**RESEARCH ARTICLE**

# Context-specific volume–delay curves by combining crowd-sourced traffic data with automated traffic counters: A case study for London

Gerard Casey[1], Bingyu Zhao[2,*] , Krishna Kumar[3] and Kenichi Soga[2]

[1]Arup, London, United Kingdom
[2]Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA
[3]Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, Texas, USA
*Corresponding author. E-mail: bz247@berkeley.edu

## Abstract

Traffic congestion across the world has reached chronic levels. Despite many technological disruptions, one of the most fundamental and widely used functions within traffic modeling, the volume–delay function has seen little in the way of change since it was developed in the 1960s. Traditionally macroscopic methods have been employed to relate traffic volume to vehicular journey time. The general nature of these functions enables their ease of use and gives widespread applicability. However, they lack the ability to consider individual road characteristics (i.e., geometry, presence of traffic furniture, road quality, and surrounding environment). This research investigates the feasibility to reconstruct the model using two different data sources, namely the traffic speed from Google Maps' Directions Application Programming Interface (API) and traffic volume data from automated traffic counters (ATC). Google's traffic speed data are crowd-sourced from the smartphone Global Positioning System (GPS) of road users, able to reflect real-time, context-specific traffic condition of a road. On the other hand, the ATCs enable the harvesting of the vehicle volume data over equally fine temporal resolutions (hourly or less). By combining them for different road types in London, new context-specific volume–delay functions can be generated. This method shows promise in selected locations with the generation of robust functions. In other locations, it highlights the need to better understand other influencing factors, such as the presence of on-road parking or weather events.

---

### Impact Statement

Volume–delay curves are widely used in traffic analysis. They form a critical part of the traffic assignment stage in the four-step modeling approach. We propose a novel data-driven approach to accurately represent the traffic behavior under congestion using emerging data sources for a better understanding of real-world context-specific behavior.

---

## 1. Introduction

When the vehicular demand for a road exceeds its free-flow threshold, a journey time-delay is incurred, resulting in congestion. Historically in traffic studies and planning, the relationship between traffic volume and time-delay has been simplified to macroscopic principles ignoring the behavior of individual

vehicles (van Wageningen-Kessels et al., 2015). The most widely adopted functional representation of such a macroscopic volume–delay relationship is the Bureau of Public Roads (BPR) relationship proposed in the 1960s (Bureau of Public Roads, 1964). Since then, local traffic agencies worldwide have calibrated the BPR curve coefficients to suit the local needs, such as in Suh et al. (1990), Kurth et al. (1996), Irawan et al. (2010), Mtoi and Moses (2014), and Kucharski and Drabicki (2017). However, the calibration of BPR coefficients to local roads requires an extensive set of volume and delay data, and gathering such data (e.g., through surveys) is often a challenge. Hence, traffic engineers often end up using a generic set of coefficients to cover a wide range of roads (e.g., one set of coefficients for all roads in the city, without considering the presence of road furniture or on-street parking that could affect the road capacity) and do not updated these coefficients to the changing demands of the city. Nevertheless, real-world conditions exhibit an enormous variability in the volume–delay characteristics, which is nearly impossible to account for entirely and certainly insufficient when representing this volume–delay relationship using a generic and static function for the entire city. Therefore, it is critical to understand the errors in volume–delay prediction based on real-world data.

This research aims to provide an efficient data-driven approach to calibrate the volume–delay curves by incorporating emerging data sources, mainly crowd-sourced travel time information from location and routing service providers, such as Google Maps (Google Inc., 2020). This type of calibration has never been done before to our knowledge. Specifically, this paper investigates the integration of two disparate data sources for this purpose: (1) the traffic speed data from Google Maps, and (2) the traffic count data from the automated traffic counter (ATC) system in Greater London. The reason to combine the two data sources is two-fold: (1) first of all, many traffic detectors in cities are single loop detectors, which do not offer accurate speed/delay measurement (Wang and Nihan, 2003); (2) secondly, even if there are speed sensors (e.g., double loop detectors or radars), they often only measure point speeds, which do not reflect the delay experienced on the road network, especially in urban settings (Wang and Liu, 2005; Zhang et al., 2015). In this study, we assess the feasibility and Google Map of obtaining site-specific volume–delay information based on these new data sources.

The paper is structured in the following way. Section 2 discusses relevant literature, including various functional forms of the volume–delay relationships, as well as some applications of such relationships in transport research and planning. Next, in Sections 3 and 4, we introduce the two data sources, namely the hourly traffic volume from the ATCs and the traffic speed from Google Maps. These data are used for estimating the road characteristics (free-flow travel time and capacity) and calibrating the volume–delay curve coefficients. Sections 5 and 6 explain the process of data cleaning and model building. In particular, three models are tested, including one base model with default volume–delay parameter values and two data-informed models with partial or full set of parameters calibrated from the real-world data. The three models are compared in terms of their performances in quantifying the level of delays at different traffic congestion levels. Section 7 offers an extensive discussion on the potential factors that could cause the variability observed in the volume–delay relationship, as well as the future prospects to adopt the proposed method at larger scales. Using fine-resolution data in the form of ATCs and aggregated device-based location-informed journey times on a range of roads, we demonstrate the capability of the new data-driven approach for efficiently capturing the volume–delay characteristics of roads in selected roads in Greater London.

## 2. Literature Review

Volume–delay functions, as the name indicates, relate two fundamental traffic parameters using nonlinear mathematical expressions. The independent parameter, volume, expresses the level of traffic demand and the dependent parameter, delay, indicates the deterioration in the traffic speed as the demand increases. The actual trend between volume and delay is a property of the road and is related to factors such as the speed limit, width, geometry, and the presence of road furniture. Even though the volume–delay relationships are highly context-specific, there exist some widely accepted functional forms to model them. The most widely used function is the BPR curve, which was developed in the 1950s for uncongested

freeways in the USA. Its simple mathematical form and minimal input requirements are attributed to its widespread adoption (Skabardonis and Dowling, 1997). The travel time $t$ on a road link computed using the BPR function has the form:

$$t = t_0 \times \left(1 + \alpha \left(\frac{v}{v_c}\right)^{\beta}\right), \tag{1}$$

where $t_0$ is the time required to traverse the road link at free-flow speed; $v_c$ is the road capacity (vehicles per unit time); $\alpha$ and $\beta$ are calibration coefficients; $v$ is the traffic volume to be modeled. The function is sometimes also expressed in terms of the vehicle speed $u$, which can be obtained from $t$ and the road length $l$.

Since the BPR function was created by fitting a polynomial equation to uncongested freeway data from the 1950s in the USA, it does not reflect the current operating conditions of the road network (Skabardonis and Dowling, 1997). As a result, many different organizations have adapted the BPR curve with various local empirical and simulated data to suit the local road conditions better (Irawan et al., 2010; Mtoi and Moses, 2014). The work of Kurth et al. (1996) focuses on obtaining refined, free-flow time and capacity ($t_0$ and $v_c$) for each road based on guidelines from the Highway Capacity Manual (HCM, 1994 version), rather than getting $t_0$ and $v_c$ from traditional lookup tables with only a few categories. Their approach is found to produce more accurate traffic speed estimations. However, it requires time-consuming identification of road characteristics (e.g., road grade, vehicle mix, and land use) and is still limited by the available adjustment factors that the HCM procedure can take into account.

Kucharski and Drabicki (2017) estimated both the calibration coefficients $\alpha$, $\beta$ and the road characteristics $t_0$, $v_c$ together based on loop detector data. They suggested transforming the volume–delay relationship to speed–density relationship for regression, as the latter remains monotonic in congested cases. However, the analysis lacks rigor due to employing linear $R$-squared metric to quantify errors in a nonlinear model and inconsistency in calculating the density. In the case of London, Transport for London (TfL) has calibrated the BPR function using observed traffic counts and defined $\alpha = 1.0$ and $\beta = 2.0$ for the area (TfL, 2010). In general, these calibrated volume–delay functions fit the local observations better.

Apart from the BPR function, several volume–delay relationships have been proposed over the years, as summarized in Mtoi and Moses (2014). Davidson (1966) proposed a general-purpose travel-time formula in 1966 and this method has undergone numerous modifications since it was first proposed (Mtoi and Moses, 2014). It has exhibited a closer match to actual volume counts and has a more robust theoretical base than the BPR (Rose et al., 1989). Among the modifications of the Davidson function (Akçelik, 1991; Tisato, 1991), the Akçelik function is the most widely used. The Akçelik method is a time-dependent modification of the Davidson model, which uses the coordinate transformation technique in an attempt to overcome the conceptual and calibration issues with the Davidson method (Akçelik, 1991). This function showed good results for certain road types, tolls roads, and signalized arterials (Mtoi and Moses, 2014). Spiess (1990) proposed the conical method, which attempts to overcome some of the limitations of BPR at both the upper and lower bounds by employing hyperbolic conical sections while maintaining a similar form to the BPR. The similarity to BPR enables a direct transfer of parameters. These alternative formulations have also been adopted in practice and research.

Volume–delay functions are typically employed in utility estimation (e.g., time cost) for static or semi-dynamic traffic assignments, and informing route choices for agent-based modeling (Suh et al., 1990; Çolak et al., 2016). The differentiable form and convex nature of volume–delay functions make them an ideal candidate for optimization-based traffic assignment, such as the assignment that satisfies Wardrop's equilibrium (Lien et al., 2016). Despite their extensive use in research and practice, the volume–delay functions have several limitations. For example, it is possible to obtain traffic volume-to-capacity ratio much higher than one, which is unrealistic (Nie et al., 2004; Chiu et al., 2011). Closely associated is the problem that the volume–delay functions can only model the hypocritical section of the traffic fundamental diagram (monotonic increase in the flow on a link with travel time and density), but not the hypercritical section (when a road is congested to a certain level, the flow will decrease despite an increase

in density and travel time). These limitations mean that volume–delay functions cannot model traffic phenomena such as spillbacks, wave propagation, and gridlocks, which would require the use of a dynamic model (Lo and Szeto, 2005; Chiu et al., 2011). These concerns are also reflected in the UK Department for Transport (DfT) Transport Analysis Guidance (TAG), which recommended to model junction delays explicitly especially for congested urban roads (DfT, 2020). Despite these shortcomings, volume–delay functions are a useful tool for regional-scale simulations and analysis requiring faster computations. Carefully calibrated volume relationships frequently show a good match with the real-world data (Kurth et al., 1996; Irawan et al., 2010).

In order to effectively use volume–delay functions in regional-scale analyses, it is crucial to maintain an up-to-date coefficient specific to local regions, which has proven to be a nontrivial task. Past studies have identified the need to inform these functions with empirical and context-specific data (Rose et al., 1989; Spiess, 1990), but also recognize the difficulty and cost associated with collecting such empirical data as being prohibitive (Rose et al., 1989). There have been studies that incorporate different forms of field sensor data for such calibration (Mtoi and Moses, 2014; Neuhold and Fellendorf, 2014; Kucharski and Drabicki, 2017). Nevertheless, the data used in these studies have been generated specifically for that application and requires specific hardware and software for use.

Recent innovations in information and communication technologies have led to an increase in the adoption of real-time crowd-sourced data feeds in transport modeling. This includes applications of location data for emissions estimations (Hirschmann et al., 2010), building origin and destination matrices (Toole et al., 2015) and general urban traffic management applications (Artikis et al., 2014). This study investigates the use of novel real-time crowd-sourced data feeds that have wider spatial coverage and are not generated specifically for estimating volume–delay functions. These sources could be used to consider some of these previously ignored characteristics and create temporally and spatially dynamic volume, speed, and saturation relationships. Such data sources can harvest data at a finer resolution over a longer (even indefinite) period of time giving a far greater understanding of the temporal variations and trends exhibited on road infrastructure.

This work attempts to combine new data sources in order to create functions that do not require a large range of survey inputs. The general and transparent methodology of harvesting crowd-sourced data enables its easy deployment to multiple sites. As a result, highly localized relationships can be obtained for a diverse range of road links, reflecting individual characteristics of the road (i.e., geometry, the presence of traffic furniture, road quality, and the surrounding land use). Such varying characteristics can result in very different vehicular behavior on roads that may be considered similar by traditional approaches.

## 3. Traffic Volume Inputs: ATCs

Two sources of data inputs are utilized to calibrate the volume–delay characteristics localized to the road link level: the link-level traffic counts (volume) from the ATCs and the link-level travel time (delay) from Google Maps' real-time information. Data from these two sources were harvested over the same time period (late February to mid March, 2016) and then paired according to the time of collection to create volume–delay observations. These observations will be used later in this paper to calibrate the context-specific volume–delay functions for road links where observations are available. In this and the next section, the two data sources will be introduced.

ATCs are magnetic induction loops embedded under the road surface. The passing of a vehicle results in an electromagnetic signal. The ATCs in Greater London count every vehicle which passes over the inductions loop. The data used here were harvested over a period of 3 weeks from February, 27 to March 21, 2016.

There are 37 DfT ATC locations distributed in Greater London (Figure 1). Among these data collecting locations, 34 roads have ATCs operate in both directions, while three roads have ATCs operate only in one direction. The ATC locations provide traffic counting information for a range of different DfT defined road classes (Table 1). The DfT (DfT, 2012a) publishes guidance on the road classification system in the UK.
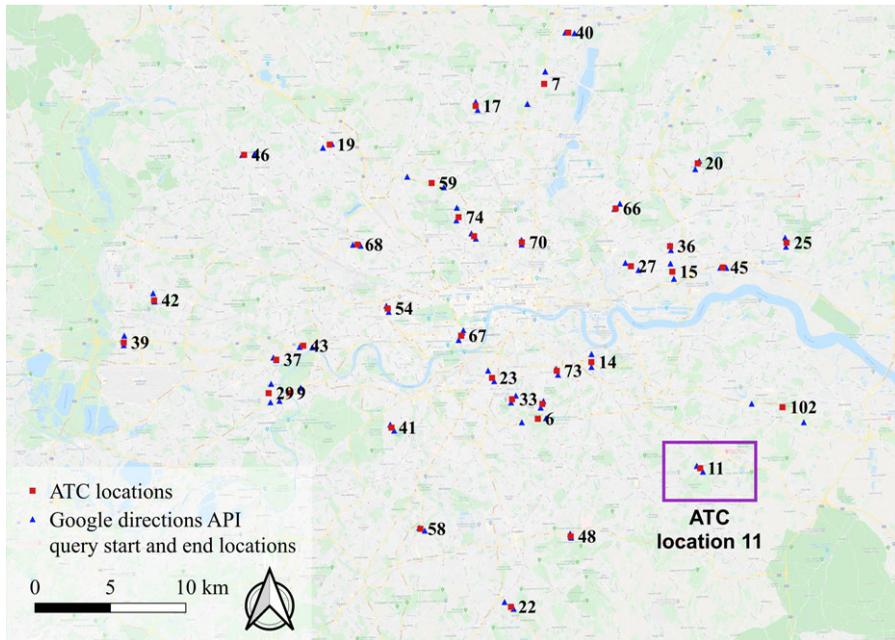
***Figure 1.*** *ATC locations in Greater London (Google Inc., 2020). The red dots illustrate the locations of the ATC. The blue triangles illustrate the origin and destination locations specified in order to harvest journey time information.*

**Table 1.** ATC locations by road class.

| Road class | Trunk (motorway or A) | Principal (A) | B | C | Unclassified | Total |
|---|---|---|---|---|---|---|
| Count | 6 | 16 | 3 | 3 | 9 | 37 |

The raw ATC data contain individual records for each vehicle that passed the counter, including its speed and the exact time (accurate to the second). Over the test period of 3 weeks, there were approximately 4.5 million recorded vehicles. An example record of a vehicle crossing the ATC Site 11 (labeled in Figure 1) on the first day of data collection reads:
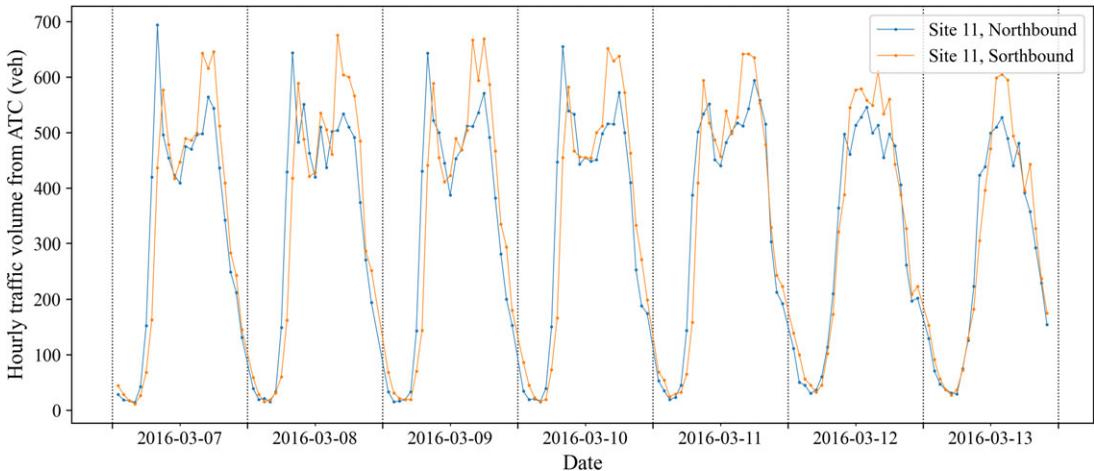
*Site: 11, Direction: Northbound, Date: February 27, 2016, Time: 00:00:28, Speed: 37.*

Individual vehicle records at each ATCs are aggregated by the hour to obtain the traffic volumes per hour (i.e., the hourly traffic flow). The total number of vehicles passing in an hour is considered as the hourly traffic volume at the measurement site. For gathering the aggregate data at each ATC, a unique identifier is defined by concatenating the location ID and the directionality of the ATC. For example, the northbound detector at ATC location 11 is identified as site "11 N." The timestamp is then rounded up to the next hour in order to quantify the hourly traffic volume up to that hour. This results in an output dataset which features traffic counts per hour for each site, along with the direction and date. A sample of the aggregated dataset at ATC location 11 N is shown in Table 2.

Figure 2 shows an illustration of the variations in the hourly traffic volume at ATC location 11 (northbound and southbound). Location 11 is situated on the Royal Parade road (A208) with the northbound direction leading to central London and the southbound direction leaving from central London. It is clear that during weekdays (March 7, 2016 to March 11, 2016), the northbound direction exhibits a higher peak during the morning rush hours (commuting trips into the city), while the southbound direction carries more traffic leaving the city during the evening peak. On weekends (March 12, 2016 and March 13, 2016), there is only one peak, which also starts later than the regular

**Table 2.** Processed ATC data record sample.

| Combined ID | ATC ID | Direction | Date | Hour | Hourly traffic volume |
|---|---|---|---|---|---|
| 11 N | 11 | Northbound | March 7, 2016 | 6 | 152 |
| N | 11 | Northbound | March 7, 2016 | 7 | 420 |
| N | 11 | Northbound | March 7, 2016 | 8 | 694 |
| N | 11 | Northbound | March 7, 2016 | 9 | 496 |



**Figure 2.** *Hourly traffic volume distribution for Site 11 (March 7, 2016 to March 13, 2016).*

morning peak observed on weekdays. This variation in traffic volume follows the general understanding of the distribution of traffic loads throughout the day.

## 4. Traffic Speed and Time Delay from Google Maps Application Programming Interface (API)

Link travel time or equivalently the inverse of the space-mean speed of vehicles passing through the link, can be collected using several methods. For example, the ATC data presented in the previous section have vehicle speed records. However, the ATC speed is a point measurement and may not be suitable to calculate the average travel time across the road link, as required by many traffic simulation studies. An alternative method to infer the travel time across the link is to use real-time, crowd-sourced location information gathered from mobile phone users. Mobile phones with location service enable harvesting of fine resolution temporal and spatial position data obtained from GPS positioning, cell tower triangulation, WiFi Service Set Identifier mapping, Bluetooth, and other technologies, either in isolation or in tandem. Such data hold a great deal of promise due to the range of possible uses it has in the transportation sector, such as understanding peak hour travel demands and inferring transport modes (Zheng et al., 2010; Çolak et al., 2016).

Anonymized crowd-sourced location data are useful for understanding real-time road traffic conditions on congested or smooth flowing roads (Barth, 2009). Crowd-source data collection is suitable for urban areas with high population density, high travel demand, and high mobile phone uptake. This information has been widely used by technology companies such as Google and Apple to inform their users of the optimum path that avoids traffic (Apple Inc., 2020; Google Inc., 2020; Microsoft, 2020; TomTom N.V., 2020). Individuals with access to these services can make an informed decision on route choice for a given mode or even a mode decision on how to get from their starting point to a desired
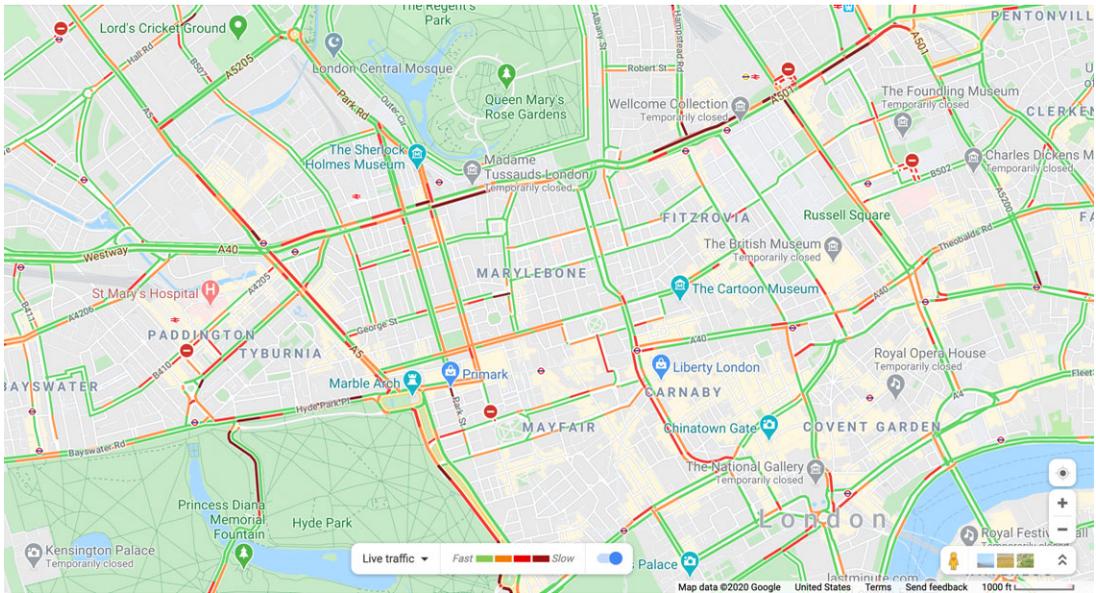
***Figure 3.*** *Google Maps traffic layer, showing live traffic in the Camden/Soho/Marylebone/Mayfair area of London on a Friday evening (Google Inc., 2020).*

destination with the lowest time cost. For example, Google provides real-time color-coded maps that offer a qualitative representation of the current traffic conditions on roads where sufficient data are available. Figure 3 shows the traffic conditions from Google Maps traffic layer of the Camden area in London on a typical Friday evening.

In this research, the traffic condition information is retrieved in batches using the Google Maps Directions API (Google Cloud, 2020). Depending on the personal settings of a mobile phone user, the freely available Google Maps app send anonymized data of their location to Google. Such data are personally and commercially sensitive, so post-processing is carried out by Google in order to ensure that no user movements can be isolated from the flows. Google's Directions API used in this paper is a service that calculates travel time and routing directions between given origins and destinations using a Hypertext Transfer Protocol (HTTP) request (Google Cloud, 2020). The use of an HTTP request allows for scheduled and bulk harvesting of journey information between a selected set of origin and destination pairs.

This research aims to combine crowd-sourced location-informed journey times with traffic counts from the DfT ATC network. The first step is to specify origins and destinations for the HTTP request to the Google Directions API to retrieve the journey times at the ATC locations. However, the DfT ATC network uses EPSG:27700 (British National Grid) coordinate system, while Google Maps employ the EPSG:4326 (WGS84) system. Hence, there is a need to convert between the two coordinate systems. Furthermore, to retrieve the journey time on a particular road, we need to optimally choose an origin and destination pair whose route would lie on the chosen ATC location section. The choice of origin and destination must be sufficiently further apart such that the journey times are meaningful while having a sufficiently short distance to exclude undesired results such as detours.

Figure 4 shows the manually defined origin and destination points for ATC location 19 Eastbound, the A5109 Deansbrook Road in Edgware, HA8. Automating the definition of these origin and destination pairs is challenging as merely taking the start/end of a given road yielded a route that is too long and distorted by other traffic, outside of the ATC consideration. Other methods based upon an idealized distance between points and the density of junctions is deemed too complicated and not durable. For bidirectional ATC locations, it is often not possible to define the Eastbound route as the inverse of the
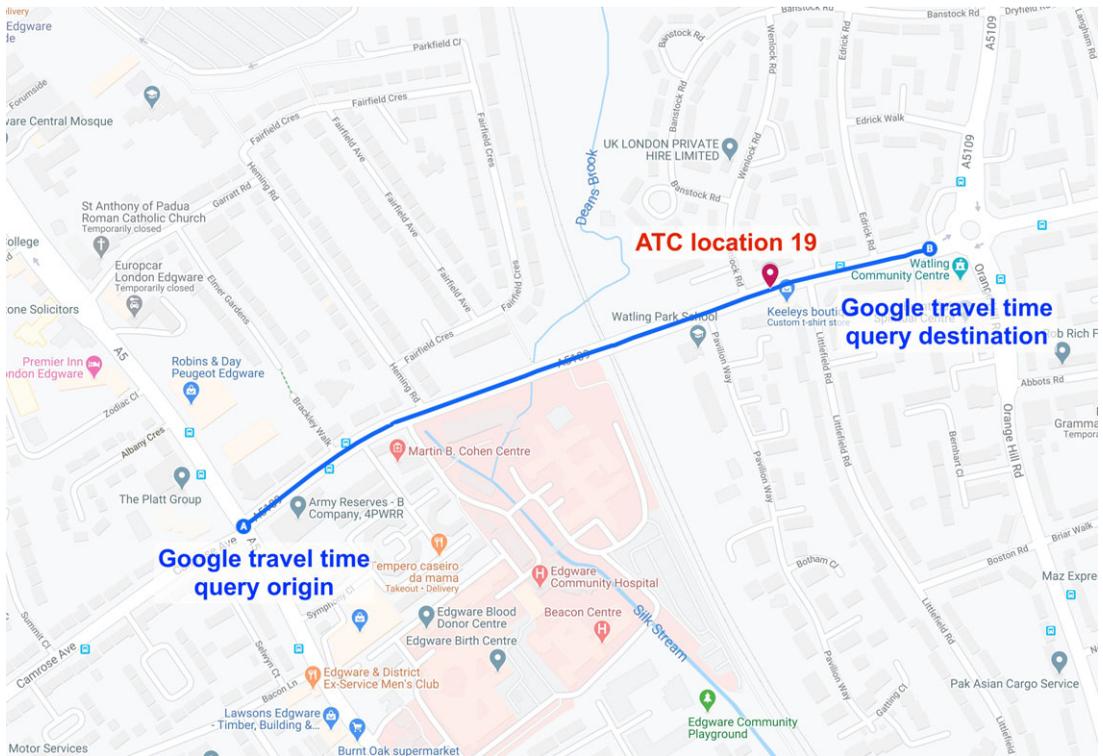
**Figure 4.** *ATC 6 Eastbound with defined origin and destination points (Google Inc., 2020).*

Westbound route as Google distinguish between different sides of the road, resulting in a route which involves a detour to navigate to the correct orientation safely. Thus a manual process is employed to visually inspect each location, the surrounding context and decide on the most appropriate origin and destination locations for querying for the travel time. Once this manual process is complete, a list of origin and destination pairs is produced containing ATC metadata that allows for the pairing of Google travel time results to its corresponding ATC traffic volume data.

During the information request process, an HTTP request is sent to Google's Directions API with the following information: origin, destination, mode (driving), and the specified departure time. In order to harvest real-time data that is informed by aggregated data of individual device-location at each hour, a cron scheduler (Crontab Documentation, 2012) is used to run the same origin and destination pairs repeatedly. In response to such an HTTP request, Directions API returns results in the JavaScript Object Notation (JSON) format, a lightweight data-interchange format (ECMA International, 2017). The JSON result has the *duration_in_traffic* field, which is Google's estimated journey time between the requested origin and destination. In the next section, we explain how these travel time data will be paired with ATC vehicle counts to produce volume–delay relationships.

Figure 5 shows a subset of Google's estimated journey time data along the Royal Parade, between Manor Park Road and Bromley Road in Southeast London (561 m in length). ATC 11 is located on the north end of this stretch of the road. Apart from the bimodal distribution of journey time with peaks associated with the morning (to work) and evening (leaving work) periods during weekdays and the single peak on weekends (Mullick and Ray, 2012), it is interesting to see a much higher travel time in the northbound direction at all times. A visual inspection of the corresponding road shows the presence of on-street parking spots in the northbound lane. Also, the ending point of the northbound lane intersects with a more major road (Bromley Road, A222) than southbound (Manor Park Road, B264). Both factors could contribute to the unexpected slowness in the northbound direction.
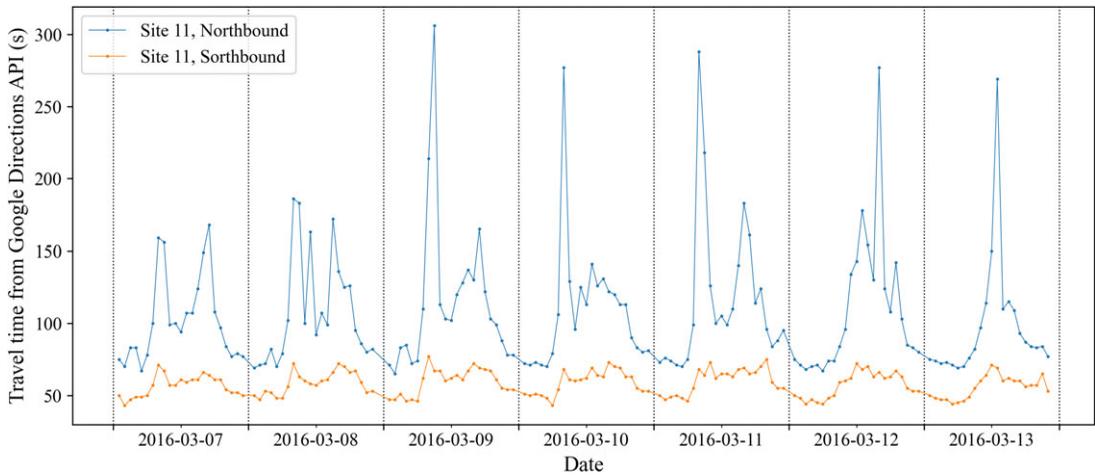
**Figure 5.** *Journey time distribution for Site 67 (March 7, 2016 to March 13, 2016) along the Royal Parade, between Manor Park Road and Bromley Road in Southeast London (561 m in length).*

## 5. Data Cleaning and Site Selection

The volume–delay relationship at different traffic conditions at each site is constructed by combining the Google travel time data and the ATC vehicle count data. Figure 6 shows an example of such data for site 9 S, 11 N, and 39 N. Outliers with high travel times can be observed at moderate traffic flow rates. These abnormal delays could occur due to external reasons (e.g., weather and road incidents), or as mentioned in the literature review, being the hyper-congested cases where flow reduces with an increase in delay (especially the cluster of points with hourly volume between 400 and 600 for site 11 N). The hyper-congested regime is not modeled by the volume–delay function. Instead, they are treated as residuals of the fitted volume–delay curve and quantifications of the magnitude of such residuals are provided in Section 6.4.

Regardless of the residual quantification, the collected data must be cleaned to remove points falling too far away from the continuous volume–delay relationship. Specifically, a manually tuned density-based spatial clustering of applications with noise (DBSCAN) algorithm is used to remove points in low-density areas in the plots (Scikit-Learn Developers, 2019; Brownlee, 2020). The DBSCAN algorithm has two parameters, $\varepsilon$, which sets the distance between points to be considered as a cluster and *minPoints*, the minimum number of points required to be considered as a core point in the cluster. Considering that the horizontal and vertical axes have different scales, the original data are first normalized to the maximum observed volume and travel time, so that the data in the transformed axes are all within 0 and 1. It was found that by specifying $\varepsilon = 0.1$ and *minPoints* $= 5$, it offers a good balance between removing outliers and keeping data points close to the main cluster. In fact, data cleaning is purposefully kept to a minimum to accommodate possible variations that may occur in real life. As a result, only 66 points are removed from the 17,745 observations.

Furthermore, when the journey time vs. traffic counts data for all the 39 original sites were plotted, it was immediately obvious that a few sites do not have valid data (very low or no variations in observed travel time through the observation period). Most of these sites only have very low traffic volume throughout the data collection period. For example, for site "74 S," Swain's Lane, a minor road in Camden, the maximum hourly volume is only 24 vehicles. The travel delays on these roads are relatively small and the data collected cannot be used to extract knowledge on the delays that might occur when the road is congested. These sites are subsequently removed from analysis and which leaves 24 sites remaining for further analysis.
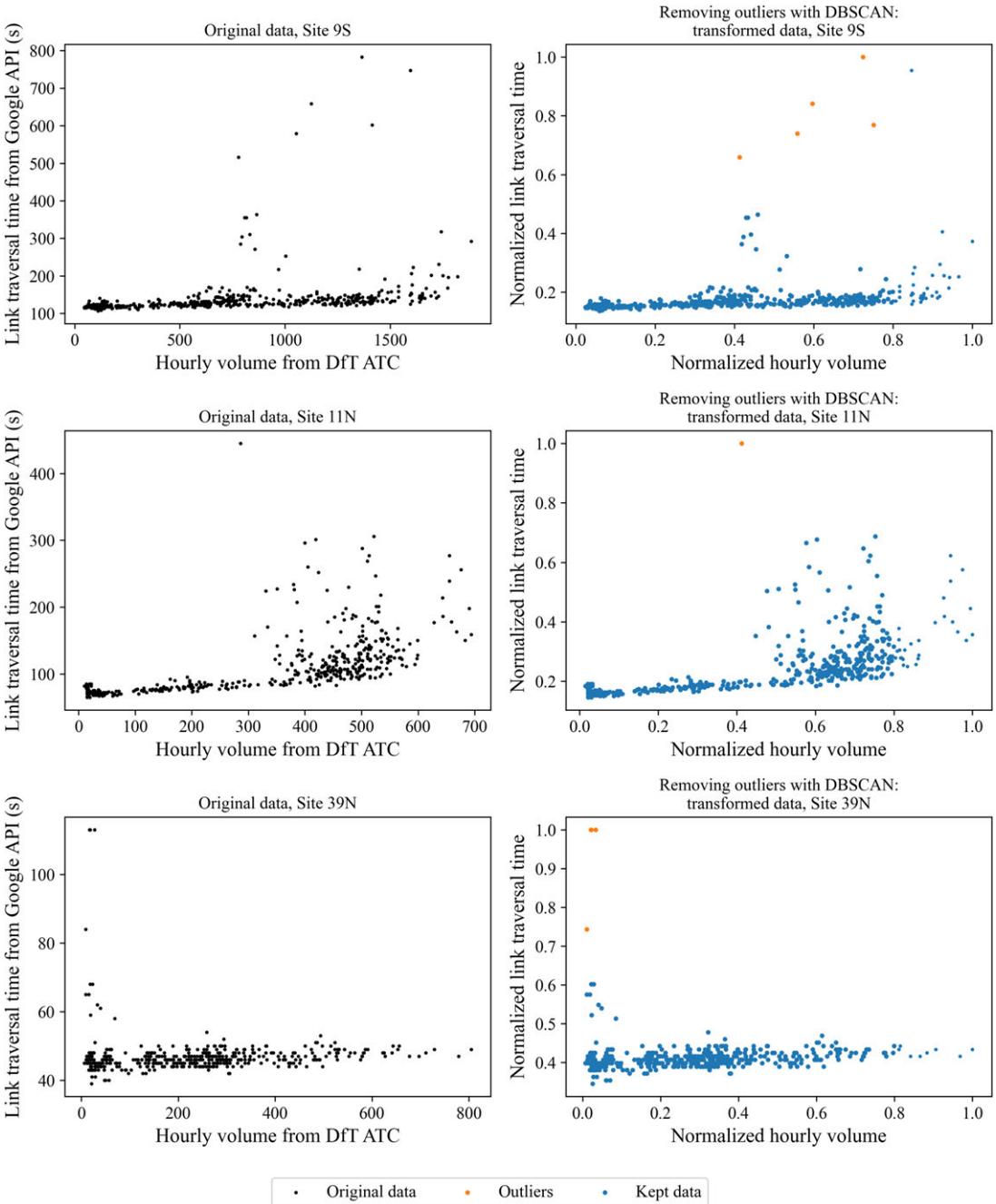
**Figure 6.** *Examples showing the original, removed, and kept observations in volume–delay scatter plot for three sites.*

## 6. Analysis and Model Building

The combination of Google travel time data and the DfT ATC vehicle count data enables various relationships to be assessed and defined. Based on these real-world datasets, the goal of this section is to derive data-driven parameters related to the traffic analysis, including (1) the free-flow travel time, or its inverse, the free-flow speed; (2) the road capacity; and (3) the volume–delay relationship.

### 6.1. Free-flow travel time

The free-flow travel time usually corresponds to the time for a vehicle to pass through a road link when no other vehicles are present, such as the travel time experienced in the early morning. In research studies and engineering practice, it is usually taken as the time to go through a link when traveling at the designated speed limit, sometimes multiplied by a slowing-down factor of roughly 1.3 to reflect the minor delays caused by stopping at intersections and other factors in an urban environment (Çolak et al., 2016). For example, the DfT defines the free-flow journey time using time delay coefficients for different road types, speed limits, link lengths, widths, gradients, and traffic junctions, which is obtained for our study sites and plotted as the horizontal axis value in Figure 7a (DfT, 2002).

In this study, it is recognized that using speed limit to calculate the free-flow travel time based on road type frequently ignores the impact of localized, irregular factors, for instance, a curve or pothole on the road. Rather than making assumptions about the free-flow time, we estimated it from the Google travel time data collected over the period of the study. For each site, the 5th percentile (fastest) travel time among all observation points is used to represent the observed free-flow travel time of that particular site. Free-flow speeds for each direction of the same road are estimated separately. Figure 7a shows the contrast between the speed limit and the observed maximum speed calculated from the observed free-flow travel time. Each point represents a specific site in the dataset. On average, the observed maximum speed of a road link is 20% lower than the specified speed limit. However, the former tends to be higher (exceeding the speed limit) on roads with the low posted speed limit. A 45-degree line is also plotted in Figure 7a for reference to compare the speed limit against the observed maximum speed.

### 6.2. Road capacity

Compared to the free-flow travel time, road link capacity is an ambiguous quantity to define. For example, the UK Design Manual for Roads and Bridges (DMRB) defines capacity vaguely as "the maximum sustainable flow of traffic passing in 1 hr, under favourable road and traffic conditions" (The UK Highways Agency, 1999). The DMRB also provides lookup tables that feature traffic capacities for a range of road types, road widths, and number of lanes. A manual survey of satellite imagery was carried out to assess the lane count, estimate the road width for each of the ATC locations and thus provide an estimated capacity by this DMRB definition. However, such a method was deemed unacceptable due to the uncertainty in what constitutes as *favorable road and traffic conditions.*
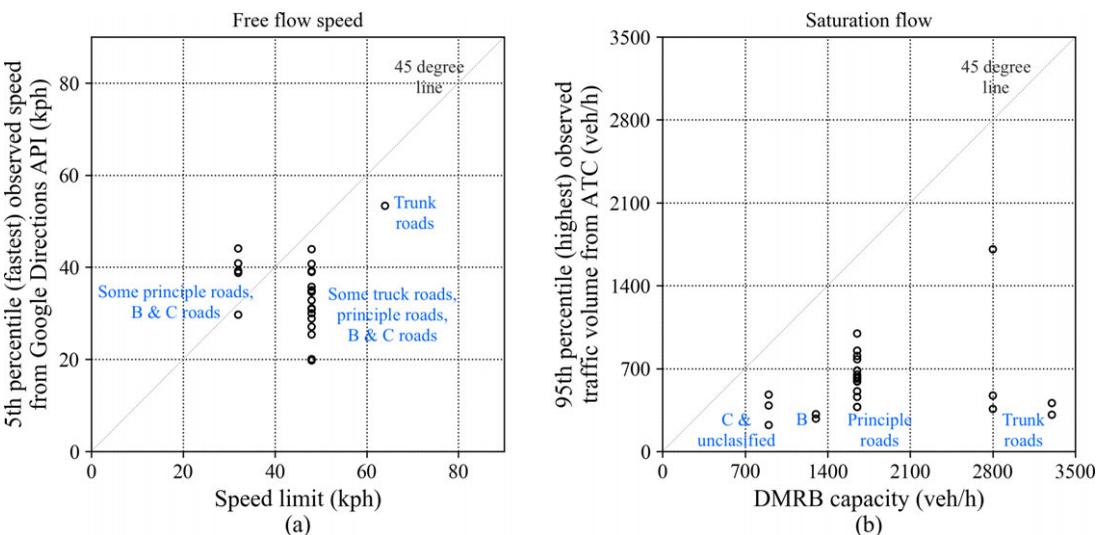


***Figure 7.*** *Comparison between the theoretical traffic parameter value and observed value at 24 ATC sites. (a) Free-flow speed; (b) capacity or saturation flow.*

Instead, a more nuanced definition from Spiess (Spiess, 1990) is adopted, which defines capacity as the volume at which congested speed is half the free-flow speed. In the combined Google and ATC data, since there is hardly any observation when the speed is exactly half of the free-flow speed, a more relaxed definition of capacity is subsequently utilized. For each site, the capacity is taken as the highest 95th percentile hourly traffic volume among all observations where the travel time falls between 1.8 and 2.2 times the free-flow speed. Capacities for each direction of the same road are estimated independently, since each direction may have different contexts, such as the presence of bus stop, street parking, or lane width. As a result, each individual road is given a capacity attribution. This capacity attribution is observed to be consistently lower than the DMRB recommended value, as can be seen in Figure 7b. On average, the observed capacities are only 34% of those recommended by the DMRB, reflecting factors such as signalized intersections, bus stops, and street parking that cannot be perfectly captured using simple lookup tables. The 45-degree line is plotted in Figure 7b for reference. Also, it can be seen that there are only a few values of the DMRB road capacity, clustered according to the road types and lane counts.

### 6.3. Volume–delay relationship

Based on observations of the road link travel time and hourly traffic counts presented above, the context-specific volume–delay relationship can be constructed. The BPR formulation of the volume–delay relationship is chosen due to its mathematical advantages (differentiable and convex) and widespread use (Equation (1)). In particular, for the baseline case, we used $\alpha = 1.0$ and $\beta = 2.0$ as per TfL recommendation, while $t_0$ and $v_c$ are based on the speed limit and DMRB recommended capacity as computed in Section 6.1 and Section 6.2. It should be noted that when the BPR curve is used in practice, $t_0$ is often corrected by a delay factor to reflect the city-specific conditions (Çolak et al., 2016) while $v_c$ is often corrected by factors such as road geometry, gradient, and signal status (Kurth et al., 1996). However, these corrections usually require labor-intensive site inspections and the correction factors are often empirical or generalized for the whole city. The aim of this analysis is to propose an efficient way to obtain alternative, site-specific function parameters from real observations. The purpose of this study is to also contrast the performance of the data-informed model parameters with the base curve in terms of their ability in calculating the time-delay at different congestion levels.

A summary of the three models used for comparison is given in Table 3. The base curve uses BPR coefficient $\alpha = 1$ and $\beta = 2$. In particular, the free-flow travel time $t_0$ is calculated from the speed limit, while the capacity $v_c$ follows the DMRB recommended value. In traffic modeling, the transport agencies may apply some adjustment coefficients to consider external influencing factors on $t_0$ and $v_c$. However, in

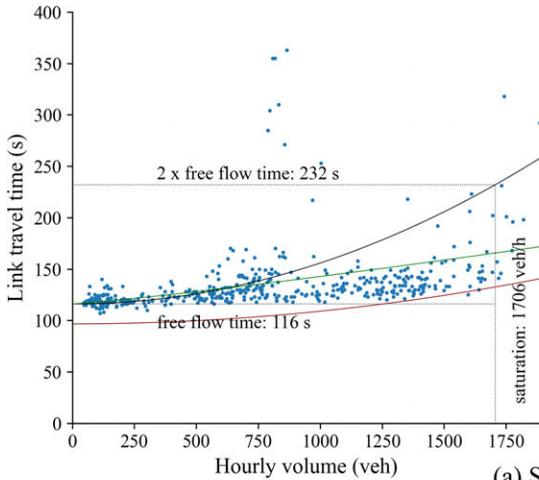**Table 3.** Summary of volume–delay relationships tested.

| Function name | Parameters |
|---|---|
| Base curve | $\alpha = 1.0, \beta = 2.0$. |
| | $t_0$: road length divided by speed limit. |
| | $v_c$: DMRB definition. |
| Data-informed 1 (DD1) | $\alpha = 1.0, \beta = 2.0$. |
| | $t_0$: 5th percentile link travel time (fastest) in observations. |
| | $v_c$: 95th percentile traffic volume (largest) for observations with travel time between 1.8 and $2.2 \times t_0$. |
| Data-informed 2 (DD2) | $\alpha, \beta$: fitted value using nonlinear least square regression. |
| | $t_0$: 5th percentile link travel time (fastest) in observations. |
| | $v_c$: 95th percentile traffic volume (largest) for observations with travel time between 1.8 and $2.2 \times t_0$. |

most of the cases, these factors are predetermined and may not be up-to-date or localized to each road segment. The alternative formulations are developed to address this issue: for the first data-informed alternative formulation (DD1), the base value of $t_0$ and $v_c$ is replaced by observed values obtained in Section 6.1 and Section 6.2. For the second data-informed alternative (DD2), not only $t_0$ and $v_c$ are replaced by the observed values, but also $\alpha$ and $\beta$ are calibrated using regression models on the real-world data. Once the data collection and processing pipeline have been built, it is relatively easy to obtain the coefficients used in the data-informed volume–delay curves, even updating weekly or monthly to take into account any temporal changes in the road characteristics (e.g., roadworks or weather conditions).

It is evident from Figure 6 that the observed relationships between the volume and delay are site-specific, nonlinear, and heteroskedastic (variance of the data increases with the independent variable). All these characteristics are expected from the traffic perspective. In particular, the nonlinearity means that the travel time increases more dramatically as the hourly traffic volume approaches the capacity. The heteroskedasticity implies more variations in vehicle speed when the road is congested due to interactions with other vehicles. Usually, to deal with such data, a nonlinear model needs to be used (which is already the case for the BPR curve). Heteroskedasticity can be accounted for by using techniques that include data transformation or using weighted least square method (Vynck, 2017). However, it is not a requirement if the goal is to estimate regression parameters, as the ordinary least square (OLS) method also produces unbiased, though inefficient, estimation of the coefficients. Standard errors of the estimation parameters are not used, so the OLS method is deemed sufficient for obtaining the $\alpha$ and $\beta$ parameters used in the DD2.
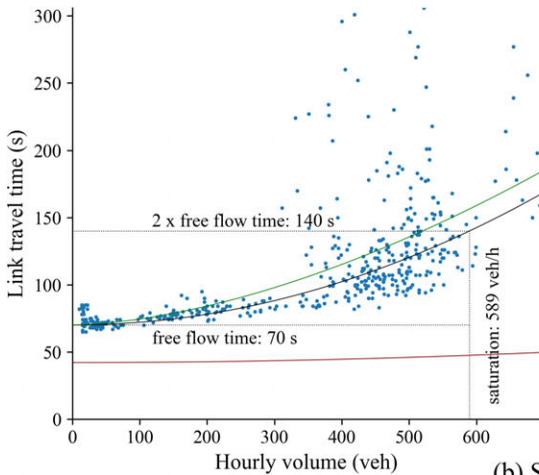
Figure 8 shows the fitting results of all candidate functions for three representative sites, each from a different road class. It can be seen that the base curves with the default speed and capacity values fit the poorest to the data across all three examples. In particular, the base curves underestimate the free-flow time for Site 9 S (Figure 8a, Trunk road) and 11 N (Figure 8b, Principal road), while overestimating the free-flow time for Site 35 S (Figure 8c, B road). The base curves overstate the capacity in all three examples, which has already been shown in Figure 7b. By simply substituting the free-flow time $t_0$ and capacity $v_c$ in the base curve with the observed values, the fitting of the DD1 improves significantly based on a visual inspection. Quantitative evaluations of the fitting performance will be given next in Section 6.4. The DD2 (green curve in Figure 8) fits the data more closely than the DD1. However, in certain cases, the DD2 may result in unrealistic $\beta$ values of less than one, due to a large number of observations at medium volume but large delay such as the hyper-congestive conditions, thereby making the volume–delay curve unreasonably concave. Hence, the DD1 is a good approach in general to improve the performance of volume–delay curves.

By dividing the length of the road link with the link travel time shown in Figure 8, the speed–flow relationship can be obtained. The DfT TAG uses linear or piece-wise linear models to predict the reduction in speed with increasing flow. Depending on the road class (e.g., whether it is urban, suburban, rural), different formulae are used. For each class, the speed–flow formulae again depend on a variety of factors (e.g., percentage of frontage development, numbers of major and minor intersections). It should be noted that, although a speed–flow formula is provided for urban roads, the preferred way of doing traffic assignment is still to model junctions and low cruise speed in congested urban areas explicitly, rather than relying on the speed–flow curves. If to be used, it should be recognized that the speed–flow curves are developed based on network-average conditions, which limits their applicability to individual links independently. Fully recognizing these limitations, a comparison plot is still made for the observed data points and the speed–flow curves in the TAG document for illustration purposes (Figure 9). The blue dots are the speed–flow observation data. The black and green curves are the same as the ones in Figure 8, only showing in the speed–flow plane. The purple lines are the speed–flow relationship made according to the TAG model. Parameters of the purple curves are given in Table 4 and these values are obtained by examining Google Street View of each site presented in Figure 9. It can be seen that the purple curves do not agree well with the local data observations. Part of it is because link observations (e.g., the percentage of frontage development for this particular street) are used in place of network-average values. While it is
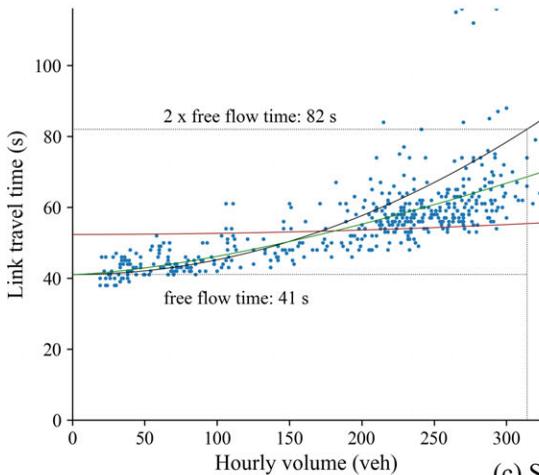
Base curve:
$$t = 96.6 \times [1 + 1.0 \times (\tfrac{vol}{2800.0})^{2.0}]$$

DD1:
$$t = 116.0 \times [1 + 1.0 \times (\tfrac{vol}{1706.8})^{2.0}]$$

DD2:
$$t = 116.0 \times [1 + 0.43 \times (\tfrac{vol}{1706.8})^{1.16}]$$

(a) Site 9S, Trunk Road

Base curve:
$$t = 42.1 \times [1 + 1.0 \times (\tfrac{vol}{1650.0})^{2.0}]$$

DD1:
$$t = 70.0 \times [1 + 1.0 \times (\tfrac{vol}{589.9})^{2.0}]$$

DD2:
$$t = 70.0 \times [1 + 1.23 \times (\tfrac{vol}{589.9})^{1.68}]$$

(b) Site 11N, Principal Road

Base curve:
$$t = 52.3 \times [1 + 1.0 \times (\tfrac{vol}{1300.0})^{2.0}]$$

DD1:
$$t = 41.0 \times [1 + 1.0 \times (\tfrac{vol}{314.3})^{2.0}]$$

DD2:
$$t = 41.0 \times [1 + 0.67 \times (\tfrac{vol}{314.3})^{1.47}]$$

(c) Site 35S, B Road

**Figure 8.** *The fitting three different models in the observed data at three sites.*

(a) Site 9S, A316 Twickenham Road    (b) Site 11N, A208 Royal Parade    (c) Site 35S, B219 Barry Road
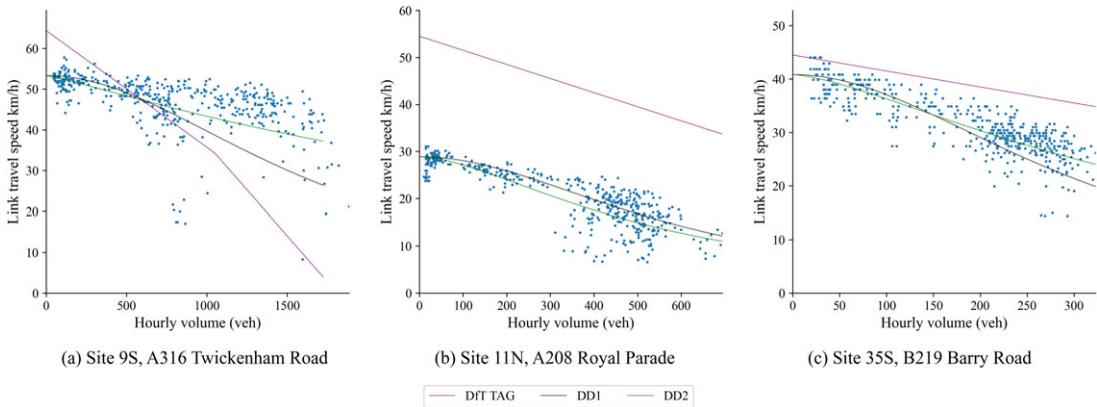
— DfT TAG    — DD1    — DD2

*Figure 9. Comparison with the DfT TAG speed–flow curves for three sites.*

also true that the curve parameters, even if to be collected at the network level, are somewhat subjective/ elastic and may not represent all context-specific information (e.g., bus stops, subtle geometry effects, or temporary closures). In comparison, the proposed crowd-sourced data collection approach can then provide a more localized, context-specific estimation of the volume–delay curves, or even used to calibrate parameter values of the TAG models.

### 6.4. Variation quantification

As it can be seen in Figure 8, the real-world traffic data are very noisy to be represented by a single curve. There have been second-order models that include a family of curves, which can reproduce the variations in travel time around the same traffic flow conditions to a certain degree (Fan, 2020). However, well-calibrated first-order models that only include the speed term (no time derivatives of it), such as the BPR curve are still beneficial for specific tasks, including analysing the impact of building new infrastructures and conducting regional-wide static and semi-static traffic simulations. After fitting the three candidate models to the data, their fitting performances are analysed in this section.

The fitting performance is evaluated with the mean absolute error (MAE, Equation 2) because the MAE metric has the same unit as the dependent variable (time in seconds or speed in km/hr) and is less sensitive to outliers compared to the frequently used root mean squared error metric. As the variations in the observed travel time are uneven and more substantial when the hourly traffic volume approaches the capacity, MAE is calculated for different subsets of the data, divided based on the quartile of the hourly traffic volume. Figure 10 shows the MAE in terms of travel time (a)–(c) and speed (d)–(f) estimations of all roads, grouped by the volume-to-capacity ratio of the observed data. For example, Q1 refers to the subset of the data where the hourly traffic volume are within 0–25% of the estimated capacity. Q5 refers to the subset of the points where the hourly traffic volume exceeds the estimated saturation flow. The Q5 partition exists because capacity in this study is defined as the 95th percentile hourly traffic volume for observations with travel times between 1.8 and 2.2 times the free-flow time (Section 6.2). It can be seen that, in terms of travel time or speed, the base curve leads to the largest errors in estimating the time delay given the traffic volume (Figure 10a,d). While for the two data-informed models, the errors of the DD2 appear to be less, but not much lower than the DD1. As mentioned before, engineers and planners would normally apply a series of factors on $t_0$ and $v_c$ for the base curve to reflect city or street characteristics (e.g., signal green time ratio or road curvature). We tried to apply these factors to scale $t_0$ and $v_c$ so that the base value matches more closely with the observed value in Figure 7, but it still produces a larger MAE compared to the two data-informed curves due to its lack of ability to adapt to individual street situations.

**Table 4.** Parameter values for curves in Figure 9.

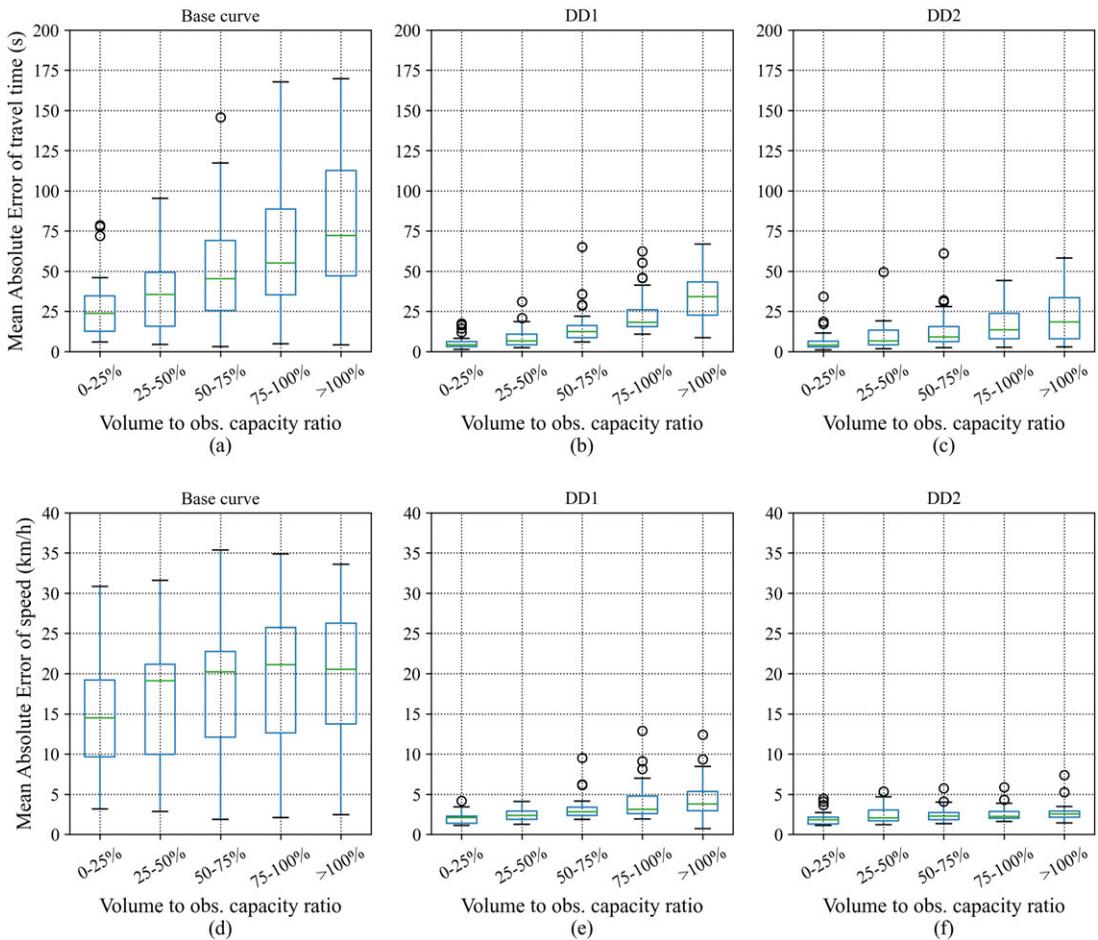| Curve | Basic form | Site 9 S | Site 11 N | Site 35 S |
|---|---|---|---|---|
| DfT TAG (purple curve) | $v = v_0 - \text{coefficient} \times \text{volume}$ | Class 10, suburban single carriageway | Class 7, urban noncentral roads | Class 7, urban noncentral roads |
| | | Percentage of heavy vehicles $PHV(\%) = 0$ | Frontage development $DEVEL(\%) = 50$ | Frontage development $DEVEL(\%) = 100$ |
| | | Maximum realistic flow $Q_C = 1{,}500*(92 - PHV/80)$ | Free-flow speed $v_0 = 64.5 - \frac{DEVEL}{5}$ | Free-flow speed $v_0 = 64.5 - \frac{DEVEL}{5}$ |
| | | Flow change point $Q_B = 0.7 \times Q_C$ | $v = v_0 - 30 \times \frac{volume}{1000}$ | $v = v_0 - 30 \times \frac{volume}{1,000}$ |
| | | Major intersections $INT = 1/\text{km}$ | | |
| | | Minor intersections $AXS = 4/\text{km}$ | | |
| | | Free-flow speed $v_0 = 70 - 5 \times INT - 3 \times AXS/20$ | | |
| | | For flow $< Q_B$ $v = v_0 - \left(12 + 50 \times \frac{INT}{3}\right) \times \frac{volume}{1,000}$ | | |
| | | For $Q_B < $ flow $< Q_C$ $v = v_0 - \left(12 + 50 \times \frac{INT}{3}\right) \times \frac{Q_B}{1,000} - 45 \times \frac{volume - Q_B}{1,000}$ | | |
| DD1 (black curve) | $v = v_0 / \left(1 + \alpha \times \left(\frac{volume}{capacity}\right)^{\beta}\right)$ | capacity $= 1{,}706.8$ | capacity $= 589.9$ | capacity $= 314.3$ |
| | | $v_0 = 53.3$ | $v_0 = 28.9$ | $v_0 = 40.8$ |
| | $\alpha = 1, \beta = 2$ | $\alpha = 1, \beta = 2$ | $\alpha = 1, \beta = 2$ | |
| DD2 (green curve) | $v = v_0 / \left(1 + \alpha \times \left(\frac{volume}{capacity}\right)^{\beta}\right)$ | capacity $= 1{,}706.8$ | capacity $= 589.9$ | capacity $= 314.3$ |
| | | $v_0 = 53.3$ | $v_0 = 28.9$ | $v_0 = 40.8$ |
| | | $\alpha = 0.43, \beta = 1.16$ | $\alpha = 1.23, \beta = 1.68$ | $\alpha = 0.67, \beta = 1.47$ |

Boxplot grouped by VOC_ratio



**Figure 10.** *Boxplots of the overall fitting MAE for all study sites for three model types, grouped by the traffic volume level. (a–c) MAE for the estimated travel time. (d–f) MAE for the estimated link speed.*

$$\text{For travel time}: MAE_t = \frac{\sum_{i=1}^{n}|t_{i,\text{model}} - t_{i,\text{obs}}|}{n},$$

$$\text{For speed}: MAE_u = \frac{\sum_{i=1}^{n}|u_{i,\text{model}} - u_{i,\text{obs}}|}{n},$$

(2)

where $t_{i,\text{model}}$ and $u_{i,\text{model}}$ are estimated travel time and speed from the model; $t_{i,\text{obs}}$ and $u_{i,\text{obs}}$ are individual travel time and speed observations; $n$ is the total number of observations.

The above results not only indicate the performance of the models, but also give engineers quantitative evidence on what amounts of variations to be expected when using the simple volume–delay function curves in their estimation. For example, Figure 10a–c shows the variations between the observed link traversal time and the estimated travel time by each individual model. For the roads selected as the study sites, the link travel time predictions can be 1–2 min off using the base curve, while the errors are usually below half a minute using the data driven models (DD1 and DD2), except in Q5, where the traffic volume is relatively high (>observed capacity) and the errors increase to 0.5–1 min. The model prediction errors are better discussed by comparing the predicted speed against the observed speed at different traffic levels,

as in Figure 10d–e since the speed measure can be compared across roads. As seen in Figure 10d–e, without the site-specific knowledge, the speed estimations using the default volume–delay curve can vary by 10–25 km/hr (Figure 10d). Nevertheless, merely fixing the speed and capacity estimations (DD1), the variation in the speed estimation can be reduced to less than 5 km/hr for low to medium traffic volume cases, and less than 10 km/hr when the volume is close to the capacity (Figure 10d). If using calibrated coefficients $\alpha$ and $\beta$ as in the DD2, the variation in the speed estimations are almost always below 5 km/hr from the observations (Figure 10e). However, it should be noted that calibrating $\alpha$ and $\beta$ without constraints may produce concave volume–delay curves in the case of the DD2.

Table 5 shows the MAE of speed (km/hr) by model, road class, and volume level, which is an extension of the information in Figure 10d–f disaggregated by the road facility type. As can be seen in the table, the Principal roads generate the highest number of observations, while the Unclassified roads have the least. Among the observations from Principal roads, the largest group of observations have the volume-to-capacity ratio around 75–100%, indicating a high demand. At the same time, less important roads (e.g., C or Unclassified) have the majority of the observations in the low traffic category (less than 75%), which means less congestion in general on these types of roads. The general trend on the MAE for the estimated speed is not much different across different road classes. For all road classes, the MAE of the estimated speed is significantly higher when using the base curve and lower when using the data-informed models. Also, the MAE for speed, in general, grows with the volume level: when the traffic volume-to-capacity ratio is higher, the MAE for speed goes up as well, which is in agreement with the visual judgements made in Figure 10.

## 7. Discussions on the Limitations of the Methodology

The empirical data gathered in this study vary significantly from the standard volume–delay curve. The context-specific saturation delay and saturation speed functions and the standard BPR volume–delay function show significant variations in the residuals. The large scatter observed in the volume–delay relationship must be due to unknown external factors that have not been included in the candidate models.

### 7.1. Example scenario and possible factors

Figure 11a presents the observed volume and delay scatter plot for ATC location 66 in both directions. It is obvious that the southbound direction consistently experiences higher delay than the northbound direction, even at same volumes of traffic. In an attempt to better understand and possibly explain this distinct directional behavior, satellite photography, and street-level photography of the road are assessed. Figure 12 shows a satellite image of ATC location 66 obtained from Google Maps (Google Inc., 2020). The red marker gives the location of the ATC counter itself, and the two blue markers illustrate the origin/destination location used in the Google Directions API request. From this satellite imagery and street view imagery, possible explanatory factors can be identified:

1. The southbound lane features road parking.
2. The southbound lane features a large bus stop and taxi lay-by (for Leytonstone Station) and a smaller bus stop lay-by.

The larger bus stop and taxi lay-by serving Leytonstone is a significant geometric feature that is likely to have a significant impact on the southbound traffic. The second smaller bus stop lay-by and on-road parking may also have an impact, albeit to a smaller extent. Such factors may explain the exhibited differences from the location-device-informed journey data and thus permit their inclusion in an informal way.

### 7.2. Challenging scenarios

Inspections of the volume–delay scatter plots reveal a subset of roads that are not suitable to be described with the BPR-style monotone function. For example, as previously explained in Section 5, 15 sites are

**Table 5.** Speed MAE (km/hr) by model form, road type, and traffic volume level.

| Volume-to-capacity ratio | | <25% | 25–50% | 50–75% | 75–100% | >100% | Total |
|---|---|---|---|---|---|---|---|
| Trunk | Obs. counts | 383 | 245 | 375 | 366 | 23 | 1,392 |
| | % obs. counts | 27.51 | 17.60 | 26.94 | 26.29 | 1.65 | 100 |
| | Base | 14.4 | 19.3 | 23.5 | 25.8 | 27.9 | 19.3 |
| | Data 1 | 2.3 | 3.3 | 2.7 | 2.4 | 3.8 | 2.7 |
| | Data 2 | 1.7 | 2.6 | 2.3 | 2.2 | 2.8 | 2.2 |
| Principal | Obs. counts | 1,641 | 1,082 | 1,996 | 2,746 | 216 | 7,681 |
| | % Obs. counts | 21.36 | 14.09 | 25.99 | 35.75 | 2.81 | 100.00 |
| | Base | 13.3 | 17.2 | 18.9 | 21.1 | 21.5 | 18.7 |
| | Data 1 | 1.9 | 2.2 | 2.9 | 2.8 | 3.4 | 2.6 |
| | Data 2 | 1.8 | 2.0 | 2.5 | 2.6 | 2.5 | 2.4 |
| B | Obs. counts | 179 | 125 | 183 | 343 | 102 | 932 |
| | % obs. counts | 19.21 | 13.41 | 19.64 | 36.80 | 10.94 | 100.00 |
| | Base | 9.0 | 6.6 | 4.3 | 3.5 | 3.8 | 4.8 |
| | Data 1 | 1.8 | 2.3 | 4.6 | 6.7 | 8.6 | 3.6 |
| | Data 2 | 1.8 | 2.1 | 2.2 | 2.1 | 2.2 | 2.1 |
| C | Obs. counts | 245 | 100 | 306 | 255 | 26 | 932 |
| | % obs. counts | 26.29 | 10.73 | 32.83 | 27.36 | 2.79 | 100.00 |
| | Base | 20.7 | 21.9 | 21.6 | 21.1 | 19.4 | 20.7 |
| | Data 1 | 1.8 | 2.1 | 2.3 | 3.4 | 4.9 | 2.4 |
| | Data 2 | 1.7 | 1.9 | 1.6 | 1.7 | 1.5 | 1.6 |
| Unclassified | Obs. counts | 126 | 120 | 48 | 26 | 9 | 329 |
| | % obs. counts | 38.30 | 36.47 | 14.59 | 7.90 | 2.74 | 100.00 |
| | Base | 24.7 | 25.5 | 26.0 | 27.0 | 30.6 | 26.0 |
| | Data 1 | 2.1 | 1.3 | 2.3 | 4.7 | 3.7 | 2.3 |
| | Data 2 | 2.0 | 1.2 | 1.3 | 2.3 | 3.4 | 2.0 |

eventually removed from the analysis due to invalid data. Among the remaining 24 sites, many have outliers even after the data cleaning process (e.g., Figure 6). Figure 11b shows another "abnormal" example, particularly with the northbound direction: the delay has a considerable variation when the hourly volume is around 200 vehicles, far from its capacity of about 400 vehicles per hour. This is probably resulted from the hyper-congested stage of the traffic, where the density has exceeded the optimum value and flow decreases despite continuous increase in the delay. Such phenomena cannot be correctly modeled by the monotone volume–delay function and should be dealt with using dynamic models if the hyper-congested stage is crucial in the analysis (which is usually the case for short-term predictions, such as evacuations). This is also reflected in the modeling guidance from the DfT, where the preferred choice of modeling congested urban roads in traffic assignment is to explicitly represent the junction delay and low cruise speed (DfT, 2020). The proposed data-driven models are useful in certain scenarios, such as traffic assignment at the city-scale, where simplicity of the volume–delay curves is
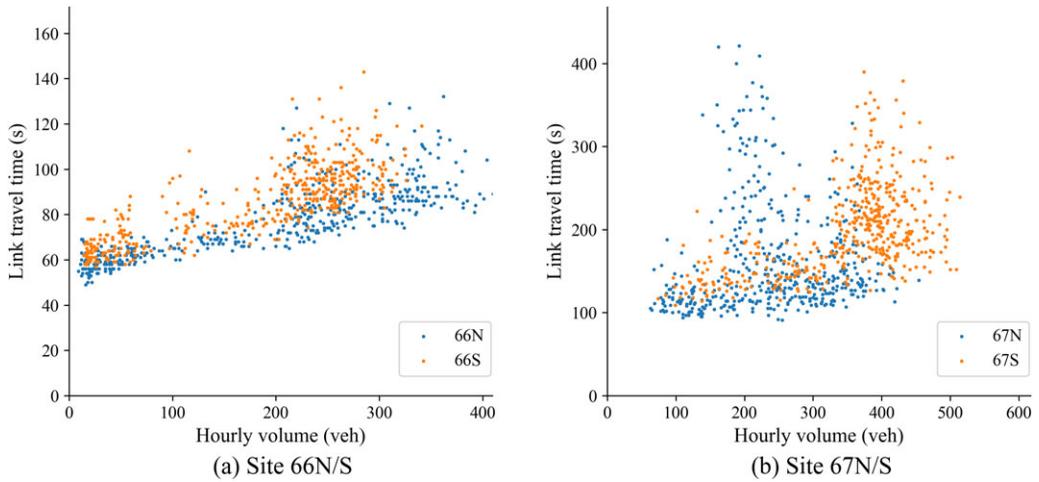
**Figure 11.** *Challenging scenarios: volume and delay observations at ATC site 66 (a) and 67 (b).*
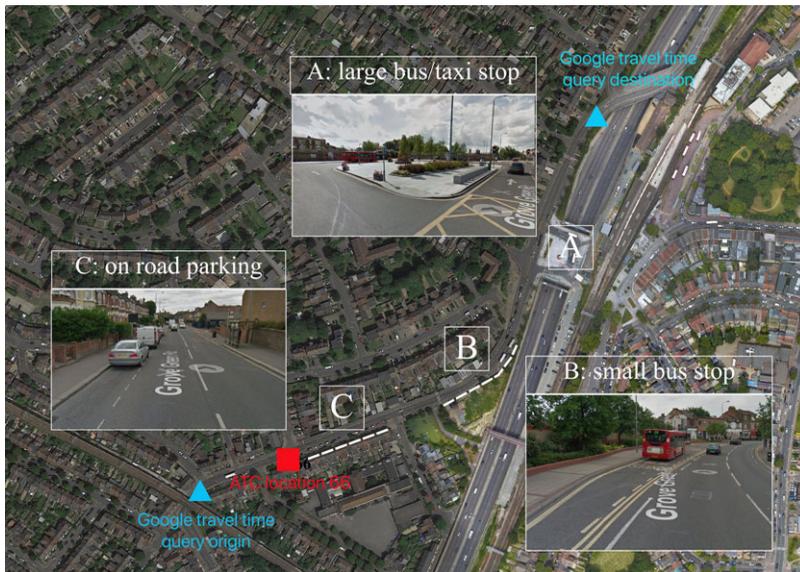


**Figure 12.** *Satellite and street view images of ATC 66 (Google Inc., 2020).*

crucial to ensure the computational tractability of the model. However, its limitation, namely allowing unrealistic high-traffic flow to be assigned to certain links, should be aware of. We plan to compare the performance of the data-driven method with the more detailed models involving junctions in our future work. The errors of using the monotonic volume–delay function in face of hyper-congested real-world observations are reflected in the MAE quantification as shown in Figure 10. Apart from these reasons, external factors could also affect the sensibility of the data, for example, a traffic collision or flooding may result in a high level of delay in a short period of time.

### 7.3. Influencing factors

In Section 7.1 the existence of public transport infrastructure and on-road parking was identified as being potentially explanations of the distinct southbound northbound behavior on that road. A range of other possible factors have been identified, which may help explain the unexplained variations:

1.  *Vehicle type*

In this study, all vehicles are considered to be the same, despite clear distinction in the interactions between two cars compared to an interaction of a car and a lorry (Vap et al., 2007). Further work is planned to assess the impact of the vehicle type. A wide variety of vehicles are observed on different road types, and such differences should be considered. The use of statistical vehicle mix sampled by road type (DfT, 2012b) was discounted for this study due to the small sample size of such statistics compared to the resolution of the data used here. The use of traffic cameras with number plate recognition and sufficient privileges to the Driver and Vehicle Licensing Agency database would enable the disaggregation of vehicle type at a similar spatial and temporal resolution to the Google Directions and DfT ATC data presented here.

2.  *Weather*

Weather events may impact the journey times by altering the functionality of the vehicle, the performance of the road, and/or the performance of the driver. Ongoing research focuses on combining a large dataset of Google Directions journey times with data from the UK's Met Office NIMROD precipitation dataset (The UK Met Office, 2003) to assess the relationship between these variables. At a microscopic level, it is known that an increase in precipitation increases journey times as a result of increased risk and the resulting decrease in vehicle speeds to compensate for this factor (Mashros et al., 2014).

3.  *Road incidents*

Road works and road traffic collisions can lead to decreased or even zero capacity on a road link, resulting in increased saturation which impacts the journey time. Depending on the warning before such an event, the vehicle traffic may have the ability to adjust to this information, resulting in a greater distribution of traffic thus reducing the mediated travel times. Alternatively, an accident may occur without any warning to other road users committed to their route choice, resulting in longer journey time and perhaps gridlocks in the extreme cases. A method incorporating different accident and roadworks databases with Google Directions data is currently being investigated.

4.  *Road geometry, type, and land use*

Different road layouts may increase the complexity of vehicle interactions. For example, the curvature of a corner and the road surface quality will impact the speed of a vehicle. The surrounding land use also has an influence on the vehicle speeds (e.g., drivers take safety precautions near a school or leisure center). The inclusion of such factors poses many challenges, including the size and complexity of the data plus the uncertainty and variability in how drivers react to the data. In Figure 12, a series of geometric factors are displayed as an attempt to explain the different behaviors on the same road in different directions. The factors shown in Figure 12 such as on-road parking and bus lay-bys could be quantitatively captured using computer vision and data sources such as Google Street View.

### 7.4. *Generalization and application*

This study proposes an efficient data collection and processing pipeline for calibrating the macroscopic volume–delay relationships for urban roads. The methodology considers the context-specific factors that might affect such relationships but are hard to quantify using conventional approaches. Apart from the challenges of using simple volume–delay curves in capturing the complex traffic behavior, another bottleneck for the proposed methodology is the availability of the data. This limitation comes in two parts. First of all, the volume data could only be obtained for a small number of fixed locations, where the ATCs are installed and can provide continuous traffic counts. The DfT also conducts manual traffic surveys at a larger number of locations. However, these manual count data usually come at a lower frequency (e.g., annual data) and may again miss the temporal network changes that have an impact on traffic (e.g., roadworks). This bottleneck could be overcome in the near future by employing alternative data sources

of traffic flow information, such as the information extracted from traffic cameras using computer vision techniques (Drake, 2018). Secondly, the traffic speed data from Google (or other mobility data providers) are not free, which may require the cost of acquiring data to be budgeted in for large-scale studies. However, this method is expected to be cheaper and more scalable than collecting the speed/delay data from conventional methods, such as the floating car data method.

Nonetheless, the proposed methodology could still be beneficial in understanding the performance of key network locations, such as bridges or major routes, when only point-based traffic measurements are available. In addition, if more types of traffic sensors are installed in the future (e.g., radar or camera), the proposed methodology could then be applied on a larger scale.

## 8. Conclusions

By combining two disparate data sources from the ATCs and Google's traffic speed data, context-specific volume–delay curves for a range of different locations and road types have been generated for selected roads in Greater London. The derived volume–delay curves informed by real observational data have shown significant improvements in capturing individual road characteristics compared to the standard model. Apart from obtaining the site-specific knowledge, in a more practical use-case, the derived functions could now be used in the traffic assignment stage of the traditional four-step model. In some cases, the data present a clear evidence that unknown external factors, such as those listed in Section 7.3, have a significant impact on the traffic behavior and warrant further investigation. In these cases, both the derived functions and the standardized curve deviate significantly from empirical data and as such their use should be considered with care.

Overall, this paper has demonstrated the feasibility of linking the ATC traffic count data and Google's traffic speed data for better characterizing the volume–delay relationship across a range of sites. These data sources have longevity and are available at real-time. In the case of the real-time traffic speed data, they can be obtained from partnership with mobility service providers such as Google or its competitors (e.g., Bing, TomTom), leveraging their mature platforms and products. The methods presented in this paper may be employed over a long time horizon and at a finer temporal resolution in order to better understand the recurrent temporal and spatial trends as well as the impacts of special influencing factors, such as sporting occasions and weather events. The automation of this method over longer time horizons may lead to better explanations for various external factors influencing traffic flow behavior and highlight areas that require investigation in order to better understand the performance of road infrastructure.

## References

**Akçelik R** (1991) Travel time functions for transport planning purposes: Davidson's function, its time dependent form and alternative travel time function. *Australian Road Research 21*(3) 49–59.

**Apple Inc.** (2020) Apple Maps. Available at https://www.apple.com/ios/maps/ (accessed 1 July 2020).

**Artikis A**, **Weidlich M**, **Schnitzler F**, **Boutsis I**, **Liebig T**, **Piatkowski N**, **Bockermann C**, **Morik K**, **Kalogeraki V**, **Marecek J**, **Gal A**, **Mannor S**, **Kinane D and Gunopulos D** (2014) Heterogeneous stream processing and crowdsourcing for urban traffic management. In *Proceedings of the 17th International Conference on Extending Database Technology (EDBT)*, Athens, Greece. volume *14*, pp. 712–723.

**Barth D** (2009) The Bright Side of Sitting in Traffic: Crowdsourcing Road Congestion Data. Available at https://googleblog. blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html. (accessed 1 July 2020).

**Brownlee J** (2020) 10 Clustering Algorithms with Python. Available at https://machinelearningmastery.com/clustering-algorithms-with-python/ (accessed 1 July 2020).

**Bureau of Public Roads** (1964) *Traffic Assignment Manual for Application with a Large, High Speed Computer*. US Department of Commerce. Washington, D.C.

**Chiu YC**, **Bottom J**, **Mahut M**, **Paz A**, **Balakrishna R**, **Waller T and Hicks J** (2011) Dynamic traffic assignment: a primer. *Dynamic Traffic Assignment: A Primer*. Transport Research Board. Available at https://onlinepubs.trb.org/onlinepubs/circulars/ec153.pdf (accessed Dec 8, 2020).

**Çolak S**, **Lima A and González MC** (2016) Understanding congested travel in urban areas. *Nature Communications 7*(1), 1–8.

**Crontab Documentation** (2012) Linux Man-Pages. Available at https://man7.org/linux/man-pages/man5/crontab.5.html (accessed 1 July 2020).

**Davidson K** (1966) A flow travel time relationship for use in transportation planning. In *Australian Road Research Board (ARRB) Conference, 3rd, 1966, Sydney*, volume 3.

**DfT** (2002) COBA 11 User Manual Part 5: speeds on links. London, UK. Available at http://www.dft.gov.uk/pgr/economics/software/coba11usermanual/part5speedsonlinks.pdf (accessed Dec 8, 2020).

**DfT** (2012a) Guidance on Road Classification and the Primary Route Network.

**DfT** (2012b) National Travel Survey: 2012.

**DfT** (2020) TAG Unit M3.1: Highway Assignment Modelling. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/888363/tag-unit-m3.1-highway-assignment-modelling.pdf (accessed 28 October 2020).

**Drake A** (2018) Opencv Traffic Counter. Available at https://github.com/alex-drake/OpenCV-Traffic-Counter (accessed 28 October 2020).

**ECMA International** (2017) The JSON Data Interchange Syntax. Available at https://www.json.org/json-en.html (accessed 1 July 2020).

**Fan S** (2020) Generic Second Order Models. Available at http://publish.illinois.edu/shimao-fan/research/generic-second-order-models/ (accessed 1 July 2020).

**Google Cloud** (2020) Google Directions API. Available at https://cloud.google.com/maps-platform/routes (accessed 1 July 2020).

**Hirschmann K**, **Zallinger M**, **Fellendorf M and Hausberger S** (2010) A new method to calculate emissions with simulated traffic conditions. In *13th International IEEE Conference on Intelligent Transportation Systems*, pp. 33–38. IEEE. Funchal, Portugal.

**Google Inc**. (2020) Google Maps. Available at https://www.google.com/maps (accessed 1 July 2020).

**Irawan MZ**, **Sumi T and Munawar A** (2010) Implementation of the 1997 Indonesian Highway Capacity Manual (MKJI) volume delay function. *Journal of the Eastern Asia Society for Transportation Studies 8*, 350–360.

**Kucharski R and Drabicki A** (2017) Estimating macroscopic volume delay functions with the traffic density derived from measured speeds and flows. *Journal of Advanced Transportation*, 2017, 4629792. https://doi.org/10.1155/2017/4629792

**Kurth DL**, **Van den Hout A and Ives B** (1996) Implementation of Highway Capacity Manual-based volume–delay functions in regional traffic assignment process. *Transportation Research Record 1556*(1), 27–36.

**Lien J**, **Mazalov V**, **Melnik A and Zheng J** (2016) Wardrop equilibrium for networks with the BPR latency function. In: Kochetov Y., Khachay M., Beresnev V., Nurminski E., Pardalos P. (eds) Discrete Optimization and Operations Research. DOOR 2016. Lecture Notes in Computer Science, vol *9869*. Springer, Cham. https://doi.org/10.1007/978-3-319-44914-2_4Vynck (2017) is a thesis. If location is needed, it will be Ghent, Belgium.

**Lo HK and Szeto WY** (2005) Road pricing modeling for hyper-congestion. *Transportation Research Part A: Policy and Practice 39*(7–9), 705–722.

**Mashros N**, **Ben-Edigbe J**, **Hassan SA**, **Hassan NA and Yunus NZM** (2014) Impact of rainfall condition on traffic flow and speed: a case study in Johor and Terengganu. *Jurnal Teknologi 70*(4), 65–69.

**Microsoft** (2020) Bing Maps. Available at https://www.bing.com/maps (accessed 1 July 2020).

**Mtoi E and Moses R** (2014) Calibration and evaluation of link congestion functions: applying intrinsic sensitivity of link speed as a practical consideration to heterogeneous facility types within urban network. *Journal of Transportation Technologies*, *4*, 141–149. doi: 10.4236/jtts.2014.42014.

**Mullick A and Ray AK** (2012) Dynamics of Bimodality in Vehicular Traffic Flows. *arXiv preprint arXiv:1205.2314.*

**Neuhold R and Fellendorf M** (2014) Volume delay functions based on stochastic capacity. *Transportation Research Record 2421*(1), 93–102.

**Nie Y**, **Zhang H and Lee DH** (2004) Models and algorithms for the traffic assignment problem with link capacity constraints. *Transportation Research Part B: Methodological 38*(4), 285–312.

**Rose G**, **Taylor MA and Tisato P** (1989) Estimating travel time functions for urban roads: options and issues. *Transportation Planning and Technology 14*(1), 63–82.

**Scikit-Learn Developers** (2019) Clustering: DBSCAN. Available at https://scikit-learn.org/stable/modules/clustering.html#dbscan (accessed 1 July 2020).

**Skabardonis A and Dowling R** (1997) Improved speed–flow relationships for planning applications. *Transportation Research Record 1572*(1), 18–23.

**Spiess H** (1990) Conical volume–delay functions. *Transportation Science 24*(2), 153–158.

**Suh S**, **Park CH and Kim TJ** (1990) A highway capacity function in Korea: measurement and calibration. *Transportation Research Part A: General 24*(3), 177–186.

**TfL** (2010) Traffic Modelling Guidelines. version 3.0.

**The UK Highways Agency** (1999) The Design Manual for Roads and Bridges (Volume 5) Assessment and Preparation of Road Schemes.

**The UK Met Office** (2003) Met Office Rain Radar Data from the NIMROD System.

**Tisato P** (1991) Suggestions for an improved Davidson travel time function. *Australian Road Research 21*(2), 85–100.

**TomTom N.V.** (2020) TomTom Maps. Available at https://www.tomtom.com/en_us/drive/maps-services/maps/ (accessed 1 July 2020).

**Toole JL**, **Colak S**, **Sturt B**, **Alexander LP**, **Evsukoff A and González MC** (2015) The path most traveled: travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies 58*, 162–177.

**van Wageningen-Kessels F**, **Van Lint H**, **Vuik K and Hoogendoorn S** (2015) Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics 4*(4), 445–473.

**Vap D and Sun C** (2007) Investigating Large Truck-Passenger Vehicle Interactions (Report No. FHWA MO-2007-00X). Missouri Department of Transportation. https://spexternal.modot.mo.gov/sites/cm/CORDT/or08005.pdf

**Vynck M** (2017) Heteroscedasticity in Linear Models: An Empirical Comparison of Estimation Methods. Master's thesis, Ghent University. Available at https://lib.ugent.be/catalog/rug01:002376288 (accessed Dec 8, 2020).

**Wang Y and Nihan NL** (2003) Can single-loop detectors do the work of dual-loop detectors? *Journal of Transportation Engineering 129*(2), 169–176.

**Wang Z and Liu C** (2005) An empirical evaluation of the loop detector method for travel time delay estimation. *Journal of Intelligent Transportation Systems 9*(4), 161–174.

**Zhang J**, **He S**, **Wang W and Zhan F** (2015) Accuracy analysis of freeway traffic speed estimation based on the integration of cellular probe system and loop detectors. *Journal of Intelligent Transportation Systems 19*(4), 411–426.

**Zheng Y**, **Chen Y**, **Li Q**, **Xie X and Ma WY** (2010) Understanding transportation modes based on GPS data for web applications. *ACM Transactions on the Web (TWEB) 4*(1), 1–36.