

Empirical Flow-Density and Speed-Spacing Relationships: Evidence of Vehicle Length Dependency

Benjamin Coifman, PhD

Associate Professor

The Ohio State University

Joint appointment with the Department of Civil, Environmental, and Geodetic Engineering, and
the Department of Electrical and Computer Engineering

Hitchcock Hall 470

2070 Neil Ave, Columbus, OH 43210

Phone: (614) 292-4282

E-mail: Coifman.1@OSU.edu

Abstract

Traffic flow theory has come to a point where conventional, fixed time averaged data are limiting our insight into critical behavior both at the macroscopic and microscopic scales. This paper develops a methodology to measure relationships of density and vehicle spacing on freeways. These relationships are central to most traffic flow theories but have historically been difficult to measure empirically. The work leads to macroscopic flow-density and microscopic speed-spacing relationships in the congested regime derived entirely from dual loop detector data and then verified against the NGSIM data set. The methodology eliminates the need to seek out stationary conditions and yields clean relationships that do not depend on prior assumptions of the curve shape before fitting the data. Upon review of the clean empirical relationships a key finding of this work is the fact that many of the critical parameters of the macroscopic flow-density and microscopic speed-spacing relationships depend on vehicle length, e.g., upstream moving waves should travel through long vehicles faster than through short vehicles. Thus, the commonly used assumption of a homogeneous vehicle fleet likely obscures these important phenomena. More broadly, if waves travel faster or slower depending on the length of the vehicles through which the waves pass, then the way traffic is modeled should be updated to explicitly account for inhomogeneous vehicle lengths.

Keywords

Traffic flow theory, fundamental relationship, fundamental diagram, loop detectors, highway traffic, macroscopic models, microscopic models, congested traffic, NGSIM

Highlights

- Empirical flow-density and speed-spacing relationships in the congested regime
- Results are derived from loop detector data and verified against NGSIM
- The work shows flow-density and speed-spacing curves depend on vehicle length
- Thus, waves travel through long vehicles faster than through short vehicles
- Past studies assuming homogeneous vehicle fleet have obscured these phenomena

1. Introduction

Traffic flow theory broadly includes macroscopic hydrodynamic models and microscopic car following modes. The *state of the art* of traffic flow theory leaves much to be desired, there are many common phenomena that have yet to be fully understood. While important advances continue to be made, most recent empirical developments rely on labor intensive data reduction; thus, limiting the scope of many studies while focusing attention on the few microscopic data sets available. Obviously this slow pace of discovery is of concern within the traffic flow theory community, but the impacts are felt beyond the theoretical domain. Traffic flow models are critical to many aspects of surface transportation, e.g., transportation planning, network design, vehicle routing, traffic management, traffic control, and traveler information all depend on models and simulation software that are based upon traffic flow theory. Unfortunately many of the *state of the practice* products require an intensive calibration. As a result, the commercial tools are often left at the default settings, thereby degrading the performance of the various applications that depend on traffic flow theory.

At the heart of the problem is a lack of the *right* data. Since 1980 most empirical traffic flow theory developments have been based on macroscopic data (e.g., 30 sec average speed, flow, and occupancy) collected from conventional vehicle detectors that were deployed for real time traffic control. These data greatly expanded the amount of empirical data compared to the earlier labor-intensive data collection efforts. The coarse macroscopic data aggregation made sense in an era where computing power was expensive; however, in the present day the granularity of conventional macroscopic detector data unnecessarily obscures the underlying processes and has lead to several misconceptions. In short, our understanding of traffic flow has come to a point where these coarsely aggregated data are now limiting our insight into critical behavior both at the macroscopic and microscopic scales.

As reviewed in Section 1.1, much of traffic flow theory depends a fundamental relationship (FR) between flow, density, and space mean speed; either explicitly, e.g., hydrodynamic models such as LWR (Lighthill and Whitham, 1955, and Richards, 1956) or implicitly, e.g., car following models (Chandler et al., 1958; Gazis et al., 1961). While conventional theory has proven to be a good first-order approximation, there are numerous unreasonable assumptions that are often employed (e.g., "near stationary traffic" and "homogeneous vehicles") that limit the applicability of many theories. Most conventional FR curves are fit to a cloud of the coarsely aggregated data and as a result, the fit is poor, requiring prior assumptions of the curve shape before fitting. Typically, many other curve shapes could easily fit the point cloud just as well if the assumptions were changed (see, e.g., Drake et al., 1967 for an early work that remains typical of contemporary, conventional efforts). The scattered cloud of empirical data points has commonly been accepted as unavoidable; yet Coifman (2014a) found evidence that much of the scatter could be attributed to large, underappreciated sampling errors arising from the conventional, fixed time aggregation process.

In response to these challenges, Coifman (2014b) developed a new data-driven approach for aggregating the individual vehicle measurements from dual loop detectors that leads to a very clean FR derived from empirical data. The start of Section 2 summarizes the details of this new method. For now it is sufficient to note that the approach accommodates inhomogeneous vehicles, does not require "near stationary" traffic, and it does not require any preconceived curve to be fit to the data. Coifman (2014b) only contemplated flow-occupancy relationships. As

discussed in the latter portion of Section 2, the current work extends the methodology to macroscopic flow-density and microscopic speed-spacing relationships derived entirely from empirical loop detector data, yielding relationships that have historically been difficult to measure. The key distinction of this advance is how vehicle length is accounted for.

Section 3 then applies the new methodology to three large data sets. The advances provide a new level of accuracy to calibrate macroscopic hydrodynamic models and microscopic car following models using commonly deployed dual loop detectors. These advances in turn should advance traffic flow theory. A key finding of this work is the fact that many of the critical parameters (e.g., the slope of the congested regime) turn out to depend on vehicle length. These relationships are obscured in conventionally aggregated data because vehicles of different lengths are arbitrarily grouped together based on arrival order. Then the paper closes in Section 4 with a discussion and conclusions.

1.1. Motivation- Conventional Traffic Flow Theory

As noted above, much of traffic flow theory depends a fundamental relationship (FR) between flow, q , density, k , and space mean speed, v , either explicitly (e.g., hydrodynamic models) or implicitly (e.g., car following models). All of the empirically generated FR's use data that average conditions over time and/or space to calculate the traffic state, (q, k, v) . Fig. 1A shows a hypothetical example of an ideal triangular flow-density FR curve (denoted $qkFR$), it is uniquely defined by any three of the four parameters: roadway capacity, q_o , free speed, v_f , jam density, k_j , and the slope of the congested regime, w . Given a triangular $qkFR$ like this, following LWR, w corresponds to the velocity at which disturbances travel within a queue. While debate continues about the correct shape of the $qkFR$ (see Coifman and Kim, 2011 for details), most researchers agree that it should either be triangular (e.g., Hall et al, 1986; Newell, 1993) or parabolic (e.g., Greenshields, 1935; Highway Capacity Manual, TRB, 2000). In any event, the current work makes no a priori assumptions about the shape of the FR. Resolving the debate is complicated by the fact that it is difficult to measure k directly. Although k remains a dominant variable in many traffic flow theories, most empirical FR studies use occupancy, occ , as a proxy for k , where: occ is the percentage of the sampling period, T , that vehicles occupy a vehicle detector. As shown in Equation 1, with a homogeneous vehicle fleet occ is proportional to k by the average effective vehicle length, L_{eff} during T ; where a given vehicle's effective length is the sum of its physical length and the size of the detection zone. Of course in reality L_{eff} varies from sample to sample, so in an effort to avoid the dependence on L_{eff} , many researchers prefer to calculate k from measured v and q via the fundamental equation (Equation 2). Unfortunately Equation 2 only masks the dependency on L_{eff} : Figure 1A shows clearly that the $qkFR$ underlying Equation 2 depends on k_j and w , while the results in Section 3 will show that these two parameters both depend on L_{eff} .

$$occ = k * L_{eff} \tag{1}$$

$$q = k * v \tag{2}$$

To date there has been no good way to measure k directly using conventional loop detectors for empirical traffic flow theory development. There are a few exceptions that fit niche applications. For example, Edie's generalized definitions of q , k , and v ensure that Equation 2 always holds, without any requirement of stationarity (Edie, 1965). Cassidy and Coifman (1997) used this method to find k over the short distance spanned by the paired loops in a dual loop detector. There are other techniques to estimate density over a link between two detector stations,

e.g., Coifman (2003). However, these methods do nothing to address the granularity problems mentioned previously. They average across changing traffic states and cannot fully address the impacts of L_{eff} , thus limiting their impact to advance traffic flow theory.

1.1.1. Troubles with the Conventional Theory

There are several problems with the data forming the empirical foundations of traffic flow theory. First, Fig. 1B shows a typical example of conventionally measured q versus occ sampled at $T=30$ s from an actual dual loop detector over a day. The point cloud shows considerable scatter, which is commonly attributed to combining several non-stationary traffic states in a given sample, e.g., Cassidy (1998). It is hard to imagine any single curve that would be representative of all of the observed data points. Yet many traffic flow theories are contingent upon the existence of some well defined FR curve through these data (or the corresponding $qkFR$ curve) and often the dependence includes higher order derivatives of the curve.

Second, unfortunately most empirical FR studies do not contemplate the impacts of inhomogeneous vehicles, or if they do, it is typically only to the depth of passenger car equivalents (PCE). PCE is a scale factor to account for long vehicles. However, as shown in Coifman (2014a), PCE is far too simplistic to address the impacts of different values of L_{eff} and the results presented in Section 3 show conclusively that trucks really do behave differently than cars. Some studies have tried to avoid this issue with one of the best examples being Persaud and Hurdle (1988). They explicitly used a lane with a truck restriction to avoid, "the complication of determining and applying passenger car equivalency conversions." However, one needs to include trucks and other long vehicles in a comprehensive traffic flow theory. More recently a few studies have explicitly contemplated multiple vehicle types, e.g., Punzo and Tripodi (2007) and Yang et al. (2013) developed multiclass models and then empirically calibrated them with two classes of vehicles: cars and trucks. While these multiclass models specifically address inhomogeneous vehicle fleets, they do not encompass all of the issues addressed in the current work, such as relaxing the need for stationary conditions and relaxing the need for prior assumptions about the shape of the FR.

Third, most empirical FR studies use traffic state measurements from conventional detectors that average vehicle measurements over fixed time sampling periods. Coifman (2014a) goes in to detail as to how and why conventional fixed time sampling is a crude methodology. That work used hypothetical microscopic models to revisit the process of generating empirical FR and uncovered several commonly underappreciated factors that result in surprisingly large, non-linear distortions of empirical traffic state measurements.

1.1.2. Car Following Models

While macroscopic traffic flow models seek to capture the aggregate behavior, another body of traffic flow theory approaches the problem from microscopic car following models to capture the individual driver behavior. Most of these models have evolved from the early work by GM, e.g., Chandler et al. (1958) and Gazis et al. (1961). These models typically assume a relationship between speed and spacing (VXp), e.g., Fig. 1C. In this case the VXp curve is uniquely defined by the distance, d , a driver will take at $v=0$; their response time, τ ; and preferred free speed, v_f . Typically there is a direct relationship between VXp and $qkFR$ curves, where the latter uses the average speed across drivers, k is the reciprocal of average spacing, and q is calculated via Equation 2. Like macroscopic traffic flow theory, there are many overly simplistic assumptions in car following models. One of the biggest problems with car following models is the lack of large quantities of in situ data for development and calibration. The

development side has been partially alleviated by the Next Generation Simulation (NGSIM) data sets, consisting of empirical vehicle trajectory data to support the development of better traffic simulation (Kovvali et al., 2007); however, NGSIM has numerous problems- there are only four data sets, consisting of a few thousand vehicles, short duration (under an hour per set), short distance (roughly 1/3 mile per set), and the extracted data are riddled with errors (see, e.g., Duret et al., 2008; Hamdar and Mahmassani, 2008; Montanino and Punzo, 2013; Punzo et al. 2011, Thiemann et al., 2008). Needless to say, there remains a need for more data and in particular a viable means to actually collect calibration data from a corridor under study.

2. Methodology

As noted in Section 1.1.1, current traffic flow theories are encumbered with poor quality empirical data due to conventional averaging over fixed time periods, e.g., as evident by the scatter in Fig. 1B. Many of the problems arise from:

- (1) a small number of vehicles per sample, especially during periods of low flow,
- (2) non-integer number of vehicle headways per sample,
- (3) a large range of vehicle lengths in a given sample due to an inhomogeneous vehicle fleet,
- (4) the need to seek out stationary conditions in congestion, and
- (5) the need for prior assumptions of the FR shape before fitting a curve to the point cloud.

To address these problems, Coifman (2014b) developed the single vehicle passage (svp) method, summarized as follows. In the svp methodology vehicles are no longer taken successively in the order in which they arrived and there is no requirement to seek out stationary traffic conditions; rather, for each and every individual vehicle passage the traffic state is measured over the headway, h (thus, by definition, avoiding the partial headways) and the vehicles are grouped by similar lengths (thus, avoiding the inhomogeneous fleet) and then grouped by speed before aggregation. Flow for each individual svp is calculated via Equation 3. The associated detector on_time is measured and is used in conjunction with h to calculate occ_{svp} via Equation 4 (note that h is measured *rear bumper to rear bumper* to ensure that the gap ahead of a vehicle is associated with its on_time). The corresponding v_{svp} and L_{svp} are then calculated via Equations 5-6 following conventional methods.

$$q_{svp} = \frac{1}{h} \quad (3)$$

$$occ_{svp} = \frac{on_time}{h} * 100\% \quad (4)$$

$$v_{svp} = \frac{detector_spacing}{traversal_time} \quad (5)$$

$$L_{svp} = v_{svp} * on_time \quad (6)$$

The vehicles are sorted into L_{svp} bins that only span 5 or 10 ft. The vehicles in each length bin are then treated separately from the other length bins. Within a given length bin the vehicles are further sorted into v_{svp} bins that only span 1 mph. To be clear, the sampling process measures each and every passing vehicle, regardless of its behavior and without regard for the presence or

absence of stationary traffic conditions (e.g., the bins combine accelerating and decelerating vehicles together). The result is a homogeneous set of vehicles and speeds in each bin. To ensure the largest possible number of similar vehicles per sample, the median q_{svp} and median occ_{svp} are found for each combined length and speed bin. Only bins containing at least 100 observations are retained, thus, ensuring a large sample size in every bin. Since fewer vehicles pass per unit time during congested periods and most vehicles are passenger cars, to achieve sufficient sample size for long vehicles during congestion it is necessary for the current study to combine data over many days before calculating the median q_{svp} and median occ_{svp} for each bin.

Unlike most conventional traffic state measurements that use the mean to find the center of the distribution, the median is far less sensitive to noisy sample data, e.g., outliers from uncommon driver behavior and/or detector errors. The points in Fig. 2A show a typical empirically measured flow-occupancy FR (denoted $qoccFR$) arising from the svp method for seven different length bins (ranging in increasing lengths from the top curve, 18-22 ft vehicles, to the bottom curve, 68-78 ft vehicles). Each data point on each curve represents the median from hundreds to thousands of single vehicle observations, and the observations underlying any one point on a $qoccFR$ curve are independent of all of the other points on the given curve. Since there is no smoothing across the different bins, the resulting curves do not depend on any prior assumptions of shape.

The published work in Coifman (2014b) only used measures that are directly available from dual loop detectors: speed, flow, occupancy, and vehicle length. The svp method sorts vehicles by length at the start, so the final curves arise from homogeneous vehicles.

The present study extends the svp methodology to density and spacing, Xp , as follows. The small range of vehicle lengths in a given $qoccFR$ curve from the svp method gives rise to special properties. In particular, the homogeneous length range in each svp length bin means that $L_{eff} \approx \text{median}(L_{svp})$ for the small range of lengths in a given bin; thus, allowing for direct application of Equations 1-2 to derive k . As noted in Section 1.1 this conversion is something that cannot be done (correctly) with conventional fixed time measurements. By extension this work calculates the average spacing for each bin from the reciprocal of k , yielding a microscopic speed-spacing curve (denoted VXp) similar to Fig. 1C.

It is in the context of VXp that one can understand how the svp method works for congested traffic. Fig. 3 shows hypothetical VXp curves for 3 different vehicles, giving rise to the following three distinct regions: strictly car following (sCF) corresponding to congested traffic, unimpeded free flow (uiFF) conditions, and mixed conditions where some drivers are car following and some are unimpeded.¹ In the sCF region, taking the median across v for a large number of vehicles with the same L will yield the central tendency of the speed and length bin (it does not matter if one starts with occ and calculates spacing, or the reverse). Sorting vehicles by L in this fashion will isolate any impacts of the physical vehicle length to the respective length bin. For a given length bin, taking the median spacing (or k or q or occ) by speed bin only makes sense where the curve is expected to fall in a small range of spacing, i.e., in the sCF region.² On the other hand, it should also be clear that taking the median by speed bin like this will lead to

¹ For simplicity this example VXp corresponds an underlying triangular $qkFR$. For the regions of interest, sCF and to a lesser degree uiFF, the key relationships necessary for this work also apply to the VXp corresponding to an underlying parabolic $qkFR$.

² Given the fact that the $qkFR$ emerges from the underlying driver behavior, the VXp form of the FR must make physical sense for any traffic model. Of course the various forms of the FR only capture the central tendency, omitting within-driver and across-driver variability. When considering the further impacts of driver variability, it is important to note that the scatter in the VXp plane is expected to be asymmetric at a given v , with longer tails to larger spacing since it remains safe to drive with an excessively large spacing while impossible to drive with an excessively small spacing.

poor performance in uiFF and mixed traffic since some or all of the drivers have achieved their desired v_f and choose a large spacing that is generally accepted to be independent of speed. In other words, the svp method derived above should only work well within the sCF region, which is still beneficial since this congested regime has traditionally been the most difficult conditions to study empirically. Further motivation for the svp method can be found in the appendix of this paper. Meanwhile, the next section goes into more details of the svp processing and demonstrates the results in qkFR and VXp.

3. Development and Application

Before getting into the analysis, Section 3.1 discusses the data used in this work. Then Section 3.2 applies the methodology to the data sets. Section 3.3 fits curves to the data and estimates the underlying parameters, in the process, the analysis reveals that the qkFR and VXp curves change shape, depending on the given L_{eff} of the underlying vehicles.

3.1. Data Sets

This research uses three independent data sets, all of which were collected from I-80 in the Berkeley Highway Laboratory, BHL, just north of Oakland, California, USA (Coifman et al. 2000). The primary data set comes from I-80 and consists of 31 days of single vehicle passage data from the eastbound dual loop detector stations 1, 4, 5, 6, and 7.³ These data were collected between Mar 30 and April 29, 2000. After excluding detector errors using the method presented in Coifman (2014b) this work combines all of the svp data over the entire 31 day period from lanes 2-5, i.e., the four general-purpose lanes, across all five stations. Lane 1 on the median was excluded because it has a time-of-day HOV restriction and exhibits behavior markedly different from the other lanes. For brevity, this data set will be referred to as *2000 eb*. The resulting data set includes over 12 Million svp measurements, with the L_{svp} frequency distribution shown in the top of Table 1.

The second data set was previously used in Coifman (2014b).⁴ It includes all of the svp data from the westbound loop detectors at BHL station 6 over 18 days in December, 1997. For brevity, this data set will be referred to as *1997 wb*. The resulting data set includes over 1.2 Million svp measurements, with the L_{svp} frequency distribution shown in Table 1. The third data set comes from the video-based vehicle trajectories in the NGSIM I-80 data set. It is used to verify that the results do not reflect some unanticipated artifact of loop detector operation or the subsequent application of the svp method. This NGSIM data set was collected from the eastbound lanes in the BHL in April 2005 and includes just over 5,600 vehicles. The NGSIM data set will be discussed in greater detail later, and for brevity it will be referred to as *NGSIM eb*.

³ Stations 2 and 3 were excluded because they exhibited vehicle length distributions suggesting that the detector sensitivity was set lower than the other stations while stations 8 and 9 were not operational at this time. Meanwhile, as of publication, the data set is available at <http://www.ece.osu.edu/~coifman/documents/>

⁴ Further details on this data set can be found in Coifman (2014b). Note that the 1997 wb data pre-dates the HOV lane present in the 2000 eb data collection. Finally, as of this writing the data set is available at <http://www.ece.osu.edu/~coifman/documents/>

3.2. Application

Fig. 2A shows the qoccFR from applying the svp method from Coifman (2014b) to the 2000 eb data set.⁵ Since each one of the length bins comes from (roughly) homogeneous vehicle lengths, the current work uses Equation 1 and the corresponding L_{eff} for the given length bin to convert the qoccFR data to the qkFR curves in Fig. 2B. It should be clear that the relationship between curves from different length bins in qkFR has noticeably changed from qoccFR, illustrating the differences between easily measured occ and traditionally difficult to measure k. By extension, this work also takes the speed and the reciprocal of k for each bin, yielding the microscopic speed-spacing curve (denoted VXp) for the sCF region in Fig. 2C. While there is no clear upper boundary of the sCF region, this work assumes that vehicles exhibiting 40 mph speeds on a freeway with a 65 mph speed limit should be within the sCF region. On the lower end, none of the curves go all the way to 0 mph because there are few observations in this speed range. The slower a vehicle moves, the smaller distance it covers and thus, decreases the likelihood that it will pass a detector while moving so slow. While there were some extremely low speed observations, the plots only show the results for combined speed-length bins that had at least 100 observations. The empirical results from the svp method in each length bin of Fig. 2B and 2C approach the form of the ideal curves shown respectively in Fig. 1A and 1C.

The location and dates included in the 2000 eb data set were deliberately chosen because the site exhibited data throughout the full range of congested speeds. One should be careful to note that not all locations will show a complete curve like Fig. 2A, e.g., if downstream capacity is constrained by a downstream bottleneck the qoccFR curve should exhibit a plateau at the bottleneck capacity and trace out the trapezoidal shape as predicted by Hsu and Banks (1993) and found in Coifman (2014b). In fact even Fig. 2C shows slight curvature above 30 mph. We suspect there are many factors that can cause the observed FR curve to deviate from the ideal form, but the pursuit of those factors is the subject of on-going research.

To verify that these results were not unique to the 2000 eb data set, Fig. 4 repeats the analysis on the 1997 wb data set and finds similar results. The VXp curves in Fig. 4C shows that the upper end of the linear region falls around 20 mph in this data set. Throughout this figure the curves for the longer vehicle length bins show more scatter than Fig. 2 because the number of svp observations is an order of magnitude smaller than the 2000 eb data set.

Of course all of the results thus far come from loop detectors. To verify the findings were not somehow dependent on an artifact of the loop detector data, the work repeated the analysis on the video-based NGSIM eb data set. The analysis uses the direct measurements of vehicle speed and spacing reported in the NGSIM data. Since the effective lengths from dual loop detectors include the size of the detection zone, 6 ft was added to the physical length recoded in the NGSIM data to emulate the effective vehicle length measurement from a loop detector. As noted in Section 1.1.2, there are many data reduction errors in the NGSIM data; however, just as the svp method is robust to transients in loop detector data, the method is also robust to the NGSIM data reduction errors.⁶ One of the chief limitations of the NGSIM data is simply the small number of vehicles observed. So unlike the loop data that used each single vehicle passage once and only once, given the small number of vehicles in the NGSIM data, the analysis used all

⁵ As noted earlier, the curves range in increasing lengths from the top curve, 18-22 ft vehicles, to the bottom curve, 68-78 ft vehicles.

⁶ As partial motivation for this assertion, the error mechanisms in the loop detector data and the NGSIM data are very different. If either or both of the error mechanisms impacted the central tendencies, then the loop detector results should differ from the NGSIM results.

vehicles with a leader, at all times and locations throughout the data set, which is reported at 10 Hz; thereby yielding dozens to hundreds of observations of a given individual vehicle. The resulting L_{svp} frequency distribution after including multiple observations of a given vehicle is shown in Table 1. Fig. 5 repeats the analysis on the NGSIM eb data set and once more, finds similar results. Since the raw NGSIM eb data consist of speed and spacing, in this case the analysis started with VXp and then projected back to the qkFR and qoccFR. Because the NGSIM data were collected over 45 min in almost exclusively congested conditions, there are few data points with speeds above 25 mph; thus, unlike Fig. 2, the curves do not exhibit a complete curve over the entire congested regime and instead, plateau above 30 mph. More scatter is evident compared to the two loop detector data sets because the number of vehicles is smaller by three to four orders of magnitude. Similar to the 1997 eb data set, Fig. 5C shows that the upper end of the linear region falls around 20 mph in this data set. Unlike loop detectors, the NGSIM measures vehicles throughout space, thus, the sampling process collects many more observations at low speeds. So the curves for all of the length bins extend down to 1 mph, but one should keep in mind that this increase in the number of low speed data points comes from many repeated observations of a given vehicle over a short span of time and distance in space.

3.3. Analysis

To quantify the relationships between length bins and across data sets, the least squares method was used to linearly fit spacing to speed in the VXp curves over the linear regions of the sCF for all three of the data sets. As noted above, the upper end of the linear region varied from 20 to 30 mph, depending on the set and this upper limit was also used for the curve fitting.⁷ The lower end of the speed range was set to 5 mph for the loop detector data because a vehicle can come to a stop over a dual loop detector for an unknown amount of time and still yield a speed measurement up to 5 mph (Wu and Coifman, 2014), while the NGSIM range extended down to 1 mph because the instantaneous speeds were measured directly from the video-based vehicle trajectories.

The left-hand column of plots in Fig. 6 shows the fitted curves while Table 1 shows the resulting d and τ (as defined in Fig. 1C) and r^2 from the VXp least squares linear curve fit for the three data sets and seven length bins over the speed ranges noted above. All cells corresponding to r^2 below 0.95 are dimmed in Table 1, but included for completeness. Likewise, the fitted curves corresponding to r^2 below 0.95 are shown with a lighter color in Fig. 6. Recall that the NGSIM data set has far fewer vehicles than the loop detector data sets; hence the r^2 also tends to be lower, especially for the long vehicles which are less common in general. All three data sets show a similar trend, with both d and τ increasing with L_{eff} .⁸ It makes sense that d increases with L_{eff} , since much of this increase is due to the physical length of the vehicle. Likewise, in retrospect it seems reasonable that τ increases with L_{eff} since the longer vehicles tend to also be heavier, and require greater reaction times. However, this finding of common dependency on L_{eff} is contrary to conventional wisdom that d and τ are independent, e.g., as expressed in Newell (2002).

The plots on the right-hand side of Fig. 6 show the VXp fitted curves projected into the corresponding qkFR curves. The bottom of Table 1 shows the calculated k_j and w (as defined in Fig. 1A and Equations 7-8) for qkFR corresponding to the given d and τ . Like the microscopic

⁷ Although not shown, changing the upper speed threshold by ± 5 mph for any one of the data sets did not change the resulting fit much.

⁸ Recall that $L_{eff} \approx \text{median}(L_{svp})$ for the small range of lengths in a given svp length bin.

statistics, k_j and w both depend on L_{eff} . The trends for k_j are not surprising: the longer L_{eff} , the fewer identical vehicles can fit per unit distance at jam density. However, the fact that wave speed depends on length should prove to be an important finding. Interpreting this result, since there are fewer vehicles at jam density over the same distance there are fewer gaps between longer vehicles. With fewer gaps per unit distance the longer vehicles are less compressible than the shorter vehicles, and thus, signals and waves will pass upstream through longer vehicles at higher velocities than through shorter vehicles.

$$k_j = \frac{1}{d} \tag{7}$$

$$w = \frac{-d}{\tau} \tag{8}$$

Fig. 6 shows that the relationships are very stable within the given site (recall that all of the points are independent of one another across the length and speed bins), but vary slightly from one data set to the next. Revisiting the underlying data, these differences appear to be due at least in part to the detector settings. Across all days: 16 of the 20 dual loop detectors used in 2000 eb had a median L_{svp} of 20.0 ft while the remaining four dual loop detectors had median L_{svp} between 20.8 and 21.2 ft; whereas three of the 4 dual loop detectors used in 1997 wb had median L_{svp} between 19.1 and 19.6 ft, with only one lane having L_{svp} of 20.0.⁹ As per Lee and Coifman (2012), at urban detector stations like these, daily median L_{svp} depends on the detector sensitivity setting and it should be very stable over years. Moreover, if the detectors all exhibited the exact same sensitivity, they should also exhibit the same median L_{svp} . So while the general relationships appear to be transferrable, e.g., the fact that w and k_j depend on L_{eff} , it is important to establish the detector calibration before generalizing the specific measured qkFR or VXp curves from one location to another.

4. Discussion and Conclusions

Traffic flow theory is an important cornerstone to a broad range of topics, but it also has to overcome the inherent difficulties of studying traffic in detail given: the number of vehicles interacting, long distances traveled, and the fact that conditions can change dramatically over short distances or time periods. Conventional methods to collect and process empirical traffic data are insufficient and inadvertently obscure critical details. These methods typically average measurements over fixed time periods in a crude methodology that was originally designed to smooth out variability across vehicles in an era when computing power was expensive. In the process the conventional methods often rely on unrealistic assumptions of homogeneous vehicles and stationary traffic conditions. As a result there is much we still do not know about how people drive or the emergent flows that arise.

Density and vehicle spacing are central to most traffic flow theories but have historically been difficult to measure empirically. This paper extends the single vehicle passage (svp) method to flow-density (qkFR) and speed-spacing (VXp) relationships; thus, providing a viable method to measure and calibrate previously hard to measure relationships, and doing so in situ using

⁹ Thus, the greater skew to shorter measured lengths in the 1997 wb data compared to the other two data sets in Table 1. In contrast, all eastbound lanes at stations 2 and 3 during the time period used for 2000 eb had L_{svp} between 17.0 and 18.5 ft, hence, as discussed in Section 3.1, these two stations were excluded.

commonly deployed dual loop detectors.¹⁰ The loop based results were verified against NGSIM, demonstrating that the loop based method yields relationships similar to those from vehicle trajectories within the congested regime (and specifically in the sCF to region). While the NGSIM analysis follows the svp processing by taking many observations, sorting them into length and speed bins, and then finds the median spacing in each bin, because the NGSIM analysis uses every available observation of every vehicle whenever that vehicle has a leader, this sampling approach is also analogous to how one would directly measure a VXp relationship, e.g., Xuan and Coifman (2012). Thus reaching the same findings independent of the loop detector data or the necessary assumptions to project from qoccFR to VXp. In the process the NGSIM analysis provides independent validation of the loop detector based analysis.

The svp method eliminates the need for (near) stationary traffic states to study the congested regime while also eliminating the need to fit a preconceived curve to the data. As a result, it is now possible to focus on understanding the factors that determine driver behavior and traffic flow. The analyses revealed several important findings that transcend the svp methodology, namely that the car following parameters d and τ depend on L_{eff} , e.g., Table 1 shows that longer vehicles exhibit larger response times, τ . The fact that both d and τ depend on L_{eff} is contrary to conventional wisdom that d and τ are independent, e.g., as expressed in Newell (2002); thus, even when the true L_{eff} is known for a given sample of inhomogeneous vehicles, one cannot simply apply Equation 1 to derive k from occ. It only becomes possible in this paper because the svp process explicitly yields samples with (nearly) homogeneous vehicle lengths.

Equations 7-8 show that both k_j and w depend on L_{eff} by way of d and τ . Thus, the shape of qkFR also depends on L_{eff} . Interpreting these results, the longer L_{eff} , the fewer identical vehicles can fit per unit distance at jam density. Interpreting the fact that wave speed depends on vehicle length: since there are fewer vehicles at jam density over the same distance there are fewer gaps between longer vehicles. With fewer gaps per unit distance the long vehicles are less compressible than shorter vehicles, and thus, signals and waves will pass upstream through long vehicles at higher velocities than through shorter vehicles.

Even if one accounts for the average vehicle length when calculating k from occ via Equation 1 or from q and v via Equation 2, mixing inhomogeneous vehicles in a sample should lead to errors in the resulting k due to the dependence on the individual vehicle lengths. Fig. 6 shows that the relationships are very stable within the given site (recall that all of the points are independent across length and speed bins), but vary slightly from one data set to the next due to the site-specific detector calibration. So while the general relationships appear to be transferrable, e.g., the fact that w and k_j depend on L_{eff} , it is important to establish the detector calibration before generalizing the specific measured qkFR or VXp curves from one location to another.

Ultimately, the most important finding of this work is the fact that the velocity at which a given wave travels upstream through congested traffic depends on the vehicle lengths through which the wave passes. If waves travel faster or slower depending on the length of the vehicles through which the waves pass, then the way traffic is modeled should be updated to explicitly account for inhomogeneous vehicle lengths.

¹⁰ While few dual loop detector stations currently report the necessary single vehicle passage data, these data exist in the controller prior to aggregation into fixed time samples. It is only a small modification for an operating agency to collect these svp data and a small minority of operating agencies already collect the svp data from their detectors. For more traditional operating agencies there are commercially available tools to collect the svp data in parallel with a traffic controller conventionally aggregating the data.

4.1. Further Points of Note

While this paper has shown that the qkFR and VXp curves depend on L_{eff} , it is important to note that even after accounting for detector calibration the curves likely depend on many other factors as well, e.g., vehicle weight is likely to be an important explanatory variable determining the driver and vehicle response. The generally low r^2 for the 38-48 ft vehicles seen across the data sets in Table 1 likely reflects the fact that this length range includes passenger vehicles pulling trailers, buses, and single unit trucks. Furthermore, for a given type of truck the vehicle performance likely changes depending on whether it is loaded or empty.¹¹ The current work also excludes non-vehicle sources of variation in traffic flow data, such as the influence of time of day and the driver population (as distinct from the vehicle population). To fully assess these other factors, data from many more sites are required. Regardless, if other explanatory variables are measured or otherwise become available, the svp method could be extended to account for them via additional dimensions in the sampling bins.

Although this work focused strictly on freeway traffic, the fact that the length dependencies are found on the freeway suggests that there is a good chance that a similar relationship exists in the interrupted flow on arterials, but investigating such a relationship on arterials is a topic left to future research.

Acknowledgements

The author would like to thank the anonymous reviewers for their constructive feedback that helped make this a better paper.

References

- Cassidy, M. (1998) "Bivariate Relations in Nearly Stationary Highway Traffic," *Transportation Research Part B*, Vol. 32, No. 1, p. 49-59.
- Cassidy, M. and Coifman, B., (1997) "Relation Among Average Speed, Flow, and Density and Analogous Relation Between Density and Occupancy," *Transportation Research Record 1591*, pp. 1-6.
- Chandler, R.E., Herman, R., Montroll, E.,W., (1958) "Traffic dynamics: studies in car following," *Operation Research*, Vol. 6, No. 2, pp. 165-184.
- Coifman, B. (2002) "Estimating Travel Times and Vehicle Trajectories on Freeways Using Dual Loop Detectors", *Transportation Research: Part A*, vol 36, no 4, pp. 351-364.
- Coifman, B., (2003) "Estimating Density and Lane Inflow on a Freeway Segment", *Transportation Research: Part A*, vol 37, no 8, pp 689-701.
- Coifman, B., (2014a) "Jam Occupancy and Other Lingerin g Problems with Empirical Fundamental Relationships," *Transportation Research Record 2422*, pp 104-112.
- Coifman, B., (2014b) "Revisiting the Empirical Fundamental Relationship," *Transportation Research- Part B*, Vol 68, pp 173-184.

¹¹ Unfortunately at the moment there are no known data sets that include individual vehicle weights during congestion, in part because weigh-in-motion (WIM) performance degrades significantly during congestion. Thus, operating agencies typically deploy WIM stations away from locations with recurring congestion.

- Coifman, B., Kim, S., (2011) "Extended Bottlenecks, the Fundamental Relationship and Capacity Drop," *Transportation Research- Part A*. Vol 45, No 9, pp 980-991.
- Coifman, B., Lyddy, D., Skabardonis, A. (2000) "The Berkeley Highway Laboratory- Building on the I-880 Field Experiment," *Proc. IEEE ITS Council Annual Meeting*, IEEE, pp. 5-10.
- Coifman, B., Wang, Y. (2005) "Average Velocity of Waves Propagating Through Congested Freeway Traffic," *Proc. of The 16th International Symposium on Transportation and Traffic Theory*, July 19-21, 2005, College Park, MD. pp 165-179.
- Daganzo, C. (1997) *Fundamentals Of Transportation And Traffic Operations*, Elsevier, 339 p.
- Drake, J., Schofer, J., May, A., (1967) "A Statistical Analysis of Speed Density Hypotheses," *Highway Research Record*, No. 154, pp 53-87.
- Duret, A., Buisson, C., Chiabaut, N. (2008) "Estimating Individual Speed-Spacing Relationship and Assessing Ability of Newell's Car-Following Model to Reproduce Trajectories," *Transportation Research Record 2088*, p. 188-197.
- Edie, L., (1965) "Discussion of Traffic Stream Measurements and Definitions," *Proc., 2nd International Symposium on the Theory of Traffic Flow*, OECD, Paris, pp. 139-154.
- Gazis, D. C., Herman, R., Rothery, R. W., (1961) "Nonlinear follow the leader models of traffic flow," *Operations Research*, Vol. 9, No. 4, pp. 545-567.
- Greenshields, B., (1935) "A Study of Traffic Capacity," *Proc. of the Highway Research Board*, Vol 14, pp. 448-477.
- Hall, F., Allen B., Gunter M. (1986) "Empirical Analysis of Freeway Flow-Density Relationships," *Transportation Research Part A: Policy and Practice*, Vol 20, No 3, pp. 197-210.
- Hamdar, S., Mahmassani, H. (2008) "Driver Car-Following Behavior: From Discrete Event Process to Continuous Set of Episodes," *Proc. of the 87th Annual Meeting of the Transportation Research Board*.
- Hsu, P., Banks, J., (1993) "Effects of Location on Congested-Regime Flow-Concentration Relationships for Freeways," *Transportation Research Record 1398*, pp 17-23.
- Kovvali, V., Alexiadis, V., Zhang, L. (2007). "Video-Based Vehicle Trajectory Data Collection," *Proc. of the 86th Annual TRB Meeting*, TRB.
- Lee, H., Coifman, B., (2012) "Quantifying Loop Detector Sensitivity and Correcting Detection Problems on Freeways," *Journal of Transportation Engineering, ASCE*, Vol. 138, No. 7, pp 871-881.
- Lighthill, M., Whitham, G., (1955) "On Kinematic Waves II. a Theory Of Traffic Flow on Long Crowded Roads," *Proc. Royal Society of London, Part A*, Vol. 229, No. 1178, pp 317-345.
- Montanino, M., Punzo, V., (2013) "Making NGSIM data usable for studies on traffic flow theory: a Multistep Method for Vehicle Trajectory Reconstruction," *Transportation Research Record: Journal of the Transportation Research Board*, Issue 2390, pp 99-111

- Newell, G., (1993) "A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks," *Transportation Research- Part B*, Vol 27, No 4, pp 289-303.
- Newell, G. (2002) "A simplified car-following theory: a lower order model," *Transportation Research Part B*, Vol. 36, 2002, pp. 195-205.
- Persaud, B., Hurdle V. (1988) "Some New Data That Challenge Some Old Ideas About Speed-Flow Relationships," *Transportation Research Record 1194*, p. 191-198.
- Punzo, V., Borzacchiello, M., Ciuffo, B. (2011) "On the Assessment of Vehicle Trajectory Data Accuracy and Application to the Next Generation SIMulation (NGSIM) Program Data," *Transportation Research Part C*, Vol 19, No 6, pp 1243-1262.
- Punzo, V., Tripodi, A. (2007) "Steady-State Solutions and Multiclass Calibration of Gipps Microscopic Traffic Flow Model," *Transportation Research Record 1999*, pp 104-114
- Richards, P., (1956) "Shock Waves on the Highway," *Operations Research*, Vol. 4, No. 1, pp 42-51.
- Thiemann, C., Treiber, M., Kesting, A. (2008) "Estimating acceleration and Lane-Changing Dynamics from Next Generation Simulation Trajectory Data," *Transportation Research Record 2088*, pp 90-101.
- TRB (2000) Chapter 7, Traffic Flow Parameters, *Highway Capacity Manual*, Transportation Research Board.
- Wu, L., Coifman, B. (2014) "Vehicle Length Measurement and Classification in Congested Freeway Traffic," *Transportation Research Record 2443*, pp 1-11.
- Xuan, Y., Coifman, B. (2012) "Identifying Lane Change Maneuvers with Probe Vehicle Data and an Observed Asymmetry in Driver Accommodation," *Journal of Transportation Engineering, ASCE*, Vol 138, No 8, pp 1051-1061.
- Yang, D., Jin, J., Ran, B., Pu, Y., Yang, F. (2013) "Modeling and Analysis of Car-Truck Heterogeneous Traffic Flow Based on Intelligent Driver Car-Following Model," *Proc. of the Transportation Research Board 92nd Annual Meeting*, 18p.

Appendix

Motivation for the svp Method

Stationary traffic requires q , k and v to be constant from one sample to the next. Yet it is very rare to find q or v constant over time during congested periods on a freeway. As a result, many researchers believe that LWR and other hydrodynamic models do not require stationary traffic conditions, e.g., as noted in Daganzo (1997), "The LWR model arises from the assumption that the stationary relationships... also apply when traffic is not stationary." In this context, if the relationships apply when traffic is non-stationary, then it should also be possible to measure these relationships when traffic is non-stationary, e.g., as Coifman and Wang (2005) did to measure wave speeds between detector stations or in the case of the current work, seeking the central tendency of the svp data. As empirical evidence to support this view that LWR does not require stationary conditions, Coifman (2002) began with the premise that the constantly varying, non-stationary evolution of congested traffic could be approximated with a piecewise

linear evolution (similar to how a Fresnel lens approximates a conventional lens in optics). Each vehicle passage at a dual loop detector station captures the average state, represented by the constant measured speed during the measured headway. This average state is assumed to propagate upstream at a constant velocity. Then applying LWR to individual vehicle passage data at real detector stations that work integrated across the resulting bands to accurately estimate vehicle trajectories over space, strictly using local conditions observed over time at the given loop detector station. There was no assumption of equilibrium or stationary traffic yet the results held over a wide range of congested traffic conditions, even when speeds changed dramatically from one vehicle to the next, i.e., during non-stationary conditions.

List of Tables

Table 1, Distribution of vehicle lengths in each data set; VXp least squares linear curve fit to svp based loop detector data or spacing based NGSIM data and the associated coefficient of determination; and finally, the calculated qkFR statistics derived from the VXp fit. All entries associated with r^2 less than 0.95 are shown in gray.

List of Figures

Figure 1, (A) Hypothetical example of an ideal triangular qkFR curve denoting the key parameters, (B) a plot of actual q versus occ from 30 sec samples at one detector on one day, (C) a hypothetical example of an ideal linear VXp curve denoting the key parameters. Typically the exact shape of the qkFR and VXp have been subject to the judgment of the researcher due to the scatter in real empirical data, as such, the curves in parts A and C are showing one possible shape.

Figure 2, (A) An example of an empirical qoccFR from the svp method applied to actual dual loop detector data from the 2000 eb data set, (B) the corresponding qkFR after deriving k from occ and L , (C) detail of the corresponding VXp after converting k to average spacing. In each plot there is one curve for each length bin. In plots A and B the vehicle length bins increase from top to bottom and in plot C the vehicle length bins increase from left to right. Since there is no smoothing across the different speed bins, all of the points are independent and the resulting curves do not depend on any a priori assumptions of shape. Since there are fewer long vehicles, those curves show more scatter than the shorter vehicles.

Figure 3, Hypothetical VXp curves for three vehicles, all the same length, with slightly different v_f , d , and τ . Below some speed all vehicles are sCF, and above some spacing all vehicles are uIFF. In between the traffic is mixed. Within the sCF region, taking the median spacing at a given v across a large number of vehicles should yield the center of the spacing distribution.

Figure 4, Repeating the analysis from Fig. 2, except now applied to the 1997 wb data set. The curves corresponding to the long vehicles show more scatter than Fig. 2 because the number of svp observations is an order of magnitude smaller than the 2000 eb data set.

Figure 5, Repeating the analysis from Fig. 2, except now applied to the NGSIM eb data set. Since these raw data are explicitly comprised of speed and spacing, unlike the loop detector data, VXp is found first and then projected to qkFR and qoccFR. These curves do not show a complete curve over all speeds, because unlike the two loop detector data sets, NGSIM eb does not see speeds throughout the entire congested regime. More scatter is evident compared to the two loop detector data sets because the number of vehicles is smaller by three to four orders of magnitude.

Figure 6, The left column of plots show the VXp least squares linear curve fit to the svp based loop detector data or spacing based NGSIM data by length bin for each of the three data sets. The right column of plots projects the given VXp curve fit to the qkFR. The top row is from 2000 eb, middle row from 1997 wb, and bottom row from NGSIM eb. Throughout this figure the raw data are shown with points and fitted curves with lines. The fitted curves are plotted in a light shade whenever the coefficient of determination is less than 0.95. The relationships between length bins remains unchanged from Fig. 2.

Table 1, Distribution of vehicle lengths in each data set; VXp least squares linear curve fit to svp based loop detector data or spacing based NGSIM data and the associated coefficient of determination; and finally, the calculated qkFR statistics derived from the VXp fit. All entries associated with r^2 less than 0.95 are shown in gray.

		18-22 ft	22-28 ft	28-38 ft	38-48 ft	48-58 ft	58-68 ft	68-78 ft
distrib. of veh lengths	2000 eb	82.8%	11.9%	1.32%	0.73%	0.49%	1.65%	1.19%
	1997 wb	88.1%	7.0%	1.03%	0.68%	0.63%	1.78%	0.75%
	NGSIM eb	80.9%	15.6%	0.86%	0.54%	0.27%	1.07%	0.75%
d (ft)	2000 eb	25.8	33.4	45.3	45.1	64.2	74.6	84.1
	1997 wb	24.0	31.2	45.1	52.2	59.7	78.3	93.2
	NGSIM eb	26.9	31.0	45.4	54.6	77.7	87.0	98.4
τ (s)	2000 eb	1.18	1.37	1.77	2.06	1.92	1.89	2.20
	1997 wb	1.49	1.80	1.94	2.15	2.52	2.32	2.37
	NGSIM eb	1.26	1.36	1.58	2.54	1.50	2.11	2.18
r^2	2000 eb	0.999	0.997	0.942	0.914	0.812	0.951	0.974
	1997 wb	1.000	0.996	0.968	0.901	0.975	0.986	0.872
	NGSIM eb	0.998	0.996	0.957	0.914	0.675	0.896	0.928
k_j (veh/mi)	2000 eb	205.0	158.1	116.5	117.0	82.2	70.8	62.8
	1997 wb	219.9	169.1	117.1	101.1	88.4	67.5	56.6
	NGSIM eb	196.5	170.2	116.4	96.7	68.0	60.7	53.6
w (mph)	2000 eb	-14.9	-16.6	-17.4	-15.0	-22.8	-26.9	-26.1
	1997 wb	-11.0	-11.8	-15.8	-16.5	-16.2	-23.0	-26.9
	NGSIM eb	-14.5	-15.6	-19.6	-14.7	-35.3	-28.1	-30.7

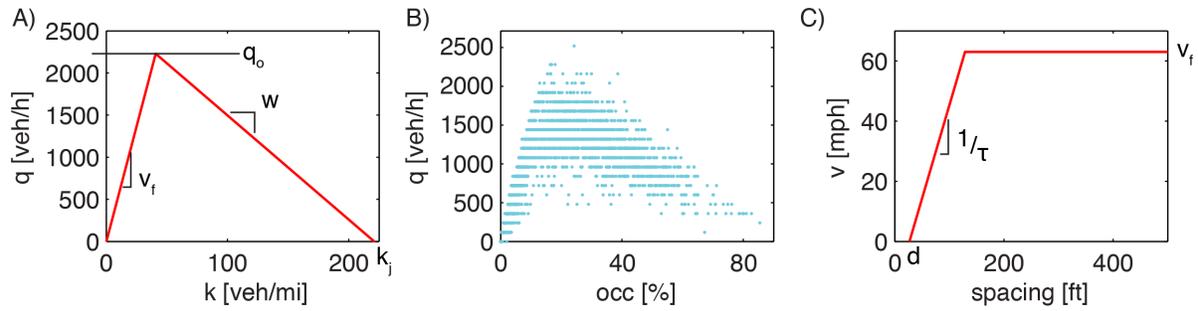


Figure 1, (A) Hypothetical example of an ideal triangular qkFR curve denoting the key parameters, (B) a plot of actual q versus occ from 30 sec samples at one detector on one day, (C) a hypothetical example of an ideal linear VXp curve denoting the key parameters. Typically the exact shape of the qkFR and VXp have been subject to the judgment of the researcher due to the scatter in real empirical data, as such, the curves in parts A and C are showing one possible shape.

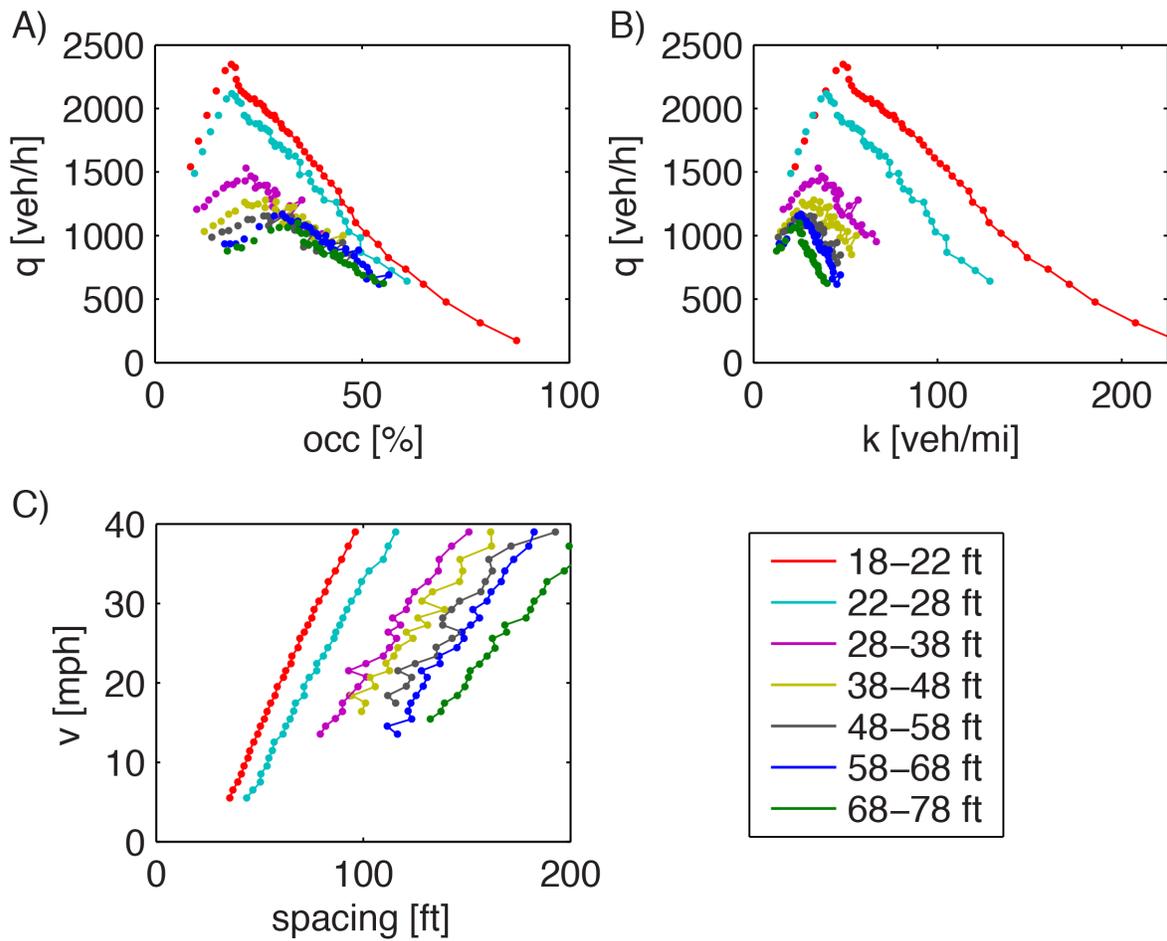


Figure 2, (A) An example of an empirical qoccFR from the svp method applied to actual dual loop detector data from the 2000 eb data set, (B) the corresponding qkFR after deriving k from occ and L , (C) detail of the corresponding VX_p after converting k to average spacing. In each plot there is one curve for each length bin. In plots A and B the vehicle length bins increase from top to bottom and in plot C the vehicle length bins increase from left to right. Since there is no smoothing across the different speed bins, all of the points are independent and the resulting curves do not depend on any a priori assumptions of shape. Since there are fewer long vehicles, those curves show more scatter than the shorter vehicles.

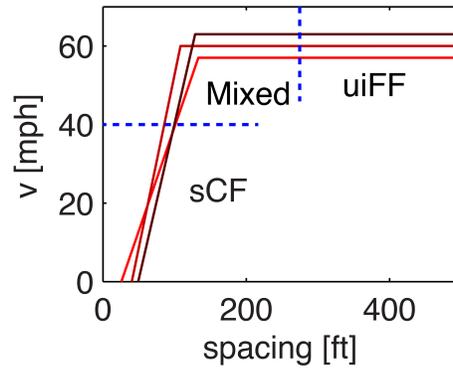


Figure 3, Hypothetical VXp curves for three vehicles, all the same length, with slightly different v_f , d , and τ . Below some speed all vehicles are sCF, and above some spacing all vehicles are uiFF. In between the traffic is mixed. Within the sCF region, taking the median spacing at a given v across a large number of vehicles should yield the center of the spacing distribution.

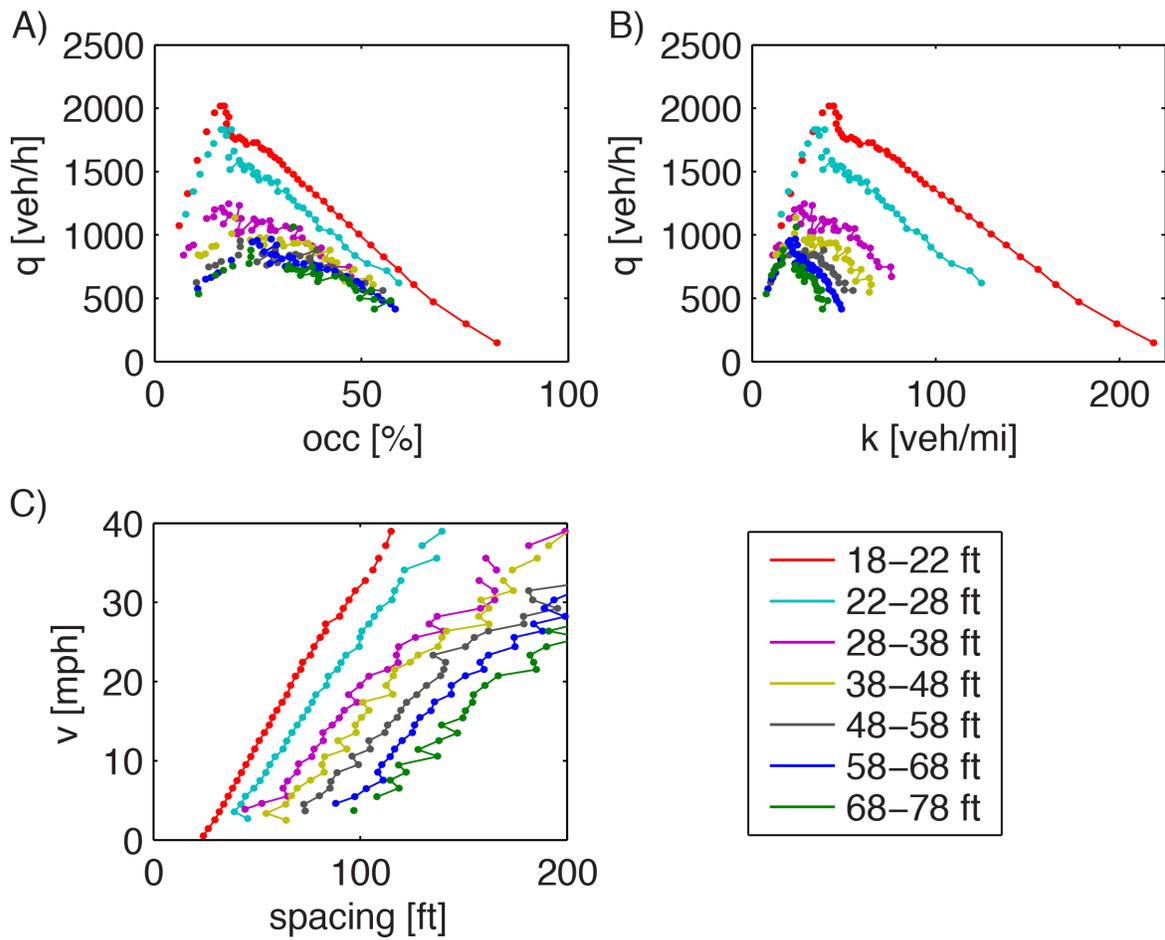


Figure 4, Repeating the analysis from Fig. 2, except now applied to the 1997 wb data set. The curves corresponding to the long vehicles show more scatter than Fig. 2 because the number of svp observations is an order of magnitude smaller than the 2000 eb data set.

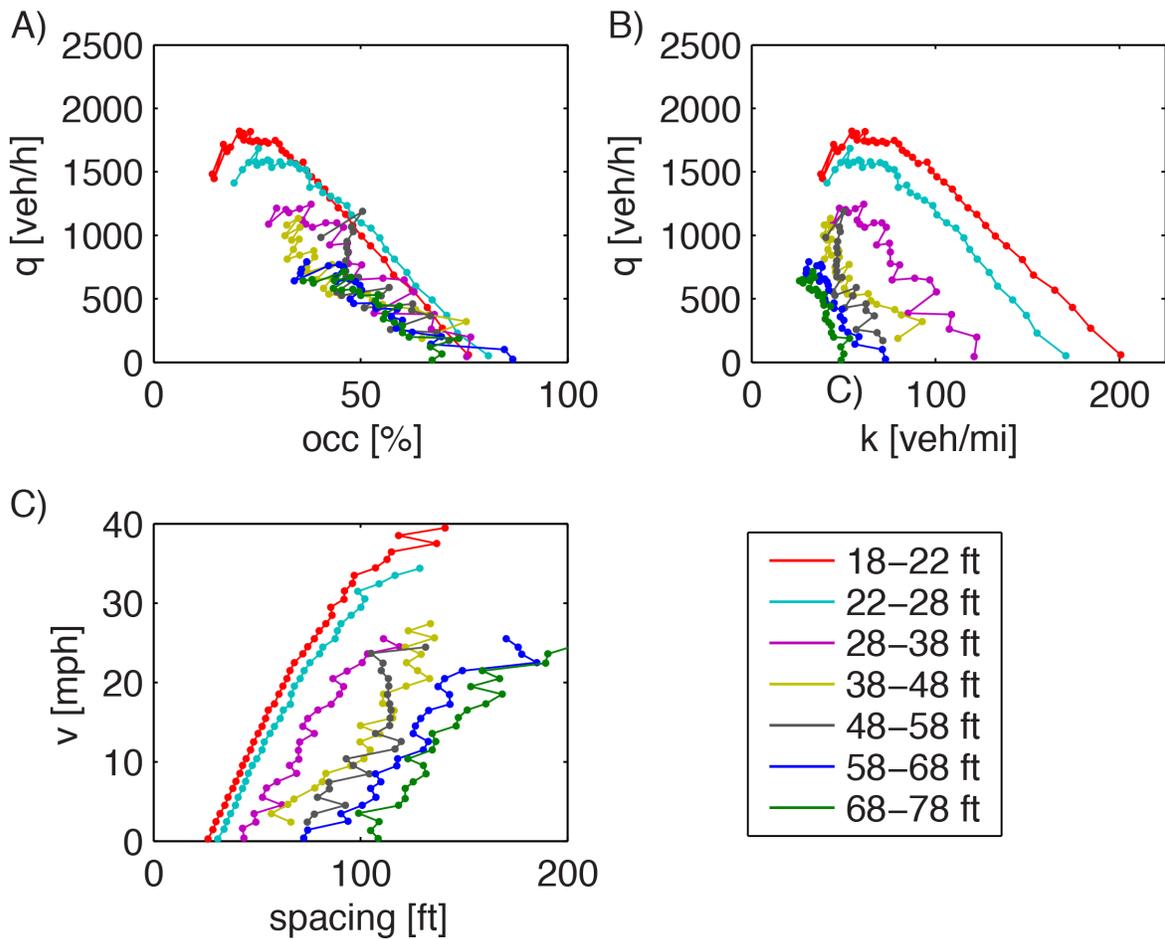


Figure 5, Repeating the analysis from Fig. 2, except now applied to the NGSIM eb data set. Since these raw data are explicitly comprised of speed and spacing, unlike the loop detector data, VX_p is found first and then projected to qk_{FR} and $qocc_{FR}$. These curves do not show a complete curve over all speeds, because unlike the two loop detector data sets, NGSIM eb does not see speeds throughout the entire congested regime. More scatter is evident compared to the two loop detector data sets because the number of vehicles is smaller by three to four orders of magnitude.

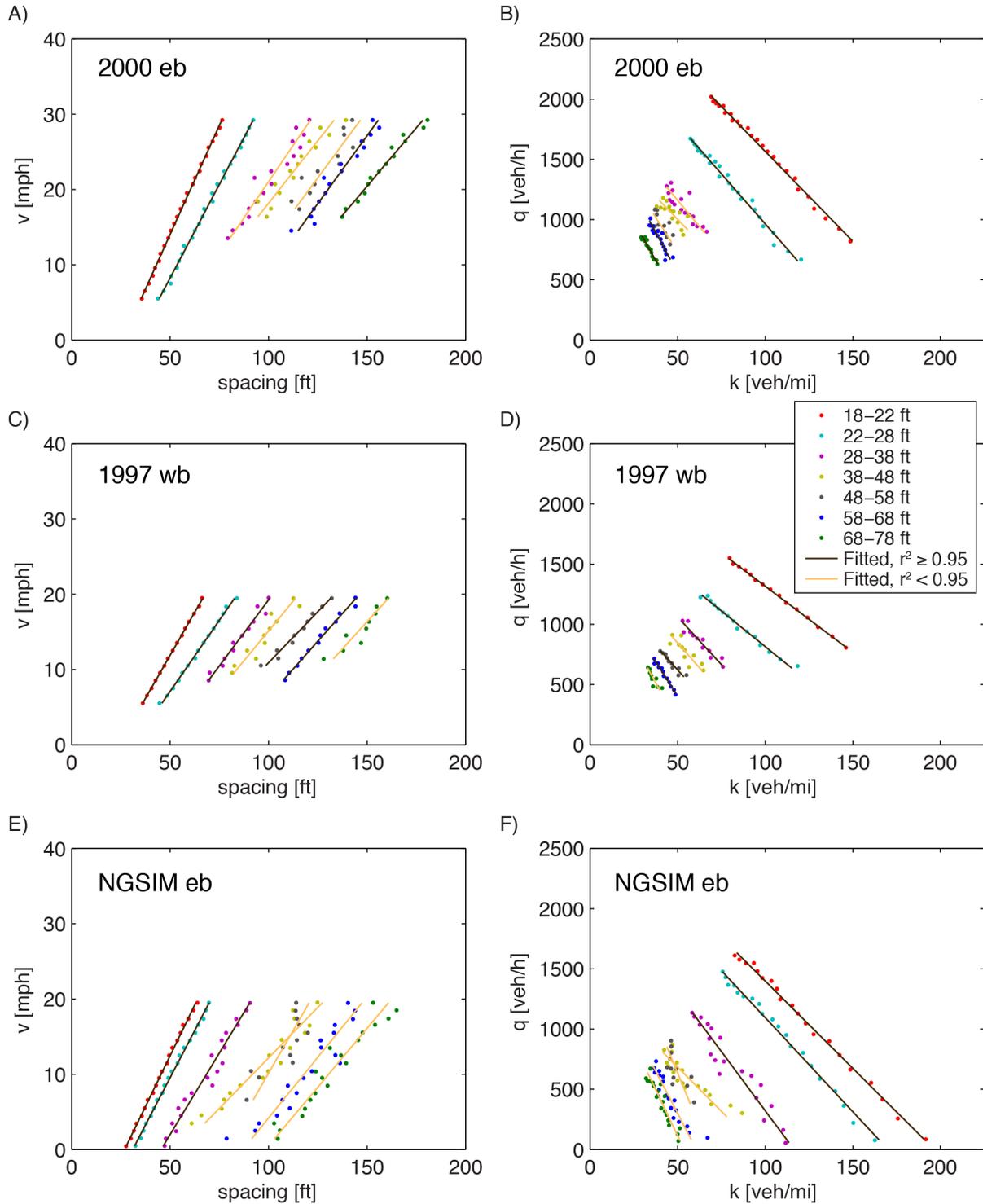


Figure 6, The left column of plots show the VXp least squares linear curve fit to the svp based loop detector data or spacing based NGSIM data by length bin for each of the three data sets. The right column of plots projects the given VXp curve fit to the qkFR. The top row is from 2000 eb, middle row from 1997 wb, and bottom row from NGSIM eb. Throughout this figure the raw data are shown with points and fitted curves with lines. The fitted curves are plotted in a light shade whenever the coefficient of determination is less than 0.95. The relationships between length bins remains unchanged from Fig. 2.