Fitting Empirical Fundamental Diagrams of Road Traffic: A Comprehensive Review and Comparison of Models Using an Extensive Data Set

Daniel M. Bramich●, Mónica Menéndez●, *Member, IEEE*, and Lukas Ambühl

Abstract— Understanding the inter-relationships between traffic flow, density, and speed through the study of the fundamental diagram of road traffic is critical for traffic modelling and management. Consequently, over the last 85 years, a wealth of models have been developed for its functional form. However, there has been no clear answer as to which model is the most appropriate for observed (i.e. empirical) fundamental diagrams and under which conditions. A lack of data has been partly to blame. Motivated by shortcomings in previous reviews, we first present a comprehensive literature review on modelling the functional form of empirical fundamental diagrams. We then perform fits of 50 previously proposed models to a high quality sample of 10 150 empirical fundamental diagrams pertaining to 25 cities. Comparing the fits using information criteria, we find that the non-parametric Sun model greatly outperforms all of the other models. The Sun model maintains its winning position regardless of road type and congestion level. Our study, the first of its kind when considering the number of models tested and the amount of data used, finally provides a definitive answer to the question "Which model for the functional form of an empirical fundamental diagram is *currently* the best?". The word "currently" in this question is key, because previously proposed models adopt an inappropriate Gaussian noise model with constant variance. We advocate that future research should shift focus to exploring more sophisticated noise models. This will lead to an improved understanding of empirical fundamental diagrams and their underlying functional forms.

Index Terms— Empirical fundamental diagrams, review, flow-density relationship, speed-density relationship, loop detector data, model comparisons.

I. INTRODUCTION

THE macroscopic description of traffic flow along a road requires the definition of quantities characterising the average properties of the vehicular traffic at a specific location x and time t; namely, traffic density $k(x, t)$ as the number of vehicles per unit length at time t (veh km^{-1}), traffic flow

$q(x, t)$ as the number of vehicles passing location x per unit time (veh h^{-1}), and space-mean speed $v(x, t)$ as the arithmetic mean of the instantaneous speeds of the vehicles over some distance (km h^{-1}) [1], [2]. Following naturally from their definitions, these three quantities are related to each other by the fundamental equation of traffic flow:

$$q(x, t) = k(x, t) \, v(x, t) \qquad (1)$$

If a speed-density model is adopted such that speed is a function of density only [3], [4], then one may write:

$$q(k) = k \, v(k) \qquad (2)$$

Haight [5] named the flow-density diagram and the relationship therein as the "Fundamental Diagram of Road Traffic".

The importance of the fundamental diagram (FD), originally conceived for uninterrupted flow facilities, lies in traffic modelling (e.g. [2], [6]–[8]) and management (e.g. [9], [10]). While there is value in understanding the theoretical FD relationship for uninterrupted flow facilities, we also believe that it is of great benefit to be able to correctly model the observed FD for any road. This relates more to the practicalities of traffic management and real-time control both at the individual link and urban network levels (e.g. [11]–[13]). For instance, the flow-density FD, which has a concave functional form (often approximated by a triangle), exhibits a peak flow q_{cap} (the capacity) at a certain density k_{crit} (the critical density), with zero flow at zero density and at jam density k_{jam}. Therefore, it is to the benefit of the road users for the traffic to be regulated on each road so as to minimise congestion (i.e. to avoid $k > k_{\mathrm{crit}}$) and avert gridlock ($k = k_{\mathrm{jam}}$). This is only possible with a clear understanding of the actual inter-relationships between traffic flow, density, and speed on any road under a wide range of conditions.

The functional form of the observed FD, referred to hereafter as the empirical fundamental diagram (EFD), usually differs from that of the theoretical FD due to the existence of flow interruptions (e.g. intersections, bottlenecks) and the effects of non-stationary traffic conditions. Even precise data gathered from hypothetical identical driver-vehicle units would not necessarily lie on the theoretical FD, but below it due to the dynamic nature of traffic flow. Since this paper deals with modelling EFDs, whenever we refer to the functional form of the FD relationship, we are referring to that of the EFD.

There is an extensive literature on the study of EFDs. Importantly, the EFD in flow-density, speed-density, and

speed-flow form has been observed to exist for different road types (i.e. highways, freeways, urban roads) in a large number of studies (e.g. [14]–[19]). Starting with Greenshields [14], this has lead to the publication of many papers proposing a plethora of different models that can be used to fit EFD data. However, most of these studies have relied on limited data sets from just a few detectors that are hardly enough to claim the superior performance of the models proposed to emulate the observations (e.g. [20], [21]). Furthermore, only small subsets of available candidate models have been compared in any single study (e.g. [22]), while in some cases, invalid statistical comparisons between models with differing numbers of parameters have been made (e.g. Wang *et al.* [23] compare model performance using mean relative error). We also note that most of the existing studies have focussed on fitting EFD data for freeways, and only very few have used urban data (i.e. from interrupted flow facilities).

It is blatantly clear then that the question as to which model(s) for the functional form of the FD are the most appropriate for EFDs, including how this may depend on road topology, is still unresolved. This is the issue that we aim to tackle with this work, and the path to an answer lies in testing the performance of as many as possible of the existing models on a very large and varied data set. The intention here is not to obtain an estimate of the theoretical FD, but instead to find the best model for representing the observed data. As to be expected, traffic observations have improved in quality and quantity over time, especially as more and more cities begin to collect and store ever increasing amounts of traffic data. Yet it is only now that the traffic research community has started to capitalise on this "big data" deluge (e.g. [24], [25]). The Urban Traffic Data 2019 (UTD19) data set [24] is particularly suitable for discriminating between FD models, since it is currently the largest and most diverse data set on EFDs, pertaining to 41 cities worldwide and 23,767 detectors located on several different types of road.

Having chosen an appropriate data set, the next logical step is to identify a comprehensive list of FD models to be tested. In our own investigation of the literature on the subject, we found that reviews such as those by Gerlough and Huber [26] and Carey and Bowers [27] were useful for guiding our reading. However, it became clear to us that the available reviews, the latest of which is now somewhat out of date, have a number of gaps and shortcomings that require addressing. For instance, none of the reviews cover all of the published FD models and some are inevitably left out. Furthermore, small errors in formulae and a failure to credit some of the most popular FD models to the correct researchers are issues that are being propagated forwards through the literature. All of this provides strong motivation for including our own improved and up-to-date review in this paper.

It is important to realise that any model for an EFD involves assumptions about the noise properties of the data, even though these assumptions are rarely stated explicitly (or even recognised by the practitioners!). Since our work in this paper is focussed on reviewing and comparing previously proposed models, we adopt the same curve fitting method as that employed in the vast majority of previous studies; namely,

the method of least squares, which is equivalent to adopting a Gaussian noise model with constant variance and applying maximum likelihood (ML) estimation (Section III-D). Even though it is conventional to use the Gaussian noise model, it is now known that it is a rather poor choice for modelling EFD data because of the complicated properties of the observed noise (e.g. due to driver behaviour, traffic dynamics, hysteresis effects, etc., [28]–[30]). Some researchers have argued that an EFD should be "cleaned" prior to fitting by discarding measurements taken during non-stationary traffic conditions (e.g. [31], [32]). While this goes some way to simplifying the noise in EFDs, the noise properties after data filtering are still far from Gaussian with constant variance (e.g. see Figs. 8 & 9 in Cassidy [31]). Furthermore, since the filtering step requires detailed analysis of high-resolution data from before aggregation, this approach is infeasible (or highly cumbersome) for many valuable data sets. Therefore, in practice, a significant number of studies on fitting EFDs use unfiltered data (e.g. [20], [23]). The results and conclusions derived from fits employing an unsuitable noise model, like those presented in Section V, should be evaluated bearing in mind this caveat. An extension of our investigation to include alternative and more sophisticated noise models that can better account for the observed scatter in EFDs (filtered or not), while sorely needed, is beyond the scope of this paper.

In Section II, we describe the subset of data from UTD19 that we select for the analysis. In Section III, we gather together and review all of the previously proposed speed-density and flow-density models for the functional form of the FD, where we also identify the gaps and shortcomings in previous reviews. We outline the methods used to fit the data and compare the models in Section IV, and we present and discuss our results in Section V. Finally, in Section VI, we provide a brief set of conclusions and recommendations for future research directions.

II. LOOP DETECTOR DATA

Loop detectors (LDs) are inductive wire loops embedded in the road surface, covering one or more lanes, that are connected to a controller cabinet at the side of the road. They work by detecting the change in loop inductance caused by the ferrous material in the vehicles passing over them. In a given time interval, an LD reports the number of vehicles that have passed over it (called the *flow* and commonly quoted as a number of vehicles per hour per lane), and the fraction of time that a vehicle was within its detection zone (called the *occupancy*, which is a number in the range 0 to 1). It is worth noting that LDs are prone to various types of malfunction (e.g. [33], [34]), and even under "normal" conditions, they may yield noisy measurements (e.g. they can be confused by multi-axle trucks [32]). Some of the main sources of LD measurement errors include pulse breakup (increases observed flow and reduces observed occupancy [35]) and splashover (increases both observed flow and occupancy [36]).

LDs are a type of point sensor in that they measure traffic states at fixed locations in a road network during specific time intervals (i.e. they provide "temporal" measurements).

While this enables LDs to perform direct measurements of traffic flow, they are unable to measure traffic density directly because density requires measurement at a single instant of time over a specific length of road (i.e. density is a "spatial" measurement). However, the occupancy measured by an LD is approximately proportional to the traffic density, despite being a temporal as opposed to a spatial measurement. The constant of proportionality is the mean (effective) vehicle length as "seen" by the LD (i.e. the length of the vehicle plus the length of the detection zone [37]–[39]). Deviations from proportionality may be observed if the mix of vehicle types and/or speeds varies significantly, especially since vehicle lengths and speeds can be correlated (e.g. a higher truck-to-car ratio during night time combined with the fact that trucks tend to travel at lower speeds than cars [39]). Furthermore, in urban networks, the occupancy measurements are biased by the location of the LD along the road. Specifically, sections of the road where vehicles tend to move slower will result in higher occupancy measurements (e.g. close to the downstream intersection [5], [17], [40]–[43]).

An accurate calibration of occupancy to density therefore requires detailed external information (e.g. vehicle lengths and speeds) that is not often available, or that is only available as a gross approximation (e.g. an estimate of the mean effective vehicle length). Fortunately, in the absence of the necessary external information for the calibration, the approximate proportionality between occupancy and density allows one to model traffic flow as a function of occupancy as a proxy for modelling the calibrated flow-density FD. This is especially important for single LDs that in general provide only flow and occupancy measurements, and it has therefore become standard practice in this situation to model flow (veh h^{-1} lane^{-1}) as a function of occupancy directly (e.g. [16], [20]). This is also the approach that we take in this paper since the data that we use consist only of flow-occupancy measurement pairs.

A. Source of the Data

We use the UTD19 data set for 41 cities and 23,767 detectors curated by Loder *et al.* [24], and we supplement it with an extra year of LD measurements data for Zurich (1st July 2017 to 30th June 2018 inclusive). The combined data set includes 271,656,127 pairs/triples of flow-occupancy, flow-speed, or flow-occupancy-speed measurements, where each measurement pair/triple corresponds to a time interval, typically of 3 or 5 minutes in duration, within which the raw LD signals have been aggregated. Note that we use the flow-occupancy-speed measurements as provided by Loder *et al.* [24] from before the application of the combined noise and outlier filtering (detailed under "Methods" in Loder *et al.* [24]) since the noise filtering introduces strong correlations between consecutive measurements.

Cassidy [31] has demonstrated that by discarding data corresponding to transient/non-stationary traffic conditions from the raw LD signals, the scatter in the aggregated measurements can be reduced. Whether this is desirable or not, this is not possible in our case given that the raw LD data have already been aggregated. In fact, this "cleaning" of the data is often not possible and it is rarely ever done. We also note that the UTD19 data set does not provide any information on the heavy vehicle fractions (and their time-dependencies) for the roads in each city, which may lead to uncontrolled distortions in the EFDs, and as such this may negatively impact our modelling results. However, the data mostly pertain to urban areas where heavy vehicles are not so common, and any impact that they may have on the EFDs should therefore be small. The UTD19 data set was collated from a different source in each city (typically the local authorities) and consequently it is somewhat heterogeneous between cities with regards to quantity, quality, and time coverage.

The data for three of the cities consist of flow measurements from LDs, but no occupancy measurements, and speed measurements from Bluetooth detectors that are not necessarily at the same location as the LDs. Hence we exclude the data for these cities (Groningen, 55 detectors; Melbourne, 1,630 detectors; Utrecht, 1,072 detectors). A further six cities use LDs to provide flow and speed measurements, but they do not provide any occupancy measurements. Speed measurements from single LDs are generally unreliable as they require calibration with information that is usually poorly approximated (i.e. LD lengths, individual car lengths [38], [44]). Furthermore, for these cities, we lack any documentation from the data providers on the procedure(s) they employed to derive speed from flow and occupancy. Hence we also exclude the data for Birmingham (66 LDs), Bolton (166 LDs), Innsbruck (16 LDs), Manchester (181 LDs), Rotterdam (259 LDs), and Torino (399 LDs). Compared to the other cities, Paris has a very long aggregation time-interval of 1 hour, which cannot capture the typical variations in traffic flow, and hence we reject these data (247 LDs). Another three cities only provide data on a single date with too few measurements in total, and we therefore reject the data for Frankfurt (112 LDs), Munich (520 LDs), and Vilnius (581 LDs). Finally, the data for Duisburg (190 LDs) cannot be used because a non-disclosure agreement is in force. This leaves a data set with 18,273 LDs across 27 cities.

B. Data Filtering

In the spirit of creating reproducible science, we explain here how we filter the data described in Section II-A. We retain only those LDs for which the road on which they are located has a classification[1] from the set:

```
{ living_street, motorway, motorway_link,
  primary, primary_link, residential,
  secondary, secondary_link, service,
  tertiary, tertiary_link, trunk,
  trunk_link, unclassified }
```

This leads to the rejection of 133 LDs. If an LD covers more than a single lane, then the occupancy that is measured is very difficult to interpret, and it is unclear that it would be proportional to the traffic density. Therefore we reject the 3,043 LDs for which the number of lanes that they cover is undefined or greater than one. We reject 2 LDs in Toulouse because they

---

[1]Road classifications in UTD19 follow the Open Street Map convention - https://wiki.openstreetmap.org/wiki/Key:highway
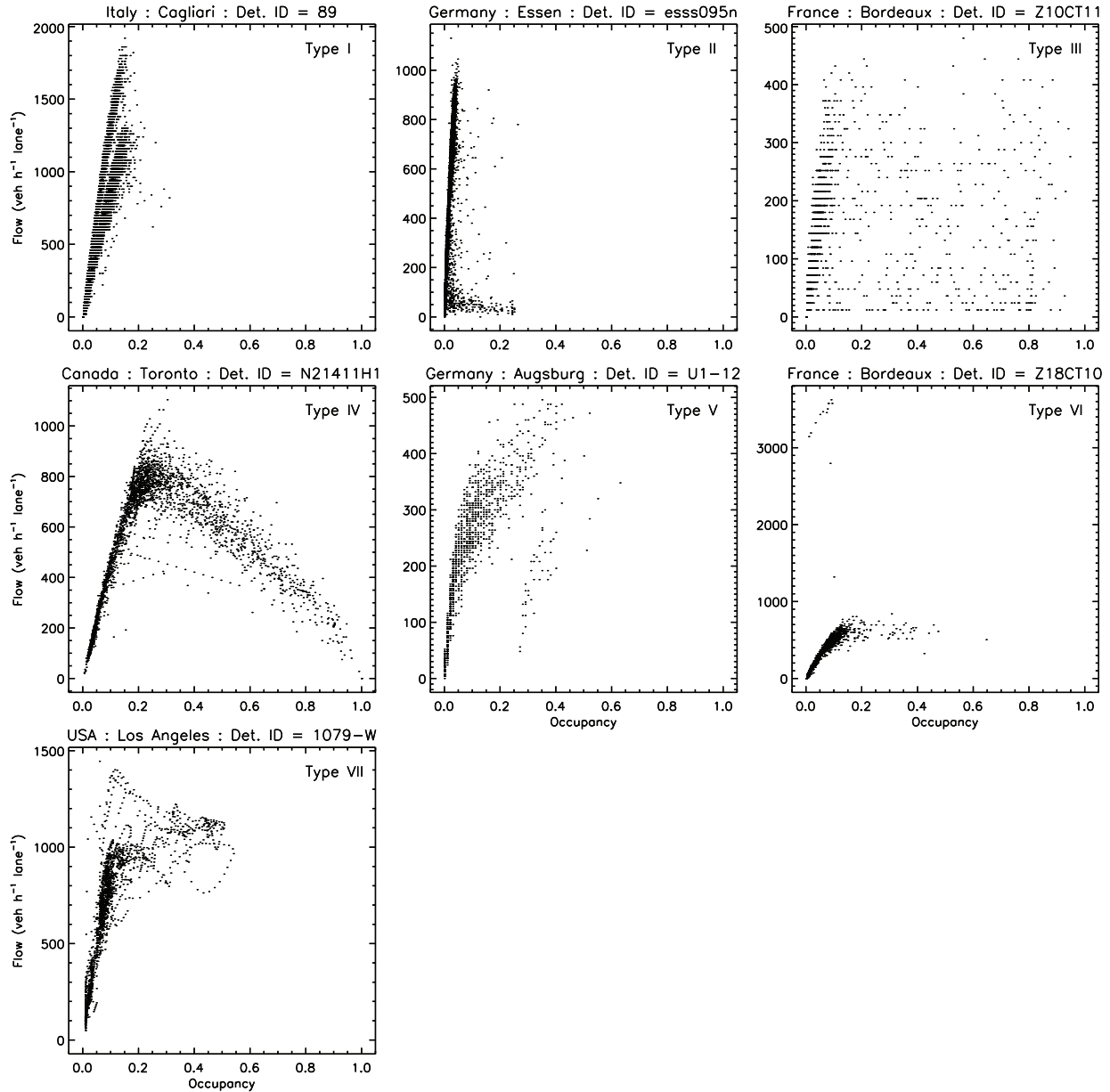
Fig. 1. Examples of EFDs that exhibit the problematic features described in Section II-B. The category into which each of these example EFDs has been classified is indicated in each plot.

have duplicated detector IDs. We find that 138 LDs do not have any associated flow-occupancy measurements, and that for 2,775 LDs, all of the flow-occupancy measurements are flagged as errors. Hence we also reject these LDs. Furthermore, we reject 924 flow-occupancy measurement pairs for which no information about the corresponding time interval has been recorded. Finally, to ensure that each LD provides sufficient data for all of our model fitting and comparison purposes (see Section IV-B), we require that an LD has at least 900 flow-occupancy measurement pairs that are not flagged as errors (i.e. 900 "good" measurements). This results in the rejection of 605 LDs.

On visual inspection of the flow-occupancy EFDs for the remaining LDs, we observe that a non-negligible number of EFDs exhibit features that will either require special modelling, or that are indicative of a malfunctioning LD.

We classify these problematic EFDs into seven broad types. Examples of EFDs for each of the below categories are displayed in Fig. 1.

I. Most likely due to changing speed limits, some EFDs exhibit two or more free-flow branches. Efforts to model such EFDs will require special treatment that is beyond the scope of this paper, and so we reject 426 LDs with EFDs that fall into this category.

II. EFDs in this category exhibit a second branch of points at very low (or zero) flow that is independent of occupancy. In some cases, intermittent periods of queueing (low flow, medium/high occupancy) could be the source of the second branch. The remainder of these EFDs can be explained by numerous other causes (e.g. a malfunctioning LD, an incident, etc.). Regardless of the underlying causes, fitting multiple branches in EFDs

TABLE I

PROPERTIES OF THE FINAL FILTERED DATA SAMPLE DESCRIBED IN SECTION II-B

Country	City	No. Of LDs	No. Of Measurements	No. Of Good Measurements	Duration Of Time Interval (min)	Start Date	End Date	No. Of Days With Measurements
Austria	Graz	223	632,160	549,281	5	2016-04-04	2016-09-23	10
Canada	Toronto	70	398,304	351,442	15	2016-09-01	2017-01-31	61
France	Bordeaux	195	392,832	347,707	5	2016-11-21	2016-11-27	7
France	Marseille	60	864,000	759,673	3	2017-06-01	2017-07-01	31
France	Toulouse	268	900,480	730,195	3	2008-05-16	2008-06-27	7
Germany	Augsburg	355	1,833,843	1,309,142	5	2017-05-06	2017-05-25	20
Germany	Bremen	419	2,790,010	2,445,416	3	2016-09-19	2016-10-02	14
Germany	Constance	84	169,344	159,217	5	2017-02-13	2017-02-19	7
Germany	Darmstadt	189	2,701,000	1,966,707	3	2015-09-21	2016-05-06	39
Germany	Essen	29	235,888	234,446	5	2017-03-27	2017-09-30	35
Germany	Hamburg	294	14,554,197	10,099,450	3	2016-08-27	2016-12-09	105
Germany	Kassel	358	400,880	383,637	5	2016-08-28	2016-09-02	6
Germany	Speyer	134	899,392	781,421	3	2016-09-19	2016-10-02	14
Germany	Stuttgart	173	385,632	352,732	5	2016-03-21	2016-07-22	8
Germany	Wolfsburg	106	712,245	606,292	3	2016-09-19	2016-10-02	14
Italy	Cagliari	58	1,353,196	1,313,489	3	2016-05-16	2016-07-29	50
Japan	Tokyo	1,295	23,123,520	19,557,780	2.5	2017-07-01	2017-07-31	31
Spain	Madrid	315	1,814,400	1,086,628	5	2016-08-29	2017-11-11	20
Spain	Santander	60	190,080	115,283	5	2016-06-17	2016-12-02	11
Switzerland	Basel	55	110,880	105,258	5	2016-10-24	2016-10-31	8
Switzerland	Bern	342	691,314	587,992	5	2016-10-24	2016-10-31	8
Switzerland	Luzern	127	21,785,925	16,913,775	3	2015-01-01	2015-12-31	365
Switzerland	Zurich	761	65,910,654	65,826,865	3	2015-10-26	2018-06-30	372
Taiwan	Taipeh	84	490,783	444,306	3	2017-09-18	2017-10-01	14
United Kingdom	London	4,096	24,842,716	20,006,512	5	2015-05-15	2016-05-22	23
Total	-	10,150	168,183,675	147,034,646	-	-	-	-

is beyond the scope of this paper, and we reject the 267 LDs with EFDs in this category.

III. We reject 57 malfunctioning LDs for which the flow-occupancy measurements form a cloud of seemingly random points in the EFD with very little or no discernable structure.

IV. The "straight line" structures of flow-occupancy measurements observed in these EFDs are the result of Loder *et al.* [24] using linear interpolation to fill in missing data over relatively large data gaps. The interpolated measurements contaminate the bona fide measurements with highly correlated samples that cannot be modelled using the methods presented in this paper. We reject the 164 affected LDs.

V. A fraction of the occupancy measurements suffer from a systematic zero-point offset that shifts them to higher occupancy values, creating a horizontally shifted "copy" of the EFD structure. This is clearly an LD controller malfunction and we reject the 293 affected LDs.

VI. A fraction of the flow measurements suffer from a systematic zero-point offset that shifts them to higher flow values, creating a vertically shifted "copy" of the EFD structure. Again, this is clearly an LD controller malfunction and we reject the 18 affected LDs.

VII. The "loop" structures observed in these EFDs are a problem that is unique to the Los Angeles LDs. No other cities are affected. The Los Angeles LD flow-occupancy measurements suffer from a large fraction of missing data, and these have been smoothly interpolated over by Loder *et al.* [24]. However, this contaminates the bona fide measurements with highly correlated samples that cannot be modelled using the methods presented in this paper. We therefore reject all 202 LDs for Los Angeles.

Notice that all of these features are important in the context of this paper as our focus is not on network aggregated metrics, but on proper fitting of individual EFDs.

Our final data sample consists of 10,150 LDs from 25 cities with 168,183,675 flow-occupancy measurement pairs, of which 147,034,646 measurement pairs are good (i.e. not flagged as errors). We provide a summary of the properties of these data in Table I, and a breakdown by road classification of the LD sample in Table II. Note that the majority of the LDs are located on interrupted flow facilities (i.e. roads with intersections) which are typically monitored for the purpose of controlling the traffic flow on to the highest level roads. This filtered version of the UTD19 data set can be downloaded from the Harvard Dataverse [45], while the code used to generate the data sample is available on GitHub at https://github.com/danlegend5/LDRD_41_Cities_Pipeline.

III. LITERATURE REVIEW ON MODELLING EMPIRICAL FUNDAMENTAL DIAGRAMS

Due to the importance of the FD, there happens to be a vast literature on modelling flow-density-speed inter-relationships. Before we delve into the review, it is important to emphasise that any model for an EFD actually consists of two main components; namely, a model component that specifies the functional form of the FD relationship, and a model component that specifies the noise properties. The model component for the functional form is what is discussed in depth in the literature. Regrettably, the model component for the noise is rarely even recognised, although it is implicitly defined by the adopted fitting procedure(s). A few notable exceptions to this are the papers by Inman [46], Heydecker & Addison [16], Wang *et al.* [47], Ni *et al.* [48], and Bai *et al.* [49].

TABLE II
BREAKDOWN BY ROAD CLASSIFICATION OF THE LDs IN THE FINAL
FILTERED DATA SAMPLE DESCRIBED IN SECTION II-B

Road Classification	Road Grouping	No. Of LDs
motorway	Arterial	2
motorway_link	Arterial	50
trunk	Arterial	297
trunk_link	Arterial	37
primary	Arterial	2,798
primary_link	Arterial	46
Subtotal	Arterial	3,230
secondary	Collector/Distributor	2,095
secondary_link	Collector/Distributor	29
tertiary	Collector/Distributor	3,471
tertiary_link	Collector/Distributor	0
Subtotal	Collector/Distributor	5,595
residential	Local	952
living_street	Local	3
service	Local	18
unclassified	Local	352
Subtotal	Local	1,325
Total	All	10,150

A. Motivation and Scope

The data that we analyse in this paper are flow-occupancy measurement pairs. Therefore we review all of the models in the literature for the functional form of the FD relationship that are stated analytically in flow-density form with q expressed as a function of k, or that can be rearranged as such. Numerically solving equations to determine the functional form is beyond the scope of this paper, and we generally do not consider any models that do not provide explicit analytical expressions between flow, density, and speed. We also include in our review any models for the functional form of the FD relationship that are stated analytically in speed-density form with v expressed as a function of k because they can be trivially transformed into flow-density form by using Equation 2. Models for the functional form that are in density-flow, density-speed, flow-speed, or speed-flow form can rarely be manipulated via a direct analytical transformation into flow-density form, while specifically the density-flow and speed-flow forms have the further problem that they are multi-valued for each flow value. Hence, we generally do not consider these models, which are covered in [50]–[59]. We have tried to make this list reasonably complete, although it is quite possible that we may have missed some relevant papers given the extent of the available literature.

There are a handful of papers in the literature on numerical simulations of traffic flow where the authors have adopted their own specific formulae for the flow-density functional form with fixed values in place of coefficients that could potentially be allowed to vary (e.g. [60], [61]). In these works, justifications for adopting these formulae are not provided, and their properties and performance when used to fit real data are not studied. It is not within the remit of our paper to generalise these functional forms for the flow-density relationship, and hence we do not include them in our review.

One of the desirable properties of a model component for the functional form $q(k)$ of the flow-density relationship is that it should predict zero flow at zero density (see property P.2 in Section III-B). We believe that this property is indispensable. Therefore we do not consider the relevant models described

in [52], [62], [63] since they do not have this property. Again, we have tried to make this list complete.

We only include the simplest multi-regime models for $q(k)$ in our review, because the sheer multitude of possibilities for multi-regime models (of which many have been proposed in the literature, especially as part of traffic simulation software) makes it nearly impossible to assess them all. Some examples of better known multi-regime models which we do not consider include [7], [64]–[73].

Most of the model components for $q(k)$ that are reviewed in Section III-C are specified with q expressed as an explicit analytical function of k. Such expressions are referred to as "parametric" in the sense that they include a fixed and clearly defined set of parameters to be estimated. Non-parametric[2] smoothing functions, on the other hand, allow the data to determine the functional form of the flow-density relationship, including the amount of smoothing involved (or equivalently the effective number of free parameters [74]). This is done without stringent prior specifications on any structure that may exist within the data. In fact, non-parametric smoothing functions can be viewed as a generalisation of multi-regime models for $q(k)$ where the number of regimes is allowed to vary in a systematic and controlled way.

The idea of employing non-parametric smoothing functions for $q(k)$ is not new. Chen *et al.* [75] proposed using the method of principal curves for generating a non-parametric representation of $q(k)$, and Locally Weighted Scatterplot Smoothing (LOWESS) was applied by Antoniou and Koutsopoulos [76] to fit speed-density data. Neither method is considered in our review. Principal curves are not guaranteed to predict a single value for the flow at each density, and the LOWESS implementation in the software that we use in Section IV-A is too restrictive for our purposes. However, we do consider the monotonically-decreasing penalised B-splines proposed by Sun *et al.* [20] (Section III-C17) for modelling the functional form of the speed-density relationship. Finally, we exclude from our review any machine learning techniques that have been used to model $q(k)$ where training is required (e.g. [77]), since tackling the extra complications involved in the training of such models is beyond the scope of this paper.

Our review concentrates in detail on flow-density relationships, and it is necessary and timely for five main reasons. Firstly, other similar reviews are now somewhat out of date, with the most recent dedicated review being nearly a decade old [27]. Secondly, the immense amount of data available in UTD19 (a factor of ~100-1000 times larger than any of the data sets used in previous studies) enables a large scale comparison of models for $q(k)$ to be carried out for the first time. Thirdly, none of the other reviews are fully comprehensive in that some models for $q(k)$ inevitably get left out. Fourthly, we have noticed that small errors in the formulae for some of the flow-density relationships are being propagated forwards through the literature. For example, the exponent

---

[2]The term "non-parametric" is actually somewhat misleading since all smoothing functions contain (hidden) parameters to be estimated including a dominant parameter that controls the amount of smoothing (i.e. the smoothing parameter).

in the so-called "Drew" model is often written incorrectly as $n + 1/2$ instead of $(n + 1)/2$. Fifthly, there has been a systematic failure in the literature to attribute some of the most important models for $q(k)$ to the correct researchers. In the paper by Gazis *et al.* [78] (hereafter GZ1961), a generalised car-following model is used to derive a family of steady-state flow-density relationships. The models for $q(k)$ now referred to as the "Drew", "Northwestern/Drake", and "Pipes" models are actually special cases of this family, and yet GZ1961 are not referenced/credited appropriately[3] by Drew [79], Drake *et al.* [80], or Pipes [81]. Furthermore, only Drew [79] provides a derivation for the special case that is proposed. Considering only reviews of flow-density relationships, which by now should have clarified this oversight, we find that the failure to attribute these models correctly seems to trace back to the monograph by Gerlough and Huber [26] and the book by May [82]. This is surprising since the author of the latter had previously worked on the GZ1961 family of models (e.g. [83]). Unfortunately, even the modern review by Carey and Bowers [27] continues to propagate this last issue. All of these shortcomings are addressed in the review presented here.

As noted by Del Castillo and Benítez [84], there have generally been two approaches to specifying FD relationships. The first approach is simply to formulate an analytical expression for the functional form that possesses at least some of the desirable properties of the FD relationship under consideration, and that can be fitted to traffic data by adjusting some free parameters. The analytical expression can be bestowed with a phenomenological meaning by relating the free parameters to useful properties of traffic flow (e.g. free-flow speed, jam density, capacity, etc.). The second approach relies on developing an underlying theory of traffic dynamics at a micro-, meso-, or macroscopic scale (e.g. car-following, cellular-automata, gas-kinetic, etc.) that can predict a functional form for the FD relationship (see [85] for a genealogy of traffic flow models). The latter approach is more desirable since it allows traffic flow theories to be tested by observations.

In Section III-B, we describe a set of desirable properties for the functional form $q(k)$ of the flow-density relationship. These will be used to help characterise the previously proposed model components for $q(k)$, which we review in Section III-C in chronological order of when they first appeared in the literature. In Section III-D, we detail the model component for the noise that is inherently assumed when using the method of least squares for fitting. Each model component (functional form or noise) is given an abbreviated name in the relevant subsection for easy reference (e.g. GS1935, GZ1961A, GaussSigCon).

B. Desirable Properties of the Flow-Density Functional Form

Del Castillo and Benítez [84] discuss desirable properties of the model components for the functional forms of the speed-density and flow-density relationships. These properties were conceived from logical considerations about traffic flow, and from observed features of EFDs. We gather the properties

specifically pertaining to the flow-density relationship into a list for the purpose of aiding in the characterisation of previously proposed model components for $q(k)$ (Section III-C). The overall flow-density model, and the desirable properties for its functional form $q(k)$, are only applicable for physically relevant density values k ranging from zero to jam density (i.e. for $0 \leq k \leq k_{jam}$). For some model components, k_{jam} is a model parameter to be estimated, or computed from the estimated model parameters/relationship. Other model components may fail to predict k_{jam}, in which case this quantity can be computed by counting the number of vehicles required to cover 1 km of a single lane of roadway when positioned bumper-to-bumper.

At this point, the following equation, derived by differentiating Equation 2, is required:

$$q'(k) = v(k) + k v'(k) \tag{3}$$

where $(')$ represents the first derivative with respect to k.

The desirable properties of the model components for $q(k)$ are:

P.1 The predicted values of the flow $q(k)$ are strictly non-negative ranging between zero and a maximum called the capacity q_{cap} (which occurs at the critical density k_{crit}). This is written concisely as $0 \leq q(k) \leq q(k_{crit}) = q_{cap}$.

P.2 As vehicle spacing tends to infinity, vehicle speeds tend to the free-flow speed v_{ff}, and density tends to zero. Hence $v(0) = v_{ff}$, and Equation 2 implies that $q(0) = 0$.

P.3 At jam density k_{jam}, vehicles are stopped (i.e. $v(k_{jam}) = 0$), which, using Equation 2, also implies that $q(k_{jam}) = 0$.

P.4 As vehicle spacing tends to infinity, vehicle speeds tend to the free-flow speed v_{ff}, and density tends to zero. Hence $v(0) = v_{ff}$, and Equation 3 implies that $q'(0) = v_{ff}$.

P.5 From a speed of v_{ff} at $k = 0$, the speed decreases with increasing density (i.e. $v(k) \leq v_{ff}$ and $v'(k) \leq 0$). From Equation 3, this implies that $q'(k) \leq v_{ff}$.

P.6 The functional form $q(k)$ of the flow-density relationship is concave for at least some of the density range from zero to jam density, which allows for the formation of back-propagating shock waves. Mathematically, this property translates into $q''(k) < 0$ for some range of k.

P.7 The functional form $q(k)$ is continuous. Whether this property is desirable or not is certainly debatable (see [84] for a discussion).

P.8 The functional form $q(k)$ is differentiable. Again, whether this property is desirable or not is certainly debatable (see [84] for a discussion).

Table III lists the previously proposed model components for $q(k)$ that are described in Section III-C along with the sets of desirable properties that they each possess.

We remind the reader that since occupancy is proportional to density (Section II), any model component for $q(k)$ is also valid when k is replaced by occupancy, although some model parameters may need rescaling appropriately. Furthermore, all of the above considerations P.1 to P.8 for $q(k)$ also apply to

---

[3]This is potentially due to the limited flow of information between researchers at the time.

the flow-occupancy relationship with rescaled v_{ff}, k_{crit}, and k_{jam}. For any model component for $q(k)$ that includes k_{jam} as a free parameter, we also consider a version with k_{jam} fixed, which we identify with "kjf" appended to the model component name. For these "kjf" model components in flow-occupancy form, we make the equivalent assumption that the jam occupancy is fixed at unity.

C. Model Components for the Flow-Density Functional Form

In this section, we gather together and review a comprehensive collection of 50 model components for $q(k)$. All of these model components have been previously proposed in the literature, although we have improved on a few of them through reparameterisation and/or generalisation (e.g. VA1995, BD1995, WG2011A, SN2014, SN2014kjf).

1) The Free-Flow Model (FF): Many of the LDs in our data sample only provide observations in the free-flow uncongested traffic regime (i.e. at low densities). Assuming a constant speed independent of density in this regime, then the corresponding flow-density relationship is given by:

$$q = v_{ff} k \tag{4}$$

where the free-flow speed v_{ff} is the only free parameter. We include this elementary linear model in our collection as a potentially competitive model component for EFDs from LDs that only ever measured traffic in the free-flow regime.

2) The Greenshields Model (GS1935): Based on his pioneering photographic study of road traffic in lightly congested conditions, Greenshields [14] postulated the existence of a linear relationship between average speed and average density.[4] In flow-density form, this relationship may be written as:

$$q = v_{ff} k - \frac{v_{ff}}{k_{jam}} k^2 \tag{5}$$

with free parameters v_{ff} and k_{jam}. For physically meaningful parameter values, this formula specifies an inverted parabola with critical density $k_{crit} = k_{jam}/2$. Although this model component possesses all of the desirable properties P.1 to P.8 (Table III), its simple form imposes some limitations such as the fact that the back-propagating wave speed at jam density is forced to match the free-flow speed. This model component is linear in the explanatory variables k and k^2.

3) The Greenberg Model (GB1959): Assuming that a traffic stream behaves like a continuous one-dimensional fluid, Greenberg [86] derived the following flow-density relationship:

$$\begin{aligned} q &= v_{bw} k \ln\left(\frac{k_{jam}}{k}\right) \\ &= (v_{bw} \ln k_{jam}) k - v_{bw} k \ln k \end{aligned} \tag{6}$$

with free parameters v_{bw} and k_{jam}, where v_{bw} is the back-propagating wave speed at jam density. A deficiency

---

[4]Interestingly enough, the data used for the formulation of this hypothesis are rather sparse (only seven speed-density measurement pairs) and have varied provenance (they come from three different roads observed on different dates). Furthermore, only one measurement pair has a medium density of $k \approx 90.5$ veh km^{-1}, while the remainder are at low densities $k < 35$ veh km^{-1}.

of this model component is that the speed tends to infinity as the density tends to zero. The critical density is given by $k_{crit} = k_{jam} \exp(-1)$. This model component is linear in the explanatory variables k and $k \ln k$. Gazis *et al.* [87] subsequently showed how to derive Equation 6 from a car-following model.

4) The Edie Multi-Regime Model (ED1961): By refining a car-following model for uncongested traffic, Edie [88] independently derived the UW1961A model (Section III-C5), and then proposed combining it with the GB1959 model to create a two-regime model for $q(k)$ that avoids the principal deficiencies of UW1961A and GB1959. Being a two-regime model, there is a break-point $0 < k_b < k_{jam}$ to be determined such that for $0 \leq k \leq k_b$, the UW1961A model is used, while for $k > k_b$, the GB1959 model is used. This model component is non-linear with five free parameters v_{ff}, k_{crit}, k_b, v_{bw}, and k_{jam}.

5) The Underwood Models (UW1961A-B): Traffic data collected from two different sites were used by Underwood [89] to postulate an exponential speed-density relationship, which in flow-density form, is given by:

$$q = v_{ff} k \exp\left(-\frac{k}{k_{crit}}\right) \tag{7}$$

This model component, which we refer to as UW1961A, has two free parameters v_{ff} and k_{crit}, and it is non-linear.

A deficiency of UW1961A is that the flow never reaches zero as the density increases. To remove this deficiency, Underwood [89] suggested modifying the formula as follows:

$$q = v_{ff} k \exp\left(-\frac{k}{k_{crit}}\right) - a k \tag{8}$$

where a is an extra free parameter to be determined with $0 < a < v_{ff}$. Doing this, however, means that v_{ff} and k_{crit} lose their physical interpretations. The jam density for this model component, which we refer to as UW1961B, is $k_{jam} = k_{crit} \ln(v_{ff}/a)$. The model component UW1961Bkjf with fixed k_{jam} can be formed by substituting the expression $a = v_{ff} \exp(-k_{jam}/k_{crit})$ into Equation 8.

6) The Franklin-Newell Model (FN1961): Newell [90] considered non-linearities in car-following models and analysed a specific example model for the functional form of the speed-density relationship, although the author admitted that "no motivation for this choice is proposed other than the claim that it has approximately the correct shape and is reasonably simple". Independently, Franklin [91] derived an equivalent steady-state flow-density relationship by adopting a car-following model where available acceleration decreases linearly with speed and a driver adjusts the actual acceleration in proportion to the relative velocity of the vehicle ahead. The flow-density formula corresponding to the Newell parameterisation is given by:

$$q = v_{ff} k \left(1 - \exp\left(-\frac{\lambda}{v_{ff}}\left(\frac{1}{k} - \frac{1}{k_{jam}}\right)\right)\right) \tag{9}$$

which is non-linear with three free parameters v_{ff}, λ, and k_{jam}. Note that λ is related to the back-propagating wave speed at jam density via $v_{bw} = \lambda/k_{jam}$. There is no simple analytical expression for k_{crit}.

7) The Gazis Models (GZ1961A-H): By generalising the sensitivity term in a car-following model, GZ1961 were able to derive the following family of steady-state solutions:

$$f_m\left(\frac{q}{k}\right) = c_1 f_l\left(\frac{1}{k}\right) + c_2 \tag{10}$$

where c_1 and c_2 are free parameters. The numbers m and l are fixed constants (not necessarily integers) that define the form of the sensitivity term in the underlying car-following model. The function $f_p(x)$ is defined as follows:

$$f_p(x) = \begin{cases} x^{1-p} & \text{for } p \neq 1 \\ \ln x & \text{for } p = 1 \end{cases} \tag{11}$$

Various combinations of m and l generate model components that have already been considered (i.e. $(m, l) = (0, 2)$ yields GS1935, $(m, l) = (0, 1)$ yields GB1959, and $(m, l) = (1, 2)$ yields UW1961A). However, there are of course an infinite number of possible model components that can be generated by choosing different combinations of m and l. Easa and May [92] provide a useful table of formulae corresponding to the different regions in the (m, l) plane.

We limit ourselves to testing a selection of the potential models that were plotted by GZ1961 in their Figs. 1-3. Each model component has two free parameters taken from the set v_{ff}, v_{bw}, k_{crit}, k_{jam}, and q_{cap}. We call the first model component GZ1961A which corresponds to $(m, l) = (0, 1/2)$:

$$q = 2 v_{bw} k_{jam}^{1/2} k^{1/2} - 2 v_{bw} k \tag{12}$$

A deficiency of this model component is that the speed tends to infinity as the density tends to zero. The critical density is given by $k_{crit} = k_{jam}/4$. This model component is linear in the explanatory variables k and $k^{1/2}$.

For GZ1961B we adopt $(m, l) = (0, 3/2)$:

$$q = v_{ff} k - \frac{v_{ff}}{k_{jam}^{1/2}} k^{3/2} \tag{13}$$

The critical density and back-propagating wave speed at jam density are given by $k_{crit} = 4 k_{jam}/9$ and $v_{bw} = v_{ff}/2$, respectively. This model component is linear in the explanatory variables k and $k^{3/2}$.

For GZ1961C we adopt $(m, l) = (0, 3)$:

$$q = v_{ff} k - \frac{v_{ff}}{k_{jam}^2} k^3 \tag{14}$$

The critical density and back-propagating wave speed at jam density are given by $k_{crit} = \sqrt{3} k_{jam}/3$ and $v_{bw} = 2 v_{ff}$, respectively. This model component is linear in the explanatory variables k and k^3.

For GZ1961D we adopt $(m, l) = (-1, 0)$:

$$q = 2 q_{cap} \left(\frac{k}{k_{jam}}\left(1 - \frac{k}{k_{jam}}\right)\right)^{1/2} \tag{15}$$

A deficiency of this non-linear model component is that the speed tends to infinity as the density tends to zero. Furthermore, the back-propagating wave speed becomes unbounded as the density tends to k_{jam}. The critical density is given by $k_{crit} = k_{jam}/2$.

For GZ1961E[5] we adopt $(m, l) = (-1, 1)$:

$$q = \sqrt{2} \exp\left(\frac{1}{2}\right) q_{cap} \frac{k}{k_{jam}} (\ln k_{jam} - \ln k)^{1/2} \tag{16}$$

A deficiency of this non-linear model component is that the speed tends to infinity as the density tends to zero. Furthermore, the back-propagating wave speed becomes unbounded as the density tends to k_{jam}. The critical density is given by $k_{crit} = k_{jam} \exp(-1/2)$.

For GZ1961F[6] we adopt $(m, l) = (1, 3)$:

$$q = v_{ff} k \exp\left(-\frac{1}{2}\left(\frac{k}{k_{crit}}\right)^2\right) \tag{17}$$

A deficiency of this non-linear model component is that the flow never reaches zero as the density increases.

It is interesting to note that the family of flow-density relationships subsequently derived by Drew [79] from generalising the fluid-flow analogy of Greenberg [86] actually forms a subset of the family of model components defined by Equations 10 & 11. Specifically, they correspond to $m = 0$ and $l > 1$. If one considers l as a free parameter (limited to $l > 1$) as opposed to a fixed constant, then one obtains the following non-linear model component with three free parameters:

$$q = v_{ff} k - \frac{v_{ff}}{k_{jam}^{l-1}} k^l \tag{18}$$

which we refer to as GZ1961G. The critical density and back-propagating wave speed at jam density are given by $k_{crit} = l^{1/(1-l)} k_{jam}$ and $v_{bw} = (l-1) v_{ff}$, respectively. As $l \to 1$ from above, this model component converges to GB1959.

Reading further forwards through the literature, one finds that the family of flow-density relationships proposed by Pipes [81] in his review of car-following models[7] also forms a subset of the family of model components defined by Equations 10 & 11. Specifically, they correspond to $m < 1$ and $l = 2$. If one considers m as a free parameter (limited to $m < 1$) as opposed to a fixed constant, then one obtains the following non-linear model component with three free parameters:

$$q = v_{ff} k \left(1 - \frac{k}{k_{jam}}\right)^{1/(1-m)} \tag{19}$$

which we refer to as GZ1961H. The critical density is given by $k_{crit} = (1 - m) k_{jam}/(2 - m)$. A deficiency of this model component is that, as the density tends to k_{jam}, the back-propagating wave speed becomes unbounded for $m < 0$ and it tends to zero for $0 < m < 1$. For $m = 0$, $v_{bw} = v_{ff}$. As $m \to 1$ from below, this model component converges to UW1961A.

---

[5]Over a decade later, El-Hosaini [93] showed how to derive this formula by assuming that traffic behaves like a compressible gas under piston action. However, no reference to GZ1961 was made.

[6]A few years later, Drake *et al.* [80] explicitly proposed this same formula with "no theoretical foundation" and without credit to GZ1961.

[7]There is no reference to the work of GZ1961 in Pipes [81], and he states incorrectly that his proposed flow-density relationship "is a purely empirical one and not based on any definite following law".

Finally we note that if m and l are both treated as free parameters in Equation 10, then one obtains the most generalised Gazis model. However, this model component is highly non-linear in its four free parameters, and fitting it to data is a very tricky process that requires careful optimisation across nine distinct regions of the (m, l) plane [92]. This task is beyond the scope of this paper.

8) The Drake Multi-Regime Models (DK1966A-B): As part of their comparison study of speed-density relationships using freeway traffic data, Drake *et al.* [80] proposed a set of new multi-regime models.[8] We consider the two simplest of these multi-regime models here. The first, which we refer to as DK1966A, has two regimes with a different linear trend in speed as a function of density in each regime. In flow-density form, this relationship may be written as:

$$q = \begin{cases} v_{\text{ff}}\, k - c\, k^2 & \text{for } 0 \le k \le k_{\text{b}} \\ v_{\text{bw}}\, k - (v_{\text{bw}}/k_{\text{jam}})\, k^2 & \text{for } k > k_{\text{b}} \end{cases} \quad (20)$$

This model component is non-linear with five free parameters v_{ff}, v_{bw}, k_{jam}, c, and k_{b}, where k_{b} is the break-point. DK1966A suffers from the deficiency that it may exhibit two peak flows.

The second multi-regime model, which we refer to as DK1966B,[9] has two regimes comprising of the FF and GB1959 models with continuity imposed at the break-point. In flow-density form, this relationship may be written as:

$$q = \begin{cases} v_{\text{bw}}\, (\ln k_{\text{jam}} - \ln k_{\text{b}})\, k & \text{for } 0 \le k \le k_{\text{b}} \\ (v_{\text{bw}} \ln k_{\text{jam}})\, k - v_{\text{bw}}\, k \ln k & \text{for } k > k_{\text{b}} \end{cases} \quad (21)$$

This model component is non-linear with three free parameters v_{bw}, k_{jam}, and k_{b}, where k_{b} is the break-point. The free-flow speed and critical density are given by $v_{\text{ff}} = v_{\text{bw}} (\ln k_{\text{jam}} - \ln k_{\text{b}})$ and $k_{\text{crit}} = \max\{ k_{\text{jam}} \exp(-1), k_{\text{b}}\}$, respectively.

9) The Munjal Multi-Regime Model (MJ1971): The earliest description that we can find of the ubiquitous triangular flow-density FD is in [95]. This multi-regime flow-density relationship is defined by:

$$q = \begin{cases} v_{\text{ff}}\, k & \text{for } 0 \le k \le k_{\text{crit}} \\ v_{\text{bw}}\, (k_{\text{crit}} - k) + v_{\text{ff}}\, k_{\text{crit}} & \text{for } k > k_{\text{crit}} \end{cases} \quad (22)$$

which is non-linear with three free parameters v_{ff}, k_{crit}, and v_{bw}. The critical density k_{crit} is also the break-point. The jam density is given by $k_{\text{jam}} = (v_{\text{ff}} + v_{\text{bw}})\, k_{\text{crit}}/v_{\text{bw}}$.

10) The Boardman Model (BM1977): Boardman and Lave [52] used highway traffic data to assess various speed-flow, flow-speed, speed-density, and flow-density relationships. One of their proposed speed-density relationships has not been considered so far in this review, and in flow-density form it is given by:

$$q = v_{\text{ff}}\, k \exp(-c_1 k) \exp(-c_2 k^2) \quad (23)$$

This model component is non-linear with three free parameters v_{ff}, c_1, and c_2. The critical density is the positive root (if it exists) of the quadratic equation $2 c_2 k^2 + c_1 k = 1$. A deficiency

---

[8]Two of the multi-regime models were attributed to an unpublished report by Ellis from 1964.

[9]This particular multi-regime model was independently proposed by Dick [94].

---

of this model component is that the flow never reaches zero as the density increases.

11) The van Aerde Model (VA1995): By proposing a relatively simple car-following model for the minimum desired separation between consecutive vehicles, van Aerde [96] derived a flow-density relationship that we have reparameterised in the form:

$$q = \alpha \left(1 - \beta\, k - \left((\gamma\, k - 1)^2 + \delta k^2\right)^{1/2}\right) \quad (24)$$

This flow-density relationship is non-linear and the four free parameters α, β, γ, and δ relate to the original parameterisation in [96] via $\alpha = (2 c_3)^{-1}$, $\beta = c_1 - c_3 v_{\text{ff}}$, $\gamma = c_1 + c_3 v_{\text{ff}}$, and $\delta = 4 c_2 c_3$. Physically meaningful non-negativity constraints on c_1, c_2, and c_3 combined with $v_{\text{ff}} > 0$ translate directly to non-negativity constraints on α, γ, and δ, while β may take on any value. The critical density can be computed by solving a rather complicated quadratic equation in k, and the jam density is given by $k_{\text{jam}} = 2 (\gamma - \beta)/(\gamma^2 - \beta^2 + \delta)$. The model component VA1995kjf with fixed k_{jam} requires an alternative reparameterisation:

$$q = \alpha \left(1 - \left(k_{\text{jam}}^{-1} - \psi - \omega\right) k \right.$$
$$\left. - \left(\left(\left(k_{\text{jam}}^{-1} - \psi + \omega\right) k - 1\right)^2 + 4\, \psi\, \omega\, k^2\right)^{1/2}\right) \quad (25)$$

with three free parameters α, ψ, and ω, where $\psi = c_2/v_{\text{ff}} \ge 0$ and $\omega = c_3 v_{\text{ff}} \ge 0$.

12) The Bando Model (BD1995): In their numerical traffic simulations, Bando *et al.* [97] proposed a simple formula that directly relates vehicle speed to vehicle headway. By including appropriate scale factors in the formula, one obtains the following steady-state flow-density relationship:

$$q = v_{\text{ff}}\, k \left(\frac{\tanh\left(c_1 k^{-1} - c_2\right) + \tanh c_2}{1 + \tanh c_2}\right) \quad (26)$$

where v_{ff}, c_1, and c_2 are the three free parameters, and $c_1 > 0$. A deficiency of this non-linear model component is that the flow converges to a positive value as the density increases. Note that there is no simple analytical expression for k_{crit}.

13) The Del Castillo Models (DC1995A and DC2012B): Del Castillo and Benítez [84] proposed a set of four speed-density formulae as examples of "generating functions" that possess certain properties desirable of a speed-density relationship. However, as the authors themselves noted, the shapes of the "rational", "double-exponential", and "reciprocal-exponential" families of curves "can be very well reproduced by the exponential generating functions". For this reason, and following [84], we also only consider the "exponential" family of speed-density curves, which have the following formula in flow-density form:

$$q = v_{\text{ff}}\, k \left(1 - \exp\left(1 - \left(1 + \frac{v_{\text{bw}}\, k_{\text{jam}}}{m\, v_{\text{ff}}}\left(\frac{1}{k} - \frac{1}{k_{\text{jam}}}\right)\right)^m\right)\right) \quad (27)$$

where v_{ff}, v_{bw}, k_{jam}, and m are the four free parameters, and $m > 0$. This non-linear model component, which we refer to

as DC1995A, is a generalisation of the Franklin-Newell model (Section III-C6) since DC1995A reduces to FN1961 when $m = 1$. We note that there is no simple analytical expression for k_{crit}.

Nearly two decades later, Del Castillo [22] revisited his earlier work with the aim of developing new flow-density formulae that include the Munjal multi-regime model (Section III-C9) as a particular limiting case. From three similar flow-density relationships that were considered, Del Castillo [22] settled on the "negative power model" as the best option for various practical reasons. This non-linear model component, which we refer to as DC2012B, has the formula:

$$q = v_{ff}\, k \left(1 + \left(\frac{v_{bw}\, k_{jam}}{v_{ff}} \left(\frac{1}{k} - \frac{1}{k_{jam}} \right) \right)^{-m} \right)^{-1/m} \tag{28}$$

with four free parameters v_{ff}, v_{bw}, k_{jam}, and m, where $m > 0$. The critical density can be computed analytically in this case using:

$$k_{crit} = k_{jam} \left(1 + \left(\frac{v_{ff}}{v_{bw}} \right)^{m/(m+1)} \right)^{-1} \tag{29}$$

As $m \to \infty$, this model component converges to MJ1971.

14) The Ghandehari Model (GD2008): In a short technical report, Ghandehari and Ardekani [98] proposed a modified Greenberg model by introducing an extra free parameter as follows:

$$q = c_1\, k\, \ln \left(\frac{k_{jam} + c_2}{k + c_2} \right) \tag{30}$$

where k_{jam}, c_1, and c_2 are the three free parameters, and $c_2 \geq 0$. The free-flow speed and the back-propagating wave speed at jam density are given by $v_{ff} = c_1 \ln((k_{jam} + c_2)/c_2)$ and $v_{bw} = c_1\, k_{jam}/(k_{jam} + c_2)$, respectively, while there is no simple analytical expression for k_{crit}. This non-linear model component is equivalent to GB1959 when $c_2 = 0$.

15) The MacNicholas Model (MN2008): Without any particular theoretical justification, MacNicholas [99] proposed the following non-linear flow-density relationship:

$$q = v_{ff}\, k\, \frac{1 - (k/k_{jam})^n}{1 + c\,(k/k_{jam})^n} \tag{31}$$

where v_{ff}, k_{jam}, $c \geq 0$, and $n > 1$ are the four free parameters. The back-propagating wave speed at jam density is given by $v_{bw} = n\, v_{ff}/(1 + c)$, while the critical density can be computed by solving a quadratic equation in $(k/k_{jam})^n$. When $c = 0$, this model component reduces to GZ1961G, and then further converges to GS1935 as $n \to 1$ from above.

16) The Wang Models (WG2011A-C): For aesthetical reasons as opposed to theoretical ones, Wang *et al.* [23] put forward a set of generalised logistic functions as candidates for representing the functional form of the speed-density relationship. We have partly reparameterised the five-parameter formula, which now reads as follows in flow-density form:

$$q = c_1\, k + c_2\, k\, (1 + \exp(c_3\,(k - k_{ref})))^{-m} \tag{32}$$

where c_1, c_2, c_3, k_{ref}, and m are free parameters that control the shape of the curve,[10] and $c_3 \neq 0$ and $m > 0$. A deficiency of this non-linear model component, which we refer to as WG2011A, is that if all of c_2, c_3, and k_{ref} are strictly positive as one might expect, then the formula only predicts a finite jam density when $c_1 < 0$. Otherwise, for $c_1 = 0$, the flow tends to, but never reaches zero, as the density increases, and for $c_1 > 0$, the flow increases without bound. We note that there is no simple analytical expression for k_{crit}.

By setting $m = 1$ in WG2011A, Wang *et al.* [23] obtained a four-parameter model component,[11] which we refer to as WG2011B. Then, by setting $c_1 = 0$ in WG2011B, the three-parameter model component WG2011C is derived. These less flexible model components still suffer from the same deficiencies as those outlined for WG2011A.

17) The Sun Model (SN2014): Sun *et al.* [20] proposed using penalised B-splines (a type of non-parametric smoothing function), constrained to be monotonically decreasing, to model the functional form of the speed-density relationship. In their work, they used a quadratic B-splines basis, varied the number of knots from 2 to 9, and kept the penalty parameter fixed at unity. In this case, it is the number of knots that acts as a smoothing parameter, and because this number is limited to a finite set of discrete values, the full flexibility of these smoothing functions is underexploited.

In our own improved implementation of the Sun model, also constrained to be monotonically decreasing, we employ a cubic B-splines basis with 11 equally spaced knots. The effective number of free parameters contributed to a fit by the penalised B-splines, estimated by computing the trace of the smoother matrix, acts as the smoothing parameter in our case [74], and it is allowed to vary continuously with a minimum value of unity. Estimation of the effective number of free parameters for a regularised model enables its comparison in terms of degrees of freedom with unregularised models, where the latter have discrete numbers of free parameters. Equal knot spacing is important for penalised B-splines [100], and the only impact of our choice for the number of knots is to limit the maximum effective number of free parameters to 13. Let $B_{pen}^-(k)$ represent the monotonically-decreasing penalised B-splines. Due to certain limitations of the modelling software that we use in Section IV-A, in order to force a prediction of zero flow at zero density, we must write the corresponding flow-density formula as follows:

$$q = k\, \exp\left(B_{pen}^-(k) \right) \tag{33}$$

This non-linear model component has no physically meaningful free parameters, and k_{crit} must be estimated numerically. The free-flow speed may be evaluated via $v_{ff} = \exp(B_{pen}^-(0))$. Since the penalised B-splines are not constrained to be *strictly* monotonically decreasing, the flow can potentially diverge in a linear fashion for densities above the maximum observed

---

[10]Contrary to what is stated in [23], the parameters of the generalised logistic function do not have any physical meaning when applied to the FD. For example, c_1 is not the speed at jam density, and $c_1 + c_2$ is not the free-flow speed.

[11]Wang *et al.* [23] are in error when they state that k_{ref} is the critical density for the four- and three-parameter models.

density. A further deficiency that is specific to our implementation of this model component is that the flow never reaches zero as the density increases.

The Sun model, regardless of implementation details, is not guaranteed to predict a jam density. To construct the model component SN2014kjf and force the prediction of zero flow at a fixed jam density, we rewrite Equation 33 as follows:

$$q = k \left(1 - \frac{k}{k_{\mathrm{jam}}} \right) \exp \left(B_{\mathrm{pen}}^{-}(k) \right) \qquad (34)$$

While it is certainly possible to treat k_{jam} as a free parameter in this formula, we have not yet managed to develop a practical implementation of this feature using the software tools available to us.

D. The Gaussian Noise Model With Constant Variance (GaussSigCon)

The noise that is observed in an EFD originates from the stochasticity inherent to the underlying system dynamics (e.g. driver behaviour, driver types, traffic composition, transient states, hysteresis effects, etc.) combined in some measure with the uncertainty introduced during data collection and/or processing. To fit curves to the noisy data, regardless of whether measurements taken during non-stationary traffic conditions have been rejected or not, all but a few of the studies on EFDs have employed the method of least squares, which consists of adjusting the curve parameters to minimise the sum of the squared residuals from the fitted curve. This is equivalent to:

  (i) assuming that the observations of the dependent variable (i.e. the flow in our case) are independent of each other and that they all follow a Gaussian distribution,
 (ii) assuming that the Gaussian distribution has a mean that matches the curve to be fit, and a variance (or standard deviation) that is constant (i.e. independent of density/occupancy in our case),
(iii) adjusting the model parameters to maximise the likelihood function (i.e. ML estimation).

In the literature on modelling EFDs, Inman [46] seems to have been one of the first authors to recognise that "implicit in the ordinary least squares regression analyses of previous studies has been a presumption of a normal error structure for the dependent variable". For this reason, the author assumes that "transformed observations of volume ... are independently normally distributed with constant variance" and then he applies ML estimation to perform the fits.

Using Q_i to denote a univariate random variable for flow corresponding to the ith flow measurement, we may write this particular model component for the noise in the flow observations as:

$$Q_i \overset{\mathrm{ind}}{\sim} \mathcal{N}(q(k_i), \sigma_{\mathrm{con}}) \quad \text{for } i = 1, 2, \ldots, N_{\mathrm{dat}} \qquad (35)$$

where $\mathcal{N}(\mu, \sigma)$ represents a Gaussian distribution with mean μ and standard deviation σ, k_i is the ith density measurement, $q(k_i)$ is the predicted flow at k_i given the adopted model component for the functional form, σ_{con} is the standard deviation, and N_{dat} is the number of flow-density measurement pairs.

This noise model component has a single free parameter σ_{con} (apart from the free parameters in $q(k)$).

The adoption of a Gaussian distribution in the noise model ensures that the predicted distribution of flow values at any given density is unimodal, although it has the drawback of predicting negative flows with a non-zero probability. The assumption of independence between observations is highly likely to be violated to some extent, while the variance in the flow is known to vary as a function of density[12] (e.g. [31], [102], [103]). Furthermore, this noise model component fails to take into account the fact that the density/occupancy measurements also suffer from noise. Clearly, the GaussSigCon noise model component is not particularly suitable for modelling flow-density (or speed-density) EFD data. However, our focus is on reviewing and comparing previously proposed models for the FD, the vast majority of which have used the method of least squares for fitting, and hence we adopt GaussSigCon as the equivalent noise model component in our study. Improving on these models by addressing the deficiencies in GaussSigCon is not part of this work.

IV. METHODS

From this point onwards, we refer to a particular FD model solely by the abbreviated name of its model component for $q(k)$ (e.g. GB1959) since all of the models that we consider in this paper have the same noise model component (i.e. GaussSigCon).

A. Model Fitting

For each of the 10,150 LDs in the data sample that we selected in Section II, we fit all of the models reviewed in Section III to the corresponding flow-occupancy EFDs using ML estimation, or maximum penalised likelihood (MPL) estimation, as appropriate. Flow-occupancy measurement pairs that are flagged as errors are ignored in the fits. The data include 3,196,239 non-flagged flow-occupancy measurement pairs where the occupancy is zero, and 2,758,156 of these also have zero flow. The data with positive flow measurements at zero occupancy are most likely erroneous, while the data with zero flow at zero occupancy do not constrain either the functional form or noise model components of any model since we have required that all functional form model components satisfy property P.2 (Section III-B). Furthermore, occupancy values of zero can cause difficulties in implementing the fitting algorithms for certain models (e.g. evaluating $k \ln k$ in GB1959). Hence we ignore all of the flow-occupancy measurement pairs with zero occupancy in the fits.

While there are many software options available for performing fits via ML/MPL estimation, we have specifically

---

[12]Note that reweighting schemes for least squares fits to EFDs like the one improvised by Qu *et al.* [101] are equivalent to modelling the flow measurements with a non-constant standard deviation in Equation 35. The motivation for the development of such schemes lies in the poor quality of the fits that are obtained with certain $q(k)$ model components when using the GaussSigCon noise model for EFD data that (typically) have a heavy imbalance in the number of flow measurements between the uncongested and congested regimes. However, this imbalance ceases to be an issue when an appropriate noise model (i.e. with a density-dependent variance) is employed.

TABLE III

PROPERTIES OF THE PREVIOUSLY PROPOSED MODEL COMPONENTS FOR THE FUNCTIONAL FORM $q(k)$ OF THE FLOW-DENSITY RELATIONSHIP THAT ARE DESCRIBED IN SECTION III-C. IN THE COLUMNS P.1 TO P.8, "YES" SIGNIFIES THAT THE CORRESPONDING DESIRABLE PROPERTY IS FULFILLED FOR A PHYSICALLY RELEVANT AND NON-TRIVIAL SUBSET OF THE MODEL COMPONENT PARAMETER SPACE

Model Component For $q(k)$	Model Component Classification	No. Of Free Parameters	P.1	P.2	P.3	P.4	P.5	P.6	P.7	P.8
FF	Parametric Linear	1	Yes	Yes	No	Yes	Yes	No	Yes	Yes
GS1935	Parametric Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GS1935kjf	Parametric Linear	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GB1959	Parametric Linear	2	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GB1959kjf	Parametric Linear	1	Yes	Yes	Yes	No	Yes	Yes	Yes	No
ED1961	Parametric Non-Linear	5	Yes	Yes	Yes	Yes	Yes	Yes	No	No
ED1961kjf	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	No	No
UW1961A	Parametric Non-Linear	2	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
UW1961B	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UW1961Bkjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FN1961	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FN1961kjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961A	Parametric Linear	2	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GZ1961Akjf	Parametric Linear	1	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GZ1961B	Parametric Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961Bkjf	Parametric Linear	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961C	Parametric Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961Ckjf	Parametric Linear	1	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961D	Parametric Non-Linear	2	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GZ1961Dkjf	Parametric Linear	1	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GZ1961E	Parametric Non-Linear	2	Yes	Yes	Yes	No	Yes	Yes	Yes	No
GZ1961Ekjf	Parametric Linear	1	Yes	Yes	No	No	Yes	Yes	No	No
GZ1961F	Parametric Non-Linear	2	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
GZ1961G	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961Gkjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GZ1961H	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
GZ1961Hkjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
DK1966A	Parametric Non-Linear	5	Yes	Yes	Yes	Yes	Yes	Yes	No	No
DK1966Akjf	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	No	No
DK1966B	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
DK1966Bkjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
MJ1971	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
MJ1971kjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
BM1977	Parametric Non-Linear	3	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
VA1995	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
VA1995kjf	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BD1995	Parametric Non-Linear	3	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
DC1995A	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DC1995Akjf	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DC2012B	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DC2012Bkjf	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GD2008	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GD2008kjf	Parametric Non-Linear	2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MN2008	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MN2008kjf	Parametric Non-Linear	3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WG2011A	Parametric Non-Linear	5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WG2011B	Parametric Non-Linear	4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
WG2011C	Parametric Non-Linear	3	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
SN2014	Non-Parametric	$\geq 1^{(a)}$	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
SN2014kjf	Non-Parametric	$\geq 1^{(a)}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

$^{(a)}$For our implementation of these model components, the maximum effective number of free parameters is 13.

opted to use the `gamlss` software[13] in the R[14] programming language for fitting Generalised Additive Models for Location, Scale, and Shape (GAMLSS [104], [105]). The GAMLSS statistical framework provides a very general approach to regression analysis in which the dependent (or response) variable is modelled using any parametric distribution, and the distribution parameters for location (e.g. mean), scale (e.g. standard deviation), and shape (e.g. skewness, kurtosis) are allowed to vary as a function of a set of explanatory variables. The GaussSigCon noise model component is one of the simplest noise models that can be specified within the GAMLSS framework, and our motivation for employing the `gamlss` software is to open up the door to exploring alternative noise model components in future work.

For a model of EFD data consisting of the functional form model component $q(k)$ and the noise model component GaussSigCon, the likelihood function $\mathcal{L}(\boldsymbol{\theta})$ is given by:

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N_{\text{dat}}} \left(2\,\pi\,\sigma_{\text{con}}^2 \right)^{-1/2} \exp\left(-\frac{1}{2} \left(\frac{q_i - q(k_i)}{\sigma_{\text{con}}} \right)^2 \right) \quad (36)$$

where $\boldsymbol{\theta}$ is a vector of model parameters (which includes σ_{con}), q_i is the ith flow measurement, and the other symbols are as previously defined. For practical reasons, the `gamlss()` function estimates the free model parameters by minimising the log-likelihood function $-2 \ln \mathcal{L}(\boldsymbol{\theta})$ over the parameter space, which is equivalent to maximising $\mathcal{L}(\boldsymbol{\theta})$. From Equation 36, we have:

$$-2 \ln \mathcal{L}(\boldsymbol{\theta}) = N_{\text{dat}} \, \ln \left(2 \, \pi \, \sigma_{\text{con}}^2 \right) + \frac{1}{\sigma_{\text{con}}^2} \sum_{i=1}^{N_{\text{dat}}} (q_i - q(k_i))^2$$

(37)

where the second term may be recognised as the chi-squared. The penalised log-likelihood function that is minimised during MPL estimation is the log-likelihood function in Equation 37 plus an appropriate penalty term (see [105]).

The fits of the models with parametric-linear model components for $q(k)$ (see Table III) are achieved by a single call to the `gamlss()` function, as are the fits for UW1961A, GZ1961F, BM1977, SN2014, and SN2014kjf, except that for these last five models we replace the identity link function for μ in Equation 35 with a log link function. For the remaining models, repeated calls to `gamlss()` are required to iteratively explore the log-likelihood surface for the parameter space of the non-linear parameters. This is typically done using the L-BFGS-B algorithm [106], and if convergence fails, a second attempt is made using the Nelder-Mead simplex or Brent algorithms as appropriate [107], [108]. A limit of 500 iterations is placed on each of these non-linear optimisation methods. However, for the models with multi-regime model components for $q(k)$, the break-point parameter is instead explored by profiling the log-likelihood over the occupancy range of the data (using 50 equally spaced occupancy values), and then iteratively refining the best fit (7 iterations), since multi-regime models are notorious for yielding multiple local likelihood maxima. Initial parameter estimates are required for a number of the models with parametric non-linear model components for $q(k)$, and these are obtained by first fitting one or more of the models GS1935, GS1935kjf, GB1959, GB1959kjf, and UW1961A to the data.

All calls to `gamlss()` use the default values for the algorithmic control parameters, except when fitting SN2014 and SN2014kjf. We found that by increasing the maximum number of cycles of the outer iteration of the GAMLSS fitting algorithm from 20 to 300, and by relaxing the convergence tolerance on the log-likelihood from 0.001 to 0.02, some of the (very few) fits for these models that were failing to converge with the default values would instead complete successfully. Note that local ML estimation [109] is used to determine the effective number of free parameters for SN2014 and SN2014kjf.

We have implemented all of the above fitting procedures in an R software package called `FitFun` (v1.0), which can be used to fit either speed-density or flow-density EFD data with any of the models reviewed in Section III. `FitFun` is publicly available for download from GitHub at https://github.com/danlegend5/FitFun.

Fits can take anything from ~1 sec to ~10 min to complete, depending on the complexity of the model and the number of flow-occupancy measurement pairs. Any fits that require more than 30 min of processing time[15] are declared as timed-out, since complicated non-linear models that take an unreasonable amount of time to fit are of little practical use to traffic engineers.

A total of 507,500 fits are required for 10,150 EFDs and 50 models. Of these, 505,291 fits complete successfully within the 30 min time limit, while 490 fail (`gamlss()` declares non-convergence or an error), and 1,719 are timed-out. The breakdown by model for the fits that either fail or do not finish within 30 min is given in Table VI in Appendix A. Some examples of fitted models for selected EFDs are plotted in Figs. 4 & 5 in Appendix B.

B. Model Comparison and Selection Using Information Criteria

Very often in scientific modelling, one faces the task of selecting an optimal (or best) model for a data sample from a set of candidate models (multiple working hypotheses). In this context, "optimality" refers both to the Principle of Parsimony, in that the best model should constitute the simplest model that provides a good fit to the data without under- or over-fitting, and to appropriate/relevant model performance measure(s). Model estimation via ML assumes a uniform prior probability density function on the model parameters. Consequently, as parameters are added to a model, the ML always increases, rendering it useless for the purpose of model selection between models with different dimensionalities. Error-based metrics for comparing model fits (such as RMSE) also suffer from the same problem since they are derived from likelihood functions. Information criteria are used as an alternative for evaluating models with different numbers of parameters.

Various information criteria have been developed from distinct statistical view-points as implementations of the Principle of Parsimony and each one may be used to automatically select a parsimonious model from a set of candidate models. They may be applied regardless of whether the models under consideration are nested or non-nested. Here we will compare model fits using the Akaike information criterion (AIC [110]) and the Bayesian information criterion (BIC [111]). It is conceivable that we could also reject any unreasonable fits by filtering out those that have unphysical curve properties or parameter values. However, doing this fairly and consistently across all 50 FD models would be unfeasible given the fact that each model has some kind of built-in deficiency (e.g. GB1959 predicts an infinite free-flow speed, UW1961A fails to predict a jam density, GZ1961E predicts an infinite back-propagating wave speed at jam density, etc.). Furthermore, the FD models from our review have not been designed to account for the effect that the LD location relative to an active bottleneck has on the shape of an EFD, and evaluating them on this basis would be unfair. Hence, we do not implement any further

[15]Each fit is run on a single core of a 14-core Intel(R) Xeon(R) 2.40 GHz CPU (E5-2680 v4).

filters or metrics on the fit properties, and we use only the AIC and the BIC to assess model performance.

1) The Akaike Information Criterion: The AIC, which applies to linear and non-linear models estimated via ML, may be computed via:

$$AIC = -2 \ln \mathcal{L}(\hat{\boldsymbol{\theta}}) + 2N_{par} \qquad (38)$$

where $\hat{\boldsymbol{\theta}}$ is a vector of ML estimators for the model parameters, and N_{par} is the total number of free parameters in the model. Model selection with the AIC is performed by minimising $-2 \ln \mathcal{L}(\boldsymbol{\theta})$ for each model, and then minimising the AIC over the full set of models under consideration.

The AIC is derived as an asymptotic approximation to the Kullback-Leibler divergence [112] which measures the distance of a candidate model from the true underlying model under the (reasonable) assumption that the true model is of infinite dimension. In this case the AIC provides an asymptotically efficient selection of a finite dimensional approximating model. However, if the true model is finite dimensional, then the AIC does not provide a consistent model selection. A model selection criterion is said to be consistent if it selects with high probability the true model from the set of candidate models whenever the true model is represented in the set of candidate models. The aim of the AIC is to evaluate models based on their prediction accuracy.

It is important to be aware that the AIC suffers from a large negative bias when the number of model parameters is a non-negligible fraction of the number of data values (e.g. for $N_{dat}/N_{par} < 40$ [113]). This bias can be corrected for, but at the cost of general applicability, since the formula for the bias correction depends on the statistical model that is being fitted. For example, Sugiura [114] derived a bias-corrected version of the AIC for Gaussian linear regression problems that is asymptotically the same as the AIC for $N_{dat} \gg N_{par}$. In Section II-B, we expressly adopted the requirement that an LD has at least 900 flow-occupancy measurement pairs that are not flagged as errors so as to avoid the need to make any model-specific bias corrections to the AIC.

Takeuchi [115] generalised the AIC with a more complicated formula to create the Takeuchi information criterion (TIC). Subsequently, Konishi and Kitagawa [116] derived a further generalisation of the AIC and TIC, called the generalised information criterion (GIC), that can also be applied to model selection for models with parameters estimated by MPL. The `gamlss` software only implements the AIC from the available AIC-like information criteria. Consequently, for GAMLSS models that employ a non-parametric smoothing function estimated by MPL (i.e. SN2014 and SN2014kjf in our case), the `gamlss` software computes the AIC via Equation 38 by using the vector of MPL estimators for the model parameters in place of $\hat{\boldsymbol{\theta}}$ and the trace of the smoother matrix as an estimate of the effective number of free parameters contributed to the fit by the smoothing function.

2) The Bayesian Information Criterion: An alternative approach to model selection is a Bayesian approach where the model with the largest Bayesian posterior probability is chosen. The BIC is derived by approximating the posterior probability of each model, and, like the AIC, it also applies to linear and non-linear models estimated via ML. The formula for the BIC is as follows:

$$BIC = -2 \ln \mathcal{L}(\hat{\boldsymbol{\theta}}) + N_{par} \ln N_{dat} \qquad (39)$$

Model selection with the BIC proceeds in the same way as for the AIC.

The BIC generally includes a heavier penalty than the AIC (and the bias-corrected AIC) for more complicated models (e.g. in the regime $N_{par} < 20$ for $N_{dat} > 50$), therefore favouring models with fewer parameters than those favoured by the AIC. Konishi *et al.* [117] performed a deeper Bayesian analysis to derive an improved BIC, along with a version that applies to model selection for models with parameters estimated by MPL. The BIC and the improved BIC are consistent model selection criteria. Again, the `gamlss` software only implements the BIC from the available BIC-like information criteria. Consequently, for GAMLSS models that employ a non-parametric smoothing function estimated by MPL, the `gamlss` software computes the BIC via Equation 39 by using the vector of MPL estimators for the model parameters in place of $\hat{\boldsymbol{\theta}}$ and the trace of the smoother matrix as an estimate of the effective number of free parameters contributed to the fit by the smoothing function.

3) Model Likelihoods: For a set of N_{mod} models indexed by j that have each been fitted to the EFD data from the lth LD, one may rescale the corresponding AIC values as follows:

$$\Delta_{AIC,jl} = AIC_{jl} - \min_{r=1,...,N_{mod}} \{AIC_{rl}\} \qquad (40)$$

This forces the model with the lowest AIC value (i.e. the best model) for the lth EFD to have $\Delta_{AIC,jl} = 0$. Applying the simple transformation $\exp(-\Delta_{AIC,jl}/2)$ converts the rescaled AIC values into model likelihoods, which can then be normalised to sum over the set of models to unity:

$$p_{AIC,jl} = \frac{\exp(-\Delta_{AIC,jl}/2)}{\sum_{r=1}^{N_{mod}} \exp(-\Delta_{AIC,rl}/2)} \qquad (41)$$

The quantity $p_{AIC,jl}$ can be interpreted as the probability that the jth model is the best model for the lth EFD. Models for which a fit fails or is timed-out are assigned $p_{AIC,jl} = 0$. For a set of N_{LD} LDs, each with an EFD, the expected fraction $F_{AIC,j}$ of EFDs from the full set of N_{LD} EFDs for which the jth model is the best model is given by:

$$F_{AIC,j} = \frac{1}{N_{LD}} \sum_{s=1}^{N_{LD}} p_{AIC,js} \qquad (42)$$

The above procedure may also be applied to the BIC values for the model fits in exactly the same way. In a Bayesian context, use of Equation 41 is equivalent to computing the posterior probabilities of the models given the data under the assumption that the prior probabilities of the models are all the same and equal to $1/N_{mod}$.

V. RESULTS AND DISCUSSION

We apply the rescaling and transformation procedure detailed in Section IV-B3 to the AIC and BIC values computed by the `gamlss` software (Sections IV-B1 & IV-B2) for the
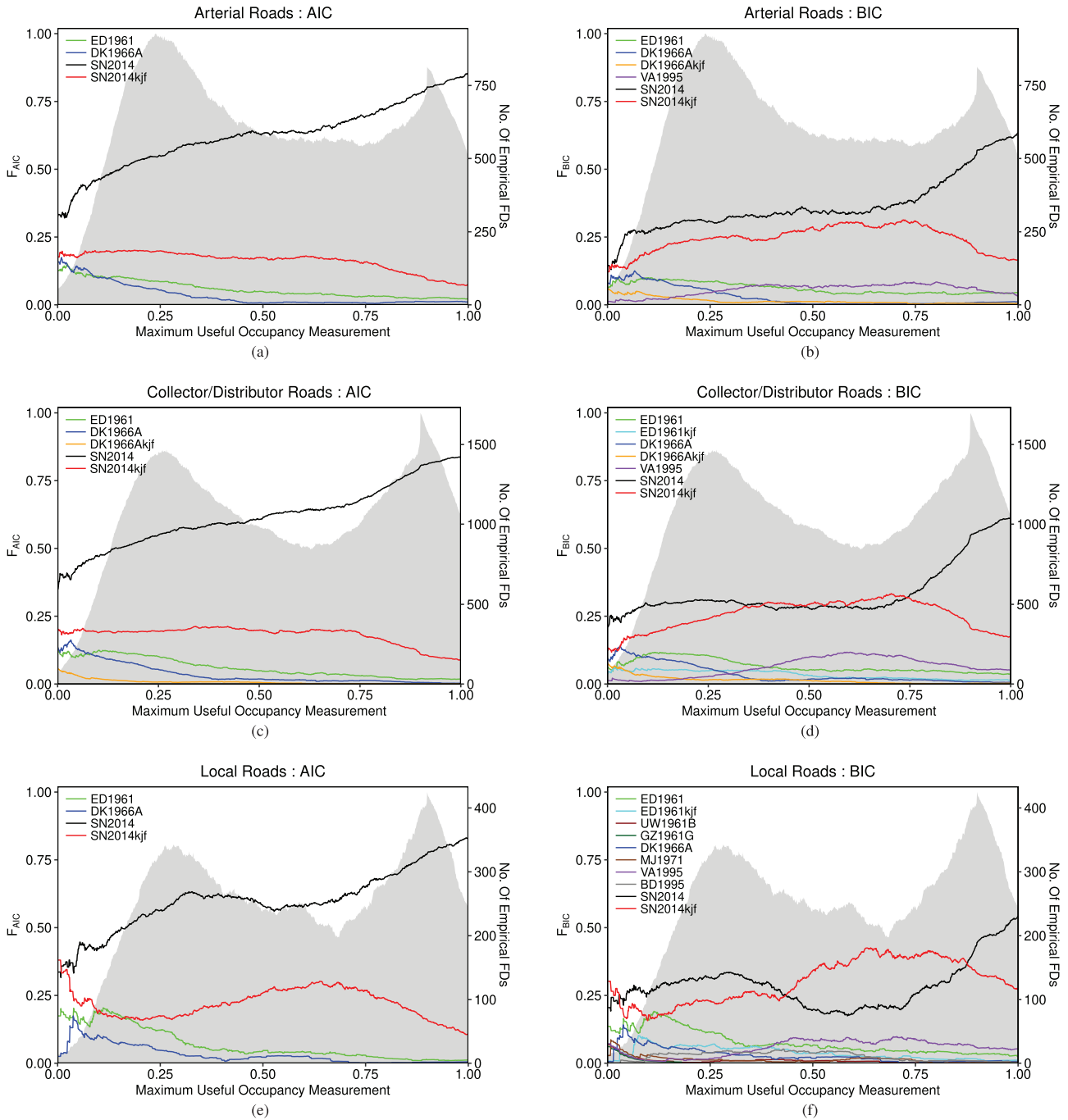
Fig. 2. Plots of curves of F_{AIC} (left-hand column) and F_{BIC} (right-hand column) for each model as a function of the maximum useful occupancy measurement. Results are further broken down by road grouping (arterial, top row; collector/distributor, middle row; local, bottom row). For clarity in each plot, a curve is only plotted for a model if the maximum value of F_{AIC} or F_{BIC} in the curve exceeds 0.05. Curves are colour coded appropriately with a legend in each plot indicating which model they correspond to. The grey region displayed in the background of each plot tracks the number of LDs (or EFDs; right-hand scale) used to compute the points on the curves.

fits that we performed in Section IV-A. This results in two $10,150 \times 50$-element arrays of model probabilities. Applying Equation 42 to each array by summing over all of the LDs in our data sample allows us to rank the models based on the full data set. In Tables IV & V, we report the top ten models ranked according to the expected fraction of EFDs from the full data set for which a model is the best model when using the AIC or BIC, respectively, for model selection.

The clear winner by far is SN2014 for both information criteria. Furthermore, the top three models in each list are the

same (SN2014, SN2014kjf, and ED1961). In fact, both lists have nine models in common with orderings that only differ slightly between them due to lower rankings for DK1966A and DC2012B in Table V. The top ten models account for the best models ~97% and ~88% of the time for the AIC and BIC, respectively, leaving very little room for the remaining models to have any impact. It is interesting to note that the models that employ non-parametric smoothing functions for $q(k)$ are the winners, followed far behind by a mix of the discontinuous multi-regime models and some of

TABLE IV

Top Ten Models Ranked According to the Expected Fraction F_{AIC} of EFDs From the Full Data Set for Which a Model Is the Best Model When Using the AIC for Model Selection

Rank	Model	Total No. Of Free Parameters	F_{AIC} (%)
1	SN2014	≥ 2	65.00
2	SN2014kjf	≥ 2	17.08
3	ED1961	6	5.44
4	DK1966A	6	2.71
5	VA1995	5	2.53
6	ED1961kjf	5	1.16
7	DC2012B	5	1.12
8	BD1995	4	0.85
9	DC1995A	5	0.67
10	WG2011A	6	0.56
		Total	97.12

TABLE V

Top Ten Models Ranked According to the Expected Fraction F_{BIC} of EFDs From the Full Data Set for Which a Model Is the Best Model When Using the BIC for Model Selection

Rank	Model	Total No. Of Free Parameters	F_{BIC} (%)
1	SN2014	≥ 2	37.05
2	SN2014kjf	≥ 2	25.69
3	ED1961	6	6.29
4	VA1995	5	5.98
5	ED1961kjf	5	3.19
6	DK1966A	6	2.66
7	BD1995	4	2.42
8	DC2012B	5	1.93
9	DC1995A	5	1.21
10	VA1995kjf	4	1.20
		Total	87.62

the single-regime models with at least four free parameters. For the winning model SN2014, one observes that $F_{\text{AIC}} > F_{\text{BIC}}$. This happens because SN2014 typically employs a greater (effective) number of free parameters than any of the other models, and hence the SN2014 fits, relative to the other model fits, receive larger penalties for the BIC than they do for the AIC. The fact that SN2014 is the best model according to both information criteria serves to highlight by just how much SN2014 outperforms all of the other models.

To explore how the model rankings depend on the type of road that an LD is located on, we split the data sample into three road groupings named "arterial", "collector/distributor", and "local" with 3,230, 5,595, and 1,325 LDs, respectively (Table II). Performing the appropriate sums over the two arrays of model probabilities using Equation 42, we find that each road grouping yields the same top five models in the same order as that listed in Tables IV & V for the AIC and BIC, respectively, with similar sets of models for the rest of the top ten in both cases. Furthermore, the values of F_{AIC} and F_{BIC} are relatively insensitive to road type. For example, for SN2014 we find $F_{\text{AIC}} \approx 64.9\%$ and $F_{\text{BIC}} \approx 39.1\%$ for the arterial roads, $F_{\text{AIC}} \approx 65.2\%$ and $F_{\text{BIC}} \approx 37.1\%$ for the collector/distributor roads, and $F_{\text{AIC}} \approx 64.6\%$ and $F_{\text{BIC}} \approx 31.8\%$ for the local roads. This is an important finding as it surprisingly reveals that model performance is mostly independent of road type, despite the presence of localised disruptions to traffic flow (e.g. intersections) on collector/distributor and local roads. We also find that different choices for the road groupings do not affect these results.

There are many ways that we could further partition the data sample into subgroupings of LDs. An important aspect

to consider, and one that we will focus on here, is the occupancy coverage of the flow-occupancy measurements in each EFD. It is conceivable that EFDs that only have data in the free-flow uncongested traffic regime (i.e. at low occupancies) will require less complicated parsimonious models than those that are required for EFDs with data extending into congested traffic conditions (i.e. at medium/high occupancies). To enable this line of investigation, for each EFD we attempt to identify the highest occupancy measurement that can meaningfully constrain any FD model. The procedure that we adopt to do this for each individual EFD is as follows:

(i) For each occupancy measurement in descending order, repeat steps (ii)-(iii).
(ii) Count the number of occupancy measurements within ± 0.1 of the current occupancy measurement.
(iii) If the number of occupancy measurements identified in step (ii) is less than 30 (chosen as a reasonable minimum number of measurements required to locally constrain any FD model), then move on to the next occupancy measurement in step (ii). Otherwise, record the current occupancy measurement as the maximum useful occupancy measurement for this EFD, and finish.

The maximum useful occupancy measurement identified in this way serves as a proxy for the effective occupancy coverage of the EFD data when fitting FD models.

By selecting a subset of LDs from a specific road grouping that have EFDs with maximum useful occupancy measurements within a certain range (or bin), and then performing the appropriate sums over the two arrays of model probabilities using Equation 42, one obtains values for F_{AIC} and F_{BIC} for each model. We use the three aforementioned road groupings and a set of bins in maximum useful occupancy of half-width 0.1, with centres following a dense arithmetic sequence of values between 0 and 1, to construct curves of F_{AIC} and F_{BIC} for each model. We plot these curves in the six panels of Fig. 2, organised into rows by road grouping and columns by information criterion. For clarity, a curve is only plotted for a model if the maximum value of F_{AIC} or F_{BIC} in the curve exceeds 0.05. The grey region displayed in the background of each plot tracks the number of LDs (or EFDs; right-hand scale) used to compute the points on the curves, which is at least 100, and up to $\sim$1600, for all but the smallest values of the maximum useful occupancy. This provides confidence in the curves down to a maximum useful occupancy of $\sim$0.1. The bin width of 0.2 was chosen as a reasonable trade-off between resolution and noise in the curves. Consequently, points on a curve that are separated by more than 0.2 in occupancy are derived from fully independent data samples.

The plots in Fig. 2 reveal a number of interesting facts. Firstly, the plots make it clear that the rankings of the top three models according to the AIC and BIC are unchanged and mostly insensitive to the effective occupancy coverage of the EFD data, except for the BIC rankings for EFDs from collector/distributor or local roads with maximum useful occupancy measurements in the range $\sim$0.4-0.85 (here the top three change to SN2014kjf, SN2014, and VA1995 in that order). This also indicates that somewhat surprisingly,
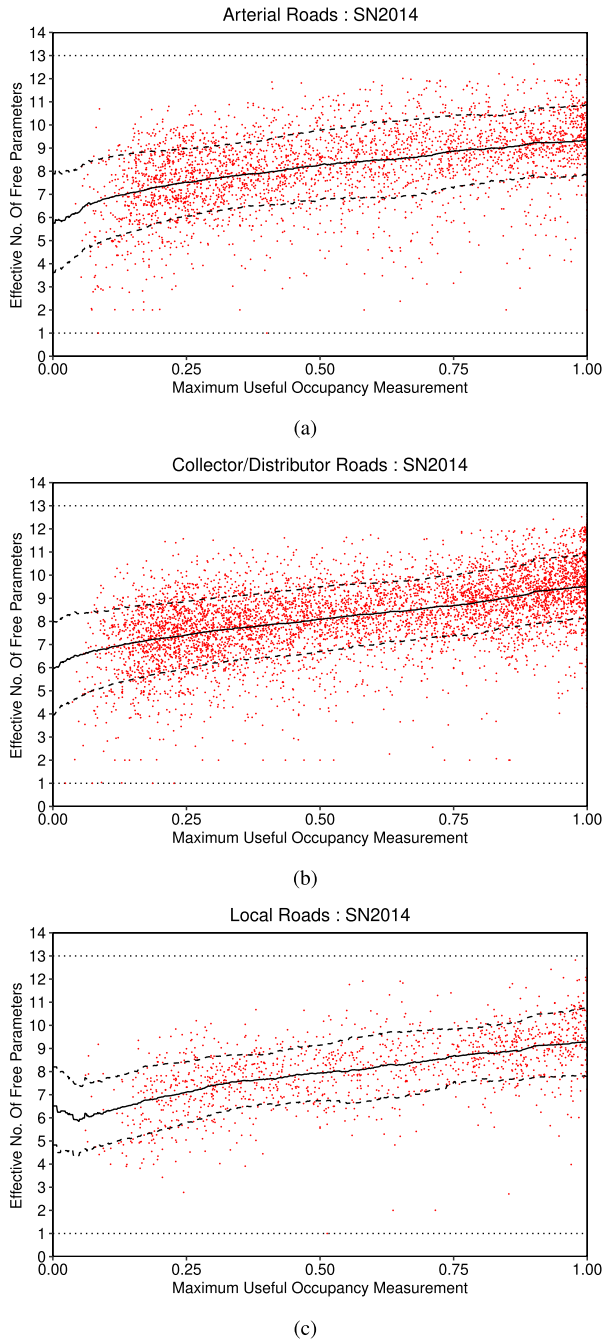
(a)

(b)

(c)

Fig. 3. Plots of (fitted) effective number of free parameters in the $q(k)$ component of the SN2014 model against the maximum useful occupancy measurement (red dots) for each road grouping. The minimum and maximum possible values for the effective number of free parameters are plotted as horizontal dotted lines. The black continuous curve in each plot is the running mean with a bin size of 0.2, while the black dashed curves represent the running mean plus/minus the running standard deviation.

for LDs with data solely in the free-flow uncongested traffic regime (i.e. with maximum useful occupancy measurements below ∼0.1), the SN2014 and SN2014kjf models still prevail as the best models according to both information criteria, whereas the simplest models such as FF and GS1935 which one might expect to be more suitable for this traffic regime do not even reach the threshold of 0.05 for F_{AIC} or F_{BIC} in order to be plotted. Secondly, the F_{AIC} and F_{BIC} values for SN2014 tend to increase with increasing maximum useful

occupancy, with a faster increase above a maximum useful occupancy of ∼0.75. Thirdly, for EFDs that cover the full range of occupancy from 0 to 1, the SN2014 model achieves F_{AIC} and F_{BIC} values of ∼80% and ∼60%, respectively, indicating that in the majority of cases this is the most parsimonious model for EFDs with the best observational coverage.

These last two facts are explained by the ability of the SN2014 model to adapt to increasingly complex functional forms of the observed flow-density relationship as one considers EFDs with greater effective occupancy coverage. In Fig. 3, for each of the three road groupings, we plot the (fitted) effective number of free parameters in the $q(k)$ component of SN2014 against the maximum useful occupancy measurement (red dots). The black continuous curve in each plot is the running mean with a bin size of 0.2, while the black dashed curves represent the running mean plus/minus the running standard deviation. There is a clear trend of increasing effective number of free parameters from ∼7 parameters at a maximum useful occupancy of ∼0.1 to ∼9.5 parameters at a maximum useful occupancy of unity. This indicates that the SN2014 model accommodates increasingly complex functional forms of EFD data by gradually increasing its effective number of free parameters. The same comments and observations also apply to SN2014kjf, although this model is slightly more restricted in that it is forced to predict zero flow at jam occupancy (unity in our case). None of the other models that we have reviewed in this paper have this adaptive ability, which undoubtedly puts them at a disadvantage when considering the full spectrum of EFDs.

Digressing briefly to consider the fixed jam density/occupancy (or "kjf") versions of the model components for $q(k)$, we note that Coifman [32] has argued that the jam occupancy should be ∼0.8. If this is correct, then by fixing the jam occupancy to unity, we may have inadvertently disadvantaged the performance of the corresponding models. However, having inspected the plots of all EFDs in our data sample, we have seen very few cases where the jam occupancy could potentially differ substantially from unity, and the EFDs shown in Figs. 4 & 5 provide unequivocal examples of where the jam occupancy is very close to unity. Interestingly, from the same figures, it can be observed that fixing the jam occupancy in SN2014kjf leads to smoothing curves that are slightly more stable (and therefore more appealing) than those for SN2014 in some cases.

The findings in this section reveal that, *given a Gaussian noise model with constant variance* (GaussSigCon), the UTD19 EFD data are not well-described by any of the parametric model components for $q(k)$ (Table III). This statement holds true regardless of the properties of the parametric model components (e.g. linear or non-linear, number of free parameters, single- or multi-regime, derived from theoretical considerations or not) or the EFD data (e.g. road grouping, effective occupancy coverage). However, as we have already discussed in Sections I & III-D, while it has always been conventional to use this simple noise model, it is now known that GaussSigCon is not particularly suitable for modelling flow-density EFD data. Specifically, the width and shape of
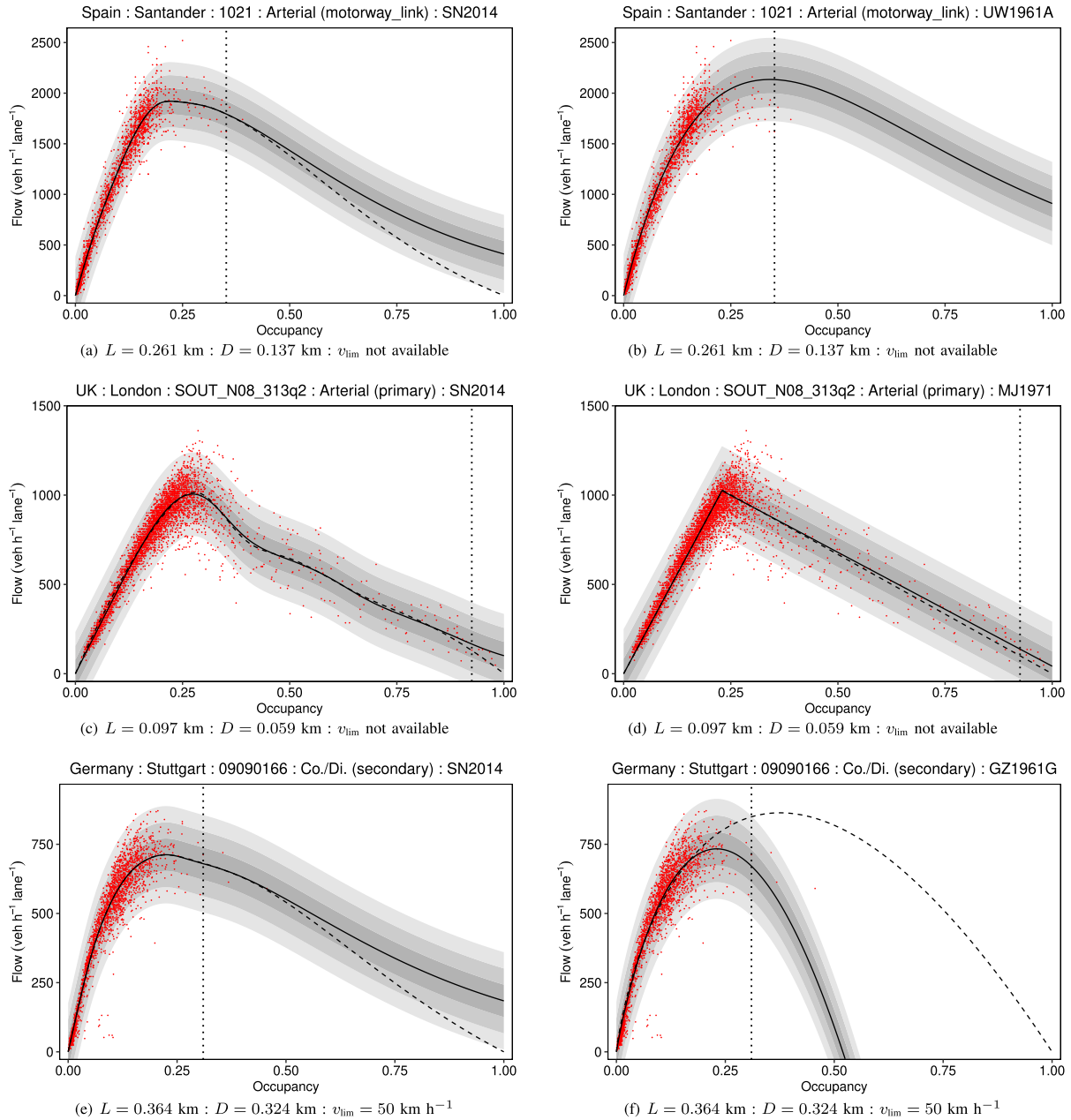
Fig. 4. Plots of EFDs (red dots). Both plots on a single row show the same EFD data but with different fitted models. The plots in the left-hand column always feature the same fitted model SN2014, whereas the plots in the right-hand column feature a different fitted model for comparison purposes. In each plot, the fitted model is represented by a black continuous curve for $q(k)$ and sequentially lighter grey regions for the $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$ standard Normal quantile regions. The fixed jam density version of the model (if it is defined) is also fitted to the data and the corresponding $q(k)$ is plotted as a black dashed curve for further comparisons. The vertical dotted line corresponds to the maximum useful occupancy measurement for the specific EFD. The plot titles are all of the same format listing the country, city, loop detector ID, road grouping and classification, and fitted model. The sub-captions provide information on the length L and speed limit v_{lim} of the road, and the distance D of the LD from the downstream intersection.

the flow distribution typically varies as a function of density, and the density/occupancy measurements also suffer from noise. Furthermore, the appearance of transient/non-stationary traffic conditions shifts the (aggregated) measurements away from their expected locations in the FD, the capacity drop phenomenon skews flow measurements to lower values, and hysteresis effects introduce time-correlations between measurements. None of these complexities in the noise properties of the data can be accounted for by GaussSigCon which only confers a single parameter to the full model. Even for EFDs constructed solely from measurements taken during

stationary traffic conditions, the noise properties of the data are still far from Gaussian with constant variance. With a non-parametric smoothing function for $q(k)$, SN2014 is the only model that is flexible enough to be able to partially compensate for the rigidity of the conventional noise model, and hence it is by far the best performing model overall. A more sophisticated noise model component (defined within the GAMLSS statistical framework for example) can account for all of the aforementioned complexities in EFD data (filtered or not), yielding unbiased fits for the functional form of EFDs. We suspect that by adopting a more appropriate noise model

(a) $L = 0.320$ km : $D = 0.270$ km : $v_{lim} = 50$ km h^{-1}

(b) $L = 0.320$ km : $D = 0.270$ km : $v_{lim} = 50$ km h^{-1}

(c) $L = 0.060$ km : $D = 0.015$ km : v_{lim} not available

(d) $L = 0.060$ km : $D = 0.015$ km : v_{lim} not available

(e) $L = 0.093$ km : $D = 0.079$ km : $v_{lim} = 40$ km h^{-1}

(f) $L = 0.093$ km : $D = 0.079$ km : $v_{lim} = 40$ km h^{-1}
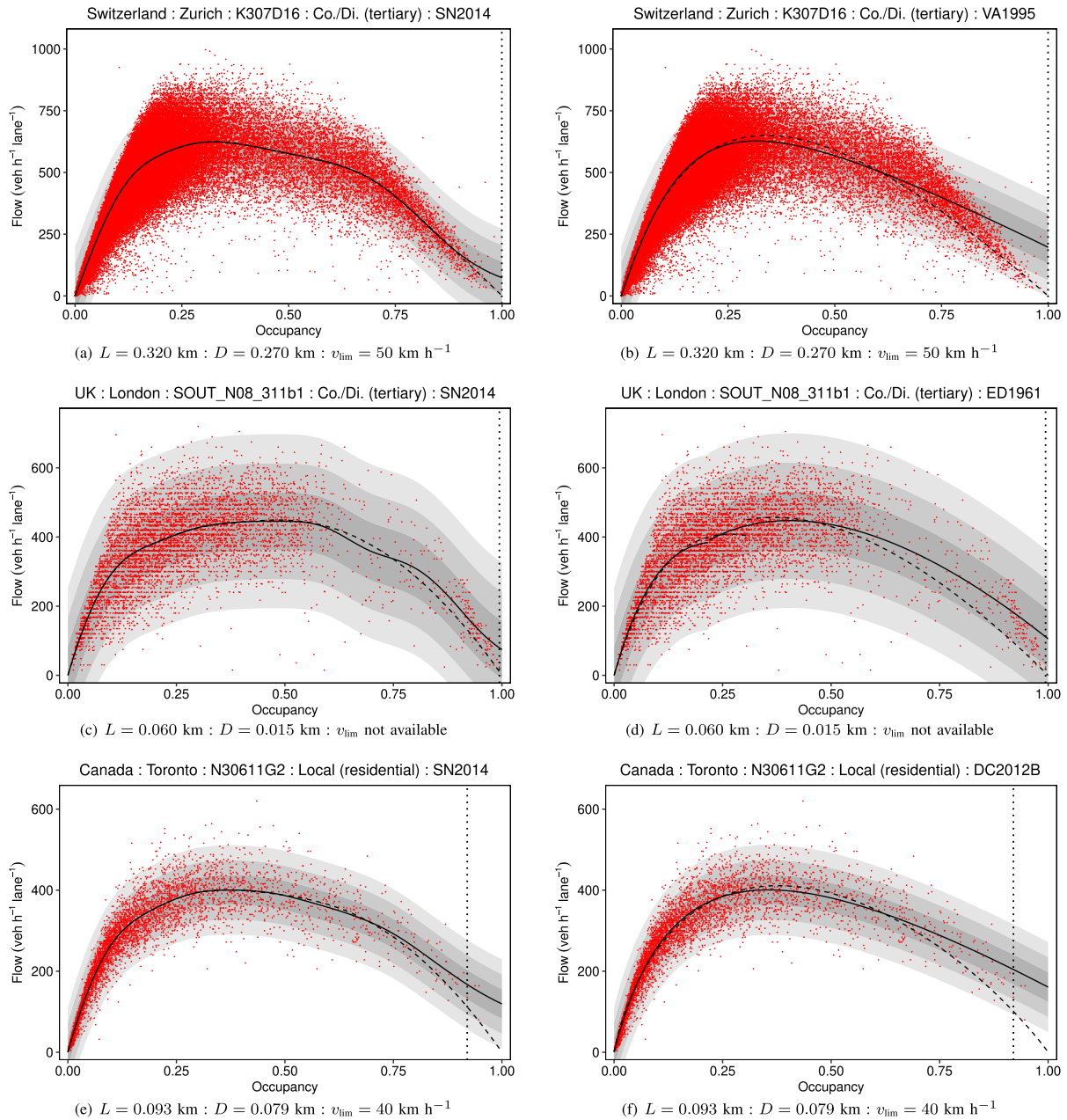
Fig. 5.   More plots of EFDs (red dots). The format of the figure is the same as in Figure 4.

component in the future, the results presented here could potentially change substantially.

VI. CONCLUSION AND RECOMMENDATIONS

Taking advantage of the ever increasing amounts of traffic data that are becoming available, we selected a high quality and well-controlled sample of 10,150 flow-density EFDs from the UTD19 data set (Section II). We then presented a detailed and comprehensive review of the literature on modelling the functional form of EFDs (speed-density and flow-density), highlighting the fact that earlier reviews, now at least a decade old, have been propagating a number of errors and misattributions forwards through the literature, and that they have inevitably left some models out (Section III). These gaps and shortcomings in previous literature reviews, along with the recent creation of UTD19, have motivated the need for a new

review and comparison study. We hope that our review, which addresses all of these issues, will now become the one-stop-shop for researchers wishing to understand the development of FD models up to the present day.

We performed fits of all 50 of the previously proposed models for the functional form of the flow-density relationship to the selected EFDs by adopting the GaussSigCon noise model and employing ML estimation (Section IV-A). This is equivalent to curve fitting using the method of least squares, which has been the method of choice by researchers up until now. We compared the fits using two different information criteria (AIC and BIC), which allow for the selection of a parsimonious model from a set of candidate models with different numbers of parameters (Section IV-B). We found that the non-parametric Sun model (and its fixed jam density version) outperforms any of the other previously proposed

parametric models by a big margin, and that this result is not affected by road type or the effective occupancy coverage of an EFD (Section V). It seems that the Sun model adapts better to the complexities of EFD data than the parametric models, and evidence for how it does this can be seen in the way that the fitted effective number of free parameters increases as the effective occupancy coverage of the data increases. Considering the types of roads to which the EFDs in our data sample belong (Table II), it can be seen that this result is mainly applicable to interrupted flow facilities (i.e. roads with intersections). Additional data would be needed to study uninterrupted flow facilities (i.e. freeways) in depth, and this is left for future studies.

Our comparison study has finally provided a definitive answer to the question that so many other researchers have previously attempted to investigate; namely, "Which model for the functional form of an empirical fundamental diagram is *currently* the best?". However, the answer that we have found suffers from the caveat that it only applies to the GaussSigCon noise model, which is a rather poor choice for EFD data (filtered or not). It is quite possible that a more sophisticated noise model could lead to substantially different results. Furthermore, the fact that a non-parametric smoothing function constitutes the best model for the observed functional form of the FD relationship may be somewhat disappointing to many traffic researchers, since this model has no physically meaningful parameters to be estimated, and no underlying traffic flow theory that can be validated, making it difficult to relate to the theoretical FD. Another shortcoming of the winning model is the fact that it does not provide any reliable flow predictions outside of the occupancy range of the data. We note that over the full data set, and also for the three different road groupings, the best model with physically meaningful parameters and an underlying traffic flow theory is ED1961 (see Tables IV & V and Section V). For some practitioners, the intended use of the fitted FD model may further influence which model is selected from our tables of rankings, in which case our results are an important first step in the model selection process.

Since the pioneering work of Greenshields [14], there has been an unbalanced focus on modelling the functional form of EFDs to the detriment of understanding the noise properties of the observations. While this may have been acceptable in the early days, the research community now needs to work on finding a better noise model. A massive amount of effort has already been expended on inventing complicated non-linear functional form models that simply cannot match the Sun model when handicapped by a restrictive noise model with just a single free parameter. The era of proposing such functional form models and fitting them using the method of least squares is over. Instead, the UTD19 data set ushers in a new era of research into EFDs where the noise model component is given just as much importance as the model component for the functional form. This balanced approach is likely to lead to a much better understanding of EFDs and the underlying functional form of the FD relationship.

On a final note, we have tried with this work to create fully reproducible science, which is still relatively hard to

TABLE VI
BREAKDOWN BY MODEL OF THE 2,209 FITS THAT EITHER FAIL OR DO NOT FINISH WITHIN 30 min. IF A MODEL IS NOT LISTED, THEN ALL OF THE CORRESPONDING 10,150 FITS COMPLETE SUCCESSFULLY

Model	No. Of Fits That Fail	No. Of Fits That Are Timed-Out
ED1961	0	377
ED1961kjf	0	345
UW1961Bkjf	1	0
FN1961	5	0
GZ1961H	54	35
DK1966B	0	356
MJ1971	0	446
BM1977	3	0
BD1995	2	0
DC1995A	10	16
DC1995Akjf	13	5
DC2012B	90	0
DC2012Bkjf	159	0
GD2008kjf	3	0
MN2008	3	122
WG2011A	26	17
WG2011B	109	0
WG2011C	11	0
SN2014	1	0
Total	490	1,719

come by even in this digital age. Towards this end, we are making the data sample that we prepared publicly available (see Section II-B), along with the `FitFun` software that we used to perform the fits (see Section IV-A). Furthermore, any of the few remaining methods or procedures that we employed that are not covered by the above have been described in this paper in sufficient detail for repeatability. Hopefully this will aid other researchers to reproduce and improve on our results by developing better (noise) models, opening up a new golden age in the study of EFDs.

APPENDIX A

From a total of 507,500 fits, 490 fail and 1,719 are timed-out. The breakdown by model of these numbers is given in Table VI.

APPENDIX B

In Figs. 4 & 5, we show some example plots of models that have been fitted to selected EFDs.

ACKNOWLEDGMENT

References

[1] J. G. Wardrop, "Some theoretical aspects of road traffic research," *Proc. Inst. Civil Eng.*, vol. 1, pp. 325–378, May 1952.

[2] L. C. Edie, "Discussion of traffic stream measurements and definitions," in *Proc. 2nd Int. Symp. Theory Traffic Flow*, Paris, France, 1965, pp. 139–154.

[3] M. J. Lighthill and G. B. Whitham, "On kinematic waves II. A theory of traffic flow on long crowded roads," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 229, no. 1178, pp. 317–345, 1955, doi: 10.1098/rspa.1955.0089.

[4] C. F. Daganzo, *Fundamentals of Transportation and Traffic Operations*. New York, NY, USA: Elsevier, 1997, pp. 66–160, ch. 4.

[5] F. A. Haight, *Mathematical Theories of Traffic Flow*, vol. 7. New York, NY, USA: Academic, 1963, ch. 3.

[6] G. F. Newell, "A simplified theory of kinematic waves in highway traffic, part I: General theory," *Transp. Res. B, Methodol.*, vol. 27, no. 4, pp. 281–287, Aug. 1993, doi: 10.1016/0191-2615(93)90038-C.

[7] C. F. Daganzo, "The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transp. Res. B, Methodol.*, vol. 28, no. 4, pp. 269–287, Aug. 1994, doi: 10.1016/0191-2615(94)90002-7.

[8] J. C. Muñoz and C. F. Daganzo, *Moving Bottlenecks: A Theory Grounded on Experimental Observation*. Transp. Traffic Theory 21st Century, Emerald Group Publishing Ltd., 2002, pp. 441–461, doi: 10.1108/9780585474601-022.

[9] M. Papageorgiou, *Applications of Automatic Control Concepts to Traffic Flow Modeling and Control*. Berlin, Germany: Springer-Verlag, 1983.

[10] J. P. Lebacque, "Intersection modeling, application to macroscopic network traffic flow models and traffic management," in *Traffic and Granular Flow*. Berlin, Germany: Springer, 2005, pp. 261–278, doi: 10.1007/3-540-28091-X_26.

[11] C. Roncoli, N. Bekiaris-Liberis, and M. Papageorgiou, "Optimal lane-changing control at motorway bottlenecks," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 1785–1791, doi: 10.1109/ITSC.2016.7795800.

[12] F. Tajdari and C. Roncoli, "Adaptive traffic control at motorway bottlenecks with time-varying fundamental diagram," in *Proc. 16th IFAC Symp. Control Transp. Syst. (CTS)*, 2021, pp. 271–277, doi: 10.1016/j.ifacol.2021.06.051.

[13] M. Keyvan-Ebkatani, A. Kouvelas, I. Papamichail, and M. Papageorgiou, "Exploiting the fundamental diagram of urban networks for feedback-based gating," *Transp. Res. B, Methodol.*, vol. 46, no. 10, pp. 1393–1403, Dec. 2012, doi: 10.1016/j.trb.2012.06.008.

[14] B. D. Greenshields, "A study of traffic capacity," *Proc. Highway Res. Board*, vol. 14, pp. 448–477, 1935.

[15] M. van Aerde and H. Rakha, "Multivariate calibration of single regime speed-flow-density relationships," in *Proc. Vehicle Navigat. Inf. Syst. Conf. (VNIS)*, Jul. 1995, pp. 334–341.

[16] B. G. Heydecker and J. D. Addison, "Analysis and modelling of traffic flow under variable speed limits," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 2, pp. 206–217, Apr. 2011, doi: 10.1016/j.trc.2010.05.008.

[17] X. Wu, H. X. Liu, and N. Geroliminis, "An empirical analysis on the arterial fundamental diagram," *Transp. Res. B, Methodol.*, vol. 45, no. 1, pp. 255–266, Jan. 2011, doi: 10.1016/j.trb.2010.06.003.

[18] K. Anuar, F. Habtemichael, and M. Cetin, "Estimating traffic flow rate on freeways from probe vehicle data and fundamental diagram," in *Proc. IEEE 18th Int. Conf. Intell. Transp. Syst.*, Sep. 2015, pp. 2921–2926, doi: 10.1109/ITSC.2015.468.

[19] Z. He, M. Chen, L. Wang, and M. Li, "Empirical analysis of urban arterial fundamental diagram," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 7719–7723, doi: 10.23919/ChiCC.2018.8483985.

[20] L. Sun, Y. Pan, and W. Gu, "Data mining using regularized adaptive B-splines regression with penalization for multi-regime traffic stream models," *J. Adv. Transp.*, vol. 48, no. 7, pp. 876–890, 2014, doi: 10.1002/atr.1232.

[21] V. L. Knoop and W. Daamen, "Automatic fitting procedure for the fundamental diagram," *Transportmetrica B, Transp. Dyn.*, vol. 5, no. 2, pp. 129–144, Apr. 2017, doi: 10.1080/21680566.2016.1256239.

[22] J. M. del Castillo, "Three new models for the flow–density relationship: Derivation and testing for freeway and urban data," *Transportmetrica*, vol. 8, no. 6, pp. 443–465, Nov. 2012, doi: 10.1080/18128602.2011.556680.

[23] H. Wang, J. Li, Q.-Y. Chen, and D. Ni, "Logistic modeling of the equilibrium speed–density relationship," *Transp. Res. A, Policy Pract.*, vol. 45, no. 6, pp. 554–566, Jul. 2011, doi: 10.1016/j.tra.2011.03.010.

[24] A. Loder, L. Ambühl, M. Menéndez, and K. W. Axhausen, "Understanding traffic capacity of urban networks," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Dec. 2019, doi: 10.1038/s41598-019-51539-5.

[25] B. Alonso, Á. Ibeas, G. Musolino, C. Rindone, and A. Vitetta, "Effects of traffic control regulation on network macroscopic fundamental diagram: A statistical analysis of real data," *Transp. Res. A, Policy Pract.*, vol. 126, pp. 136–151, Aug. 2019, doi: 10.1016/j.tra.2019.05.012.

[26] D. L. Gerlough and M. J. Huber, "Traffic flow theory: A monograph," Special Rep., TRB, Nat. Res. Council, Washington, DC, USA, Special Rep. 165, 1975, vol. 165.

[27] M. Carey and M. Bowers, "A review of properties of flow–density functions," *Transp. Rev.*, vol. 32, no. 1, pp. 49–73, Jan. 2012, doi: 10.1080/01441647.2011.608135.

[28] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, "Single-vehicle data of highway traffic: Microscopic description of traffic phases," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 5, May 2002, doi: 10.1103/PhysRevE.65.056133.

[29] W. Guan and S. He, "Statistical features of traffic flow on urban freeways," *Phys. A, Stat. Mech. Appl.*, vol. 387, no. 4, pp. 944–954, Feb. 2008, doi: 10.1016/j.physa.2007.09.036.

[30] M. M. Zefreh and Á. Török, "Distribution of traffic speed in different traffic conditions: An empirical study in Budapest," *Transport*, vol. 35, no. 1, pp. 68–86, Mar. 2020, doi: 10.3846/transport.2019.11725.

[31] M. J. Cassidy, "Bivariate relations in nearly stationary highway traffic," *Transp. Res. B, Methodol.*, vol. 32, pp. 49–59, Jan. 1998, doi: 10.1016/S0191-2615(97)00012-X.

[32] B. Coifman, "Jam occupancy and other lingering problems with empirical fundamental relationships," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2422, no. 1, pp. 104–112, Jan. 2014, doi: 10.3141/2422-12.

[33] H. J. Payne, E. D. Helfenbein, and H. C. Knobel, "Development and testing of incident detection algorithms," Federal Highway Admin., Santa Monica, CA, USA, Tech. Rep. FHWA-RD-76-20, 1976.

[34] P. J. Bickel, C. Chen, J. Kwon, J. Rice, E. van Zwet, and P. Varaiya, "Measuring traffic," *Stat. Sci.*, vol. 22, no. 4, pp. 581–597, Nov. 2007, doi: 10.1214/07-STS238.

[35] H. Lee and B. Coifman, "Identifying and correcting pulse-breakup errors from freeway loop detectors," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2256, no. 1, pp. 68–78, Jan. 2011, doi: 10.3141/2256-09.

[36] H. Lee and B. Coifman, "Identifying chronic splashover errors at freeway loop detectors," *Transp. Res. C, Emerg. Technol.*, vol. 24, pp. 141–156, Oct. 2012, doi: 10.1016/j.trc.2012.02.005.

[37] P. Athol, *Interdependence of Certain Operational Characteristics Within a Moving Traffic Stream*, no. 72. Washington DC, USA: HRB, National Research Council, Highway Research Record, 1965, pp. 58–87.

[38] B. Coifman, "Improved velocity estimation using single loop detectors," *Transp. Res. A, Policy Pract.*, vol. 35, no. 10, pp. 863–880, Dec. 2001, doi: 10.1016/S0965-8564(00)00028-8.

[39] Y. Kim and F. L. Hall, "Relationships between occupancy and density reflecting average vehicle lengths," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1883, no. 1, pp. 85–93, Jan. 2004, doi: 10.3141/1883-10.

[40] F. L. Hall, V. F. Hurdle, and J. H. Banks, "Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways," *Transp. Res. Rec.*, no. 1365, pp. 12–18, 1992.

[41] M. Papageorgiou and G. Vigos, "Relating time-occupancy measurements to space-occupancy and link vehicle-count," *Transp. Res. C, Emerg. Technol.*, vol. 16, no. 1, pp. 1–17, 2008, doi: 10.1016/j.trc.2007.06.001.

[42] T. Courbon and L. Leclercq, "Cross-comparison of macroscopic fundamental diagram estimation methods," *Proc. Social Behav. Sci.*, vol. 20, pp. 417–426, Jan. 2011, doi: 10.1016/j.sbspro.2011.08.048.

[43] L. Ambühl, A. Loder, N. Zheng, K. W. Axhausen, and M. Menéndez, "Approximative network partitioning for MFDs from stationary sensor data," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 6, pp. 94–103, Jun. 2019, doi: 10.1177/0361198119843264.

[44] F. L. Hall and B. N. Persaud, "Evaluation of speed estimates made with single-detector data from freeway traffic management systems," *Transp. Res. Rec.*, no. 1232, pp. 9–15, 1989.

[45] D. M. Bramich, M. Menéndez, and L. Ambühl, 2022, "Road traffic time-series measurements of flow and occupancy from 10,150 loop detectors in 25 cities," Harvard Dataverse, V1. [Online]. Available: https://dataverse.harvard.edu, doi: 10.7910/DVN/FSGDQM.

[46] R. P. Inman, "A generalized congestion function for highway travel," *J. Urban Econ.*, vol. 5, pp. 21–34, Jan. 1978, doi: 10.1016/0094-1190(78)90034-7.

[47] H. Wang, D. Ni, Q.-Y. Chen, and J. Li, "Stochastic modeling of the equilibrium speed-density relationship," *J. Adv. Transp.*, vol. 47, no. 1, pp. 126–150, Jan. 2013, doi: 10.1002/atr.172.

[48] D. Ni, H. K. Hsieh, and T. Jiang, "Modeling phase diagrams as stochastic processes with application in vehicular traffic flow," *Appl. Math. Model.*, vol. 53, pp. 106–117, Jan. 2018, doi: 10.1016/j.apm.2017.08.029.

[49] L. Bai, S. C. Wong, P. Xu, A. H. F. Chow, and W. H. K. Lam, "Calibration of stochastic link-based fundamental diagram with explicit consideration of speed heterogeneity," *Transp. Res. B, Methodol.*, vol. 150, pp. 524–539, Aug. 2021, doi: 10.1016/j.trb.2021.06.021.

[50] W. Vickrey, "Optimization of traffic and facilities," *J. Transp. Econ. Policy*, vol. 1, pp. 123–136, May 1967.

[51] R. J. Smeed, "Traffic studies and urban congestion," *J. Transp. Econ. Policy*, vol. 2, pp. 33–70, Jan. 1968.

[52] A. E. Boardman and L. B. Lave, "Highway congestion and congestion tolls," *J. Urban Econ.*, vol. 4, pp. 340–359, Jul. 1977, doi: 10.1016/0094-1190(77)90016-X.

[53] D. O'Mathuna, "Highway traffic kinematics and the characteristic relation," *Transp. Planning Technol.*, vol. 4, no. 2, pp. 89–96, Jan. 1978, doi: 10.1080/03081067808717124.

[54] K. H. Stappert and T. J. Theis, "Aktualisierung der verkehrsstärke-/verkehrsgeschwindigkeitsbeziehungen des BVWP-netzmodells," Heusch/Boesefeldt, Aachen, Germany, Tech. Rep. VU18009V89, (in German), 1990.

[55] M. Zhou and F. L. Hall, "Investigation of speed-flow relationship under congested conditions on a freeway," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1678, no. 1, pp. 64–72, Jan. 1999, doi: 10.3141/1678-09.

[56] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 62, no. 2, pp. 1805–1824, 2000, doi: 10.1103/PhysRevE.62.1805.

[57] J. Einbeck and G. Tutz, "Modelling beyond regression functions: An application of multimodal regression to speed–flow data," *J. Roy. Stat. Soc., C Appl. Statist.*, vol. 55, no. 4, pp. 461–475, Aug. 2006, doi: 10.1111/j.1467-9876.2006.00547.x.

[58] D. Ni, "A unified perspective on traffic flow theory, part I: The field theory," in *Proc. 11th Int. Conf. Chin. Transp. Professionals (ICCTP)*, 2011, pp. 4227–4243, Nanjing, China, doi: 10.1061/41186(421)420.

[59] D. Ni, J. D. Leonard, C. Jia, and J. Wang, "Vehicle longitudinal control and traffic stream modeling," *Transp. Sci.*, vol. 50, no. 3, pp. 1016–1031, Jul. 2016, doi: 10.1287/trsc.2015.0614.

[60] B. S. Kerner and P. Konhäuser, "Structure and parameters of clusters in traffic flow," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 50, no. 1, pp. 54–83, Jul. 1994, doi: 10.1103/PhysRevE.50.54.

[61] H. Y. Lee, H.-W. Lee, and D. Kim, "Origin of synchronized traffic flow on highways and its dynamic phase transitions," *Phys. Rev. Lett.*, vol. 81, pp. 1130–1133, Aug. 1998, doi: 10.1103/PhysRevLett.81.1130.

[62] R. E. Chandler, R. Herman, and E. W. Montroll, "Traffic dynamics: Studies in car following," *Oper. Res.*, vol. 6, no. 2, pp. 165–184, Apr. 1958, doi: 10.1287/opre.6.2.165.

[63] K. M. Kockelman, "Changes in flow-density relationship due to environmental, vehicle, and driver characteristics," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1644, no. 1, pp. 47–56, Jan. 1998, doi: 10.3141/1644-06.

[64] J. E. Tolle, "Composite car following models," *Transp. Res.*, vol. 8, pp. 91–96, Apr. 1974, doi: 10.1016/0041-1647(74)90035-5.

[65] A. Ceder, "A two-regime traffic flow model and the consistency of its parameters," *Appl. Math. Model.*, vol. 2, pp. 261–270, 1978, doi: 10.1016/0307-904X(78)90019-7.

[66] S. M. Easa, "Selecting two-regime traffic-flow models," *Transp. Res. Rec.*, no. 869, pp. 25–36, 1982.

[67] E. De Romph, "A dynamic traffic assignment model: Theory and applications," M.S. thesis, Technische Universiteit Delft, Delft, The Netherlands, 1994.

[68] P. Nelson, D. D. Bui, and A. Sopasakis, "A novel traffic stream model deriving from a bimodal kinetic equilibrium," *IFAC Proc.* vol. 30, pp. 771–776, Jun. 1997, doi: 10.1016/S1474-6670(17)43915-2.

[69] B.-H. Wang, L. Wang, P. M. Hui, and B. Hu, "Analytical results for the steady state of traffic flow models with stochastic delay," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 58, no. 3, pp. 2876–2882, Sep. 1998, doi: 10.1103/PhysRevE.58.2876.

[70] Y. Lu, S. C. Wong, M. Zhang, C.-W. Shu, and W. Chen, "Explicit construction of entropy solutions for the Lighthill–Whitham–Richards traffic flow model with a piecewise quadratic flow–density relationship," *Transp. Res. B, Methodol.*, vol. 42, no. 4, pp. 355–372, May 2008, doi: 10.1016/j.trb.2007.08.004.

[71] J. Yao, H. Rakha, H. Teng, V. Kwigizile, and M. Kaseko, "Estimating highway capacity considering two-regime models," *J. Transp. Eng.*, vol. 135, no. 9, pp. 670–676, Sep. 2009, doi: 10.1061/(ASCE)0733-947X(2009)135:9(670).

[72] A. H. F. Chow, Y. Li, and K. Gkiotsalitis, "Specifications of fundamental diagrams for dynamic traffic modeling," *J. Transp. Eng.*, vol. 141, Sep. 2015, Art. no. 04015015, doi: 10.1061/(ASCE)TE.1943-5436.0000781.

[73] D. Wang, X. Ma, D. Ma, and S. Jin, "A novel speed–density relationship model based on the energy conservation concept," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1179–1189, May 2017, doi: 10.1109/TITS.2016.2602839.

[74] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2013.

[75] D. Chen, J. Zhang, S. Tang, and J. Wang, "Freeway traffic stream modeling based on principal curves and its analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 246–258, Dec. 2004, doi: 10.1109/TITS.2004.838226.

[76] C. Antoniou and H. N. Koutsopoulos, "Estimation of traffic dynamics models with machine-learning methods," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1965, no. 1, pp. 103–111, Jan. 2006, doi: 10.1177/0361198106196500111.

[77] M. Hadiuzzaman, M. R. K. Siam, N. Haque, T. H. Shimu, and F. Rahman, "Adaptive neuro-fuzzy approach for modeling equilibrium speed–density relationship," *Transportmetrica A, Transp. Sci.*, vol. 14, no. 9, pp. 784–808, Oct. 2018, doi: 10.1080/23249935.2018.1435589.

[78] D. C. Gazis, R. Herman, and R. W. Rothery, "Nonlinear Follow-the-Leader models of traffic flow," *Operations Res.*, vol. 9, no. 4, pp. 545–567, Aug. 1961, doi: 10.1287/opre.9.4.545.

[79] D. R. Drew, "Theoretical approaches to the study and control of freeway congestion," Freeway Surveillance Control Res. Rep., Texas Transp. Inst., Texas A&M Univ., College Station, TX, USA, Res. Rep. 24-1, 1964.

[80] J. S. Drake, J. L. Schofer, and A. D. May, "A statistical analysis of speed density hypotheses," *Highway Res. Rec.*, no. 154, pp. 53–87, 1966.

[81] L. A. Pipes, "Car following models and the fundamental diagram of road traffic," *Transp. Res.*, vol. 1, pp. 21–29, Nov. 1967, doi: 10.1016/0041-1647(67)90092-5.

[82] A. D. May, *Traffic Flow Fundamentals*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1990.

[83] A. D. May and H. E. M. Keller, "Non-integer car-following models," *Highway Res. Rec.*, no. 199, pp. 19–32, 1967.

[84] J. M. Del Castillo and F. G. Benitez, "On the functional form of the speed-density relationship—I: General theory," *Transp. Res. B, Methodol.*, vol. 29, pp. 373–389, Oct. 1995, doi: 10.1016/0191-2615(95)00008-2.

[85] F. van Wageningen-Kessels, H. van Lint, K. Vuik, and S. Hoogendoorn, "Genealogy of traffic flow models," *EURO J. Transp. Logistics*, vol. 4, no. 4, pp. 445–473, 2015, doi: 10.1007/s13676-014-0045-5.

[86] H. Greenberg, "An analysis of traffic flow," *Oper. Res.*, vol. 7, no. 1, pp. 79–85, Feb. 1959, doi: 10.1287/opre.7.1.79.

[87] D. C. Gazis, R. Herman, and R. B. Potts, "Car-following theory of steady-state traffic flow," *Oper. Res.*, vol. 7, no. 4, pp. 499–505, Aug. 1959, doi: 10.1287/opre.7.4.499.

[88] L. C. Edie, "Car-following and steady-state theory for noncongested traffic," *Operations Res.*, vol. 9, no. 1, pp. 66–76, Feb. 1961, doi: 10.1287/opre.9.1.66.

[89] R. T. Underwood, "Speed, volume and density relationships," in *Quality and Theory of Traffic Flow*. Bureau of Highway Traffic, Yale University, New Haven, Connecticut, 1961, pp. 141–188.

[90] G. F. Newell, "Nonlinear effects in the dynamics of car following," *Oper. Res.*, vol. 9, no. 2, pp. 209–229, Apr. 1961, doi: 10.1287/opre.9.2.209.

[91] R. E. Franklin, "The structure of a traffic shock wave," *Civil Eng. Public Works Rev.*, vol. 56, pp. 1186–1188, 1961.

[92] S. M. Easa and A. D. May, "Generalized procedure for estimating single- and two-regime traffic-flow models," *Transp. Res. Rec.*, no. 772, pp. 24–37, 1980.

[93] S. El-Hosaini, "Transport behaviour at high densities," *Transp. Res.*, vol. 7, pp. 421–425, Jan. 1973, doi: 10.1016/0041-1647(73)90027-0.

[94] A. C. Dick, "Speed/flow relationships within an urban area," *Traffic Eng. Control*, vol. 8, pp. 393–396, Jul. 1966.

[95] P. K. Munjal, Y. S. Hsu, and R. L. Lawrence, "Analysis and validation of lane-drop effects on multi-lane freeways," *Transp. Res.*, vol. 5, pp. 257–266, Dec. 1971, doi: 10.1016/0041-1647(71)90037-2.

[96] M. van Aerde, "A single regime speed-flow-density relationship for freeways and arterials," in *Proc. 74th Annu. Meeting Transp. Res. Board, Paper*, Washington, DC, USA, 1995, pp. 1–19.

[97] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama, "Dynamical model of traffic congestion and numerical simulation," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, pp. 1035–1042, Feb. 1995, doi: 10.1103/PhysRevE.51.1035.

[98] M. Ghandehari and S. Ardekani, "A modified greenberg speed-flow traffic model," Dept. Math., Univ. Texas Arlington, Arlington, TX, USA, Tech. Rep. 357, 2008.

[99] M. J. MacNicholas, "A simple and pragmatic representation of traffic flow," in *Proc. 75 Years Fundamental Diagram Traffic Flow Theory, Greenshields Symp., Transp. Res. Circular*, 2008, pp. 161–177.

[100] P. H. C. Eilers, B. D. Marx, and M. Durbán, "Twenty years of P-splines," *Statist. Oper. Res. Trans.*, vol. 39, no. 2, pp. 149–186, 2015.

[101] X. Qu, S. Wang, and J. Zhang, "On the fundamental diagram for freeway traffic: A novel calibration approach for single-regime models," *Transp. Res. B, Methodol.*, vol. 73, pp. 91–102, Mar. 2015, doi: 10.1016/j.trb.2015.01.001.

[102] D. Helbing, "Empirical traffic data and their implications for traffic modeling," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 55, no. 1, pp. R25–R28, Jan. 1997, doi: 10.1103/PhysRevE.55.R25.

[103] D. Helbing, "Fundamentals of traffic flow," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 55, no. 3, pp. 3735–3738, Mar. 1997, doi: 10.1103/PhysRevE.55.3735.

[104] R. A. Rigby and D. M. Stasinopoulos, "Generalized additive models for location, scale and shape," *J. Roy. Stat. Soc. C, Appl. Statist.*, vol. 54, no. 3, pp. 507–554, 2005, doi: 10.1111/j.1467-9876.2005.00510.x.

[105] M. D. Stasinopoulos *et al.*, *Flexible Regression and Smoothing: Using GAMLSS in R*. Boca Raton, FL, USA: CRC Press, 2017.

[106] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, Sep. 1995, doi: 10.1137/0916069.

[107] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965, doi: 10.1093/comjnl/7.4.308.

[108] R. P. Brent, "An algorithm with guaranteed convergence for finding a zero of a function," *Comput. J.*, vol. 14, no. 4, pp. 422–425, 1971, doi: 10.1093/comjnl/14.4.422.

[109] R. A. Rigby and D. M. Stasinopoulos, "Automatic smoothing parameter selection in GAMLSS with an application to centile estimation," *Stat. Methods Med. Res.*, vol. 23, no. 4, pp. 318–332, Aug. 2014, doi: 10.1177/0962280212473302.

[110] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974, doi: 10.1109/TAC.1974.1100705.

[111] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978, doi: 10.1214/aos/1176344136.

[112] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951, doi: 10.1214/aoms/1177729694.

[113] C. M. Hurvich and C.-L. Tsai, "Regression and time series model selection in small samples," *Biometrika*, vol. 76, no. 2, pp. 297–307, 1989, doi: 10.1093/biomet/76.2.297.

[114] N. Sugiura, "Further analysts of the data by akaike' s information criterion and the finite corrections," *Commun. Statist. Theory Methods*, vol. 7, no. 1, pp. 13–26, Jan. 1978, doi: 10.1080/03610927808827599.

[115] K. Takeuchi, "Distribution of informational statistics and a criterion of model fitting," (in Japanese), *Suri Kagaku Math. Sci.*, vol. 153, pp. 12–18, 1976.

[116] S. Konishi and G. Kitagawa, "Generalised information criteria in model selection," *Biometrika*, vol. 83, no. 4, pp. 875–890, Dec. 1996, doi: 10.1093/biomet/83.4.875.

[117] S. Konishi, T. Ando, and S. Imoto, "Bayesian information criteria and smoothing parameter selection in radial basis function networks," *Biometrika*, vol. 91, no. 1, pp. 27–43, Mar. 2004, doi: 10.1093/biomet/91.1.27.

Daniel M. Bramich received the B.A. degree in mathematics and astrophysics from the University of Cambridge in 2000 and the Ph.D. degree in astronomy and astrophysics from the University of St. Andrews in 2004. He has worked in various universities and institutions around the world in both research and support roles, and he is currently working as a Research Scientist at New York University Abu Dhabi. His research interests are varied and include extra-solar planets, microlensing, variable stars, fundamental diagrams of road traffic, statistical modeling, and big data.

Mónica Menéndez (Member, IEEE) received the dual B.S. degrees *(summa cum laude)* in civil and architectural engineering from the University of Miami, FL, USA, in 2002, and the M.S. and Ph.D. degrees (focusing on transportation) from the University of California at Berkeley, Berkeley, CA, USA, in 2003 and 2006, respectively. Between 2010 and 2017, she was the Director of the Traffic Engineering Research Group, ETH Zürich, and prior to that, a Management Consultant at Bain & Company. She is currently the Associate Dean of engineering for graduate affairs and a Professor of civil and urban engineering at New York University Abu Dhabi; as well as a Global Network Professor of civil and urban engineering at the New York University Tandon School of Engineering. She is also the Director of the Research Center for Interacting Urban Networks (CITIES). She has authored or coauthored over 80 journal articles and over 180 conference contributions and technical reports in the area of transportation. Her research interests include multimodal transportation systems paying special attention to new technologies and information sources. She is an active reviewer for over 20 journals, and a member of multiple editorial boards for top journals in transportation.

Lukas Ambühl received the B.Sc. degree in civil engineering and the M.Sc. degree in management, technology, and economics from ETH Zürich (ETHZ), Switzerland, in 2012 and 2015, respectively, and the Ph.D. degree from the Research Group of Traffic Engineering, Institute for Transport Planning and Systems, ETHZ. He is currently a Post-Doctoral Researcher at ETHZ.