A three-phase fundamental diagram from three-dimensional traffic data

Maria Laura Delle Monache, Karen Chi, Yong Chen, Paola Goatin, Ke Han, Jing-Mei Qiu, Benedetto Piccoli

A three-phase fundamental diagram from three-dimensional traffic data

Maria Laura Delle Monache[*], Karen Chi[†], Yong Chen[‡], Paola Goatin[§], Ke Han [¶], Jing-Mei Qiu[‖], Benedetto Piccoli [**]

Abstract

This paper uses empirical traffic data collected from three locations in Europe and the US to reveal a three-phase fundamental diagram with two phases located in the uncongested regime. Model-based clustering, hypothesis testing, and regression analyses are applied to the speed-flow-occupancy relationship represented in the three-dimensional space to rigorously validate the three phases and identify their gaps. The finding is consistent across the aforementioned different geographical locations. Accordingly, we propose a three-phase macroscopic traffic flow model and a characterization of solutions to the Riemann problems. This work identifies critical structures in the fundamental diagram that are typically ignored in first- and higher-order models, and could significantly impact travel time estimation on highways.

Keywords: Fundamental diagram ; Traffic data ; Gap analysis ; Multi-phase models

1 Introduction

The *fundamental diagram* (FD), a term first coined by Haight in 1963 ([33]), usually refers to the empirically observed flow-occupancy curve, where the flow indicates number of vehicle passing through a fixed point per unit of time and the occupancy the number of vehicles per unit of space. Similarly, in traffic

[*]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France, email: ml.dellemonache@inria.fr

[†]University of Pennsylvania, PA, USA, email: karenchi@sas.upenn.edu

[‡]University of Pennsylvania, PA, USA, email: ychen123@mail.med.upenn.edu

[§]Université Côte d'Azur, Inria, CNRS, LJAD, Sophia Antipolis, France, email: paola.goatin@inria.fr

[¶]Imperial College, London, UK, email: k.han@imperial.ac.uk

[‖]University of Delaware, DE, USA, email: jingqiu@udel.edu

[**]Rutgers University-Camden, NJ, USA, email: piccoli@camden.rutgers.edu

modeling the same term indicates the functional relationship between flow and density (modeling counterpart of occupancy) or between average speed of vehicles and density. Because of the physical relationship *flow* = *density* × *speed*, one can use either of the definitions of FD to determine completely a traffic model.
The FD forms the foundation of some transportation applications and permeates in all level of mathematical description of traffic flow. Two of the most commonly used macroscopic/mesoscopic models are the celebrated Lighthill-Whitham-Richards (briefly LWR), ([42, 48]) and its discrete counterpart the Cell Transmission Model (briefly CTM ([15])). In both cases the FD provides a closure of the evolution equations, thus allowing a well-posed theory and well-grounded simulation tools, see [25]. For macroscopic fluid-dynamic models, there is a rich discussion on FD, see for instance [25, 11, 13, 12, 29, 10]. It is also well known that the generic flow-occupancy relationship is featured not only in the characteristics of traffic flows on single roads but also on networks. In the latter case one refers to the macroscopic fundamental diagram (MFD), see [36, 9, 27, 14, 18]. The exploitation of the MFD allows decision makers to devise robust traffic control policies on a regional scale to balance demand and supply to improve mobility.

In this article we focus on the FD for single roads by proposing a new approach to study the fundamental relationship between flow, density and speed. We propose novel statistical methodologies to analyze traffic data from fixed sensors focusing on the three-leg relationships among the flow, density and speed. In particular, rather than considering the FD as a two-quantity relationship (flow-density or speed-density), we analyze data in the three-dimensional space represented by flow, density and speed. This allows us to better exploit the statistical tools, in particular for the analysis of traffic regimes.

Conventionally, observed traffic exhibits *free* and *congested* phases, with the first corresponding to stable and regular traffic, while the second reflects delays and congestion. Moreover, in the early 2000, Kerner ([38]) introduced a tree-phase traffic theory, based on the distinction among *free flow*, *synchronized flow* and *wide-moving jam*. The last two phases are associated with congested traffic. In particular, the wide-moving jam is a phase in which there is a moving jam that maintains the mean velocity of the downstream jam front, even when the jam propagates through other traffic phases or bottlenecks.

In this paper, using clustering methodologies, we are able to identify three traffic regimes, which are distinct in a statistically significant fashion. Interestingly, two regimes appear in what is commonly referred to as the free flow traffic and the third corresponds to the congested phase. This analysis does not contradict Kerner's theory but rather points out that the static/stationary free-flow condition in the FD could exhibit two distinct phases, while the distinction of phases in congested traffic (e.g. Kerner's model) is mainly dynamic.

The second main empirical result of our paper is the clear evidence of existence of a gap between the two phases of free flow and the congested one. While the appearance of such gap is best visualized in the 3D representation of the FD relationships, we use the classical flow-density relationship to statistically prove

the existence of the gap. The main purpose is to prove the ubiquity (w.r.t. data collected at different geographical location and on different road types) of the gap in the classical setting and to enable a simpler analysis.

Building on the empirical evidence illustrated so far, we propose a new three-phase macroscopic model. The LWR model and the CTM are very popular in the traffic literature due to their simple mathematical representation. However, they have certain modeling limitations especially when it comes to describing complex wave structures such as stop and go waves, phantom jam and capacity drop. These are primarily caused by the instability of congested traffic and heterogeneous driving behavior that are not captured by the FD used in the LWR model. Motivated by these needs, in the 70s second-order models (systems of conservation laws) were developed, see [45, 55]. To overcome various limitations pointed out by Daganzo ([17]), Aw-Rascle ([5]) and independently Zhang ([58]) proposed a new model with conservation of a modified momentum. This so-called Aw-Rascle-Zhang (ARZ) model can be interpreted as part of a general family called General Second Order Models (briefly GSOM, see [40]). Such models consist of the usual conservation of mass and the advective transport of a Lagrangian (or single driver) variable, which can represent, for instance, the desired speed of drivers. A recently proposed model of this category is the Collapsed Generalized ARZ model (briefly CGARZ ([22])), where the driver speed depends on the Lagrangian variable only in the congested phase. Another line of research focuses on models showing two distinct phases, called the phase transition models ([11, 29, 12]). Such extensions were also motivated by empirical data: as shown in Figure 2, the congested phase shows a cluster of data points that, instead of forming a one-to-one relationship between density and flow (or velocity), are scattered in a two-dimensional region.

Building on the experience provided by the aforementioned modeling efforts, our proposed model is a combination of the features offered by the ARZ, CGARZ and phase transition models. Our three-phase model not only has the characteristics of a CGARZ model with a gap among phases when analyzed in the flow-density space, but also exhibits the newly discovered phase when analyzed in the speed-density space. After showing how our model performs in data fitting, we provide a complete characterization of the characteristic curves and the solutions of the Riemann problems. The latter are the building block for solutions to Cauchy problems, see [25].

To sum up, the main novelties and contributions of our paper are as follows:

- Unlike most studies that focus on traffic data from a single source, we use data from multiple geographic locations in Europe and the US, and analyze the fundamental relationships among flow, density and speed in the 3-D space instead of the commonly adopted two-variable representation of the FD. In addition, we use a set of statistical tools including model-based clustering, hypothesis testing, and regression to analyze the traffic data.

- Following the above exercise, we discover three data clusters representing three traffic regimes, two of which are contained in the free-flow phase and the third corresponds to the congested phase. Moreover, we are able

to detect a statistically significant gap between the first two regimes and the third one. These findings are validated using multiple data sources, and the main features (regimes and gaps) are consistent across different geographical areas.

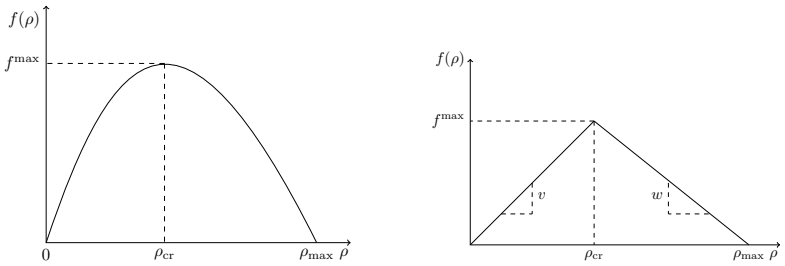
- Building on the first two, we propose a new three-phase macroscopic traffic flow model, which exhibits all the characteristics showed by our data analyses, and combines the features of the ARZ, CGARZ and phase transition models. A complete characterization of solutions of the Riemann problems is provided.

The revelation of the two distinct phases in the uncongested branch of the FD has a significant implication on highway traffic modeling. In particular, the first phase, coined the *free choice phase*, exhibits considerable speed variation for very low traffic density, which suggests that the driving behavior is characterized by free choice with little interaction with other traffic, as opposed to car-following behavior. While this new feature has relatively low impact on congestion or throughput, it could drastically influence travel time estimation in light traffic due to the huge speed variation, which could further impact traffic prediction ([35]) or dynamic traffic assignment ([46]). The proposed three-phase model has the potential to model highway traffic more accurately with refined and rigorous treatment of the uncongested phase, which is rarely seen in the literature as most existing studies tend to simplify the treatment of uncongested traffic with piecewise linear or concave FDs ([26, Chapter 8], [8, 47]). The complete characterization of solutions to the Riemann problems serves as the foundation for solving initial-value (Cauchy) problems; more elaborate treatment of initial and boundary problems, as well as network extensions will be investigated in subsequent studies.

The article is organized as follows. In Section 2 we discuss the existing literature about fundamental diagram, in Section 3 and 4 we introduce the datasets, their statistical analysis and the results obtained. Moreover, we describe the impacts of these results on traffic modeling and lastly in Section 5 we propose a new three-phase macroscopic model.

2 Basics of the Fundamental Diagram

The first traffic experiments for data collection were done in the thirties by Greenshields who carried out tests to measure traffic flow, traffic density and speed using photographic measurement for the first time ([31]). He then postulated a linear relationship between speed and traffic density. Using the relation $\textit{flow} = \textit{density} \times \textit{speed}$, the linear speed-density relation converts into a parabolic relation between speed and traffic flow, see Figure 1a. Later on, Lighthill Whitham and Richards ([42, 48]) preached a triangular function as flow-density curve, see Figure 1b. Subsequently, Edie ([20]) showed that empiric flux density data in the area of maximum traffic flow presents a discontinuity, therefore proposing a discontinuous fundamental diagram.



(a) Greenshields FD. (b) Triangular FD

Figure 1: Flux function, commonly referred to as fundamental diagram in the transportation literature.

Early works ([53, 51]) discussed the development of macroscopic models for arterials, which were later on extended to networks. In [50], Smeed modeled the number of vehicles that could enter the center city as a function of the area of the town, the fraction of area devoted to roads. In [52] data from central London were used to develop a linear speed-flow model. Later on, in [54], a similar relation between average speed and flow was developed by directly incorporating average street width and average signal spacing into the model. In [57] relations between different regions in different cities between traffic intensity, road density and weighted space mean speed were developed using data from UK and USA. The idea of a Macroscopic Fundamental Diagram (MFD in short) is quite old and it is due to [30]. Some works looked for MFD pattern in data from lightly congested real networks: [30, 4, 44]. In particular, [56, 43] investigated through simulation, network level relationship between the three fundamental variables of traffic: speed, flow, and density. To derive these relationships they assume an additional relation between average fraction of vehicles stopped and the network concentration from the two-fluid theory introduced in [34]. In [16], Daganzo suggested that traffic can be modeled dynamically in large urban regions at an aggregated level, with the aim of monitoring and controlling aggregate number of circulating vehicles at the neighborhood level to improve city mobility. Later on Daganzo and Geroliminis proved the existence of MFD in [27]. Subsequently, in [18] they explore the connection between network structure and network's MFD for urban setting. In 2009, Bussion and Ladier (in [9]) examined the differences between arterial and freeway network and they showed that heterogeneity has a strong impact on the shap/scatter of a MFDs.In [28] properties of the MFD under congestion for urban traffic are explored. MFD estimation with different type of data (loop detectors, probe vehicles or a combination) is investigated in [41, 19] and others. Moreover, clustering has been widely used to study the impact of traffic data on the fundamental diagram, see for example [37, 49]. However, to the authors' knowledge, all the previous works use spatial clustering.

In the early 2000, Kerner ([38]) introduced a tree-phase traffic theory, based on the distinction among *free flow*, *synchronized flow* and *wide-moving jam*. The

last two phases are associated with congested traffic. In particular, the wide-moving jam is a phase in which there is a moving jam that maintains the mean velocity of the downstream jam front, even when the jam propagates through other traffic phases or bottlenecks.
The FD is the building brick of the LWR model. In fact, the model introduced by Lighthill and Whitham ([42]) and independently by Richards in ([48]) heavily relies on the relation $f = \rho \cdot v$ where f is the flow rate, ρ is the car density and v is the mean traffic speed. The equation describing the evolution in time of the traffic density is given by a hyperbolic partial differential equation (PDE) describing the conservation of mass:

$$\partial_t \rho + \partial_x f(\rho) = 0. \tag{1}$$

Later on, Daganzo ([15]) presented a simplified and discrete-time version of the LWR model, called the cell-transmission model (CTM). The CTM relies on the triangular FD and it is consistent with the Godunov discretization of the LWR ([39]). The main equations of this model are the following:

$$n_j(i+1) - n_j(i) = y_j(i) - y_{j+1}(i) \tag{2}$$

$$y_j(i) = \min\left\{n_{j-1}(i), Q_j(i), \frac{w}{v}(N_j(i) - n_j(i))\right\} \tag{3}$$

with $n_j(i)$ is the number of vehicles in the cell j at time i, $y_j(i)$ is the number of vehicles entering the the cell j from the cell $j-1$, $Q_j(i)$ is the maximum number of vehicles that can flow into the cell j during the time interval i; $N_j(i)$ is the holding capacity of the cell j at time i. The positive constants w and v are the forward and backward wave speeds, which are associated with the triangular fundamental diagram.

3 Data analysis

In this section, we describe the data analyzed in the paper and then present in detail the statistical analysis performed.

3.1 Experimental data

We consider traffic data collected by static sensors (magnetic coils or radars) located on urban and extra-urban roads and highways. Sensors capture these traffic data regularly over a period of time. The sensor data provide the following aggregated quantities over a short time interval (3 to 10 minutes).

- Flux (denoted as f): also known as flow or volume, is the number of vehicles passing through a fixed location per unit of time.
- Velocity (denoted as v): the average speed of vehicles per unit of time.

- Occupancy (denoted as o): the percentage of time that a vehicle covers the sensor over the unit time of data collection.

Occupancy acts as a surrogate for the true density of traffic, as true density is practically difficult to capture, though there is some measurement error involved with its calculation.
The data were collected from three different locations: Rome (Italy), Las Vegas (Nevada, USA) and Sophia Antipolis (France). Rome's data were provided by ATAC S.p.A. ([1]) (the municipal society for traffic monitoring and control of Rome), and refer to a road in the city of Rome, Viale del Muro Torto, which links the historical center with the northern area of the city. Data were collected over a period of a week on three sensors. Each collected quantity (occupancy, flow and speed) was aggregated on 1 minutes intervals. The data from Las Vegas were collected by the Regional Transportation Commission of Southern Nevada (RTC), Freeway and Arterial System of Transportation (FAST) division ([3]). The data were collected from 50 urban and freeway sensors over a period of 5 years and aggregated on 10 minutes intervals. The data from Sophia Antipolis were collected by the Département des Alpes Maritimes ([2]) on two extra-urban sensors over a period of 8 months and aggregated over 6 minutes intervals.

To sum up, our dataset sources are:

1. *Rome*: 3 sensor over 1 week (June 2006) aggregated every 1 minute
2. *Las Vegas*: 50 sensors over 5 years (2010-2015) aggregated over every 10 minutes.
3. *Sophia Antipolis*: 4 sensors for 8 months (January - August, 2014) every 6 minutes.

We use the information collected in Rome as the primary example to illustrate the data structure. The Rome data set contains data of each minute of the entire day for one week; thus 10,080 observations were collected. Since our primary data set from Rome consists of one week of observations, we analyzed data from only one week's worth of data in the other locations as well. The other data sets from Las Vegas and Sophia Antipolis were used to validate the results from the Rome data.

Figure 2 illustrates the pairwise plots among the three measured variables based on the dynamic data collected from a sensor located in the road Viale del Muro Torto in the city of Rome on a Monday. These plots can provide useful insight on the functional relationship between these variables in two-dimensional space. For instance, the plot of flux against occupancy suggests a linear relationship with small variation when occupancy is less than a threshold (known as the **free phase**), and *much larger variation* when occupancy is larger than the threshold (known as the **congestion phase**). Furthermore, *both flux vs. speed and flux vs. occupancy plots suggest a possible "gap" between free and congestion phases, which corresponds to phase transition. These are important features that need to be taken into consideration in the mathematical modeling.*

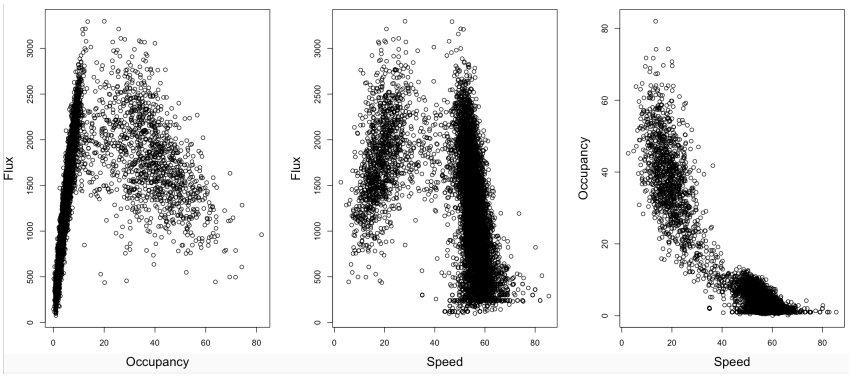


Figure 2: Pairwise scatterplots of Rome Data in the road Viale del Muro Torto: Flux vs Occupancy (left), Flux vs Speed (center) and Occupancy vs Speed (right).

Such pairwise plots are useful to generate data-driven hypotheses that need to be formally tested statistically and validated across different datasets.

Since we primarily focused on the three traffic characteristics of flux, velocity, and occupancy, we were conveniently positioned to analyze the data in three dimensions, a novel concept and approach that will be described in the Statistical Tools section.

3.2 Statistical tools

Cluster Analysis. Cluster analysis is the classification of data with a previously unknown structure and the partitioning of a data set into meaningful subsets. Clustering sheds light on hidden or non-intuitive relationships between that data and its attributes. Each cluster contains a group of objects that are more closely related to each other than they would be as objects of other clusters. *The concept of distance is thus inherently crucial in the process of cluster analysis, as clusters are grouped based on the results of this measure.* Distance serves as a way to evaluate the closeness, and likewise dissimilarity, of pairs of observations. There are at least two options to conduct cluster analysis for this traffic data: model-based clustering (e.g., mixture of normals) and non-parametric clustering (e.g., k-means). Although k-means is popular for complex

and high-dimensional data, it is generally used for data involving variables of the same scale (hence more suitable for data with spherical clusters, e.g., Euclidean distance in 3D), whereas our data consist of three variables of different scales. For this reason, model-based clustering has more flexibility in the shape of clusters; for instance, mixture models ([24]) can identify clusters in the traffic data that were ellipsoidal.
Empirical evaluations on the distributions of the three traffic variables through quantile-quantile (Q-Q) plot, Shapiro test, and Box-Cox transformation have suggested that normal distributions are appropriate. Here we propose the use of a finite mixture model with G multivariate normals ([24]). Specifically, denote data $\mathbf{y}$ with independent trivariate observations (flux, velocity and occupancy) $\{\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_n}\}$, the likelihood for a mixture model with G components is

$$\ell(\theta_1, \theta_2, \ldots, \theta_G; \pi_1, \pi_2, \ldots, \pi_G | \mathbf{y}) = \prod_{i=1}^{n} \sum_{k=1}^{G} \pi_k f_k(\mathbf{y_i}|\theta_k),$$

where i stands for ith observation, $f_k(\cdot)$ and θ_k are the density function and model parameters of the kth cluster in the mixture and π_k is the probability that an observation belongs to the kth cluster, subject to the simplex constraint $\{\pi_k \geq 0; \sum_{k=1}^{G} \pi_k = 1\}$. Such a model can be fitted by the expectation-maximization (EM) algorithm and is implemented by the R package 'mclust'.

Figure 3 provides a 3D visualization and cluster analysis result on the Rome dataset, where observations in different clusters are marked by different colors. Previous knowledge assumed that traffic involves two clusters: free flow and congestion. Free flow corresponds to steady traffic flow at high speeds (and low densities), while congestion is characterized by low flux and reduced speeds. From this new 3D visualization of data, we can identify a third phase, which we called the **"free choice phase"**, which corresponds to the situation of a relatively empty road whereby drivers choose their speed independently without influence from or interaction with other vehicles.
In the free choice phase, the flow of cars is low while the speed is variable. Model selection procedures (e.g., Bayesian information criterion - BIC - or adjusted BIC) have been used to select the number of clusters, and the datasets from Rome, Nevada and Sophia Antipolis have consistently suggested the existence of the third phase. *Such additional phase will be incorporated into the mathematical modeling.*

Notice that our three phases are different from those indicated by Kerner ([38]). Indeed we have two sub-phases in the free phase cluster opposed to Kerner's model with two sub-phases in the congestion phase cluster.

Gap analysis. We have developed and applied a rigorous hypothesis testing procedure to the datasets to formally investigate the presence of phase transitions. Specifically, investigating the presence of phase transition can be formulated as testing the existence of a "gap region" at the upper portion of occupancy in the free phase and its proximity to the lower portion of occupancy in congestion. As shown in the left panel of Figure 2, such gap region can be

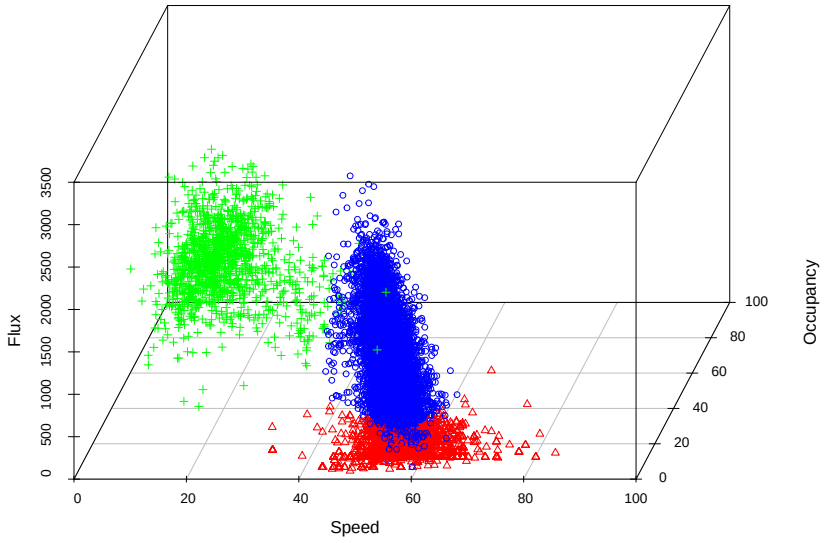


Figure 3: 3D visualization and cluster analysis results of Rome's data suggest the existence of third phase (red) in addition to the free flow (blue) and congestion (green) phases.

potentially masked by isolated points in the gap, which could be in fact due to measurement errors or random variations in flux and occupancy. To reduce the impact of these isolated points, we propose to take the upper quantile of the free phase (say, 95 percentile, denoted as ρ_{FP}) and the lower quantile of the congested phase (say, 5 percentile, denoted as ρ_C), and formally test for $H_0: \ \rho_{FP} \geq \rho_C$, i.e., there is no gap, against $H_a: \rho_{FP} < \rho_C$, i.e., there is a gap. Figure 4 illustrates these two scenarios of H_0 and H_a.

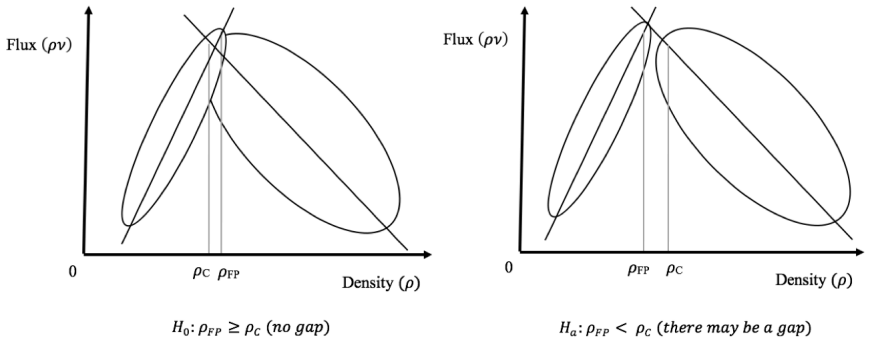


Figure 4: Illustration of hypothesis testing procedure for phase transition.

For given pre-specified percentiles q_{FP} and q_C (e.g., 95% and 5%), denote $\hat{\rho}_{FP}$ and $\hat{\rho}_C$ as the corresponding quantiles in the two clusters. The existence of phase transition can be formally tested by a *one-sided test* based on the Wald statistic

$$T = \min\{\hat{\rho}_{FP} - \hat{\rho}_C, 0\} / \sqrt{\text{var}(\hat{\rho}_{FP} - \hat{\rho}_C)}.$$

The variance $\text{var}(\hat{\rho}_{FP} - \hat{\rho}_C)$ can be approximated by

$$\text{var}(\hat{\rho}_{FP} - \hat{\rho}_C) = \text{var}(\hat{\rho}_{FP}) + \text{var}(\hat{\rho}_C) \approx \frac{q_{FP}(1-q_{FP})}{n_{FP}\{f(F^{-1}(q_{FP})\}^2} + \frac{q_C(1-q_C)}{n_C\{f(F^{-1}(q_C)\}^2},$$

where $f(\cdot)$ and $F(\cdot)$ stand for the estimated distribution function and cumulative distribution function of ρ, respectively, and n_{FP} and n_C stand for the number of data points in the free and congestion phases, respectively. The first equality in the calculation of variance is due to the independence of $\hat{\rho}_{FP}$ and $\hat{\rho}_C$ as they are estimated from different sets of observations. The approximation is due to a standard result for asymptotic variance of percentile estimates; see [32].

Table 1 presents the results from the described gap analysis on the Rome's, Nevada's and Sophia's datasets. After "trimming" a very small percent of data (e.g., 3%) and considering (97%, 3%) quantiles of free and congestion phases, the test statistics suggested a strong evidence of a gap (indicating phase transition) in the three datasets.

Dataset	Percentile (%)	Phase	Density	Test Statistic	P-value
Rome	97.5	FP	10	0	1
	2.50	C	10		
	97	FP	9	-1.79	**0.037**
	3	C	10		
	95	FP	9	-3.15	<**0.001**
	5	C	11		
Las Vegas	97.5	FP	12	-0.91	0.18
	2.50	C	13		
	97	FP	12	-2.17	**0.015**
	3	C	14		
	95	FP	11	-5.75	<**0.001**
	5	C	16		
Sophia	97.5	FP	5	-2.82	**0.002**
	2.50	C	10		
	97	FP	5	-6.44	< **0.001**
	3	C	12		
	95	FP	4	-5.79	<**0.001**
	5	C	12		

Table 1: Results of model-based cluster analysis using Rome, Las Vegas and Sophia Antipolis data. Phase: estimated phase using cluster analysis. FP: free phase; C: congestion phase. Density: estimated density value at the percentile.

4 Results and Discussion

We have conducted model-based cluster analyses on three datasets. We found statistically significant gaps between free and congested phases in all three datasets. We also used regression analysis to demonstrate the existence of free choice phase in explaining the variability in observed flux.

4.1 Results of Cluster Analysis

Graphical representations of the results are illustrated below. To clarify the color code for the graphs, the region colored in red corresponds to the Free Choice phase, blue to the Free Flow phase, and green to Congestion. The Free Flow phase in this three cluster model corresponds to the remainder of the original conception of Free Flow without Free Choice. The Free Choice and Free Flow phases from the three cluster model are collectively referred to as the Free Phase from here on. Figure 5 illustrates the clustering performed by 'mclust' of

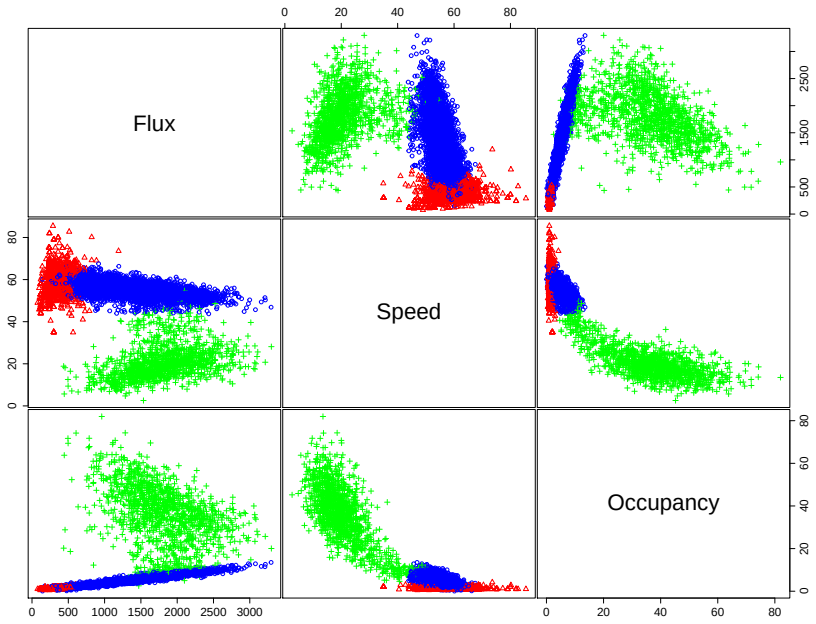


Figure 5: Pairs plot of clustered Rome data

Rome's data on a two dimensional level. Pairs plots present the data according to each pair of variables: velocity and flux, occupancy and flux, and velocity and occupancy. This type of plot provides insight on the shape and characteristics of the data in 2D. For instance, we observed some sort of gap between Free Phase and Congestion, which can be most easily viewed in the pairs plot of the variables occupancy and flux. Figure 3 is the three dimensional representation

of the Rome data, a novel approach to visualizing traffic data. Through this 3D plot of data, the proposed gap between Free Phase and Congestion is even more noticeable, reinforcing our observations from the 2D case. Specifying 'mclust' to filter through the data for two or three clusters has indicated there is some margin of difference between the original Free Flow phase in the two cluster model and the Free Phase in the three cluster model; the latter model is generally neater than the former. The disparity is minimal and perhaps insignificant, though it is worth noting.

Cluster analyses using the Las Vegas and Sophia Antipolis's data corroborate the results obtained with the Rome's data. Las Vegas data collected from highway sensors 25 and 99, which we we will refer to as Nevada 25 and Nevada 99 respectively, clusters into 9 phases through 'mclust'. However, we forced R to choose only 2 and 3 clusters to match the original 2 phase model of traffic flow proposed by the field and the 3 phase model discovered in this study. The results are reported in Figures 6 and 7. The data and analyses suggested that the 3 cluster model was preferred using Bayesian information criterion (BIC) for model selection ([24]). The data from Sophia Antipolis also demonstrate

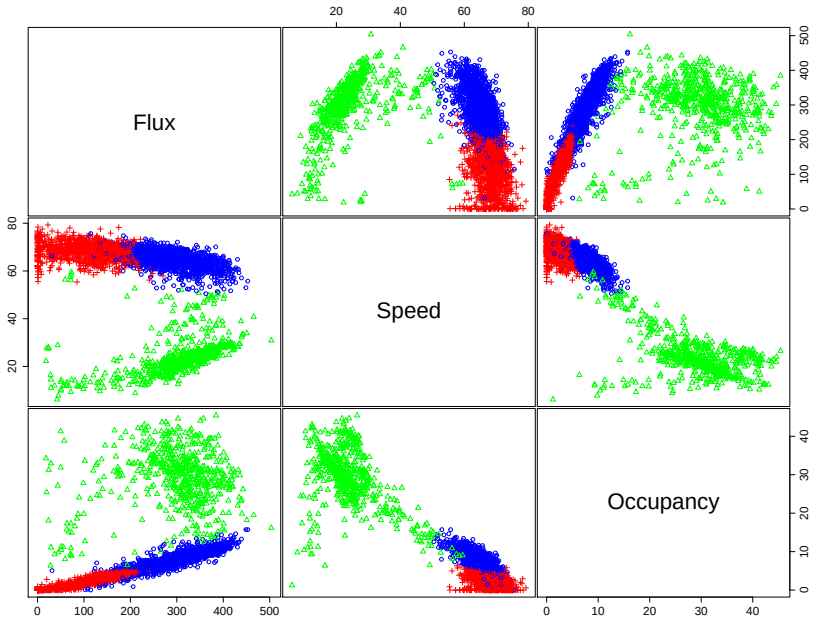


Figure 6: Pairs plot of clustered Las Vegas data

the existence of the Free Choice phase. These two data sets both have a wide range of observed speeds at low levels of occupancy, as seen in the pairs plots and 3D plots in Figures 8 and 9.

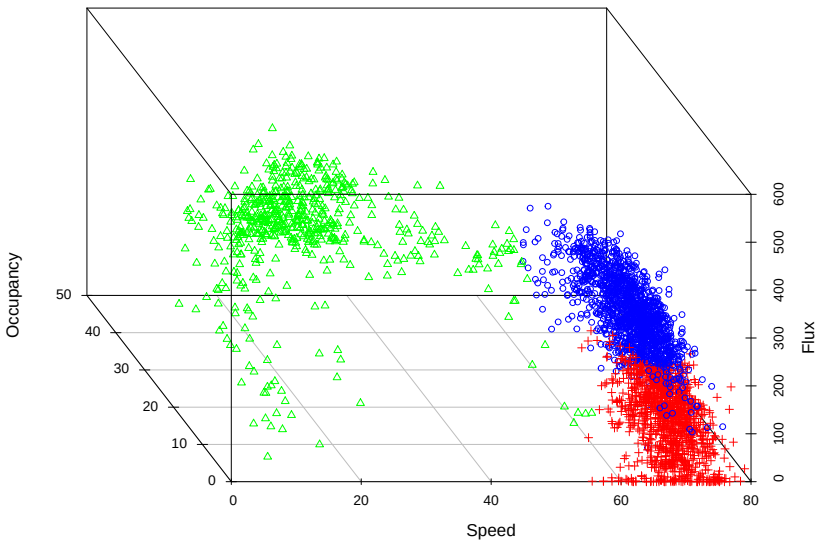


Figure 7: 3D plot of clustered Las Vegas data

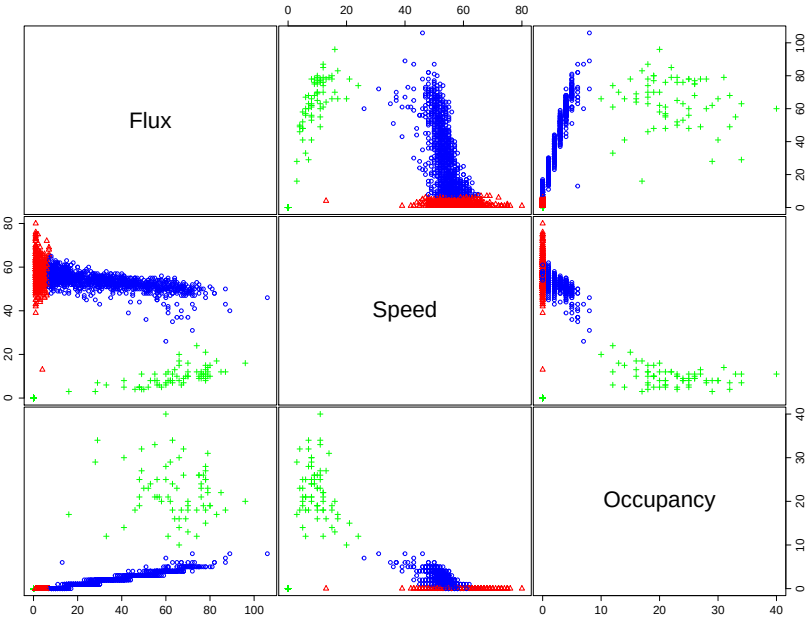


Figure 8: Pairs plot of clustered Sophia Antipolis data

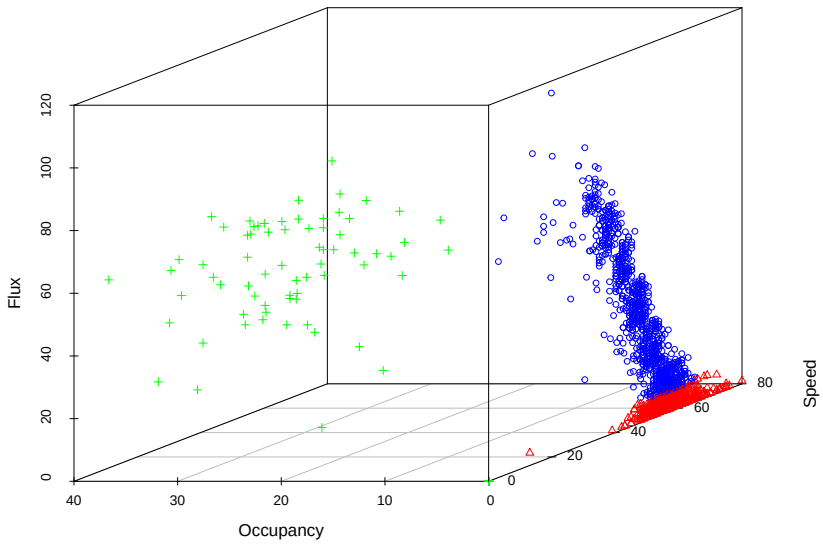


Figure 9: 3D plot of clustered Sophia Antipolis data

4.2 Quantifying the improved goodness of fit through RSS comparisons

We conducted further analyses by using residual sum of squares (RSS) values to compare various models considered in this paper. The RSS value was calculated as the sum of differences between the observed and fitted values of flux, where the fitted values of flux were obtained from two or three cluster models with or without speed as an additional predictor in addition to the occupancy. The RSS is an objective measure of the remaining variability of the flux that has not been explained by a particular model.

With Rome data, we considered a baseline model with two-phase and occupancy as the only predictor. In other words, this is a 2D and two-phase model as in the existing literature. We calculated the RSS value for this model, and compared with the RSS values from more complex models. Specifically, we found that adding speed into the model (i.e., a 3D and two-phase model) explained additionally 1.2% of the variability in observed flux, on top of the baseline 2D and two-phase model. Furthermore, adding the free choice phase alone (i.e., a 2D and three-phase model) reduced 13.6% of the remaining variability. Adding both speed and the free choice phase to the model (i.e., a 3D and three-phase model) further reduced 16.0% of the variability in observed flux, compared to the baseline model.

The results of these comparisons suggested that the three cluster model is

indeed superior to the two cluster model and that 3D rendering of the data is appropriate. The percent change for adjustment of number of clusters as well as dimensions both increases indicates a clear improvement; the three cluster model is better than the two cluster model, 3D analysis of data is more informative than that in 2D, and the 3D three cluster model provides the more favorable RSS value overall. These results have been consistent across our data sets and the calculations are deferred to the Appendix (Supplemental Materials, see Figures 12 through 15 and Tables 2 through 6). These results have important implications in understanding traffic flow. They confirm the utility of analyzing this type of data in three dimensions and reveal the presence of a third phase.

5 A macroscopic second-order model accounting for the 3 phases

Following the approach of [13, 22], we propose a new macroscopic model accounting for the 3 phases. In conservation form, the model can be expressed as

$$\begin{aligned} \partial_t \rho + \partial_x(\rho\, v(\rho, y/\rho)) &= 0, \\ \partial_t y + \partial_x(y\, v(\rho, y/\rho)) &= 0, \end{aligned} \tag{4}$$

where the velocity function is chosen such that

$$v(\rho, y/\rho) = \begin{cases} v_{FC}(\rho, y/\rho), & \text{if } 0 < \rho \le \rho_{FC}, \\ v_{FP}(\rho), & \text{if } \rho_{FC} < \rho \le \rho_{FP}, \\ v_C(\rho, y/\rho), & \text{if } \rho_{FP} \le \rho \le \rho_{\max}, \end{cases} \tag{5}$$

for some $0 < \rho_{FC} < \rho_{FP} < \rho_{\max}$, and it is continuous at ρ_{FC} and ρ_{FP}. In (4)-(5), the quantity $w = y/\rho \in [w_{\min}, w_{\max}]$ may represent various traffic characteristics, such as vehicles classes ([23]), aggressiveness ([21]), desired spacing ([59]), or perturbation from equilibrium ([7]), which are transported with the traffic stream. We will refer to the variable $y = \rho w$ as a *total property* ([22]). The function v defined in (5) must be:

1. Non-negative: $v(\rho, w) \ge 0$ for all $\rho \in [0, \rho_{\max}]$, $w \in [w_{\min}, w_{\max}]$;
2. Continuous: $v_{FC}(\rho_{FC}, w) = v_{FP}(\rho_{FC})$ and $v_C(\rho_{FP}, w) = v_{FP}(\rho_{FP})$ for all $w \in [w_{\min}, w_{\max}]$;
3. Vanishing at maximal density: $v(\rho_{\max}, w) = v_C(\rho_{\max}, w) = 0$ for all $w \in [w_{\min}, w_{\max}]$;
4. Non-decreasing with respect to w: $\dfrac{\partial v}{\partial w}(\rho, w) \ge 0$ for $\rho \in [0, \rho_{\max}]$.

To have a differentiable fundamental diagram the following regularity must be imposed:

$$\frac{\partial v_{FC}}{\partial \rho}(\rho_{FC}, w) = v'_{FP}(\rho_{FC}) \quad \text{for all } w \in [w_{\min}, w_{\max}]$$

and

$$\frac{\partial v_C}{\partial \rho}(\rho_{FP}, w) = v'_{FP}(\rho_{FP}) \quad \text{for all } w \in [w_{\min}, w_{\max}].$$

With the above assumptions, the corresponding flux function $q(\rho, w) = \rho\, v(\rho, w)$ has the following properties:

1. $q(0, w) = q(\rho_{\max}, w) = 0$ for all w;
2. Strict concavity:

$$\frac{\partial^2 q}{\partial \rho^2}(\rho, w) = 2\,\frac{\partial v}{\partial \rho}(\rho, w) + \rho\,\frac{\partial^2 v}{\partial \rho^2}(\rho, w) < 0; \tag{6}$$

if $\rho\,\dfrac{\partial^2 v}{\partial \rho^2}(\rho, w) < -2\,\dfrac{\partial v}{\partial \rho}(\rho, w)$ for all (ρ, w).

To take in to account the possible presence of a gap, as suggested by Table 1, we fix the value $v_C^{\max} \leq v_{FP}^{\min} := v_{FP}(\rho_{FP})$ of the maximal speed in congestion, and let $\rho_C \in [\rho_{FP}, \rho_{\max}[$ be the density value such that

$$v_C(\rho_C, w_{\min}) = v_C^{\max}.$$

Defining the velocity function, see Figure 10, as

$$v_g(\rho, w) = \begin{cases} v_{FC}(\rho, w), & \text{if } 0 < \rho \leq \rho_{FC}, \\ v_{FP}(\rho), & \text{if } \rho_{FC} < \rho \leq \rho_{FP}, \\ \min\left\{v_C^{\max}, v_C(\rho, w)\right\}, & \text{if } \rho_C \leq \rho \leq \rho_{\max}, \end{cases} \tag{7}$$

the corresponding flux function $q_g(\rho, w) := \rho v_g(\rho, w)$ displays the desired gap between the free-flow and congested phases, see Figure 11.

5.1 Riemann Solver

System (4) is defined on the invariant domain

$$\Omega = \left\{(\rho, \rho w) \in [0, \rho_{\max}] \times [0, \rho_{\max} w_{\max}] \colon w \in [w_{\min}, w_{\max}]\right\}.$$

We note that, under the above assumptions on the velocity function v, $(\rho, y) \in \Omega$ if and only if $w \in [w_{\min}, w_{\max}]$ and $v(\rho, y/\rho) \in [0, v(0, w_{\max})]$. The eigenvalues are given by

$$\lambda_1(\rho, y/\rho) = v(\rho, y/\rho) + \rho\,\frac{\partial}{\partial \rho} v(\rho, y/\rho) \quad \text{and} \quad \lambda_2(\rho, y/\rho) = v(\rho, y/\rho), \tag{8}$$

so the system is strictly hyperbolic for $\rho > 0$ as long as $\partial\, v(\rho, y/\rho)/\partial \rho \neq 0$ (often times this quantity is assumed to be negative). We note that the second characteristic field is linearly degenerate, giving origin to contact discontinuity waves, while the first characteristic field is genuinely non-linear if (6) holds. Moreover, the Riemann invariants of the systems are given by w and v. In

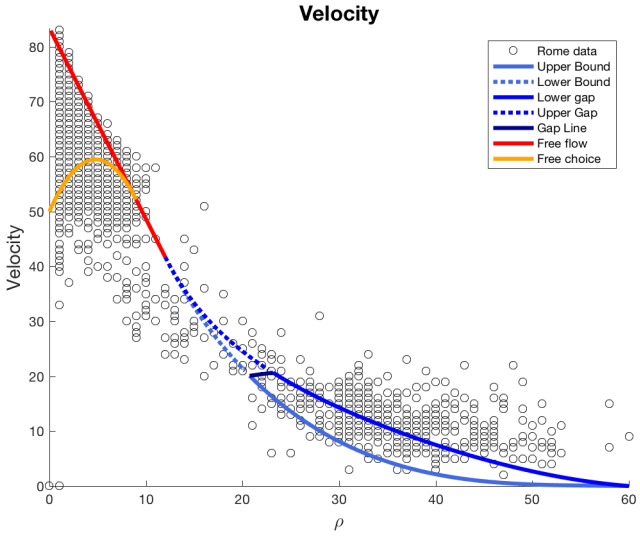


Figure 10: Speed function considered

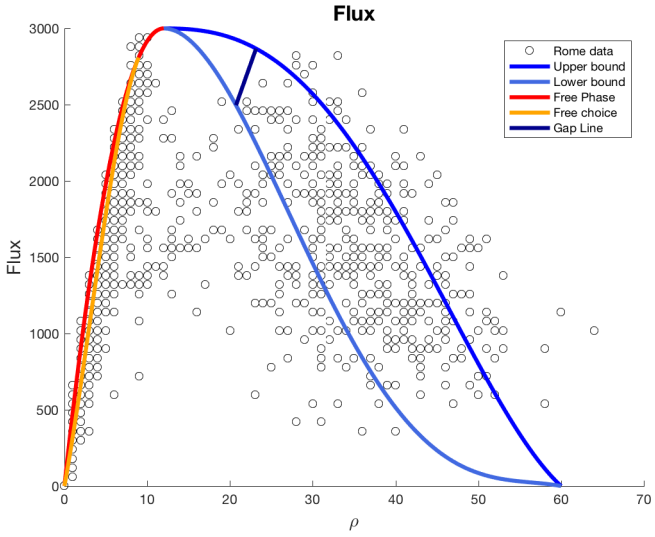


Figure 11: Flux function considered

particular, the iso-values $w = const$ correspond to waves of the 1st family (we recall that the system belongs to Temple class, i.e. shock and rarefaction curves coincide) and the contact discontinuities verify $v = const$. More precisely, in the strictly concave case (6), the elementary waves are constructed as follows.

- **1-rarefaction waves.** Two points $(\rho_l, \rho_l w_l)$ and $(\rho_r, \rho_r w_r)$ are connected by a 1-rarefaction wave if and only if

$$w_l = w_r \qquad \text{and} \qquad \lambda_1(\rho_l, w_l) < \lambda_1(\rho_r, w_r).$$

- **1-shock waves.** Two points $(\rho_l, \rho_l w_l)$ and $(\rho_r, \rho_r w_r)$ are connected by a 1-shock wave if and only if

$$w_l = w_r \qquad \text{and} \qquad \lambda_1(\rho_l, w_l) > \lambda_1(\rho_r, w_r).$$

  In this case, the jump discontinuity moves with speed

$$\sigma = \frac{\rho_l v(\rho_l, w_l) - \rho_m v(\rho_m, w_m)}{\rho_l - \rho_m}.$$

- **2-contact discontinuity.** Two points $(\rho_l, \rho_l w_l)$ and $(\rho_r, \rho_r w_r)$ are connected by a 2-contact wave if and only if

$$v(\rho_l, w_l) = v(\rho_r, w_r).$$

In the general (non-concave) case, the 1-waves consist of a concatenation of shocks and rarefactions (see [6, Sec. 1])

Based on the above elementary waves, the solution corresponding to general Riemann data $(\rho_l, \rho_l w_l)$, $(\rho_r, \rho_r w_r)$ can be constructed as follows. Let $(\rho_m, \rho_m w_m)$ be the intermediate point defined by

$$w_m = w_l, \qquad v(\rho_m, w_m) = v(\rho_r, w_r).$$

Setting $v_{w_l}(\rho) = v(\rho, w_l)$, ρ_m is given by

$$\rho_m = \begin{cases} v_{w_l}^{-1}(v(\rho_r, w_r)), & \text{if } v(\rho_r, w_r) < v(0, w_l), \\ 0, & \text{otherwise.} \end{cases}$$

In the latter case a vacuum zone appears between the sector

$$v(0, w_l)\, t < x < v(\rho_r, w_r)\, t.$$

The complete solution is then given by a 1-wave connecting $(\rho_l, \rho_l w_l)$ and $(\rho_m, \rho_m w_m)$, followed by a 2-contact discontinuity between $(\tilde{\rho}_m, \tilde{\rho}_m w_m)$ and $(\rho_r, \rho_r w_r)$ (eventually separated by a vacuum zone if $v(\rho_r, w_r) > v(\rho_m, w_l)$ and $\rho_m = 0$).

The presence of the gap between $v_C^{\max}$ and $v_{FP}^{\min}$ does not modify the procedure, since the definition domain

$$\Omega_g = \left\{(\rho, \rho w) \in \Omega \colon v(\rho, w) \in [0, v_C^{\max}] \cup [v_{FP}^{\min}, v(0, w_{\max})]\right\}$$

is still invariant. We set $\Omega_g = \Omega_{FP} \cup \Omega_C$ with

$$\Omega_{FP} = \left\{(\rho, \rho w) \in \Omega \colon v(\rho, w) \in [v_{FP}^{\min}, v(0, w_{\max})]\right\},$$
$$\Omega_C = \left\{(\rho, \rho w) \in \Omega \colon v(\rho, w) \in [0, v_C^{\max}]\ \right\}.$$

We can distinguish the following cases:

- If $(\rho_l, \rho_l w_l)$ and $(\rho_r, \rho_r w_r)$ belongs both to Ω_{FP} or Ω_C, the Riemann solver is defined as above.
- If $(\rho_l, \rho_l w_l) \in \Omega_C$ and $(\rho_r, \rho_r w_r) \in \Omega_{FP}$, the intermediate point $(\rho_m, \rho_m w_m)$ belongs to Ω_{FP}. Let $(\rho_c, \rho_c w_c) \in \partial\Omega_C$ the point defined by

  $$w_c = w_l, \qquad v(\rho_c, w_c) = v_C^{\max}.$$

  The solution is composed by 1-waves connecting (ρ_l, w_l) and $(\rho_c, \rho_c w_l)$, a phase-transition jump between $(\rho_c, \rho_c w_l)$ and $(\rho_{FP}, \rho_{FP} w_l)$ moving with speed

  $$\sigma = \frac{\rho_c v_C^{\max} - \rho_{FP} v_{FP}^{\min}}{\rho_c - \rho_{FP}},$$

  followed by 1-waves connecting $(\rho_{FP}, \rho_{FP} w_l)$ and $(\rho_m, \rho_m w_m)$ and eventually a 2-contact from $(\rho_m, \rho_m w_m)$ to $(\rho_r, \rho_r w_r)$.
- If $(\rho_l, \rho_l w_l) \in \Omega_{FP}$ and $(\rho_r, \rho_r w_r) \in \Omega_C$, the intermediate point $(\rho_m, \rho_m w_m)$ belongs to Ω_C. Therefore, the solution always contains a 1-wave (shock phase-transition) from $(\rho_l, \rho_l w_l)$ to $(\rho_m, \rho_m w_m)$, followed by a 2-contact discontinuity. Notice that the solution may also contain an intermediate 1-wave in the congested phase.

Acknowledgments. M.L. D. M. acknowledges that this article was developed in the framework of the Grenoble Alpes Data Institute, supported by the French National Research Agency under the "Investissements d'avenir" program (ANR-15-IDEX-02).
Y.C. research is supported in part by the National Institute of Health grants 1R01LM012607 and 1R01AI130460.
B.P. acknowledges the support of the NSF project Grant CNS # 1446715, the NSF project KI-Net Grant DMS # 1107444 and the Lopez chair endowment.
The authors thank ATAC S.p.a. for providing the traffic data from the city of Rome and the Départment des Alpes Maritimes for providing data from Sophia Antipolis.

Appendix: supplemental material

In this Appendix we present the outputs of statistical softwares for the RSS analysis of Section 4.2

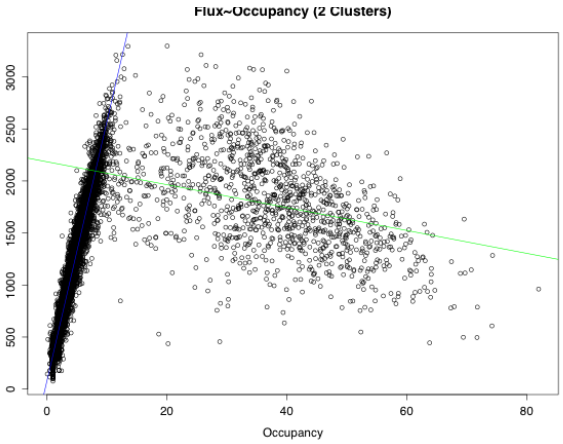


Figure 12: Flux Vs Occupancy: RSS analysis

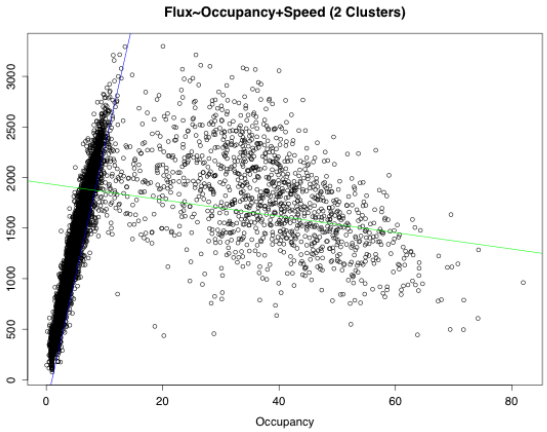


Figure 13: Flux Vs Occupancy+Speed: RSS analysis

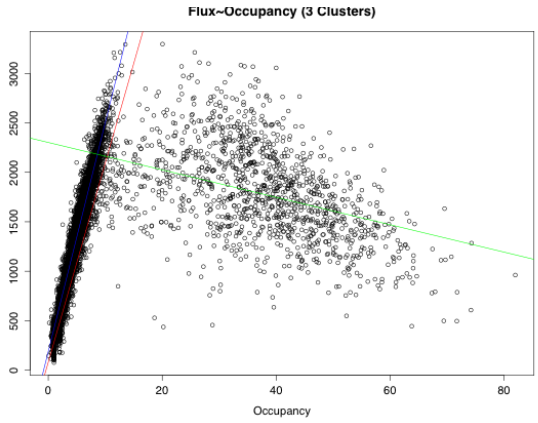


Figure 14: Flux Vs Occupancy: RSS analysis

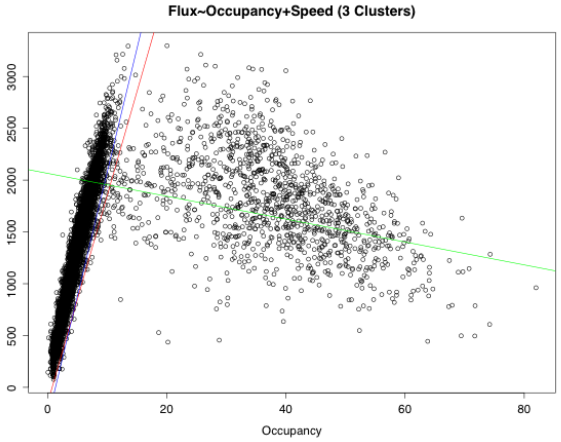


Figure 15: Flux Vs Occupancy+Speed: RSS analysis

	β_0	β_1	R^2	Adj. R^2	RSS
Free Phase	77.71	249.18	0.952	0.952	157,765,974
Congestion	2184.09	-11.00	0.09746	0.09676	283,255,197

Table 2: RSS analysis with 2 clusters with flux and occupancy. β_0 is the value of the intercept, β_1 the occupancy.

	β_0	β_1	β_2	R^2	Adj. R^2	RSS
Free Phase	-248.7	253.88	5.48	0.953	0.953	154,308,365
Congestion	1941.54	-8.11	6.64	0.1032	0.1018	281,448,225

Table 3: RSS analysis with 2 clusters with flux, occupancy and speed. β_0 is the value of the intercept, β_1 the occupancy, β_2 the speed.

	β_0	β_1	R^2	Adj. R^2	RSS
Free Choice	79.56	201.3	0.563	0.5628	14,269,933
Free Flow	187.68	231.36	0.9175	0.9175	126,714,084
Congestion	2302.55	-13.87	0.1618	0.1611	240,324,480

Table 4: RSS analysis with 3 clusters with flux and occupancy. β_0 is the value of the intercept, β_1 the occupancy.

	β_0	β_1	β_2	R^2	Adj. R^2	RSS
Free Choice	-153.88	200.36	4.03	0.6033	0.603	12,951,197
Free Flow	-316.91	239.01	8.43	0.9193	0.9192	123,984,911
Congestion	2065.61	-11.05	6.5	0.1676	0.1663	283,641,319

Table 5: Residual sum squared analysis with 3 clusters with flux, occupancy and speed. β_0 is the value of the intercept, β_1 the occupancy, β_2 the speed.

References

[1] ATAC S.p.A.

[2] Département des Alpes Maritimes.

[3] Regional Transportation Commission of Southern Nevada (RTC), Freeway and Arterial System of Transportation (FAST) division.

[4] S. Ardekani and R. Herman. Urban network-wide traffic variables and their relations. *Transportation Science*, 21(1):1–16, 1987.

% RSS improvement		
	2 clusters	3 clusters
2D	FP: 13.2% C: 15.8%	FC: 9.2% FF: 2.2% C: 0.7%
3D	FP: 11.3% C: 15.2%	FC: - FF: - C: -

Table 6: RSS improvement

[5] A. Aw and M. Rascle. Resurrection of "second order" models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000.

[6] S. Bianchini. The semigroup generated by a Temple class system with non-convex flux function. *Differential Integral Equations*, 13(10-12):1529–1550, 2000.

[7] S. Blandin, D. Work, P. Goatin, B. Piccoli, and A. Bayen. A general phase transition model for vehicular traffic. *SIAM journal on Applied Mathematics*, 71(1):107–127, 2011.

[8] M. C. Bliemer and M. P. Raadsen. Continuous-time general link transmission model with simplified fanning, part i: Theory and link model formulation. *Transportation Research Part B: Methodological*, 2018.

[9] C. Buisson and C. Ladier. Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. *Transportation Research Record: Journal of the Transportation Research Board*, 2124:127–136, 2009.

[10] G. Coclite, M. Garavello, and B. Piccoli. Traffic flow on a road network. *SIAM J. Math. Anal.*, 36(6):1862–1886, 2005.

[11] R. M. Colombo. Hyperbolic phase transitions in traffic flow. *SIAM Journal on Applied Mathematics*, 63(2):708–721, 2003.

[12] R. M. Colombo, P. Goatin, and B. Piccoli. Road networks with phase transitions. *Journal of Hyperbolic Differential Equations*, 7(1):85–106, 2010.

[13] R. M. Colombo, F. Marcellini, and M. Rascle. A 2-phase traffic model based on a speed bound. *SIAM Journal on Applied Mathematics*, 70:2652–2666, 2010.

[14] T. Courbon and L. Leclercq. Cross-comparison of macroscopic fundamental diagram estimation methods. *Procedia-Social and Behavioral Sciences*, 20:417–426, 2011.

[15] C. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B*, 28:269–287, 1994.

[16] C. Daganzo. Urban gridlock: macroscopic modeling and mitigation approaches. *Transportation Research Part B*, 41(1):49–62, 2007.

[17] C. F. Daganzo. Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, 29(4):277–286, 1995.

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

[19] J. Du, H. Rakha, and V. V. Gayah. Deriving macroscopic fundamental diagrams from probe data: Issues and proposed solutions. *Transportation Research Part C: Emerging Technologies*, 66:136–149, 2016.

[20] L. C. Edie. Car-following and steady-state theory for noncongested traffic. *Operations Research*, 9(1):66–76, 1961.

[21] S. Fan, M. Herty, and B. Seibold. Comaprative model accuracy of a data-fitted generalized aw-rascle-zhang model. *Networks & Heterogeneous Media*, 9(2):239–268, 2014.

[22] S. Fan, Y. Sun, B. Piccoli, B. Seibold, and D. B. Work. A Collapsed Generalized Aw-Rascle-Zhang Model and Its Model Accuracy. *ArXiv e-prints*, Feb. 2017.

[23] S. Fan and D. B. Work. A heterogeneous multiclass traffic flow model with creeping. *SIAM Journal on Applied Mathematics*, 75(2):813–835, 2015.

[24] C. Fraley and A. E. Raftery. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, 97(458):611–631, 2002.

[25] M. Garavello, K. Han, and B. Piccoli. *Models for vehicular traffic on networs*, volume 9 of *Series on Applied Mathematics*. American Institute of Mathematical Sciences, 2016.

[26] G. Gentile et al. The general link transmission model for dynamic network loading and a comparison with the due algorithm. *New developments in transport planning: advances in Dynamic Traffic Assignment*, 178:153, 2010.

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

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

[29] P. Goatin. The Aw–Rascle vehicular traffic flow model with phase transitions. *Mathematical and computer modelling*, 44(3):287–303, 2006.

[30] J. W. Godwey. The mechanics of a road network. *Traffic engineering and control*, 11(7):323–327, 1969.

[31] B. D. Greenshields, J. Thompson, H. Dickinson, and R. Swinton. The photographic method of studying traffic behavior. In *Highway Research Board Proceedings*, volume 13, 1934.

[32] A. J. Gross and V. Clark. *Survival distributions: reliability applications in the biomedical sciences*. John Wiley & Sons, 1975.

[33] F. A. Haight. *Mathematical theories of traffic flow*. Elsevier, 1963.

[34] R. Herman and I. Prigogine. A two-fluid approach to town traffic. *Science*, 204:148–151, 1979.

[35] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen. Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies*, 18(4):568–583, 2010.

[36] Y. Ji, W. Daamen, S. Hoogendoorn, S. Hoogendoorn-Lanser, and X. Qian. Investigating the shape of the macroscopic fundamental diagram using simulation data. *Transportation Research Record: Journal of the Transportation Research Board*, 2161:40–48, 2010.

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

[38] B. S. Kerner. Three-phase traffic theory and highway capacity. *Phys. A*, 333(1-4):379–440, 2004.

[39] J.-P. Lebacque. The Godunov scheme and what it means for first order macroscopic traffic flow models. In J. Lesort, editor, *Proceedings of the 13th ISTTT*, pages 647–677, 1996.

[40] J.-P. Lebacque and M. M. Khoshyaran. A variational formulation for higher order macroscopic traffic flow models of the gsom family. *Transportation Research Part B: Methodological*, 57:245 – 265, 2013.

[41] L. Leclercq, N. Chiabaut, and B. Trinquier. Macroscopic fundamental diagrams: A cross-comparison of estimation methods. *Transportation Research Part B: Methodological*, 62:1–12, 2014.

[42] M. J. Lighthill and G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A*, 229:317–346, 1955.

[43] H. S. Mahmassani, J. C. Williams, and R. Herman. Performance of urban traffic networks. In Elsevier, editor, *10th Int. Symp. on Transportation and traffic theory*, 1987.

[44] P. Olszewski, H. Fan, and Y. W. Tan. Area-wide traffic speed-flow model for the singapore cbd. *Transportation Research Part A*, 29(4):273–281, 1995.

[45] H. J. Payne. Models of freeway traffic and control, in mathematical models of public systems. In *Simulation Council Proceedings*, volume 1, 1971.

[46] S. Peeta and A. K. Ziliaskopoulos. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and spatial economics*, 1(3-4):233–265, 2001.

[47] M. P. Raadsen and M. C. Bliemer. Continuous-time general link transmission model with simplified fanning, part ii: Event-based algorithm for networks. *Transportation Research Part B: Methodological*, 2018.

[48] P. I. Richards. Shock waves on the highway. *Operations Research*, 4:42–51, 1956.

[49] M. Saeedmanesh and N. Geroliminis. Clustering of heterogeneuous networks with directional flows based on "Snake" similarities. *Transportation Research Part B: Methodological*, 91:250–269, 2016.

[50] R. J. Smeed. Road capacity of city centers. *Traffic engineering and control*, 8(7), 1966.

[51] R. J. Smeed. Traffic studies and urban congestion. *Journal of tranport economics and policy*, 2(1):33–70, 1968.

[52] J. M. Thomson. Speeds and flows of traffic in central London: 1. Sunday traffic survey. *Traffic engineering and control*, 8(11), 1967.

[53] J. G. Wardrop. Some thoeretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(2):325–362, 1952.

[54] J. G. Wardrop. Journey speed and flow in central urban areas. *Traffic engineering and control*, 9(11), 1968.

[55] G. B. Whitham. *Linear and nonlinear waves*, volume 42. John Wiley & Sons, 2011.

[56] J. C. Williams, H. S. Mahmassani, and R. Herman. Urban traffic network flow models. *Transportation Research Record: Journal of the Transportation Research Board*, 1112:78–88, 1987.

[57] Y. Zahavi. Traffic performance evaluation of road networks by the α-relationship. *Traffic engineering and control*, 14(5), 1972.

[58] H. M. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, 36(3):275–290, 2002.

[59] P. Zhang, S. Wong, and S. Dai. A conserved higher-order anisotropic traffic flow model: description of equilibrium and non-equilibrium flows. *Transportation Research Part B: Methodological*, 43(5):562–574, 2009.