCA results. Experiment 3 was conducted based on random link selection and Experiment 4 was conducted based on geographical location of city centre in the network, referred as "Radius Strategy". Here, it is intended to determine whether sensors placed in particular network nodes aid in improving NFD approximation. Figure 4 shows the network is divided into five equal-sized areas (A, B, C, D, and E) with the same number of sensors for a fair comparison by using the same procedure [13] used, and region B was utilized in this work.

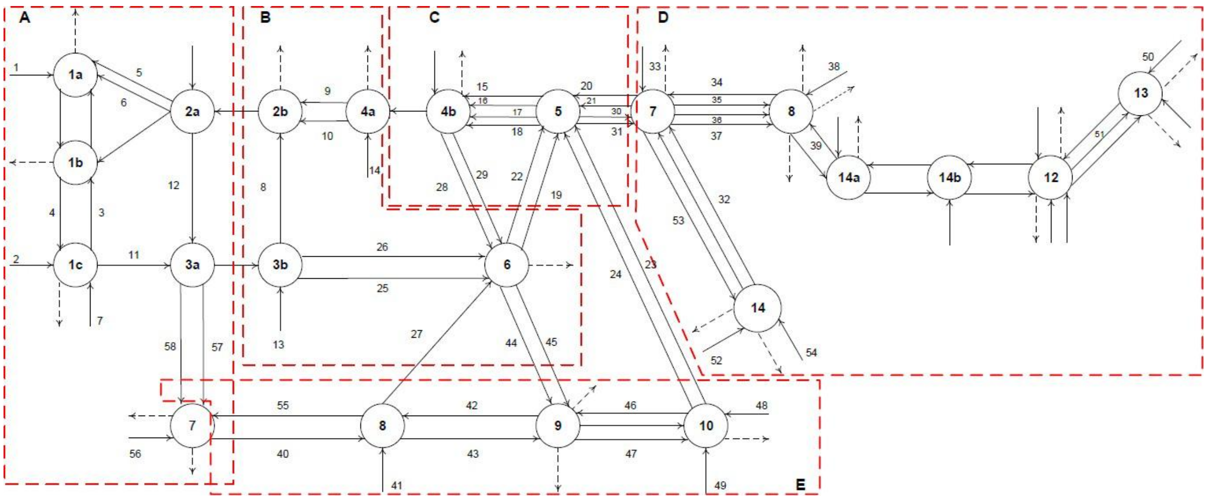


Fig. 4. The partition of region for Chania Urban Road Network [13]

Table 1 describes the details of all four experiments.

Table 1
Experiments conducted to construct subset NFDs

Experiment 1	Experiment 2	Experiment 3	Experiment 4
The flow data was used, and 14 variables (links) with the top 14 highest loading values from first PC were selected.	The flow data was used, and 14 variables (links) with the top 14 lowest loading values from first PC were	Randomly pick 14 links from the network in Figure 2.	The links were chosen from the region B from the network partitioned by the previous study [13].

selected

After identifying the respective links for each experiment, the measurements associated with the links, including flow and occupancy, were collected. Subsequently, the subset NFD was constructed based on this collected data. To perform validation measures, two approaches were used. First is to compute the Root Mean Squared Error, RMSE adapted from [20] for each experiment to assess the goodness of fit and to validate if the outcome able to minimize the estimation error between subset NFD and full NFD. Three different measures were used to assess the precision of the estimated NFDs and quantify how close the subset NFD were to the actual NFD. RMSE for network average flow and network average occupancy for T time intervals were calculated as follows:

$$RMSE(Q) = \sqrt{\frac{\sum_{t=1}^{T}\left(\widehat{Q}_t - Q_t\right)^2}{T}} \quad (1)$$

$$RMSE(O) = \sqrt{\frac{\sum_{t=1}^{T}\left(\widehat{O}_t - O_t\right)^2}{T}} \quad (2)$$

where $\widehat{O}_t$ and O_t represent the true and estimated average network occupancy in time interval t, respectively, and $\widehat{Q}_t$ and Q_t represent the true and estimated average network flow in time interval *t*, respectively. The *RMSE* which capture both flow and occupancy is calculated as follows:

$$RMSE_{total} = RMSE(Q) + \ RMSE(O) \quad (3)$$

The second assessment involves observing the subset NFD to assess its ability to offer a satisfactory approximation of NFD while preserving the shape and essential characteristics (critical flow and occupancy, dispersion) of the complete NFD.

3. Results and Discussions

The traffic conditions are varying by day, Figure 5 shows the shape of full NFD for the Friday data set based on 58 links, the full measurements.

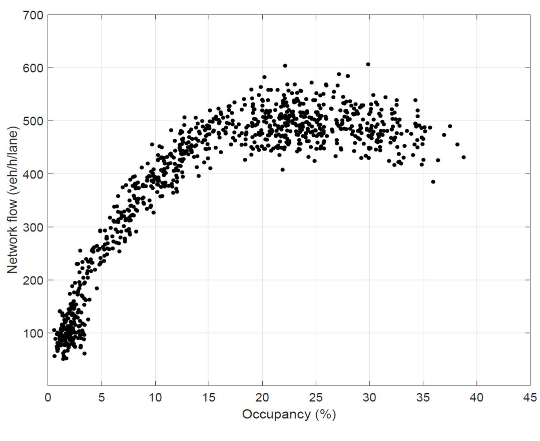


Fig. 5. Full-measurement NFDs for Friday

Table 2 shows the capacity of maximum flow and the occupancy values for Friday. As can be seen the maximum flow is around 500 veh/h/lane, while the corresponding critical occupancy values are between 20 - 25%.

Table 2
Capacity flow and critical occupancy for Friday

Day	Capacity flow (veh/h/lane)	Range of critical occupancy (%)
Friday	500	20 - 25

A scree plot is commonly used in summarizing the finding of a PCs analysis in order to decide how many PCs should be retained. Based on the screen plot in Figure 6, it shows the first component explained 50% of the total variance. Furthermore, the most noticeable change in slope in the scree plot occurs at component 4, which is the "elbow" of the scree plot.

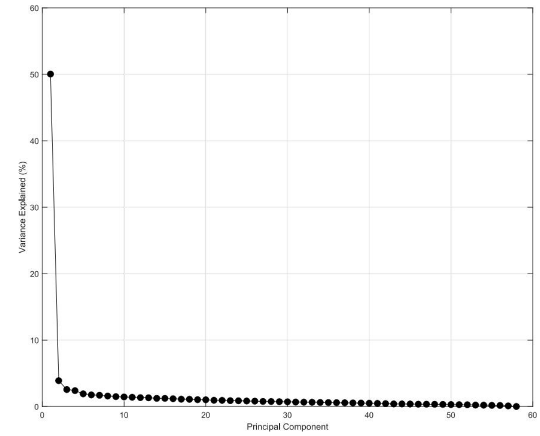


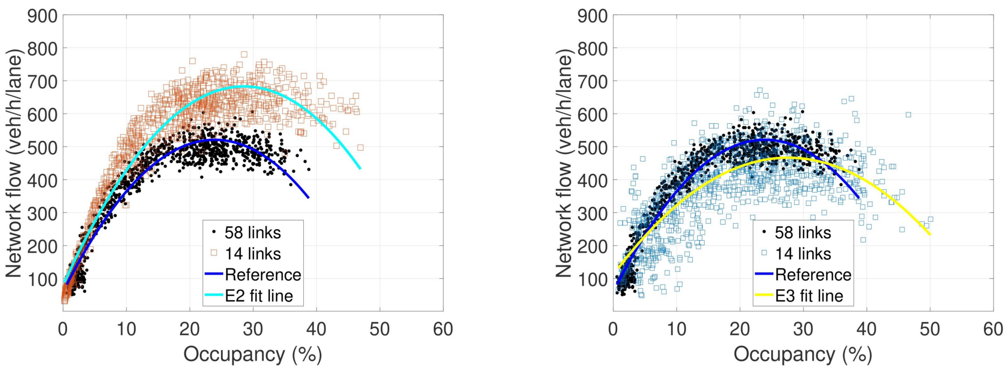
Fig. 6. Scree plot for Friday Flow

The summaries of the obtained result for all experiments were shown in Table 3. It can be seen on Table 3, PCA- based approach which are Experiment 1 and 2 managed to preserve the occupancy as the full NFD since the value lies within acceptable range of 20 - 25 %. Meanwhile, occupancy value of experiment 3 and 4 which are referring to blind and radius strategy, respectively fall outside the desired characteristic value. However, only experiment 2 was able to sustain the flow capacity of 490 vehicles per hour per lane, staying within acceptable value of flow (450 to 550 vehicles per hour per lane), which is +/- 10% from 500 vehicles per hour per lane of full NFD. Furthermore, experiment 2 also produces the lowest $RMSE_{total}$ of 79, in comparison to all experiments. This suggested experiment 2 contain more information and able to minimize the estimation error between subset NFD and full NFD. This also means experiment 1 and 3 contains less information due to high $RMSE_{total}$. The results of Experiment 4, radius strategy exhibits fair $RMSE_{total}$ with the full NFD. After all, it can be concluded that the adopted PCA-based strategy using flow data set ensures the accuracy of the information for various network coverage levels. Specifically, using only the first PC provide fair amount of variable reduction (link selection). To our surprise, first PC with low loading value gave a good approximation, instead of PC with high loading value. This finding contradicts with results from [20] as they claimed top k highest loading value from multiple PCs will provide good approximation. The limitation of our study maybe due to the usage of only the first PC. Considering a higher number of PCs maybe can lead to a better approximation. Another reason may be due to the nature of data used in this study where the data was collected from loop detector instead of using probe vehicle data.

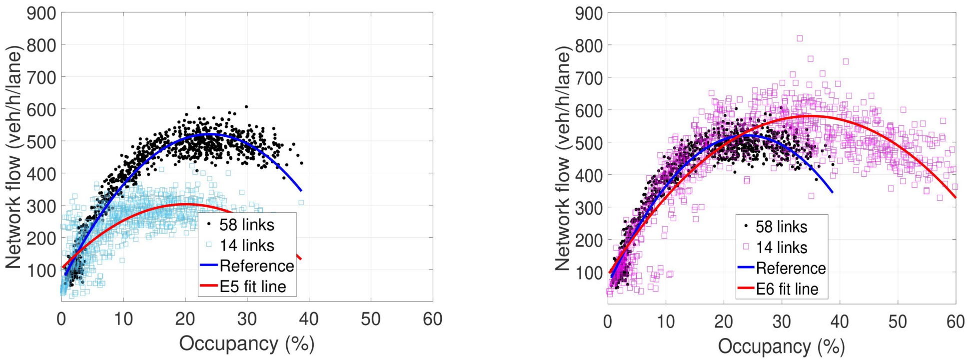
Table 3
Summary of all experiments involved

No	Parameter	Full NFD	Experiment 1 (PCA)	Experiment 2 (PCA)	Experiment 3 (Blind Strategy)	Experiment 4 (Radius B)
1	Loading Value	-	Highest	Lowest	-	-
2	Selected Links	58 links	{1, 5, 6, 16, 17, 19, 20, 26, 39, 41, 43, 46, 49, 54}	{2, 4, 8, 10, 14, 22, 23, 24, 30, 37, 44, 52, 57, 58}	{2, 3, 7, 8, 13, 14, 23, 27, 30, 37, 38, 40, 48, 52}	{8, 9, 10, 13, 14, 19, 22, 24, 25, 26, 27, 28, 44, 45}
3	Critical Occupancy (%)	20-25	24	25	20	30
4	Flow Capacity (veh/h/lane)	450-550	630	490	250	570
5	$RMSE(O)$	-	3.28	4.40	6.81	11.91
6	$RMSE(Q)$	-	117.78	75.11	180.15	68.27
7	$RMSE_{total}$	-	121.04	79.51	186.96	80.18

Figure 7 shows the subset NFD for the four experiments conducted. It can be seen that the Figure 7(b) provide a good approximation among all, although has wide scatter. Experiment 2 able to preserve the shape and characteristics of full NFD. Experiment 1 and 3 in Figure 7(a) and (c) provide higher and lower approximation, respectively where it is observed in large $RMSE_{total}$ value. On the other hand, Figure 7(d) displays a significantly greater scatter with a wide dispersion of points. This highlights the fact that installing sensors in the busiest area of the network does not guarantee a good approximation of the NFD.



(a) Experiment 1
(b) Experiment 2



(c) Experiment 3 (d) Experiment 4

Fig. 7. Fitted line for subset NFD of a smaller number of links for Friday according to different experiments

Meanwhile, Figures 8 shows the location of the links chosen in the network, for each Experiment 2. The highlighted links show where the sensors should remain.

Based on the findings of the four experiments, 14 links that were selected from the first PC with the low loading value (Experiment 2) was successful in maintaining the general shape and characteristics of the full-measurement diagram. To ensure experiment accuracy, a comparison of results was conducted using validation measurements based on the RMSE value. Experiment 2 displays a good fit to the full NFD and maintains the critical occupancy between 20 and 25%. On the other hand, the experiment based on high loading value showed poor approximation. Generally loading values represent the correlation between the original variables and the principal components, with higher loading values indicating a stronger relationship between the variable and the component. However surprisingly for our study, variables (links) with low loading values were more relevant and provide unique contribution to the analysis, especially in estimating subset NFD. Meanwhile, it is evident that the blind and radius technique, provides a worse approximation. Both approaches will become less successful in a more complicated traffic network.

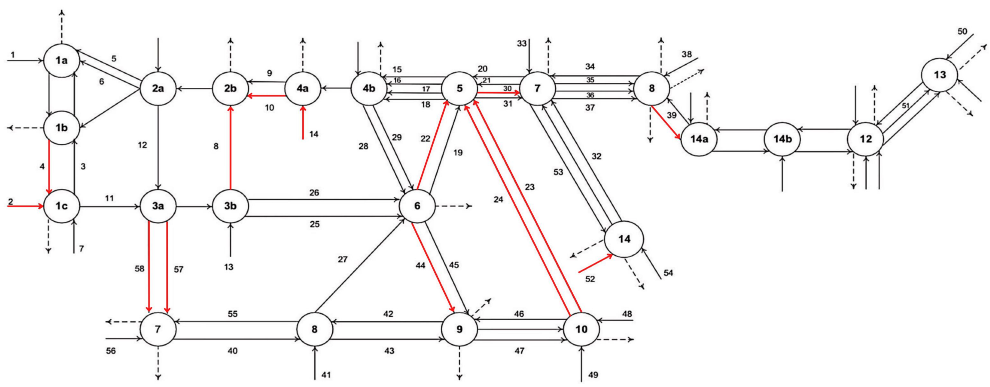


Fig. 8. Link (sensor) selection of Experiment 2

4. Conclusions

Analyses were carried out using real loop detector data from CBD of Chania to estimate the initial fundamental diagram (100% of links). Then PCA-based method was implemented to obtain the subset fundamental diagram. The goal is to determine the level of accuracy retained from subset NFD. The significance of this study is for identification of the ideal detector or sensor placement in large scale networks. With reduced fundamental diagram, an allocation of adopting a plan for monitoring and managing traffic congestion would be decreased.

According to the research conducted, a network essential diagram with little scatter exists for Chania's CBD. Further on, investigation takes place in finding ideal location of detector in the network in order to find reduction of the number of links used. PCA strategy was used and few experiments involved denoted by Experiment 1,2,3 and 4. In Experiment 1, links were selected based on variable with highest loading value, from first principal component in Friday flow dataset. In Experiment 2, links were chosen based on the variable in the flow dataset's first principal component with the lowest loading value. Then the PCA-based technique was compared with other 2 strategies namely blind and radius strategy, denoted by Experiment 3 and 4.

It can be seen from the results during the analyses that the PCA strategy of Experiment 2 gave result which was more favourable than the results obtained by other strategies or experiments. It was able to:

i. preserve the shape of the full measurements NFD
ii. retain the network capacity and critical occupancy (lies within range) as well as
iii. proven to have lowest total RMSE value, in comparison with other experiments.

Using the findings of this analysis for 14 links, or roughly 23% of all links, the reduced NFD for Chania was created, and it can be shown that it was able to preserve the shape of the full measurements NFD. However, because fewer links were chosen, there was more scatter in the diagram, as indicated by having high RMSE value. The other PCA-based strategies of experiment 1 where links were selected based on high loading value from flow dataset, coming from first principal component, gave poor estimation which exhibit high RMSE along with unable to preserve the shape of full measurement NFD.

The blind technique used in Experiment 3 was less preferable since it produces poor approximations. Finally, Experiment 4 using the radius technique, which used links in the CBD's centre, produced results that were generally satisfactory despite having a widely scatter diagram. Even so, the analyses were valuable since it helped to identify the areas of the urban network where traffic congestion is occurring, which will be helpful for any future research.

However, this research limitation is due to dataset since it is using only Friday data set. It is recommended for future research to test this adopted PCA-based strategy on an alternate day, preferably during the weekend, as the variations in traffic conditions present an intriguing aspect to explore. Also to consider using multiple number of PCs and select representative links which can capture average network traffic conditions. The presented approach can also be applied to data from other fields, such as meteorology or forestry, in order to find and choose a subset of variables that can precisely capture the behaviour of the whole system.

Acknowledgement
Acknowledgment is given to the support received, particularly from UPNM, and notably from the Faculty of Defence Science and Technology. Gratitude is also extended towards the Short-Term Grant Scheme (UPNM/2020/GPJP/SG/4).

References

[1] Aboudolas, K. M. A. E., Markos Papageorgiou, Anastasios Kouvelas, and E. Kosmatopoulos. "A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks." *Transportation Research Part C: Emerging Technologies* 18, no. 5 (2010): 680-694. https://doi.org/10.1016/j.trc.2009.06.003

[2] Daganzo, Carlos F., and Nikolas Geroliminis. "An analytical approximation for the macroscopic fundamental diagram of urban traffic." *Transportation Research Part B: Methodological* 42, no. 9 (2008): 771-781. https://doi.org/10.1016/j.trb.2008.06.008

[3] Geroliminis, Nikolas, Jack Haddad, and Mohsen Ramezani. "Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach." *IEEE Transactions on Intelligent Transportation Systems* 14, no. 1 (2012): 348-359. https://doi.org/10.1109/TITS.2012.2216877

[4] Godfrey, J. W. "The mechanism of a road network." *Traffic Engineering & Control* 8, no. 8 (1969).

[5] Ha, Chin Yee, Terh Jing Khoo, and Zheng Yik Koo. "Current status of green building development in Malaysia." *Progress in Energy and Environment* (2023): 1-9. https://doi.org/10.37934/progee.25.1.19

[6] Hayin, N. A. M. and Ismail, M. R. "Klang valley employees spend 44 hours a month sitting in traffic." *The Guardian*. (2022).

[7] Hotelling, Harold. "Analysis of a complex of statistical variables into principal components." *Journal of educational psychology* 24, no. 6 (1933): 417. https://doi.org/10.1037/h0071325
[8] Ibrahim, Muhammad Shafiq, Seri Rahayu Kamat, and Syamimi Shamsuddin. "The role of brain wave activity by electroencephalogram (EEG) in assessing cognitive skills as an indicator for driving fatigue: A review." *Malaysian Journal on Composites Science and Manufacturing* 11, no. 1 (2023): 19-31.
[9] Johari, Mansour, Mehdi Keyvan-Ekbatani, Ludovic Leclercq, Dong Ngoduy, and Hani S. Mahmassani. "Macroscopic network-level traffic models: Bridging fifty years of development toward the next era." *Transportation Research Part C: Emerging Technologies* 131 (2021): 103334. https://doi.org/10.1016/j.trc.2021.103334
[10] Mat Jusoh, Ruzanna, and Konstantinos Ampountolas. "Optimal Selection of Traffic Sensors: an Information-Theoretic Framework." (2019): 3297-3302. https://doi.org/10.23919/ACC.2019.8815159
[11] Keyvan-Ekbatani, Mehdi, Markos Papageorgiou, and Ioannis Papamichail. "Urban congestion gating control based on reduced operational network fundamental diagrams." *Transportation Research Part C: Emerging Technologies* 33 (2013): 74-87. https://doi.org/10.1016/j.trc.2013.04.010
[12] Lim, I. “Study shows kl drivers lost 159 hours and rm1,023 in fuel to peak hour traffic, could have read 31 books instead.” *Malay Mail*. (2023).
[13] Mat Jusoh, Ruzanna. "Monitoring and control of transport networks using parsimonious models." PhD diss., University of Glasgow, 2019.
[14] Mazlan, Ain Dzarah Nafisah, Muhammad Asraf Hairuddin, Nooritawati Md Tahir, Nur Dalila Khirul Ashar, and Mohamad Huzaimy Jusoh. "Comparative analysis of PCA and ANOVA for assessing the subset feature selection of the geomagnetic Disturbance Storm Time." *Journal of Electrical and Electronic Systems Research (JEESR)* 17 (2020): 8-16. https://doi.org/10.24191/jeesr.v17i1.002
[15] Musolino, Giuseppe, Antonio Polimeni, and Antonino Vitetta. "Freight vehicle routing with reliable link travel times: a method based on network fundamental diagram." *Transportation Letters* 10, no. 3 (2018): 159-171. https://doi.org/10.1080/19427867.2016.1241040
[16] Ortigosa, Javier, Monica Menendez, and Hector Tapia. "Study on the number and location of measurement points for an MFD perimeter control scheme: a case study of Zurich." *EURO Journal on Transportation and Logistics* 3, no. 3 (2015): 245-266. https://doi.org/10.1007/s13676-013-0034-0
[17] Owais, Mahmoud. "Traffic sensor location problem: Three decades of research." *Expert Systems with Applications* 208 (2022): 118134. https://doi.org/10.1016/j.eswa.2022.118134
[18] Pearson, Karl. "LIII. On lines and planes of closest fit to systems of points in space." *The London, Edinburgh, and Dublin philosophical magazine and journal of science* 2, no. 11 (1901): 559-572. https://doi.org/10.1080/14786440109462720
[19] Ramezani, Mohsen, and Mehdi Nourinejad. "Dynamic modeling and control of taxi services in large-scale urban networks: A macroscopic approach." *Transportation Research Part C: Emerging Technologies* 94 (2018): 203-219. https://doi.org/10.1016/j.trc.2017.08.011
[20] Saffari, Elham, Mehmet Yildirimoglu, and Mark Hickman. "A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks." *Transportation Research Part C: Emerging Technologies* 119 (2020): 102743. https://doi.org/10.1016/j.trc.2020.102743
[21] Samsudin, Muhammad Syazwan Nizam, Md Mizanur Rahman, and Muhamad Azhari Wahid. "Sustainable power generation pathways in Malaysia: Development of long-range scenarios." *Journal of Advanced Research in Applied Mechanics* 24, no. 1 (2016): 22-38.
[22] Shlens, Jonathon. "A tutorial on principal component analysis." *arXiv preprint arXiv:1404.1100* (2014).
[23] Yildirimoglu, Mehmet, Mohsen Ramezani, and Nikolas Geroliminis. "Equilibrium analysis and route guidance in large-scale networks with MFD dynamics." *Transportation Research Procedia* 9 (2015): 185-204. https://doi.org/10.1016/j.trpro.2015.07.011
[24] Zockaie, Ali, Meead Saberi, and Ramin Saedi. "A resource allocation problem to estimate network fundamental diagram in heterogeneous networks: Optimal locating of fixed measurement points and sampling of probe trajectories." *Transportation Research Part C: Emerging Technologies* 86 (2018): 245-262. https://doi.org/10.1016/j.trc.2017.11.017