# Framework for the Selection of Loop Detectors for Macroscopic Fundamental Diagram Estimation

Syed Muzammil Abbas Rizvi



June 20, 2023

Transportation Research Procedia 00 (2023) 000–000

World Conference on Transport Research - WCTR 2023 Montreal 17-21 July 2023

# Framework for the Selection of Loop Detectors for Macroscopic Fundamental Diagram Estimation


**Abstract**

The Macroscopic Fundamental Diagram (MFD) represents an increasingly established model for assessing the quality of traffic flow in networks. However, the uniqueness of an empirically estimated MFD cannot be guaranteed due to the problem of detector selection. Instationarity and varying flow patterns make it difficult to select the link flows that are representative of the traffic state in the whole network. This paper developed a new method for selecting loop detectors that represent a particular traffic state of a road network. The method relies on a metric of heterogeneity characterizing the role of a network link over the time of a day. The dispersion indicates the heterogeneity in traffic conditions and the dynamic role of each time interval. The heterogeneity-weighted saturation level of links is used to determine a ranking of links. The high-ranked links in the ranking represent the most homogenous sample of subset links.

The study used the loop detector data of Zurich and London and a simulated network to compare both equal (classical) and dynamic weights (proposed) by selecting the sample links based on different saturation levels. Moreover, associating the saturation level with the heterogeneity level specified the links creating the heterogeneity in the road network primarily.


*Keywords:* Macroscopic Fundamental Diagram; Heterogeneity; Loop detectors; Entropy weights; TOPSIS

## 1 Introduction:

Nowadays, congestion has become an integral part of the urban road network. Managing road networks to operate nearer to capacity is an optimal approach. However, at the optimum point, an increase in the traffic flow will initiate congestion in the road network. Therefore, efficient traffic demand management is required to avoid delays and congestion. Geroliminis and Daganzo (2008) developed the relationship between traffic flow, density, and speed in Yokohama, Japan which provides information about traffic conditions required for traffic demand management. The relationship is known as the Macroscopic Fundamental Diagram (MFD). MFD can evaluate the performance of the road network at a large scale (Mahmassani, Saberi et al. 2013). Moreover, MFD can also be used to control the inflow of the road network to avoid the density exceeding the critical point. Later, MFD developed for many cities such as California, Amsterdam, and Brisbane (Gayah and Dixit 2013, Tsubota, Bhaskar et al. 2014, Knoop, van Erp et al. 2018). MFD consists of free flow and congested regions, and the slope of the regions is known as free-flow and shockwave speeds, respectively. These parameters along with capacity and critical density, are required to develop control strategies, traffic modeling, and vehicle guidance in the road network (Daganzo, Gayah et al. 2011, Aboudolas and Geroliminis 2013, Leclercq and Geroliminis 2013).

Analytical and empirical methods are used to estimate the MFD. Many studies are present in the literature that estimate the MFD analytically by Variational Theory (VT) (Daganzo 2005), Method of Cuts (MoC) (Daganzo and Geroliminis 2008), Probability Density Function (PDF) (Geroliminis and Sun 2011), and Linear Program (LP) (Daganzo and Lehe 2016). The analytically developed MFD is independent of the impact of demand evolution, and the estimated parameters show the maximum output of the system. However, the analytical method is limited to isolated links and signalized intersections. The literature has still not developed any analytical method for the MFD estimation of the entire road network. The primary obstacles are the limited capabilities of capturing network heterogeneity, signal control schemes, and hierarchical network structures (Zhang, Yuan et al. 2020). Furthermore, optimal traffic management of an isolated section of the road network does not ensure improvement of the traffic conditions of the entire road network (Zhang, Pei et al. 2022). Therefore, the researchers are focused on the empirical estimation of the MFD. The empirical estimation of MFD uses the traffic data collected from the links of the road network.

Traffic data is the primary input for the empirical MFD. Traffic conditions are time-dependent, and the collected traffic data must cover the dynamic traffic conditions of the road network. The distribution of dynamic traffic conditions is inhomogeneous in urban road networks (Ramezani, Haddad et al. 2015). The gap between supply and demand and the road network hierarchy is the main reason for the dynamic traffic conditions in the road network (Xie and Levinson 2007). Consequently, each link in the road network has a different traffic flow pattern during the day, making the road network conditions heterogeneous. The heterogeneity increases during peak hours when congestion occurs in a few links compared to the total number of links in the road network. Simultaneously, other links in the road network are in a free-flow state. Thus, the standard deviation of the volumes is higher with a higher average road network density (Mazloumian, Geroliminis et al. 2010). From the large number of links in an urban road network, data from only a few links are typically sampled to estimate the MFD.

The heterogeneity in the road network makes it difficult to determine the subset sample of links for traffic data collection. Dynamic traffic conditions and different flow patterns of links make the sample unrepresentative of the road network and introduce uncertainties in the estimated MFD. The uncertainty in the traffic data will account for inaccurate estimation of MFD (Ambühl, Loder et al. 2018). Furthermore, the accuracy of the traffic flow parameters required for road network management will also be affected. Hence, heterogeneity in the road network plays a crucial role in selecting the sample links and the accuracy of the MFD estimation. Limited studies are available in the literature to determine the sample links for traffic data collection. Keyvan-Ekbatani, Papageorgiou et al. (2013) collected traffic data from a subset of links of the road network. The authors assumed that the subset links represent the entire road network. Nevertheless, the authors used a visual interpretation to select links in simulation that is not applicable in the real world. This method was further modified to develop mathematical models to find the optimal subset of links for estimating the MFD of the road network (Ortigosa, Menendez et al. 2015, Zockaie, Saberi et al. 2018). Ortigosa, Menendez et al. (2015) formulated four blind strategies for link selection and urged for developing a systematic method to select representative links. Ji, Xu et al. (2018) selected the 30% congested and 30% non-congested links of Changsha, China, and estimated the MFD. Saffari, Yildirimoglu et al. (2020) used the Principal Component Analysis (PCA) to identify critical links in the road network. The study collected data from local loop detectors (LLD) of the critical links and floating car data (FCD) of all links. The Principal Component (PC) of each link and LLD data of critical links reconstruct the data of the entire road network. The study used the traffic data from the FCD for calculating the PC of links that require full road coverage of the FCD, which is difficult to achieve in urban road networks.

Ambühl, Loder et al. (2018) introduced the re-sampling method to estimate the traffic parameters with minimal data. The authors randomly selected the links based on the different sample sizes and estimated the MFD. This process continues until the MFD plane has a clear upper bound. The study determines the additional capacity obtained by the subset links as a measure of the level of heterogeneity of the road network. Previous studies that collected the data from the subset links for the MFD estimation ignored the level of heterogeneity of the road network. Although Ambühl, Loder et al. (2018) calculated the level of heterogeneity, no study in the literature considers the effect of heterogeneity in the selection of the subset links. The heterogeneity of a road network is not constant. Mostly, the heterogeneity is higher during peak hours. The variation in heterogeneity is a key factor in the accuracy of MFD estimation. Similarly, the same heterogeneity level in the morning and evening peaks is not possible because travel patterns are different in every city and different road hierarchy. Moreover, some networks are heterogeneous for a certain period while some road networks remain heterogeneous most of the time during the day.

The application of MFD-based strategies depends on the accuracy of the MFD estimation, and no strategy exists that is robust enough in the application based on inaccurate MFD. On the other hand, the subset links are prone to inhomogeneity, which could affect the accuracy of the MFD estimation. Data collected from a sample (subset) of links may not represent the characteristics (traffic conditions) of the system (road network), resulting in an inaccurate estimate of the MFD. The studies in the literature on link selection have quantitative outcomes, i.e., the accuracy of the MFD is associated with selecting a higher number of links (Saffari, Yildirimoglu et al. 2020). The subset of links is homogeneous and representative of the road network when the estimated MFD is closer to the theoretical upper bound (Daganzo and Geroliminis 2008). Representing the traffic conditions of the road network from subset link data and obtaining an upper bound in the MFD plane is a challenge. Considering the temporal variation of the heterogeneity of the links and their saturation level (traffic flow rate w.r.t given capacity), the role of each time interval is dynamic. However, the literature ignores the dynamic role of each time interval which implies that all time intervals are considered equal.

This study addresses the gap by selecting the most homogeneous sample of subset links based on the variation in heterogeneity that characterizes the dynamic role of each time interval. With the weighted saturation level, a ranking is compiled by calculating a Performance Score ($P_i$) for each link. Weights represent the dynamic role of time intervals according to traffic conditions. First, weighting the time intervals of traffic flow data according to traffic conditions maintains temporal homogeneity at the time interval level. The weights are then used to determine the saturation level of each link. Links having high $P_i$ score have high saturation and similar traffic conditions during the day. As a result, selecting links with higher score have a higher probability of representing the road network; the data points are closer to the theoretical upper bound in the estimated MFD plane. Moreover, the highly saturated links provide information of different traffic states in the road network.

This study uses the traffic flow data obtained from a simulated road network and the LLD of links in the CBD of Zurich and London for MFD estimation. Data of sample links (respective detectors) selected from two different methods, i.e., weighted saturation level (proposed methodology) and unweighted saturation level, are used to estimate the MFD. The difference between the MFDs shows that the dynamic role of time intervals decreases the susceptibility of sample links to heterogeneity and increases their representativeness. Finally, the better representation of the traffic state of the network provided by the sample links increases the accuracy of the MFD estimation. The measure developed for calculating the heterogeneity level indicates the links primarily involved in the heterogeneity of the road network. The method of Ambühl, Loder et al. (2018) obtained the upper bound in the MFD by the subset data. However, the computational effort is significantly higher because both flow and density are required at the initial level, and re-sampling depends on the number of links. As a result, the computation time for a large road network would become much higher (Ortigosa, Menendez et al. 2015). Whereas the proposed methodology is relatively inexpensive, it requires only flow data for link selection and then uses the density of the selected links for MFD estimation; increasing the number of links will only linearly increase the computational time. Moreover, the study also measure the level of herterogenoty and relates it with the saturation level that specified the links primarlity invloved in creating the heterogeneity in the road network. Ultimately, the application of the methodology on real and simulated road nework highlight the transferability of the study.

The structure of this paper is organized as follows. The following section discusses the methodology for selecting the subset links. Section 3 explains the data used for the application of the method. Sections 4 and 5 discuss the results of the ranking of the detectors, the estimation of MFD, and the heterogeneity level of the road network of a real data and a simulated network, respectively. The last section concludes the study and provides insights for future extension.

## 2 Methodology:

The traffic flow pattern is different at each link of the road network. The traffic state in a road network is homogeneous when most links have similar flow patterns, and it becomes heterogeneous when the pattern varies at most of the links. The more heterogeneous the traffic state in the network, the less likely a selected sample of links represents the traffic state of the network. Following a specific flow pattern by every link to reduce heterogeneity is a subjective approach. The objective approach is identifying the dominant traffic flow pattern in most road network links. The dominant pattern gives an overall profile of the dynamic traffic conditions throughout the day (homogeneous to heterogeneous). Ultimately, the profile shows that the time intervals differ from each other based on traffic conditions and cannot be considered equal.

The primary objective of this study was to develop a ranking of the links for the selection of the sample. The ranking methodology was based on the weighted saturation level. Finding the dominant traffic flow pattern was the main task, as the flow profile of any link could not be favored to avoid subjectivity in the methodology. This study used the Entropy Weight Method (EWM) and the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) to calculate weights and saturation levels, respectively. These two models were used simultaneously and called Entropy-based TOPSIS (E-TOPSIS).

### *2.1 Selection of Links:*

An MFD estimated using data from sample links is usually susceptible to inhomogeneity of traffic conditions of the road network. Typically the data points characterizing traffic flow over density lie below the upper bound, although the primary aim is to determine the upper bound of these parameters. The heterogeneity in the road network varies with time. Road networks tend to have homogeneous traffic conditions in the morning, and heterogeneity increases during peak hours. In peak hours, only a few links have higher flow concentrations; the remaining links are in a free-flow state. Therefore, a lower dispersion is found in the data when the flow concentration is higher in a few links. Similarly, the dispersion is higher in the flow data when the flow concentration is lower in the links in homogeneous traffic conditions. As discussed, the main task in the methodology was to find the traffic flow pattern that is dominant in most of the links. Therefore, the dispersion in the traffic flow data at each time interval on given traffic demand indicates the effect of dynamic traffic conditions during the day. To take this effect into account, we applied the EWM, which calculates the dispersion in the data at each time interval and assigns weights accordingly.

If data points from the subset sample data lie close to the upper bound in the MFD plane, then sample links should have higher saturation. The TOPSIS method calculates the saturation level of each link at each time interval. Moreover, the TOPSIS method requires weights for each parameter (time intervals) and uses the Entropy Weights (EW) from EWM to include the effect of dynamic traffic conditions. The weighted saturation level ensures the formation of an upper bound in the MFD without affecting the homogeneity of the sample links. Using the combination of these two methods (E-TOPSIS), we calculated the $P_i$ of each detector. Given the $P_{i,}$ the ranking of the links could be developed. Data of the high-ranked links have homogeneous traffic conditions and a higher saturation level in the links, collectively forming a curve nearer to the upper bound in the MFD plane.

In the known research work, the data of the different time intervals were considered with an equal weight for each time interval. To evaluate the proposed method, we compared the proposed EW approach with the conventional method by selecting the sample links at different saturation levels using both the EW ranking and the equal weights approach. However, equal weighting considers time intervals with different traffic conditions similarly. Hence, the ranking created differs from the ranking developed using the EW. The difference between the two rankings was the consideration of the dynamic traffic conditions. This difference is also evident in the estimated MFD.

#### *2.1.1 E-TOPSIS:*

The E-TOPSIS is a productive method in decision-making (Behzadian, Otaghsara et al. 2012, Zavadskas, Mardani et al. 2016). The TOPSIS method requires weights for each system property. For weighing the properties of the system in the TOPSIS, the literature extensively uses EW from EWM. Although other methods are available, calculating EW is relatively straightforward and inexpensive (Chen 2021).

##### *2.1.1.1 Entropy Weight Method (EWM):*

The EWM was developed by Shannon (Shannon 1948). EWM is also a useful decision-making model (Zhang and Wang 2015, Wu, Xue et al. 2017). The advantage of EWM is the objectivity in calculating the weights (Ding, Chong et al. 2017). Entropy analyzes the dispersion in the information, and EWM gathers the data according to the dispersion by weighing the system's responses. EWM calculates the weights using the following equations

The application of the methodology requires a decision matrix for each city. The columns represent the time interval, and the row represents the loop detectors, as shown in Figure 1. The decision matrix shows the total flow of the road network over a day. In the next step, according to the EWM, Equation 1 calculates the probability of each flow value. Next, the entropy is determined by Equation 2. The entropy indicates the level of uncertainty in traffic flow data at each time interval. Later, Equation. 3 determines the divergence based on the entropy values. Finally, the entropy weights calculated using Equation 4 indicate the degree of dispersion in flow at each time interval of a day.

| $T_i$ \ $D_i$ | $T_1$ | $T_2$ | $T_3$ | $T_4$ | $T_5$ | $T_6$ | $T_7$ | …. | $T_t$ |
|---|---|---|---|---|---|---|---|---|---|
| $D_1$ | $F_{11}$ | $F_{12}$ | . | . | . | . | . | . | $F_{1t}$ |
| $D_2$ | $F_{21}$ | | | | | | | | . |
| $D_3$ | . | | | | | | | | . |
| . | . | | | | | | | | . |
| . | . | | | | | | | | . |
| . | . | | | | | | | | . |
| $D_i$ | $F_{i1}$ | $F_{i2}$ | . | . | . | . | . | . | $F_{it}$ |

Figure 1 Decision matrix, $D_i$ represents the detectors, $T_t$ represents the time intervals, and $F_{it}$ represents the corresponding flow.

$$\boldsymbol{Pr_{it}} = \frac{\boldsymbol{F_{it}}}{\sum_{i=1}^{m} \boldsymbol{F_{it}}} \quad ; \forall \boldsymbol{i}, \boldsymbol{t} \tag{1}$$

$$\boldsymbol{En_t} = -\boldsymbol{Y} \sum_{i=1}^{m} \boldsymbol{Pr_{it}} \, \boldsymbol{log_e}(\boldsymbol{Pr_{it}}) \tag{2}$$

$$where \ Y = \frac{1}{log_e(\mathrm{m})}$$

$$\boldsymbol{Div_t} = \boldsymbol{1} - \boldsymbol{En_t} \tag{3}$$

$$\boldsymbol{Ew_t} = \frac{\boldsymbol{Div_t}}{\sum_{t=1}^{n} \boldsymbol{Div_t}} \quad ; \forall \boldsymbol{t} \tag{4}$$

where m and n show the number of detectors and number of time intervals, respectively

### *2.1.1.2 TOPSIS:*

The TOPSIS method was developed by Hwang, Chen et al. (1992). It calculates the distance of a property of a system from the ideal and worst solutions and determines a $P_i$. The $P_i$ indicates how close the property of a system is from the ideal solution and how far from the worst solution at each interval. The links having higher saturation during the day have a higher probability of representing the traffic state of a road network. In this study, the capacity was assumed to be an ideal solution, and an empty link was considered as the worst/negative solution. The $P_i$ of each link is calculated using the distance between both solutions and used to develop the ranking of the detectors.

$$\boldsymbol{R_{it}} = \frac{\boldsymbol{F_{it}}}{\sqrt{\sum_{i=1}^{m} \boldsymbol{F_{it}^2}}} \quad ; \forall \boldsymbol{i}, \boldsymbol{t} \tag{5}$$

$$\boldsymbol{V_{it}} = \boldsymbol{Ew_t} * \boldsymbol{R_{it}} \quad ; \forall \boldsymbol{i}, \boldsymbol{t} \tag{6}$$

$$\boldsymbol{D_i^+} = \sqrt{\sum_{t=1}^{n} (\boldsymbol{V_{it}} - \boldsymbol{V_t^+})^2} \quad ; \forall \boldsymbol{i} \tag{7}$$

$$\boldsymbol{D_i^-} = \sqrt{\sum_{t=1}^{n} (\boldsymbol{V_{it}} - \boldsymbol{V_t^-})^2} \quad ; \forall \boldsymbol{i} \tag{8}$$

$$where \ \mathrm{V_t^+} = \max \mathrm{V_{it}} \ , \ \mathrm{V}_t^- = \min \mathrm{V_{it}} \ ; \forall t$$

$$\boldsymbol{P_i} = \frac{\boldsymbol{D_i^-}}{\boldsymbol{D_i^-} + \boldsymbol{D_i^+}} \quad ; \forall \boldsymbol{i} \tag{9}$$

where $V_t^+$, $V_t^-$, $D_i^+$, $D_i^-$ and $P_i$ show the ideal, worst solution, distances from the solutions, and performance score, respectively.

The TOPSIS method determines the $P_i$ using Equations 5-9. Equation 5 creates the normalized decision matrix $R_{it}$ using the traffic flow from the decision matrix in Figure 1. Equation 6 uses the EW of EWM and $R_{it}$ to develop a weighted normalized decision matrix. With Equations 7 and 8, the Euclidean distance of the weighted value from the ideal and worst solution is calculated at each time interval for each detector. The $P_i$ of each detector indicates nearness to the ideal solution, i.e., capacity. Ultimately, the $P_i$ represents the weighted saturation level over a day of each detector.

The E-TOPSIS method calculates the $P_i$ of each detector. Since the calculation of $P_i$ is based on the EW, the EW represents the heterogeneity in the road network at each time interval. The ranking developed using Equation 9 includes the effect of heterogeneity in calculating the $P_i$ of each link. Finally, the methodology uses the ranking to select the sample links, and the MFD is estimated using their loop detector data. Equation 10 uses the data of sample links for the MFD estimation.

$$\boldsymbol{q_{t,LLD}} = \frac{\sum_i^m q_{it} l_i}{\sum_{i=1}^m l_i}, \boldsymbol{o_{t,LLD}} = \frac{\sum_i^m o_{it} l_i}{\sum_{i=1}^m l_i} \quad (10)$$

The variable heterogeneity level of the road network makes the role of the time interval dynamic. Only a fraction of the total links represent the network; out of them, only a few are representative most of the time during the day. The ranking developed by the E-TOPSIS ranks the links according to their representation property over the course of the day.

### *2.2 Level of Heterogeneity:*

In the literature, the additional capacity obtained in the MFD estimated using the data of subset links (Ambühl, Loder et al. 2018) and variance in link densities (Ji and Geroliminis 2012) determined the level of heterogeneity. Mazloumian, Geroliminis et al. (2010) related the standard deviation of the network density and the number of congested links with the average flow of the road network. This study highlighted the difference in MFD estimated by sample links selected using the two different approaches. Therefore, the measure for calculating the heterogeneity level of the road network represents the additional capacity obtained in the MFD plane. Ambühl, Loder et al. (2018) also determined the heterogeneity level as an additional capacity for different sample sizes of links. The heterogeneity level specified in this study represents the additional capacity at different saturation levels. The former criteria of heterogeneity level only gave the information of heterogeneity at different sample sizes but could not specify which links were creating heterogeneity.

Identifying links that create heterogeneity is beneficial in reducing heterogeneity in the road network. The approach of this study by relating the heterogeneity level to the saturation level, highlighted the links responsible for the heterogeneity in the road network. Equation 9 calculates the $P_i$ of the links that indicate the saturation level of the links. The $P_i$ is used to determine the ranking of the links. (Ambühl, Loder et al. 2018) (Ambühl, Loder et al. 2018) (Ambühl, Loder et al. 2018) (Ambühl, Loder et al. 2018) (Ambühl, Loder et al. 2018) The equal weights and EW calculated different saturation levels of the links. Sample links were selected at different saturation levels i.e., 50%, 40%, 30%, 20%, and 10%. Similarly, the number of sample links selected using both methods differed for the respective saturation level bins. Therefore, the additional capacity obtained at each bin of saturation level highlighted the role of the respective bins in the heterogeneity of the road network.

The additional capacity was a percentage change in the capacities of MFDs, i.e., MFD estimated by sample links of equal weights and EW. The capacity of each MFD was the 85th percentile of the data to avoid noisy data points. Equation 11 calculated the percentage change in the capacities for each saturation level bin.

$$\boldsymbol{Additional\ Capacity} = \left(\frac{C_{EW} - C_{Equal}}{C_{EW}}\right) * \mathbf{100} \quad (11)$$

where $C_{EW}$ and $C_{Equal}$ are the capacities of the MFD estimated by sample links selected by EW and equal weights, respectively.

## 3 Macroscopic Traffic Data:

Macroscopic traffic parameters i.e., traffic flow and occupancy, are primary components of the methodology. First, the traffic flow was used to select the sample links. After selecting the sample, the occupancy of sample links was used to estimate the MFD. Loop detector data fulfilled the requirement of data in our case. Although FCD provides the required data, penetration rate could be another variable affecting the outcome. Data from loop detectors installed in the city center and its vicinity, excluding the ring roads, expressways, and motorways, were used. The data from two urban cities (Zurich and London) and a simulated urban network (Braunschweig) were used to apply the proposed methodology.

## 4 Cities of Zurich and London

Data from two different cities, Zurich, Switzerland, and London, UK, were used in this study (Loder Allister 2020).

Cities are different in demographics and traffic conditions. The population of London is much higher than that of Zurich and has an extensive transportation system. The loop detectors installed in both cities are primarily used for traffic control. The traffic data of loop detectors represent the traffic conditions of the CBD of Zurich and London, as shown in Figure 2. Table 1 gives information about the data used in the study.

Both cities are different in size. Therefore, for comparison and similarity, the study selected the detectors from the city center to apply the methodology. OpenStreetMap and spatial information of the loop detectors were used for the filtration of data. Geo-referenced the loop detectors' locations and removed the freeway, trunk, and residential roads. After the filtration, the city center of Zurich has 260 detectors, and London has 686 detectors.

TABLE 1 Data overview

| City | Zurich | London |
|---|---|---|
| Number of Detectors | 260 | 686 |
| Road Network Coverage (Km) | 56.68 | 114.88 |
| Aggregation Interval (Min) | 3 | 5 |
| Data Duration (Hr) | 19 | 19 |

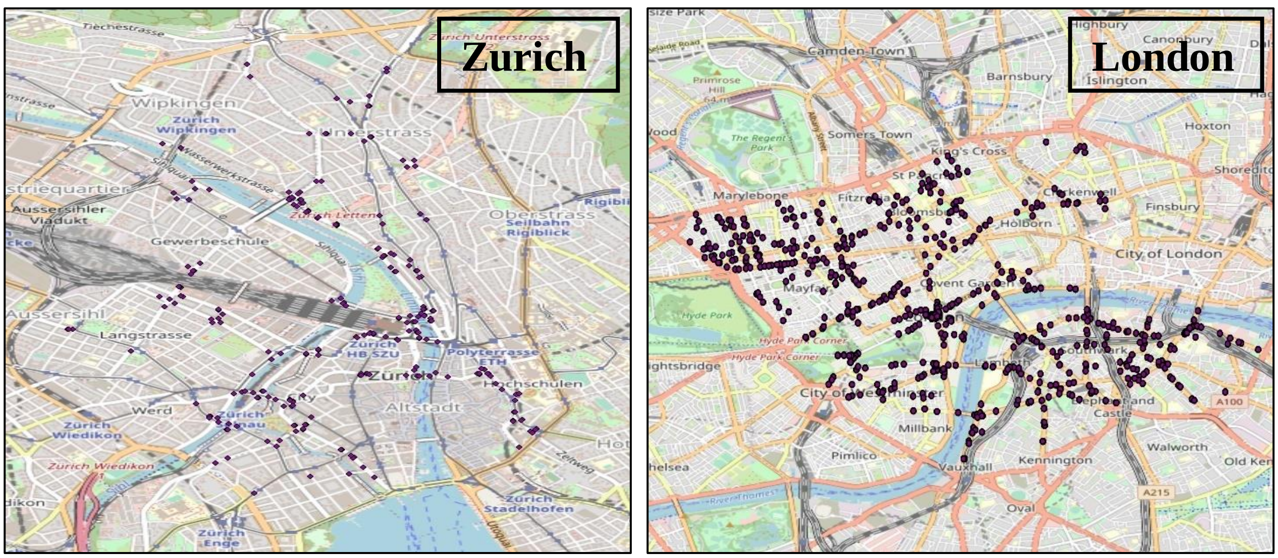


Figure 2 Bird eye view of city center of Zurich and London

### *4.1 Ranking of Links*

First, we calculated the weights using the traffic flow data of Zurich and London to determine the weighted saturation level. After forming the decision matrix for each city, Equation 4 calculated the EW for each city, as shown in Figure 3. The time aggregation interval of data is different for both cities, as shown in Table 1. Therefore, the weight profiles of both cities are plotted in the same figure but at the secondary axis. Figure 3 also includes the equal weight profile along with EW for comparison. The equal weights underestimate the starting time intervals, i.e., homogeneous traffic conditions, as shown in Figure 3. In contrast, in the peak hours when the traffic state of the road networks is heterogeneous, equal weights overestimate the traffic conditions in both cities.

In the next step to determine the weighted saturation level, there are two different weights for each city, so Equation 9 developed two different rankings of detectors for each city based on $P_i$. Figure 4 illustrates the $P_i$ of Zurich and London. The equal weights overestimate the traffic conditions at peak hours, resulting in a higher $P_i$ of the detectors in both cities. At the same time, the EW calculates a lower $P_i$. E-TOPSIS requires a capacity (ideal solution) and an empty link (worst solution) as a reference for calculating the saturation level of each detector. The capacity used was 1000 and 1400 veh/hr/ln for Zurich and London, respectively. The ranking of the detectors indicated that only very few links in the road network have higher saturation during the day. $P_i$ for the capacity flow equals one, and zero for an empty link. However, traffic conditions and link capacities are different in both cities.

Zurich has a higher difference between the Pi values compared to London. The higher difference shows that the traffic conditions have a higher variation during the day, and considering them equal results in higher scores than actual. The higher number of low saturation links in both cities indicates that few links have a high flow concentration. Unfortunately, the author has access to insufficient traffic data to analyze the cause of flow concentration in the study area. The $P_i$ in Figure 4 of both cities indicates that both road networks have very few links that are representative of the road network most of the time during the day. However, the results are relative to the number of detectors and flow in each link. Moreover, the city centers of Zurich and London have high accessibility, as shown in Figure 2. Therefore, a smaller number of representative links and any subset links selection without considering the road network's heterogeneity make them vulnerable to inhomogeneity.

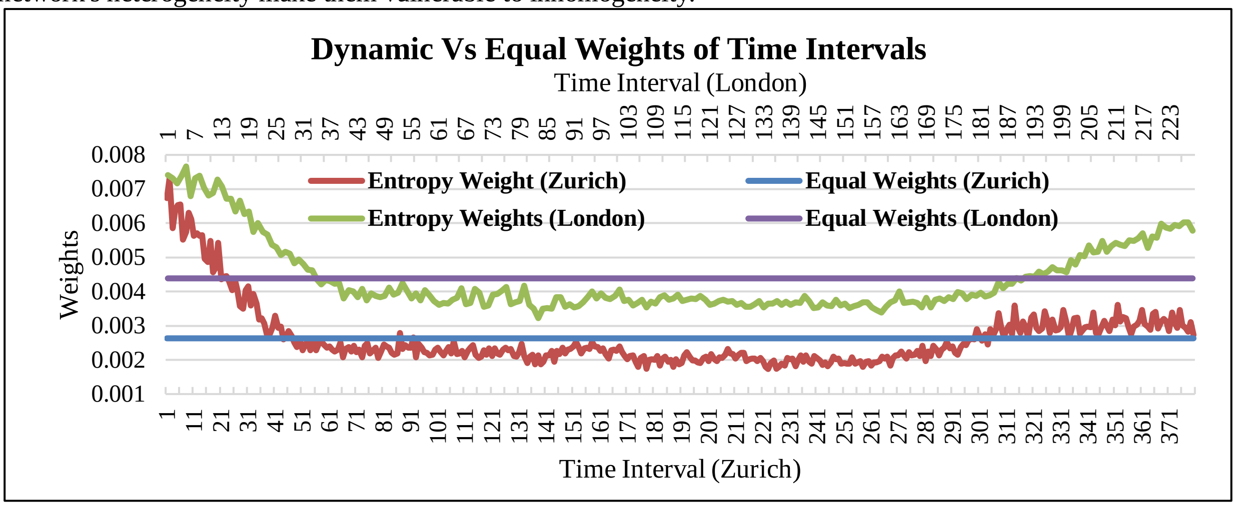


Figure 3 Weight profile of Zurich and London

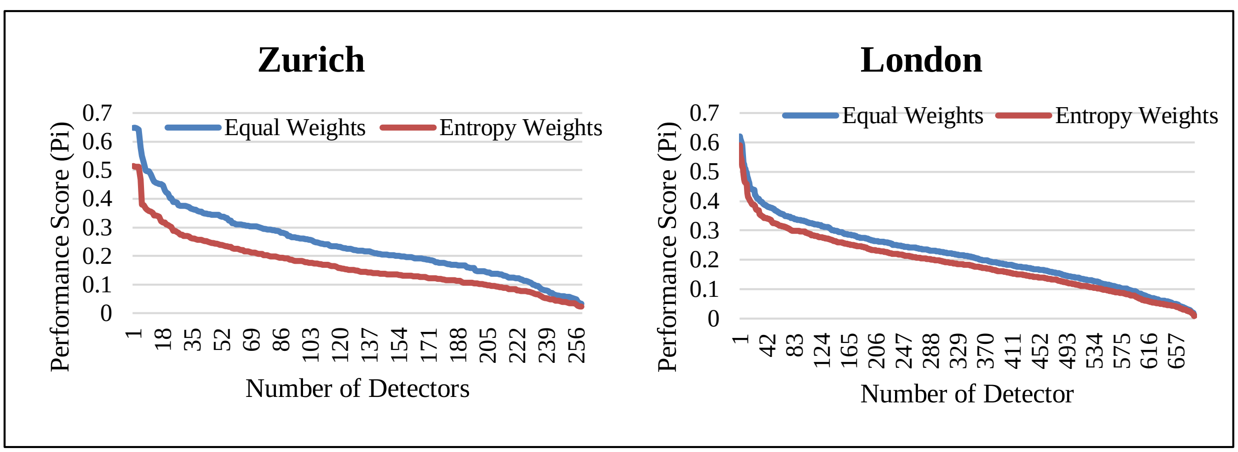


Figure 4 Performance Score of Zurich and London

### *4.2 MFD Estimation:*

The ranking developed by $P_i$ shows that only a few links are representative of the road network by both approaches (EW and equal weights). Therefore, we selected sample links based on saturation-level bins, as discussed in Section 2.2. The data from sample links estimated the MFD using Eq. 10. In Figure 5 MFD estimated by sample links having saturation level a) 50% (Zurich) b) 40% (Zurich) c) 50% (London) d) 40% (London). Both MFDs have adequate points in free flow and congested regions. A non-linear model was fitted on the traffic data to highlight the difference between both approaches. Additionally, the shape of the MFD gives information about the usability of the infrastructure and traffic control (Daganzo and Geroliminis 2008, Ambühl, Loder et al. 2020).

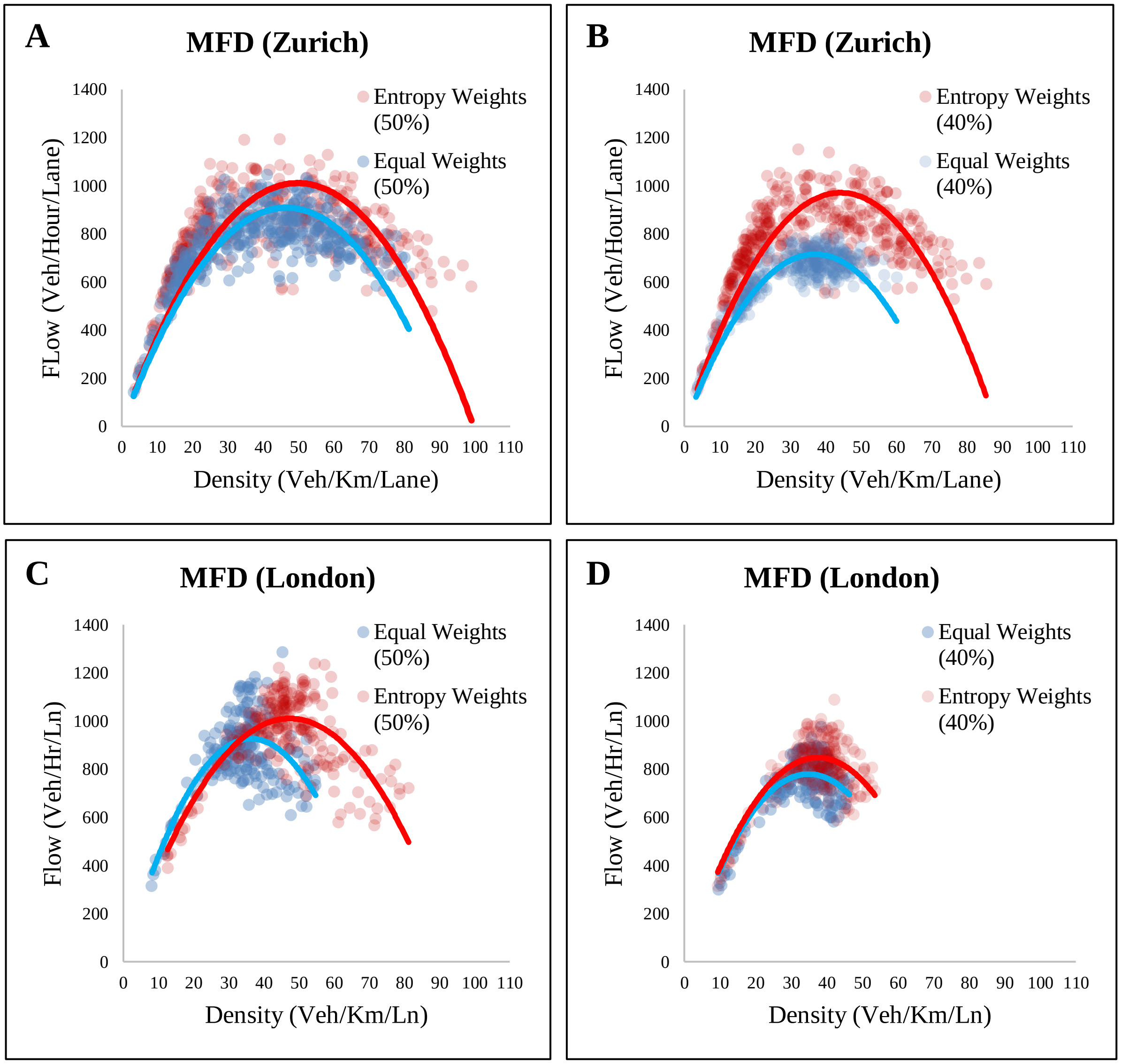


Figure 5 MFD estimated by sample links having saturation level a) 50% (Zurich) b) 40% (Zurich) c) 50% (London) d) 40% (London)

The ranking developed by EW is different from the ranking produced by equal weights. Similarly, the selected detectors differ for the MFD estimation, and the effect is evident in the MFD plane. In Section 4.1, we explained that the equal weights consider all the time intervals equal whether the traffic conditions are similar or not. The equal consideration of time intervals underestimates the homogeneous traffic conditions and overestimates the traffic conditions when the traffic states of the road network are heterogeneous in Figure 3. Both approaches select the detectors at the same saturation level. The equal weights calculate a higher saturation level of links that selects more links, and the estimated MFD lies below the MFD curve compared to the MFD estimated from the links selected using EW. The difference in the estimated MFDs indicates that EW considered the effect of heterogeneity in the selection of links, and the estimated MFD by EW is least affected by the heterogeneity of the road network. The EW estimated the MFD with more accuracy and considered the dynamic traffic conditions that are obvious in every road network, and no such road networks exist that have constant traffic conditions during the day.

The ranking is based on the saturation level of each link, using the data of top-ranked links indicating that they also have the highest flow compared to lower-ranked links. The saturation level check resulted in traffic data points in both regions of MFD. Including the effect of dynamic traffic conditions (heterogeneity), the selected detectors are minimally affected by the inhomogeneity of the road network. The data of any other link would lie below the curve, and the selected subset links would have homogeneous traffic conditions. Moreover, the saturation levels of links were calculated using the capacity values indicated in Section 3.1. We discussed the impact of change in capacity on the link selection and MFD estimation in Section 4.2.1. The estimated MFD by the proposed method lies nearer to the capacity value provided for each city, meaning that the curve is close to the upper bound and less affected by the heterogeneity of the road network.

The methodology is based on the traffic data and uses the data of a typical weekday. Including more data of different days could give more information about the traffic conditions. The density was calculated for estimating the MFD using the proportionality of occupancy and density. The occupancy data was converted into the density by a scalar conversion using mean vehicle length (Hall and Persaud 1989, Bickel, Chen et al. 2007). Mean vehicle length is the sum of the vehicle length and detector length. However, the variance in the traffic mix affects the proportionality of the occupancy and density. Therefore, the truck-to-car ratio relates to the deviation in the occupancy-density relation (Kim and Hall 2004). However, the study used the data collected from the CBD of the urban road network. Heterogeneity in the traffic stream of the CBD is minimum. Therefore, the mean vehicle length used was 6.3 m for Zurich (Ambühl, Loder et al. 2017) and 6.0 for London (Ambühl, Loder et al. 2018).

### *4.3 Sensitivity of Capacity*

In the previous section, the saturation level of links was calculated using the capacity value of 1000 and 1400 veh/hr/ln for Zurich and London, respectively. The capacity values were selected based on the flow data of loop detectors. However, the actual capacity values for both cities may differ. Therefore, the capacity was changed by +/- 20% to determine the impact of capacity variation on link selection and, ultimately on the MFD estimation. The saturation level of the links varies with the provided capacity value. Figure 6 shows link selection at 40% and 30% saturation because links were present at the stated saturation levels, even at a higher capacity. The 50% saturation level results were excluded because there were significantly fewer links or no links were present when capacity was increased by 20%; hence, the result could not be compared. However, the blue curve is an MFD estimated by links selected by the proposed method, and the green curve is calculated using the link selection by equal weights. The area above and below the respective curves is the variation in the MFD curve when capacity varies by +/- 20%.

The MFD curve of Zurich estimated by using EW was above the green MFD curve even when the capacity was increased for equal weights. Increasing the capacity led to the selection of a few links, as there were fewer links with higher saturation in the road network. But equal weights ignored the heterogeneity of the road network and selected a non-homogenous sample that resulted in inaccurate MFD. Similarly, decreasing the capacity by 20% increased the saturation level of the links; a higher number of links were selected. Nevertheless, the EW took into account the heterogeneity of the road network; the estimated MFD is above the green curve.

Figures 6c and 6d show the variation of the MFD curve with the change in capacity value in London. Compared to the MFDs of Zurich in Figures 6a and 6b, the difference between the curves is lower. The difference between the curves highlights the level of heterogeneity in the road network. Besides the difference between the curves, London follows the same trend as Zurich. The results in Figure 6 show that the proposed methodology of using EW in the link selection outperformed the equal weights approach independent of the capacity value provided. Moreover, if the capacity value used in the analysis deviates from the actual capacity of the links, the applicability of the method is not affected, and a homogeneous sample is selected.

*4.4 Level of Heterogeneity:*

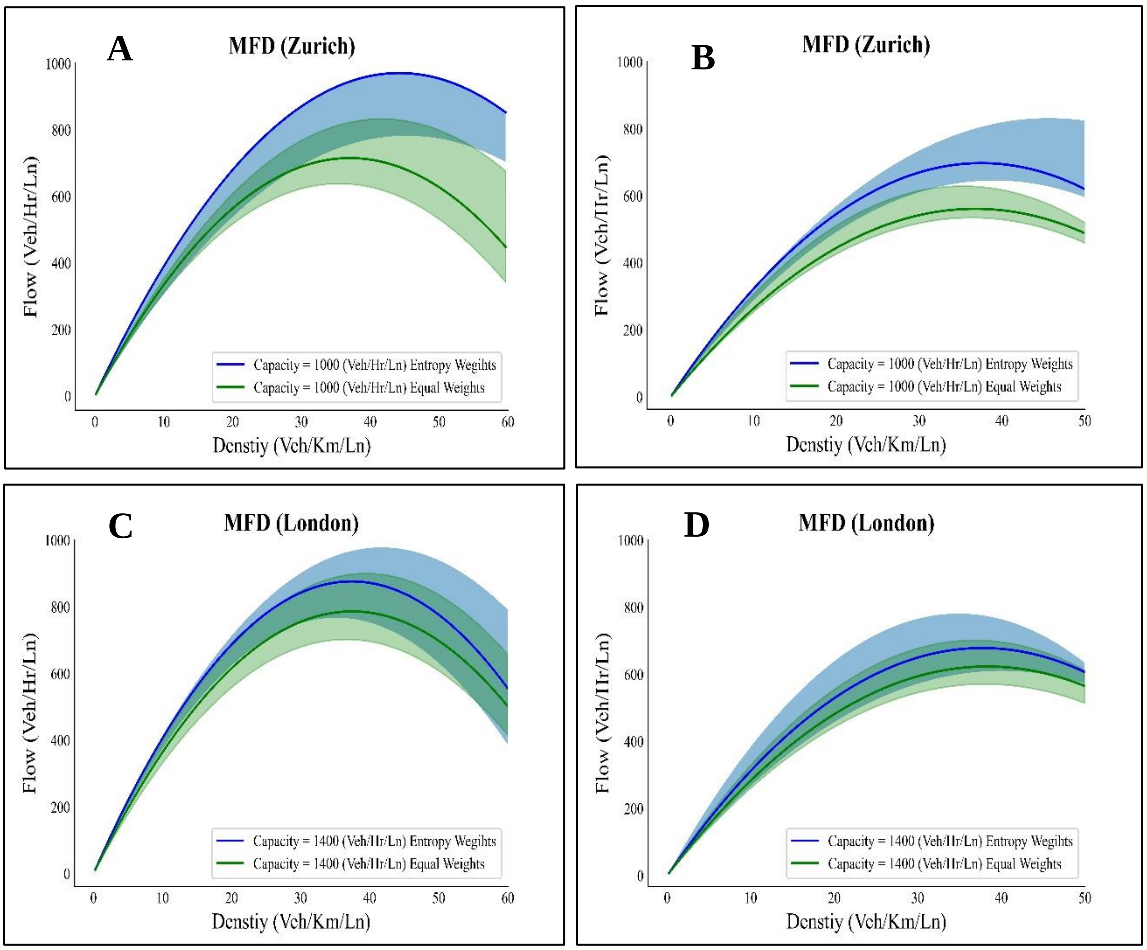


Figure 6 MFD estimated by sample links using different capacity values at saturation level a) 40% (Zurich) b) 30% (Zurich) c) 40% (London) d) 30% (London)

The study specifies that selecting the links without considering the dynamic traffic conditions is affected by the heterogeneity of the road network. As a result, the estimated MFD lies below the upper bound, as shown in Figure 5 MFD estimated by sample links having saturation level a) 50% (Zurich) b) 40% (Zurich) c) 50% (London) d) 40% (London). The EW includes the effect of heterogeneity in the link selection, and the estimated MFD is closer to the upper bound. The difference between the capacities of the MFDs is called the additional capacity obtained in a homogeneous road network (Ambühl, Loder et al. 2018). Since the MFD estimated by EW is least affected by the heterogeneity, the difference from the MFD estimated using equal weights is the level of heterogeneity of the road network.

The additional capacity is determined using Equation 11 for each saturation level. Figure 7 illustrates the additional capacity as a level of heterogeneity of Zurich and London. The road network of Zurich is more heterogeneous than London; it covers more area under the curve. Both road networks have few links with a saturation level of more than 50%; they have smaller and similar differences. Similarly, including the links of saturation level 10% or higher would obtain only 7.1% and 3.6% of additional capacity, respectively. The low saturation links have lower flow during the day, affecting the MFD least.

The links having saturation levels between 20-40% are more responsible for the heterogeneity in the road network, as shown in Figure 7. These links are higher in number than other links, as shown in Figure 4. The additional capacity obtained in Zurich is much higher than in London. Therefore, the sample link selection in Zurich is more prone to heterogeneity. Whereas the overall heterogeneity level in London is lower than in Zurich, and links with a saturation level of at least 30% are more responsible for the heterogeneity in the road network of London. When the traffic conditions of high and low-ranked links behave similarly, the additional demand in the congested links shifts towards the low-saturated links, and the road network attracts more demand. As a result, the road network would have a higher average flow in homogenous traffic conditions (Knoop, van Erp et al. 2018).

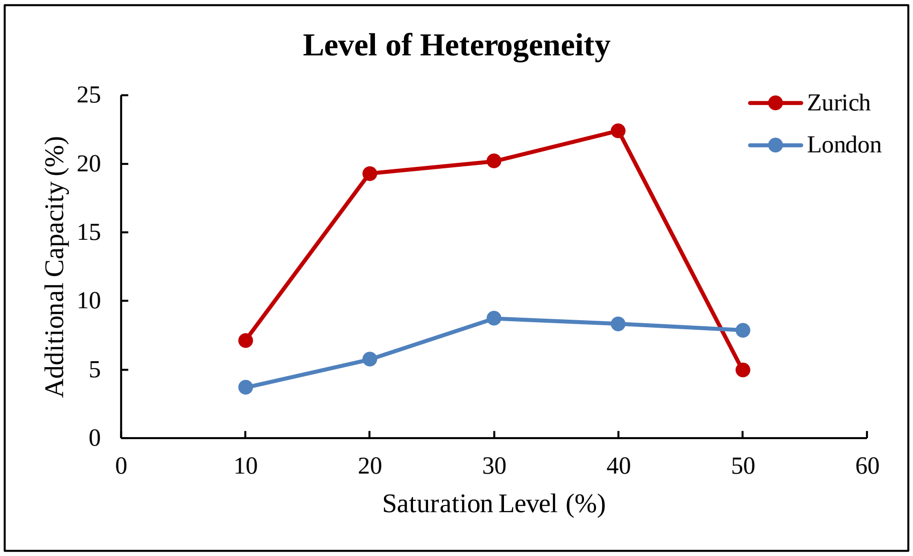


Figure 7 Level of heterogeneity in Zurich and London at different saturation levels

The transferability of the method to other road networks is still a major limitation of research studies in traffic engineering. The studies of link selection used different methods to select the links but were unable to characterize the selected links. Characterizing the links generalizes the method and could be applied to other road networks. Figure 7 highlights the role of links having different saturation levels in the heterogeneity of the road network. The saturation level characterizes the links and highlights the links which are primarily responsible for the heterogeneity in the urban road network. Relating the saturation level to the heterogeneity not only increases the transferability potential of the study to other road networks but also helps in maintaining the homogeneity in the road network.

## 5 Network Simulation

### *5.1 City of Braunschweig*

The road network of Braunschweig, Germany, is shown in Figure 8. The road network (80 $Km^2$) was simulated in SUMO. The road network comprises motorway, primary, secondary, and residential roads. Loop detectors were placed near signalized intersections at primary and secondary roads. Cycle time at each intersection was set to 90s. 33s green time for through traffic, 6s green time for turning vehicles, and 3s lost time for each phase. The demand for the model was taken from Armellini, Banse Bueno et al. (2021). The study formulated the traffic demand model in SUMO for Braunschweig, Germany, using Travel Activity PAttern Simulation (TAPAS) (Hertkorn 2005, Heinrichs, Krajzewicz et al. 2016). The demand model was calibrated using traffic data from 129 counters installed on different road types of Braunschweig.

Figure 8 Braunschweig, Germany layout used in SUMO

### *5.2 MFD Estimation by Sample Links*

The previous sections applied the method to the loop detector data of urban cities. Similarly, the method was applied to the simulated traffic data to highlight the transferability of the method to other road networks. The weighted saturation level of the links was calculated using Eqs 1-9. The pattern of the weight profile in Figure 9 is similar to Zurich and London; the homogeneous and heterogeneous traffic conditions are underestimated and overestimated by the equal weight approach, respectively.

The $P_i$ in Figure 10 was calculated using EW and equal weights. The difference is quite evident between the scores, and very few links have higher scores. Selection of links without consideration of the dynamic role of time intervals will affect the sample representativeness and accuracy of estimated MFD. However compared to both cities, the links in the simulated network have lower scores. Since the score was calculated from provided capacity value, the links' capacity depended on several factors that were not investigated and were out of the scope of the study.

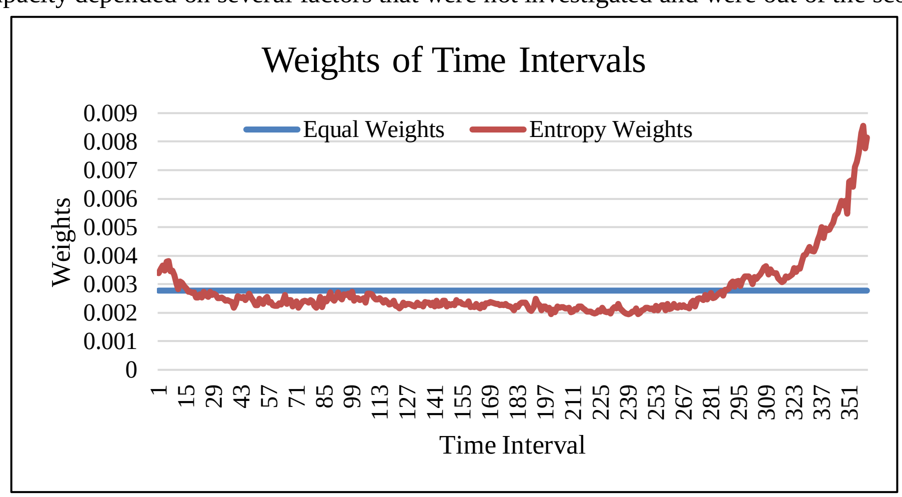


Figure 9 Weight profile of Braunschweig

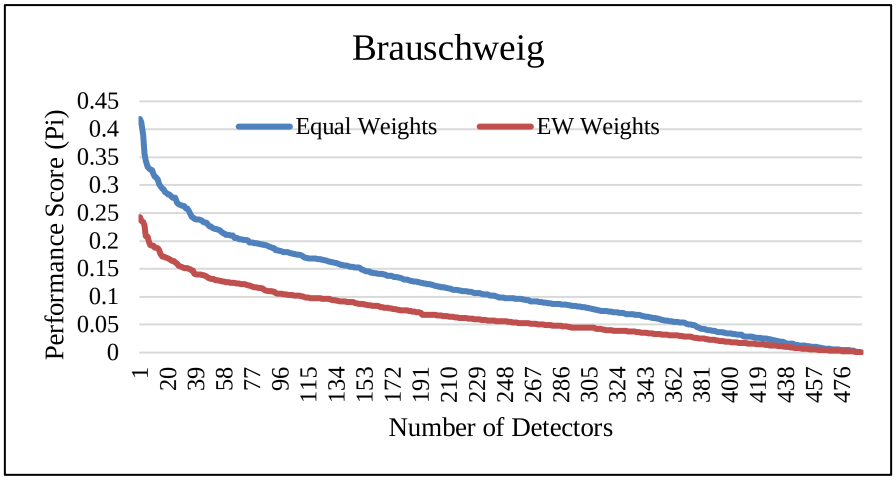


Figure 10 Performance Score (Pi) of links in Braunschweig

The $P_i$ of the links in Figure 10 indicates the presence of high saturation links is low, and selecting a higher value will select a few links that are not enough for the MFD estimation of a road network. Therefore we set 20% $P_i$ value for the sample since it will select enough links, irrespective of both approaches for the MFD estimation of Braunschweig. Figure 11 illustrates the MFD estimated by sample links having a saturation level of 20%. The difference in the MFD curve is as expected and similar to previous results. The equal weights approach selects more links than the proposed method on the same saturation level. The MFD curve of the sample links selected by the proposed method lies above the MFD of equal weights, which shows that consideration of dynamic time interval decreases the effect of heterogeneity on the MFD curve and increases the representativeness of the selected sample links.

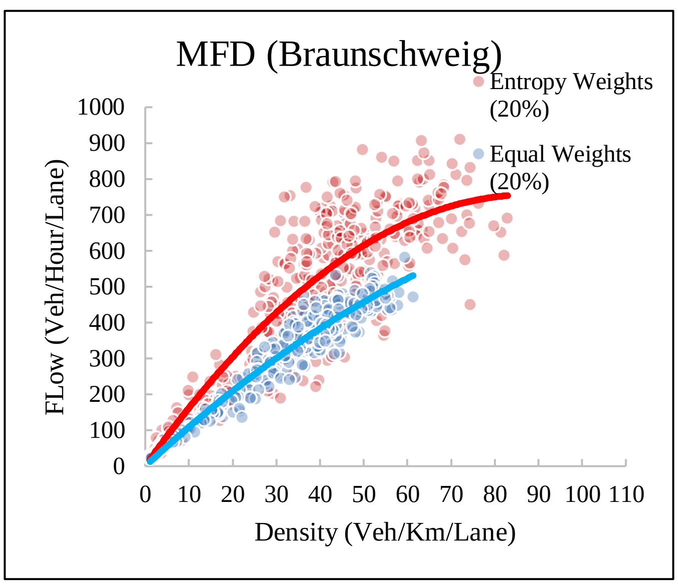


Figure 11 MFD estimated by the sample links having saturation level of 20%

The sample links were selected using the capacity value 700 (Veh/Hr/Lane). However, the actual capacity may differ, and the capacity used was taken by observing the flow data. For determining the effect of change in capacity on the MFD estimation, we changed the capacity value by +/- 20% and estimated the MFD by sample links selected by both approaches. Figure 12 shows the MFD curves estimated by 20% saturated links at different capacity values. The MFD curve estimated by the proposed method lies above the equal weight MFD even at a lower capacity. The MFD of equal weights also remains beneath the blue curve after increasing the capacity. The effect of variation in

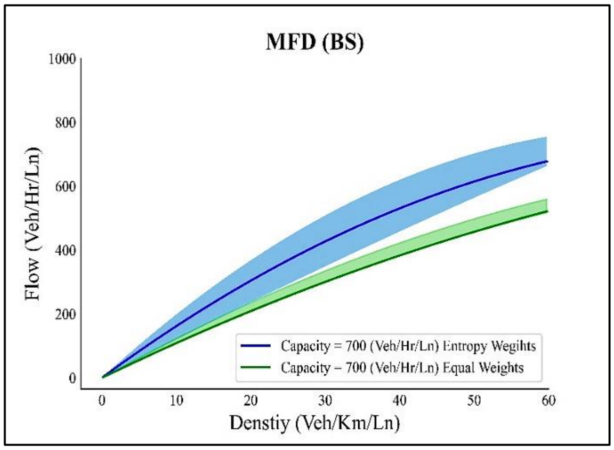


Figure 12 MFD estimated at different capacity values by 20% saturated links

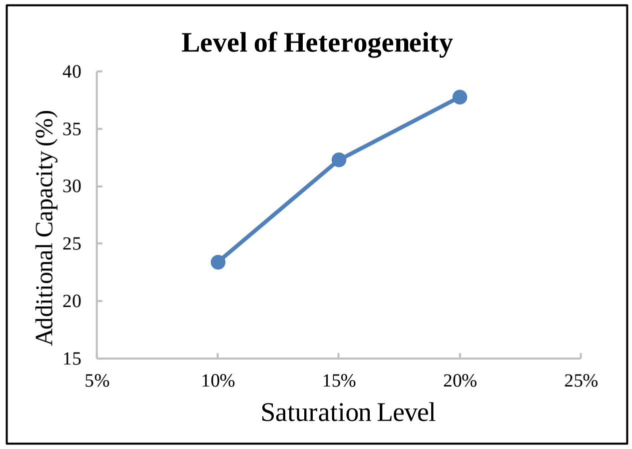


Figure 13 Level of heterogeneity in Braunschweig

capacity on the MFD estimation proves that the proposed method estimates accurate MFD compared to equal weights, and a change in capacity would not affect the accuracy of the MFD estimation.

Based on the flow data, the saturation level calculated is lower compared to previous results. Therefore, the additional capacity was calculated at lower saturation levels. However, the increase in additional capacity in Figure 13 with the increase in saturation level is similar to Zurich and London; less additional capacity is obtained at low saturation levels, and the highest value is observed between 30% and 40% saturated links. Since the maximum saturation attained in the simulation was 20%, a comparison between the case studies at a higher saturation level is not possible.

The results obtained by applying the method in a simulation environment are similar to real traffic data of urban cities. The similarity in results indicates that the proposed methodology of the study can be applied to any urban road network. Moreover, the size of the road network in the simulation is comparatively bigger than the area taken for Zurich and London. The number of detectors placed was higher and the capacity value used in the analysis was substantially lower compared to both cities. Though the road network of both cities and the simulation network are quite different but the results are similar. This implies that the proposed method is independent of the size of the road network, the number of detectors placed, and the capacity of the links. Therefore, the transferability of the method to other road networks increases the applicability of the study.

## 6 Conclusion:

MFD estimation requires extensive traffic data of the urban road network. The traffic data should highlight the actual network performance on a given demand. Mainly the data collected from the subset links. The dynamic traffic conditions in the road network affect the representation by the subset links and the accuracy of MFD estimation. Inaccurate estimates limit the parameterization of MFD. The higher demand at peak hours increases the chances of congestion, and the inhomogeneous distribution of congestion increases the heterogeneity of the road network. The susceptibility of a subset of links to inhomogeneity decreases by including the effect of heterogeneity in the selection of the links. In this study, the heterogeneity is determined at each time interval by calculating the dispersion in flow data and including the effect by weighing the time intervals using EWM. Weighing the data maintained the homogeneity between the subset of links at the time interval level. Maintaining homogeneity at a lower level implies that the subset of links has minimum susceptibility to inhomogeneity.

The proposed method uses the E-TOPSIS method to select a subset of links. The method requires the weights that EWM calculates. Once the EWM maintains homogeneity in the subset links, the TOPSIS method analyzes the saturation level of links. The weighted saturation level calculates the $P_i$ and develops the ranking simultaneously. Estimated MFD by the subset of links is least affected by the inhomogeneity when the curve lies closer to the upper bound, and the selection of highly saturated links leads to capturing the free flow and congested traffic conditions. Therefore, this study uses both approaches (EW and equal weights) to select the sample links on different saturation levels to determine the effect of heterogeneity on the estimated MFD. The literature ignores the role of the time interval, which means that all the time intervals are equal in the analysis. But in reality, the traffic conditions are dynamic, implying that each time interval's contribution is different. Comparing the proposed methodology with the equal role of time intervals reveals the better estimation capability of our approach. However, considering the equal role of time intervals could be detrimental to traffic planning and management.

The MFD estimated from the sample links of EW lies above the MFD estimated by the links selected by equal weights. Since the former MFD is less affected by heterogeneity, the additional capacity is the difference between the capacities of the two MFDs, and the percentage change is the level of heterogeneity of the road network. The MFD is estimated by selecting the sample links at different saturation levels. Earlier, the additional capacity was determined based on the sample size. The sample size gives only the number of links related to heterogeneity level but does not give any characteristics of links in the sample. This study addresses the limitation of the previous measure by relating the level of heterogeneity to the saturation level. It will identify the links that are primarily involved in the heterogeneity of the road network.

The methodology is applied to the simulated road network of Braunschweig and loop detector data of Zurich and London placed in CBD, and the selected sample represents the CBD area. The traffic conditions are different in both case studies, but based on the results, we have found that the different road networks give similar results. Therefore, the proposed method of the study can be applied to any urban road network. The study uses the loop detector data for the MFD estimation. The occupancy data is directly used, although it may have a bias due to queue formation at the signalized intersection. For this reason, the study did not report any value of the estimated traffic flow parameters but highlighted the links with a higher tendency to represent the road network. Accurate density estimation on these links from other reliable sources could be helpful in the accurate parameterization of MFD.

This study highlights the importance of consideration of dynamic traffic conditions. Including dynamic traffic behavior instead of a constant role makes the sample representative of the road network and gives a homogeneous sample of links. The planners could employ the data collection resources on selected sample links to collect reliable speed and density data for estimating accurate traffic flow parameters using data fusion techniques. Moreover, this is the first study to include the effect of heterogeneity in link selection using traffic flow data only. The flow data of loop detectors free it from any bias that can incur inaccuracy in the estimation. Measuring the heterogeneity at different saturation levels can help to maintain homogeneity in the road network by traffic managers and planners. This study calculates the entropy-weighted saturation level based on the constant capacity of the links due to the availability of the data. However, the link capacity varies with time. Using the variable capacity in determining the entropy-weighted saturation level could give more interesting results and could be the future extension of this study.

**Acknowledgment:**

The authors are pleased to thank Dr. Allister Loder and ETH Zurich for providing the loop detector data (UTD19) used in the study for the application of methodology.

**References**

Aboudolas, K. and N. Geroliminis (2013). "Perimeter and boundary flow control in multi-reservoir heterogeneous networks." Transportation Research Part B: Methodological **55**: 265-281.

Ambühl, L., et al. (2018). "Introducing a re-sampling methodology for the estimation of empirical macroscopic fundamental diagrams." Transportation Research Record **2672**(20): 239-248.

Ambühl, L., et al. (2020). "A functional form with a physical meaning for the macroscopic fundamental diagram." Transportation Research Part B: Methodological **137**: 119-132.

Ambühl, L., et al. (2017). Empirical macroscopic fundamental diagrams: New insights from loop detector and floating car data. TRB 96th Annual Meeting Compendium of Papers, Transportation Research Board.

Armellini, M. G., et al. (2021). Brunswick simulation scenario for virtual-stops based DRT services with SUMO. Proceedings of the 10th International Congress on Transportation Research.

Behzadian, M., et al. (2012). "A state-of the-art survey of TOPSIS applications." Expert Systems with applications **39**(17): 13051-13069.

Bickel, P. J., et al. (2007). "Measuring traffic." Statistical Science: 581-597.

Chen, P. (2021). "Effects of the entropy weight on TOPSIS." Expert Systems with applications **168**: 114186.

Daganzo, C. F. (2005). "A variational formulation of kinematic waves: basic theory and complex boundary conditions." Transportation Research Part B: Methodological **39**(2): 187-196.

Daganzo, C. F., et al. (2011). "Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability." Transportation Research Part B: Methodological **45**(1): 278-288.

Daganzo, C. F. and N. Geroliminis (2008). "An analytical approximation for the macroscopic fundamental diagram of urban traffic." Transportation Research Part B: Methodological **42**(9): 771-781.

Daganzo, C. F. and L. J. Lehe (2016). "Traffic flow on signalized streets." Transportation Research Part B: Methodological **90**: 56-69.

Ding, X., et al. (2017). "Fuzzy comprehensive assessment method based on the entropy weight method and its application in the water environmental safety evaluation of the Heshangshan drinking water source area, three gorges reservoir area, China." Water **9**(5): 329.

Gayah, V. V. and V. V. Dixit (2013). "Using mobile probe data and the macroscopic fundamental diagram to estimate network densities: Tests using microsimulation." Transportation Research Record **2390**(1): 76-86.

Geroliminis, N. and C. F. Daganzo (2008). "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings." Transportation Research Part B: Methodological **42**(9): 759-770.

Geroliminis, N. and J. Sun (2011). "Properties of a well-defined macroscopic fundamental diagram for urban traffic." Transportation Research Part B: Methodological **45**(3): 605-617.

Hall, F. L. and B. N. Persaud (1989). "Evaluation of speed estimates made with single-detector data from freeway traffic management systems." Transportation Research Record(1232).

Heinrichs, M., et al. (2016). "Disaggregated car fleets in microscopic travel demand modelling." Procedia Computer Science **83**: 155-162.

Hertkorn, G. (2005). Mikroskopische Modellierung von zeitabhängiger Verkehrsnachfrage und von Verkehrsflußmustern, Deutsches Zentrum für Luft-und Raumfahrt, Forschungsbericht 2004-29.

Hwang, F., et al. (1992). Fuzzy multiple attribute decision making: Methods and applications, Springer Berlin/Heidelberg.

Ji, Y. and N. Geroliminis (2012). "On the spatial partitioning of urban transportation networks." Transportation Research Part B: Methodological **46**(10): 1639-1656.

Ji, Y., et al. (2018). "Determining the macroscopic fundamental diagram from mixed and partial traffic data." Promet-Traffic&Transportation **30**(3): 267-279.

Keyvan-Ekbatani, M., et al. (2013). "Urban congestion gating control based on reduced operational network fundamental diagrams." Transportation Research Part C: Emerging Technologies **33**: 74-87.

Kim, Y. and F. L. Hall (2004). "Relationships between occupancy and density reflecting average vehicle lengths." Transportation Research Record **1883**(1): 85-93.

Knoop, V. L., et al. (2018). Empirical mfds using google traffic data. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), IEEE.

Leclercq, L. and N. Geroliminis (2013). "Estimating MFDs in simple networks with route choice." Procedia-Social and Behavioral Sciences **80**: 99-118.

Loder Allister, A. L., Menendez Monica , Axhausen Kay W. (2020, 20 September 2020). "Utd19: Understanding traffic capacity of Urban Networks." Retrieved 15 January 2022, 2022, from http://hdl.handle.net/20.500.11850/437802.

Mahmassani, H. S., et al. (2013). "Urban network gridlock: Theory, characteristics, and dynamics." Procedia-Social and Behavioral Sciences **80**: 79-98.

Mazloumian, A., et al. (2010). "The spatial variability of vehicle densities as determinant of urban network capacity." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **368**(1928): 4627-4647.

Ortigosa, J., et al. (2015). "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**(3): 245-266.

Ramezani, M., et al. (2015). "Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control." Transportation Research Part B: Methodological **74**: 1-19.

Saffari, E., et al. (2020). "A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks." Transportation Research Part C: Emerging Technologies **119**: 102743.

Shannon, C. E. (1948). "A mathematical theory of communication." The Bell system technical journal **27**(3): 379-423.

Tsubota, T., et al. (2014). "Macroscopic fundamental diagram for Brisbane, Australia: empirical findings on network partitioning and incident detection." Transportation Research Record **2421**(1): 12-21.

Wu, J., et al. (2017). "Lake water quality assessment: a case study of Shahu Lake in the semiarid loess area of northwest China." Environmental Earth Sciences **76**(5): 232.

Xie, F. and D. Levinson (2007). "Measuring the structure of road networks." Geographical analysis **39**(3): 336-356.

Zavadskas, E. K., et al. (2016). "Development of TOPSIS method to solve complicated decision-making problems—An overview on developments from 2000 to 2015." International Journal of Information Technology & Decision Making **15**(03): 645-682.

Zhang, J.-Y. and L.-C. Wang (2015). "Assessment of water resource security in Chongqing City of China: what has been done and what remains to be done?" Natural Hazards **75**(3): 2751-2772.

Zhang, J., et al. (2022). "Analysis of cooperative driving strategies at road network level with macroscopic fundamental diagram." Transportation Research Part C: Emerging Technologies **135**: 103503.

Zhang, L., et al. (2020). "Recent developments in traffic flow modeling using macroscopic fundamental diagram." Transport reviews **40**(4): 529-550.

Zockaie, A., et al. (2018). "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**: 245-262.