Variability of Traffic Flow Measures Across Freeway Lanes

Mark Carter[1], Hesham Rakha[2] and Michel Van Aerde[2]

Department of Civil Engineering, Ellis Hall
Queen's University, Kingston, Ontario
Canada K7L 3N6

Correspondence Author:
Hesham Rakha
Center for Transportation Research
Virginia Polytechnic Institute and State University
1700 Kraft Drive, Suite 2000 (0536)
Blacksburg, VA 24061
Phone: (540) 231-2261 - Fax: (540) 231-5214
E-mail: rakha@ctr.vt.edu

ABSTRACT

On most freeways, a number of factors interact to produce lane-to-lane variations in speed and volume that are both site- and volume-dependent. The following paper explores and statistically verifies these variations using detector data and a combination of complementary techniques based on data collected for 30 days at 27 detector stations in May of 1994 along the Queen Elizabeth Way freeway near Toronto, Ontario.

The analysis indicates that considerable volume dependencies exist at each site, and that these dependencies are site-specific. In addition to their independent variations, speed and flow are also shown to interact differently across different lanes, and result in different underlying speed-flow relationships. The findings are intended to be relevant to the calibration of microscopic traffic simulation models and the calibration of automatic incident detection algorithms. As such, the paper does not attempt to specifically identify the underlying causes for the variations, but rather attempts to recognize the aggregated effects of these causes in a fashion that would be useful to real-time Freeway Traffic Management System control strategies relying solely on loop detector inputs.

Key words: Traffic flow theory, traffic modeling, traffic simulation, incident detection algorithms.

BACKGROUND

Most macroscopic and microscopic traffic simulation models and highway design approaches have historically analyzed freeways in terms of traffic flow parameters, such as speed, volume, and density, that represent the average across all the available lanes. Alternatively, microscopic simulation models and other operational tools, such as incident detection algorithms, can operate on lane-specific data. The use of lane-by-lane analysis reflects that most traffic flow parameters do not have uniform volumes across all available lanes, and that the deterministic component within such non-uniformity

---

[1] Current Affiliation: Science Applications International Corporation.

[2] Current Affiliation: Center for Transportation Research, Virginia Polytechnic Institute and State University.

may be of considerable traffic engineering consequence. This paper attempts to quantify the difference in traffic behavior across lanes in order to demonstrate the need for such a detailed approach of analysis.

This background discussion provides a review of related topics, namely, reasons for lack of lane uniformity of traffic flow parameters, prior findings related to non-uniformity of traffic flow parameters, and a brief description of the study objectives and approach.

Lack of Uniformity of Traffic Flow Parameters

The primary source of lane-by-lane variability of speed and flow on a basic freeway section with a homogeneous vehicle fleet is the requirement for vehicles to pass to the left and the complementary requirement that slower vehicles should stay to the right. These two rules should interact to typically produce speeds that decrease from median to shoulder, and volumes that increase from median to shoulder.

As this assumption of homogenous traffic flow is often invalid, a secondary source of non-uniformity across lanes is the presence of trucks and other heavy vehicles in the traffic stream. These vehicles tend to congregate in the shoulder lanes due to either their slower average speeds and/or the regulatory requirements in some areas that trucks remain in the shoulder lanes. The existence of trucks should typically result in speeds that decrease from median to shoulder lanes.

The third source of non-uniformity is the presence of on- and/or off-ramps. In the first instance, vehicles are required to be in the shoulder lane near off- or on-ramps in order to exit or enter the freeway. This typically results in an increased flow, and usually also a decreased speed, near interchanges. In the second instance, through-moving vehicles may move away from the shoulder lanes in order to avoid conflicts with vehicles entering or exiting the freeway.

Prior Studies

A number of studies have analyzed typical traffic deviations, including the variation from one location to another based on geometry (Transportation Research Board, 1994), the variation from one direction to the other (May, 1990), and the variation from one day to the next (Rakha and Van Aerde, 1995). One variation that has received relatively less attention, however, is the distribution of traffic among several available lanes at the same location. This oversight may be due to a number of factors, as indicated next.

First, prior to the installation of Freeway Traffic Management Systems (FTMS), the great wealth of data necessary to study this situation was both extremely difficult and very expensive to obtain. Secondly, until the recent development of microscopic traffic simulation models and automatic incident detection algorithms, there has not been a great need for the added level of accuracy afforded by lane-by-lane considerations. As a result, only a limited number of studies have been undertaken to explore this lane-by-lane variation.

Pignataro (1973) examined a number of studies of American three-lane freeways. From these studies, a typical variation in traffic flows was observed to exist, as illustrated in Fig. 1. Specifically, for freeways with three lanes in one direction, and under conditions of low flow, the center lane was observed to carry the bulk of the total flow, or approximately 47 percent. The shoulder lane also carried a significant portion of the total traffic, namely 31 percent, while the median lane was observed to carry the lowest percentage, approximately 22 percent. As the total volume increased, Pignataro observed that the flows in both the shoulder and center lanes dropped, while the percentages of total flow in the median lane increased. When the arrival rate was sufficiently high, around 1670 veh/h/lane, the percentage of flow in both the median and center lanes were observed to be equal, each carrying approximately 37 percent of the total flow. At these higher flows, the shoulder lane was underutilized, carrying only 26 percent of the total traffic flow.

May (1990) supported these observations for U.S. three-lane freeways and expanded the focus to include the observed traffic flow variations for a German autobahn. For the German freeway, the distribution was similar to that of the American structures in that the percentages of total flow in the shoulder and center lanes continually decreased, while the median lane flow increased. However, the German data differed from the American studies in that the flows on the median and center lanes were not equal under high flows. Specifically, under high flows, the median lane was observed to carry significantly higher flows than both the shoulder and center lanes.

With reference to the variation in traffic flow parameters, such as free-speed and capacity, the Highway Capacity Manual (HCM, Transportation Research Board, 1994) indicates that under most situations, it is sufficient to effectively ignore the variation and take the average of these values for each of the available lanes. While this has been the common approach undertaken to date, the HCM recognizes that there is some inherent *error* in this approach (Transportation Research Board, 1994). The manual also recognizes that in situations where multiple lanes are available in the same direction, variations in free-speeds and capacity will exist. These variations are a function of a number of factors, including local traffic regulations (e.g. slower traffic must keep to the right), traffic composition (number of trucks), the presence of on- and off-ramps, and the particular driving habits of local drivers. Due to the wide variation in these factors, the manual does not attempt to predict *typical* lane distributions for capacity and free-speed, although it does indicate that the shoulder lane is often under-utilized and may be expected to exhibit lower capacities than other lanes.

Allen et al. (1985) conducted one of the first detailed studies of the variability of these parameters across lanes. They applied HCM style speed-flow curves over time-connected and event-based traces of traffic conditions by lane on the Queen Elizabeth Way in Ontario. While they considered only speed-flow relationships and raised some questions as to the appropriateness of the curve-fitting technique used to extrapolate from the recorded data, they did demonstrate a qualitative difference in both speeds and flows across lanes. Specifically, they demonstrated that capacity appears to be lower for the shoulder lane as compared to the median and center lanes.

In a later study, Hall and Gunter (1986) also used time-traced plots of the QEW, but expanded the analysis of Allen et al. to examine the variation of flow-occupancy across lanes. As with Allen et al., they observed a qualitatively different relationship in the shoulder lanes as compared to the median and center lanes. Unfortunately, they did not have enough data to conclude whether it was simply that the parameters (in particular, maximum flow rates and critical occupancy) differed in the shoulder lane, or whether the overall shape of the shoulder lane flow-occupancy relationship differed for that lane.

Hurdle, Merlo, and Robertson (1997) further investigated these earlier observations in a study that considered lane-by-lane speed-flow data from Highway 401. In addressing some of the concerns raised by Allen et al., the authors of this study moved away from the HCM curves and applied cubic equations to the speed-flow pairs recorded in their data set. As with the earlier studies, Hurdle, Merlo, and Robertson also demonstrated that speed-flow relationships appear to differ by lanes. Their observations match those of the earlier studies in finding substantially lower capacities for the shoulder lanes as compared to the median and center lanes. They also demonstrated that this general trend was consistent over the length of the freeway, despite changes in magnitude over this length.

While all of these sources qualitatively agree on the variability of traffic flow parameters across lanes, none of them statistically verifies these variations or examines the magnitude of this lane variability relative to observed variabilities along the length of the roadway or over time.

Study Objectives and Approach

This paper seeks to expand on the literature findings. Specifically, it applies a refined curve-fitting technique to a large well-defined data set in order to statistically explore the relative impacts of lane, location, and daily variability on common traffic flow parameters, such as capacity, speed-at-capacity, and jam density.

This study initially conducts a series of Analyses of Variance (ANOVA) on measured loop detector data along a major freeway section west of Metropolitan Toronto, as will be described in the following section. The objective of this initial analysis is to determine if speeds and volumes, when considered independently, demonstrate statistically significant variation across lanes, as is suggested by the literature studies. The study then conducts another series of ANOVA tests in order to study the impact of the interaction of the individual speed and flow variations on the major traffic flow parameters, namely free-speed, speed-at-capacity, capacity, and jam density. This set of tests statistically verifies whether the traffic flow parameters differ depending on the lane of travel, however, it does not verify whether the parameters differ from the traffic flow parameters for the average traffic conditions across all lanes. As such, the final stage of this study involves the use of a *Normal-test* type of analysis in order to verify that the estimated lane-specific parameters are statistically different from the average estimated station traffic flow parameters.

In summary, this study not only statistically verifies the trends that have been reported in other literature, but also expands the analysis to study the impact of the above mentioned traffic distribution on the fundamental traffic flow parameters.

The above explorations of lateral or lane-by-lane variations in speed and volume on freeways are performed with two specific objectives in mind. The first is to demonstrate that the use of lane-specific information is an important factor that should be incorporated within microscopic traffic simulation models. Secondly, the study demonstrates that the use of lane-specific information is important for automatic incident detection algorithms. Some of the current algorithms, such as the McMaster Algorithm, utilize lane-specific templates in order to determine if congestion has occurred and to determine if this congestion is of a recurring or non-recurring nature, while others do not. This study gives some insight into the potential benefits of including lane-specific information.

STUDY DESCRIPTION

It should be noted that the subsequent analyses are confined to sections providing three lanes of travel in each direction. Lane-by-lane variability in speed and volume on facilities providing two and four lanes per direction are also important, but are beyond the scope of the current paper.

Study Area

This study focused on the operation of the eastbound direction of the Queen Elizabeth Way (QEW) freeway as illustrated in Fig. 2. The QEW is located in Central Ontario, Canada, to the west of Metropolitan Toronto, and serves as an east-west link between Hamilton to the west and Toronto to the east. The AM peak occurs in the eastbound direction, with a corresponding PM peak in the westbound direction. The freeway is fairly flat in both directions and contains few elevated sections.

Congestion along this section of freeway is characterized by a recurring bottleneck forming at the Dixie Road off-ramp (station 24) at roughly 6:30 AM each weekday. The shock-wave from this bottleneck extends back as far as station 11, east of Southdown Road, and clears at around 9:50 AM, lasting for roughly two hours.

An FTMS is currently operating on an approximately 18-km section of this highway. At the time of the study, detector coverage was provided by 27 stations from Royal Windsor Road in the west to Highway 427 in the east. These detectorized sections traverse the cities of Oakville, Mississauga, and Etobicoke, and are numbered from west to east.

The QEW FTMS utilizes dual loop detectors at all eastbound stations, with the exception of stations 16 and 22, which are single loop detectors. The dual loop detectors allow the Traffic Management Center (TMC) to measure and log the flow, occupancy, and space-mean-speed for each of the three lanes at 30-second intervals, 24 hours a day. The single loop detectors at stations 16 and 22 do not allow for the direct measurement of speed, but do allow for the recording of flow and occupancy data.

These detector stations are spaced at approximately 700 meters and are located near on- and off-ramps in some locations, but far removed from such ramps in other locations. All stations in the study area were comprised of three lanes in each direction, with the exception of stations 6 and 26, which serve four lanes in each direction.

Study Period

Data were obtained from the loop detector records from these stations for the month of May, 1994. May was selected for two reasons: first, it avoided the effects of winter weather, and second, it avoided the impacts of summer construction.

The data set that was examined covered the eastbound lanes from 05:00 hours to 12:00 hours. This time frame allowed for the examination of both the high flow AM peak, and of conditions of relatively low flow prior to and following the peak. Furthermore, the selection of this time frame ensured that the study focused on stations that had experienced heavy congestion (LOS D and E), as well as those that had remained in free-flow state (LOS A and B). Detector measurements of individual lane speeds, flows, and occupancies were recorded at 30-second intervals, but were aggregated to five-minute data for the purpose of this study. This aggregation allowed for a compromise that limited the amount of random noise and volume of data to be handled, while still capturing most of the variability in the traffic conditions.

LANE-BY-LANE VARIATIONS IN MEASURED SPEED AND FLOW

The first step in this section of the study is to compare the traffic flow measurements along the QEW freeway with those presented in the literature. Furthermore, this study qualitatively examines the relationship between various combinations of speeds and volumes as functions of average station flows and average station occupancies. The second step is to support these findings with statistical verification of the qualitatively observed lane-by-lane variability. The statistical analyses are conducted utilizing a series of ANOVA tests (Crow et al., 1978) on the five-minute data measurements using the SYSTAT computer program (1992).

Data Aggregation

As discussed earlier, detector measurements were recorded every 30 seconds, and were aggregated at 5-minute intervals for purposes of the analysis. In addition, an average five-minute station flow and occupancy was computed for each of the 27 detector stations (average across all lanes at a station). The station flow and occupancy were then sorted into ten bins of size 250 veh/h/lane and 10 percent, respectively. For each five-minute interval, the bin corresponding to the average station flow and occupancy was selected, and the corresponding individual lane flows and speeds were placed in their respective flow and occupancy bins. Finally, the number of observations, the mean, and the standard deviation for each lane were computed for each of the ten occupancy and flow bins. The aggregation approach is best described using the example that is presented in Table 1. In the example, ten five-minute observations are recorded for a station. The average station occupancy is computed across the three lanes (column 2 of Table 1). Using bin sizes of 10 percent, each observation is assigned a bin number (column 3 of Table 1). Finally, the average speed for each lane/bin combination is computed. In the example that is illustrated in Table 1, the number of observations in the shoulder lane in the first bin is 3, and the average speed is 104.67 km/h ([100+105+109]/3).

In summary, three parameters, namely the number of observations, the mean, and the standard deviation, were computed four times for each lane/bin combination. The four computations result from the following: flow estimates based on flow bins (q_q), flow estimates based on occupancy bins (q_o), speed estimates based on flow bins (u_q), and speed estimates based on occupancy bins (u_o).

Qualitative Analysis

The first stage of the analysis of the lane-by-lane variations in measured speeds and flows was to compare findings from the QEW freeway with those presented in previous literature. Limitations in the available reference data did not allow for an extensive quantitative comparison, and as such, this analysis was limited to a qualitative examination of the data.

From the data aggregation process, it was possible to create a series of figures, such as Fig. 1, that directly compare the observed average lane flows versus average station flow for both the QEW freeway and those presented in the literature sources. Figure 1 demonstrates a strong correlation between the QEW data and the literature, and illustrates significant variability in flow per lane across the various lanes, with the shoulder lane being largely under-utilized even at high flows approaching capacity.

Recognizing that flow does not increase monotonically with speed but that occupancy does, the examination was extended beyond that of the literature by considering the average flow per lane as a function of the average station occupancy (Fig. 3). Again, it was noted that considerable variability existed across lanes, with the shoulder lane continuing to be under-utilized for uncongested conditions (occupancy less than 25 percent). However, unlike Fig. 1, Fig. 3 does demonstrate that all lanes were similarly utilized under congested conditions, with the median lane being utilized the most.

The equal utilization of lanes for congested conditions (defined by high occupancy), together with the unequal utilization of lanes at high flows, could have resulted from the presence of trucks in the two right-most lanes.

In order to illustrate this hypothesis, consider a two-lane section of freeway 1 km in length. Further assume that the median lane contains 100 passenger cars and the shoulder lane contains 50 trucks, each being the length of two passenger cars. If we further assume that the system is at or near capacity and that the speeds across both lanes will be the same, it is reasonable to assume, since the vehicle lengths are roughly equal, that the occupancies will be roughly equal. However, it is evident that considerable deviation will exist in the flows between the two lanes.

Turning the examination to consider the average speeds per lane as a function of the average station flow (Fig. 4), it was noted that while considerable lane-by-lane variability persisted, there were two significant deviations from the comparisons of flow versus flow and flow versus occupancy. First, for this comparison, the order of variability was noted to be consistent. That is, the median lane demonstrates consistently higher average speeds than the center lane, which in turn displays consistently higher values than the shoulder lane. Second, in contrast to the average flows per lane, the average speeds per lane tend to converge, minimizing their variability as capacity is approached. In fact, at capacity, the variability is essentially reduced to zero. One thing to note from the figure is that the average lane speed tended to decrease to an average station flow of approximately 1875 veh/h/lane, and then increase again at an average flow of 2125 veh/h/lane. The increase in speed probably results from the fact that there are two speeds for a specific flow rate (uncongested and congested speed). Near capacity, it appears that the number of speed observations below the nose of the speed/flow relationship tends to reduce the average lane speed, while at capacity the speed is higher (nose of relationship). This interpretation is further demonstrated in Fig. 5. .

Fig. 5. Figure 5 illustrates a more pronounced reduction in average lane speed as a function of the average station occupancy. Fig. 5. also demonstrates a reduction in lane-by-lane variability as capacity is approached. As in the previous analysis, it is noted that this variability essentially reduces to zero at conditions representing capacity. However, unlike the case of Fig. 4, Fig. 5. illustrates a monotonic reduction in speed as the occupancy increases. The monotonic reduction in lane speed in the case of occupancy (Fig. 5.) versus flow (Fig. 4) could be caused by the fact that the speed/density relationship (occupancy can be converted to density) provides a unique speed for each density measurement, which is not the case for the speed/flow relationship.

Whatever the cause, it is evident from a qualitative standpoint that significant variations in speed and flow do exist across the various lanes of the QEW freeway. It should be noted, however, that while this lane-by-lane variability is present for all sections of the freeway, the particular trends mentioned above tend to break down in areas within the influence of access and/or egress points. This result is not unexpected, as these areas experience lane-specific demands, such as preparing to exit the freeway, that are not present in areas characterized in the literature and in Fig. 1.

Statistical Analysis

The next stage of the analysis procedure involved verifying statistically the significance of the lane-by-lane variability identified previously. This verification was carried out using a series of non-factorial analyses of variance (ANOVA). While the authors acknowledge that not all of the assumptions of analysis of variance apply directly to the data (for example, the requirement that the errors be independent with a common variance), it is felt that the technique is nonetheless the most appropriate for this particular analysis. Furthermore, it has also been shown that 5-minute measurements of flow and speed do not statistically differ from the normal distribution (Rakha and Van Aerde, 1995), and thus the normality requirement of the data could be satisfied.

Recognizing these concerns for each of the four categories of interest, namely q_q, q_o, u_q, and u_o, a series of models were developed. These models were constructed using combinations of four parameters. Two classical parameters, day-of-the-week and station, were utilized in order to capture the variability in flow and speed associated with temporal and spatial factors. Furthermore, a bin parameter was used in order to capture the variability of lane flow and speed as a function of the total arrival volume. Finally, the lane parameter was examined in order to investigate the hypothesis that speeds and flows do vary significantly from one lane to another.

For each of the four categories mentioned previously, seven separate non-factorial ANOVA were undertaken. Initially, a single-replication ANOVA was performed on the bin variable alone. Following this, two-factor ANOVAs were performed on the combinations of the bin and station, bin and day-of-the-week, and bin and lane. The next stage involved the ANOVA of two three-factored models. In the first of these, the effects of lane on variability were examined, with spatial effects removed by analyzing a model combining bin, station, and lane. In the second model, spatial variations were left intact, and the contribution of the lane variable to the overall variability was examined with the aggregated temporal variations being accounted for in the inclusion of the day-of-the week parameter. Finally, an ANOVA was performed on a model that included all four variables. This analysis allowed the statistical significance of all four variables to be examined and provided insight into the relative effects of each variable on the overall observed variability.

In terms of statistical significance, it was found, as summarized in Table 2, that for all 28 ANOVA tests performed, the proposed variables were demonstrated to be significant.

In order to identify the relative effects of each variable upon the observed variation, an examination of the percentage of the total mean sum of errors was undertaken as is summarized in Table 3. As was expected, the bin variable accounted for a substantial portion of the overall variability, since it is the largest factor in the q_q, q_o, and u_o analyses. More interesting was the large effect of the lane factor, especially as compared to the classical parameters of station and day-of-the-week. In fact, the lane factor accounts for the greatest variability of all variables other than bin for three of the analyses, and is the largest factor for the fourth, u_q.

Thus, from both a qualitative and quantitative perspective, it is evident that significant lane-by-lane variability exists in both speeds and flows as measured against average flow and occupancy. What remains to be examined is whether the existence of these substantial variations translates into a variation in the traffic parameters of free-speed, speed-at-capacity, and capacity.

LANE-BY-LANE VARIATIONS IN TRAFFIC FLOW PARAMETERS

The previous section demonstrated that considerable volume dependencies exist at each site, and that these dependencies are site-specific. The objective of this section is to investigate whether speed and flow interact differently across the different lanes, and thus result in different underlying speed-flow templates.

Estimation of Traffic Flow Parameters

Speed-flow relationships are critical in any attempt to quantify the nature of traffic flow for a particular set of conditions. There are four parameters that are traditionally used to characterize the nature of a specific speed-flow relationship: free-speed, speed-at-capacity, capacity, and jam-density. Three of the four parameters, free-speed, speed-at-capacity, and capacity, are analyzed in this study. The jam-density was not analyzed because efficient estimates of the jam-density require a significant number of observations in the highly congested regime, which was not the case for many of the stations analyzed on the QEW, as will be discussed later.

The selection of a particular shape for a speed-flow relationship has been a topic of research for more than 50 years. May (1990) provides an excellent discussion and comparison of the various single and multi-regime models and describes their respective strengths and limitations in the context of producing *reasonable* free-speed, speed-at-capacity, capacity and jam-density estimates. In response to these limitations, a new single regime speed-flow relationship was developed (Van Aerde, 1995). While Van Aerde (1995) describes the logic behind the proposed relationship, Van Aerde and Rakha (1995) describe the calibration of the proposed multivariate procedure (the INTCAL model). Van Aerde and Rakha (1995) demonstrate the flexibility of the proposed model in terms of representing different types of roads using different levels of data aggregation (1-minute vs. 5-minute). Furthermore, Van Aerde and Rakha demonstrate the fit of the model to data from the literature (May, 1990), and compare the traffic flow parameters to other single- and multi-regime fits. Van Aerde and Rakha also validated the proposed model using field data from both European and North American facilities.

In order to introduce the reader to the calibration procedure that was developed by Van Aerde and Rakha (1995), a brief description is provided. The technique considers speed (*s*), flow (*v*) and density (*d*) to each be dependent variables, and that relative error in speed, density or volume of equal magnitude should be considered to be of equal importance. The technique involves the optimization of a quadratic objective function that is subject to two sets of non-linear constraints, as demonstrated in Equation 1. This optimization, which requires the use of an incremental optimization search, was demonstrated using several sample data sets.

$$\text{Min}E = \sum_i \left\{ \left(\frac{s_i - \hat{s}_i}{\tilde{s}}\right)^2 + \left(\frac{v_i - \hat{v}_i}{\tilde{v}}\right)^2 + \left(\frac{d_i - \hat{d}_i}{\tilde{d}}\right)^2 \right\} \qquad [1]$$

$$\text{S.T.}\hat{d}_i = \frac{1}{c_1 + \dfrac{c_2}{s_f - \hat{s}_i} + c_3 \hat{s}_i} \forall i$$

$$\hat{v}_i = \hat{d}_i \times \hat{s}_i \forall i$$

$$\hat{v}_i, \hat{d}_i, \hat{s}_i \geq 0 \forall i$$

Where:
d = density (veh/km) or the inverse of the vehicle headway (km/veh)
s = speed (km/h)
s_f = free speed (km/h)
c_1 = fixed distance headway constant (km)
c_2 = first variable headway constant (km^2/h)
c_3 = second variable distance headway constant (h^{-1})
$\hat{x}$ = estimated parameter
$\tilde{x}$ = average parameter

Focusing on the QEW and using both speed and flow detector data, Fig. 6 and Fig. 7 were produced to demonstrate an example fit to the 5-minute data along the freeway. The fits were performed for all lanes at station 12 utilizing the respective lane detector measurements; in addition, an average station fit was generated utilizing average station measurements. Fig. 6 demonstrates how the fit replicates the measured data parameters in the speed-flow domain. Furthermore, Fig. 7 demonstrates that in addition to capturing the data trend in the speed-flow domain, the fit also captures the trend in the speed-density domain. One can observe in Fig. 7 the S-shaped speed-density relationship that allows for a reasonable free-speed estimate. It is also of interest to note from this plot the difficulties inherent in attempting to estimate jam-densities for data sets with few, if any, points in the highly congested regime. It is for this reason that the analysis of jam-density was left out of this study. It is evident from Fig. 6 that a significant variation exists between the speed-flow parameters for each of the three lanes. Specifically, the free-speed, speed-at-capacity, and capacity are highest along the median lane, lower in the center lane, and lowest along the shoulder lane.

In the study presented in this paper, the INTCAL model was utilized to estimate four traffic parameters (free-speed, speed-at-capacity, capacity, and jam-density) for each lane at each station during the analysis period. These parameters, except for the jam-density, were then analyzed for any statistical differences as discussed in the following section.

This process of fitting the speed-flow parameters was repeated for every lane and every station for each of the 30 days in the study period. As later sections will verify, significant variations were noted by both station and lane, while day-to-day variations appeared quite minor, possibly because the analysis only included weekday data due to a lack of congestion over the weekend periods.

The two principle parameters free-speed and capacity were averaged over the 30-day period and examined by both location and lane as indicated in Fig. 8. With reference to the station variation, it was noted that there was no consistent trend from one location to the other, however, the extent of the overall variation appeared to be quite significant. When examining the parameters by lane, the variability is even more pronounced and, in fact, a consistent trend is also observed. For all three parameters considered (free-speed, speed-at-capacity, and capacity), the highest values were noted in the median lane, with lower values in the center lane, and the lowest values in the shoulder lane. Furthermore, the differences in magnitude between the center lane and the median lane, and the shoulder lane and the median lane were observed to be of a similar magnitude for most stations. As a result, the values for the center lane showed excellent correlation with the average values taken across all lanes.

Statistical Analysis

Having identified qualitatively the extensive variation in the fundamental speed-flow parameters, it was necessary to examine possible sources for this distribution and their relative contribution. This section of the paper describes some of the statistical tests that were conducted in order to quantify these traffic flow variations.

It is generally accepted by traffic engineers that speed-flow parameters are significantly affected by geometry and thus by the location of the stations. As such, the analysis began by examining variations in the speed-flow parameters on the basis of location. Furthermore, based on the visual differences in the speed-flow parameters across lanes, the lane factor was also included in the statistical analysis. Finally, temporal variability was accounted for through the inclusion of a day-of-the-week factor.

Data Set

A data set was developed using the loop detector data from the QEW freeway FTMS. The study was restricted to those days and stations that had sufficient levels of congestion (observations in the congested regime) to allow their speed-flow parameters to be estimated reliably. Furthermore, stations 6 and 26 were removed from the analysis because they each measured four lanes of data as opposed to the three lanes common to the other stations. Finally, stations 16 and 22 were removed because they were single loop detectors and could not provide the speed measures necessary for the calculation of the speed-flow parameters. The final analysis utilized a total of 523 data points comprising all three lanes, each day-of-the-week with the exception of weekends, and over 15 stations. The speed-flow measurements contained in this data set served as input to the INTCAL computer program, and the results were then analyzed as indicated below.

ANOVA Tests

In order to perform this analysis, a series of ANOVA tests were carried out as indicated in Table 3. The objective of these tests was to statistically verify that the basic traffic flow parameters differed across lanes. It must be noted that the tests were not factorial tests and thus did not consider the dual interaction of the parameters under consideration; rather, each parameter was considered separately.

The first step in each of the three analyses was to perform a one-way analysis on each of the three factors that were considered to potentially have a bearing on the variation of the parameter under consideration, namely, location, day-of-the-week, and lane. From this one-way analysis, it was determined whether these factors had any statistical contribution to the variation of the traffic flow parameters when considered in isolation.

In the second stage of the analysis, the combined effects of these three parameters were studied using a three-way non-factorial ANOVA with replication. From this analysis, the factors that contributed in a statistically significant fashion to the observed variations were reported (see Table 4). Furthermore, the sum of the squared errors was utilized to quantify the relative impact of each significant factor, and was compared to the amount of error that could not be attributed to any of the tested factors.

Analysis 1: ANOVA of Free-Speed

The data were first grouped by lane in order to test the significance of the lane factor on free-speed (Analysis 1a). The results of the one-way ANOVA showed that the free-speed was significantly different, at the 95 percent confidence level, across lanes. When the data were examined by location (Analysis 1b), it was discovered that the variation from station to station also caused significant differences in the free-speed. The variation in free-speed for different days-of-the-week (Analysis 1c) was, however, not found to be significant.

Finally, a combined ANOVA of all three factors (station, day-of-the-week, and lane) was examined (Analysis 1d). This analysis was conducted using a non-factorial three-way ANOVA with replication. Once again, the day-of-the-week did not appear to be significant within the established confidence limit because the analysis only included weekdays. Furthermore, both the station and lane factors continued to exhibit statistical significance, with the greatest amount of variance in free-speed (as indicated by the sum of squared errors) being accounted for by the lane factor. In fact, the lane factor accounted for 76.2 percent of the squared error, with the station accounting for 22.8 percent of the squared error. It therefore appeared that for the calculation of free-speeds, it would be important to consider both the lane and the stations. The effects of day-of-the-week could be effectively ignored with only a minor loss in accuracy.

Analysis 2: ANOVA of Speed-at-Capacity

For the purposes of exploring the variation in speed-at-capacity, the impacts of lane, station, and day-of- the-week were each examined separately. As with free-speed, both lane (Analysis 2a) and station (Analysis 2b) were observed to contribute statistically significant effects to the variation in the speed-at-capacity. Also, as in the results from free-speed, the effects of day-of-the-week (Analysis 2c) were once again determined to lack statistical significance.

Again, all three factors were examined using a three-way ANOVA with replication (Analysis 2d). This analysis indicated that under these conditions, the effects of day-of-the-week on the variability of speed-at-capacity had become statistically significant with a probability of significance of 99.9 percent, well within the 95 percent confidence limit. Despite its significance, however, the effect of day-of-the-week was quite small. By examining the sum of squared errors, it was observed that the day-of-the-week contributed to only 5.0 percent of the total variation. The effect of location, indicated by the station factor, contributed to 21.2 percent of the variation. Once again, the effect of the lane factor was the most pronounced, contributing to 48.7 percent of the total variation, or over twice the impact of the station factor and nine times the impact of the day-of-the-week factor. In fact, both the station factor and the day-of-the-week factor contributed less to the total variation than did the unaccounted for error of 30.0 percent. Thus, the merit of determining the speed-at-capacity by lane is quite clear. The case for considering the effects of station location is also quite strong, however, the need for examining the day-of-the-week is much less pronounced.

Analysis 3: ANOVA of Capacity

As was undertaken for free-speed and speed-at-capacity, the first step in examining the sources of variation in capacity was to examine the statistical significance of each of the three parameters (lane, station, and day-of-the-week) separately. Once again, the effects of lane (Analysis 3a) and station (Analysis 3b) factors were found to be statistically significant. As was the case for free-speed, the effects of the day-of-the-week factor on the capacity could not be demonstrated to be significant within the 95 percent confidence limit.

However, after removal of the lane and station factors in a three-way ANOVA (Analysis 3d), the day-of-the-week factor was found to contribute significantly to the observed variations in capacity. While the impact of the day-of-the-week factor was significant, it was not large. In fact, the day-of-the-week factor was determined to account for only 1 percent of the total observed variation in capacity as measured by the sum of squared errors. In keeping with the findings for free-speed and capacity, the impact of the lane factor was also found to be significant and quite large, accounting for 35.5 percent of the total observed variation. The effect of the station factor on the total variability was even greater, though not by much, with location accounting for 39.3 percent of the total observed variation. The unaccounted error in this case amounted to 23.5 percent of the total error, much larger than the effect contributed by day-of-the-week. Thus, in the case of capacity, the results indicate that consideration should be allocated to both the location and lane factors. The effects of day-of-the-week variation can likely be ignored without compromising a great level of accuracy.

Normal Test

The final analysis that was conducted was a comparison of the speed-flow parameters for each lane with the station average parameters. In order to reduce estimation bias, two data sets were created. The first, created in the previous section, included a series of parameter estimations for various lanes, stations, and days based upon lane-by-lane raw speed and flow data. The second data set contained a series of parameter estimations created using the INTCAL curve-fitting procedure, and a collection of station-averaged flows and flow-weighted average speeds.

Having developed both data sets, a normal-distribution test was performed for each station and lane, with the null hypothesis that no significant difference existed between the station-average and individual lane values.

Table 5 indicates that in the majority of cases, the null hypothesis was rejected, with a 95% confidence level, indicating significant deviation between the individual lane and average values. Of particular note is the tendency of the center lane to support the null hypothesis for a substantial number of stations. This suggests that for these stations, the values computed for the center lane may be used as representative of the average station values, a suggestion that combines well with common intuition for a three-lane roadway. Caution must be exercised, however, for this is not always the case. It is noted that this trend does not hold for stations within the influence of on- and off-ramps, nor does it necessarily apply to all stations removed from such influence.

Thus it has been demonstrated that for the majority of cases, the estimated parameter values for individual lanes show significant deviation from those parameter estimations based on station-averaged raw data. Furthermore, while the center lane parameter values may serve as good substitutions for the station-averaged values in a number of cases, this trend is tenuous at best, especially in areas within the influence of on- or off-ramps.

CONCLUSIONS

This study indicated that for the freeway study, significant lane-by-lane variation existed for both the traffic flow demands and for the traffic flow supply characteristics, namely, free-speed, speed-at-capacity, and capacity. The lane factor contributed the largest effect to the variation in the free-speed and speed-at-capacity, and the second largest effect in the capacity parameter. The effect of location maintained its historical importance and accounted for the greatest degree of variability in the capacity parameters, and the second greatest variation in the speed-at-capacity and free-speed factors. The impact of day-of-the-week was found to have only a minor effect, if any, on the variability of the three speed-flow parameters considered.

Thus, this paper suggests that the speed-flow parameters; free-speed, speed-at-capacity, and capacity should be examined with respect to both location and lane. The effects of the day-of-the-week may be effectively ignored without a large loss in accuracy.

It is felt that the variations in flow and flow parameters attributable to the lane factor are significant enough to adversely affect the operation of microscopic models and incident detection algorithms if not considered. It is hoped that their consideration may lead to a better simulation of in-situ traffic conditions and thus aid in the successful operation of Freeway Traffic Management Systems.

REFERENCES

Allen, B.L., Hall, F. L., and G. Margot (1985). "Another Look at Identifying Speed-Flow Relationships on Freeways." In *Transportation Research Record 1005*, TRB, National Research Council, Washington, D.C., pp. 54-64.

Crow, E., Davis, F., and M. Maxfield. *(1978). Statistics.* Coles Publishing Co. Ltd., Toronto, Canada.

Hall, F. L., and G. Margot (1986). "Further Analysis of the Flow-Concentration Relationship." In *Transportation Research Record 1091*, TRB, National Research Council, Washington, D.C., pp. 1-9.

Hurdle, V.F., Merlo, M.I., and D. Robertson (1997). "Study of Speed-Flow Relationships on Individual Freeway Lanes." In *Transportation Research Record 1591*, TRB, National Research Council, Washington, D.C., pp. 7-13.

May, A.D. (1990). *Traffic Flow Fundamentals.* Prentice Hall, Englewood Cliffs, New Jersey.

Pignataro, L.J. (1973). *Traffic Engineering: Theory and Practice.* Prentice-Hall, Englewood Cliffs, New Jersey.

Rakha, H., and M. Van Aerde (1995). "A Statistical Analysis of Day-to-Day Variations in Real-Time Traffic Flow Data." Presented at the 74th TRB Annual Meeting, Washington, D.C., Paper No. 950805.

SYSTAT *for Windows: Statistics, Version 5 Edition.* SYSTAT Inc., Evanston, Illinois, 1992.

Transportation Research Board (1994)., *Highway Capacity Manual.* National Research Council, Washington, D.C.

Van Aerde, M. (1995). "Single Regime Speed-Flow-Density Relationship for Congested and Uncongested Highways." Presented at the 74th TRB Annual Meeting, Washington, D.C., Paper No. 950802.

Van Aerde, M. and H. Rakha (1995). "Multivariate Calibration of Single Regime Speed-Flow-Density Relationships." Vehicle Navigation and Information Systems (VNIS) Conference, Seattle, Washington, August, 1995.

Table 1. Example illustration of aggregation process

5-Minute Observation	Station Occupancy (Percent)	Occupancy Bin	Speed (km/h)		
			Shoulder Lane	Center Lane	Median Lane
1	10	2	101	105	107
2	9	1	99	100	103
3	15	2	95	97	102
4	24	3	90	92	95
5	36	4	75	76	80
6	29	3	80	82	84
7	21	3	92	95	99
8	14	2	96	98	100
9	7	1	104	105	110
10	5	1	106	109	117

Table 2. ANOVA tests for observed data

Analysis *	Parameters Used	Significance
1	bin	yes
2	bin,station	both
3	bin,dow	both
4	bin,lane	both
5	bin,station,lane	all
6	bin,dow,lane	all
7	bin,dow,station,lane	all

* Note: Each analysis was repeated for each of the four bin combinations, q_q, q_o, speed-flow, and speed-occupancy, and the parameters for all 28 analyses were statistically significant

Table 3. Summary results for four-way ANOVA tests of observed data

Analysis	Percent of Total Mean Sum of Errors				
	Bin	Station	Dow	Lane	Error
Flow-Flow	88.2	0.7	< 0.1	11.1	< 0.1
Flow-Occ.	62.1	7.9	8.9	21.0	< 0.1
Speed-Flow	26.9	13.5	6.5	53.1	< 0.1
Speed-Occ.	91.2	2.0	0.6	6.2	< 0.1

Table 4. ANOVA tests of estimated traffic flow parameters

Analysis	Speed-Flow Parameter	ANOVA	Significance
1a	Free-Speed	1-way - Lane	Yes
1b	Free-Speed	1-way - Location	Yes
1c	Free-Speed	1-way - Day-of-week	No
1d	Free-Speed	3-way - All	Lane, Location significant, DOW not
2a	Speed-at-Capacity	1-way - Lane	Yes
2b	Speed-at-Capacity	1-way - Location	Yes
2c	Speed-at-Capacity	1-way - Day-of-week	No
2d	Speed-at-Capacity	3-way - All	All Significant
3a	Capacity	1-way - Lane	Yes
3b	Capacity	1-way - Location	Yes
3c	Capacity	1-way - Day-of-week	No
3d	Capacity	3-way - All	All Significant

Table 5. Results of three-way ANOVA tests of estimated traffic flow parameters

	Percent of Total Mean Sum of Errors				Probability of Null Set		
Analysis	Station	Lane	DOW	Error	Station	Lane	DOW
1. Free-speed	22.8 %	76.2 %	0.4 %	0.6 %	0.000	0.000	0.599
2. Speed-at-capacity	21.2 %	48.7 %	5.0 %	30 %	0.000	0.000	0.001
3. Capacity	39.3 %	35.5 %	1.0 %	23.5 %	0.000	0.000	0.001

Table 6. Results of normal test for estimated traffic flow parameters

	Statistical Difference from Global Mean (95% Confidence Level)									
	Free Speed			Speed at Capacity			Volume at Capacity			
	1	2	3	1	2	3	1	2	3	Influence
Sta. 9	no	yes	yes	no	yes	yes	yes	no	yes	off-ramp
Sta. 11	yes	no	yes	yes	no	yes	yes	no	yes	none
Sta. 12	yes	no	yes	yes	no	yes	yes	yes	yes	none
Sta. 13	yes	no	yes	yes	no	yes	yes	no	yes	none
Sta. 14	yes	no	yes	yes	no	yes	yes	no	yes	none
Sta. 17	yes	yes	no	yes	yes	yes	yes	no	yes	none
Sta. 20	yes	yes	yes	yes	no	yes	yes	no	yes	none
Sta. 23	yes	no	yes	yes	no	yes	yes	no	yes	none
Sta. 24	yes	yes	yes	yes	yes	no	yes	yes	yes	off-ramp
Sta. 25	yes	no	yes	no	yes	yes	no	no	no	on-ramp

*** Note:** lane 1 = median lane, lane 2 = center lane, lane 3 = shoulder lane

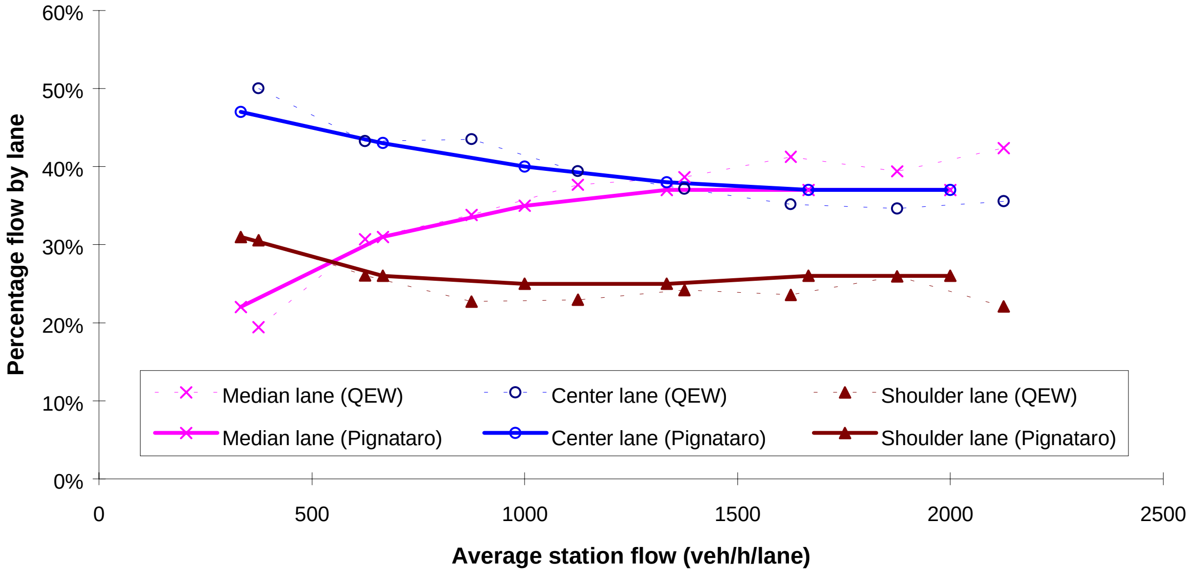

Fig. 1. Lane flow distributions as predicted by Pignataro (1973)

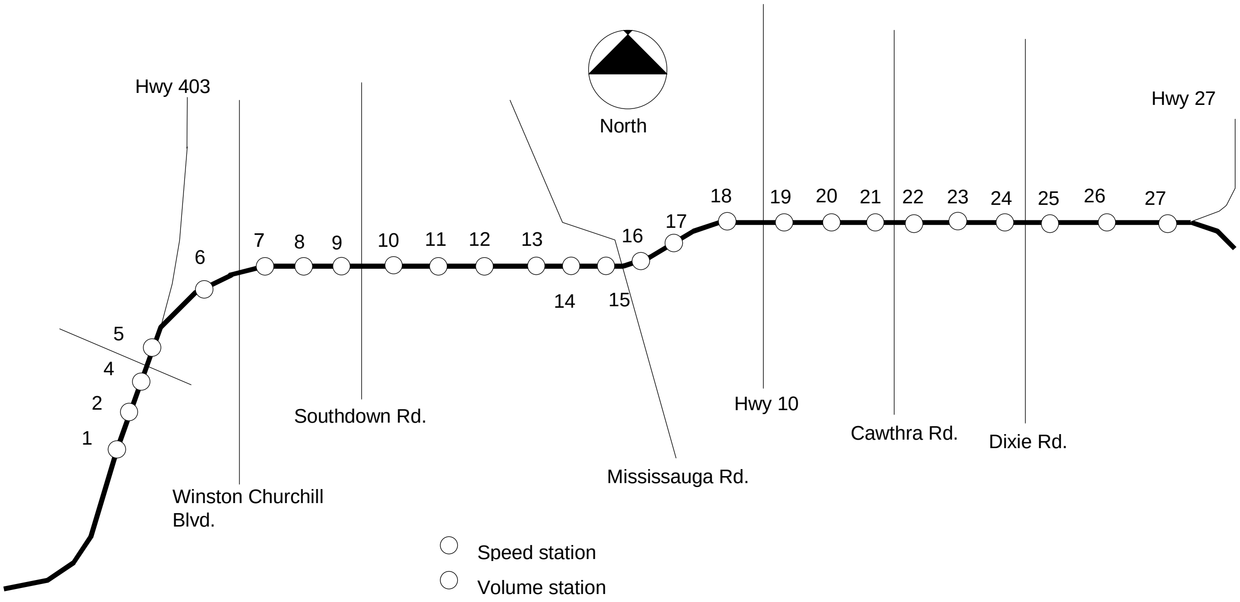

Fig. 2. Configuration of the QEW FTMS

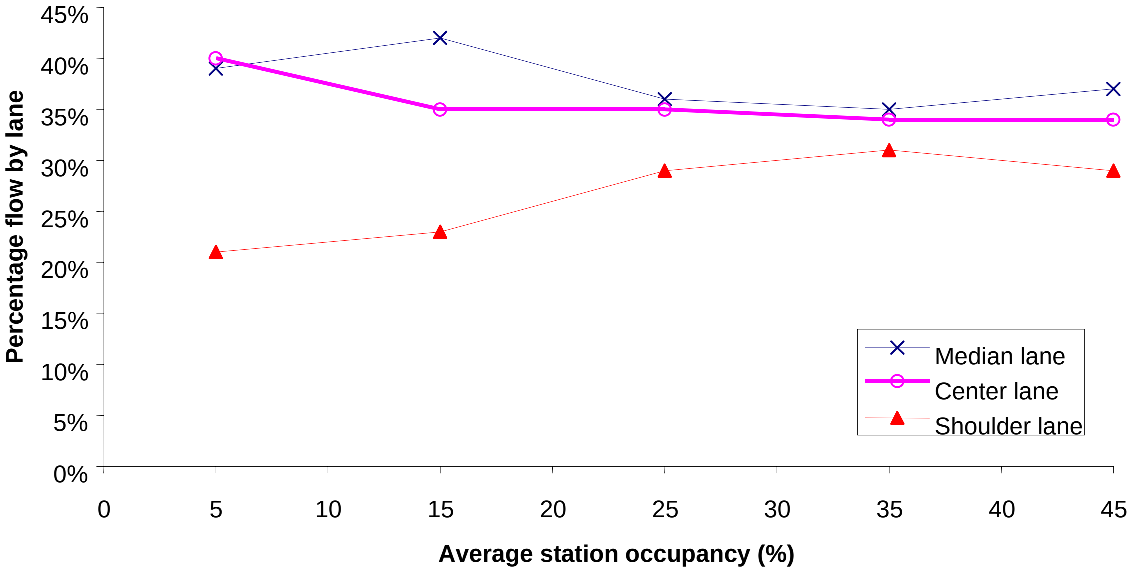

Fig. 3. Percentage flow per lane as a function of average station occupancy

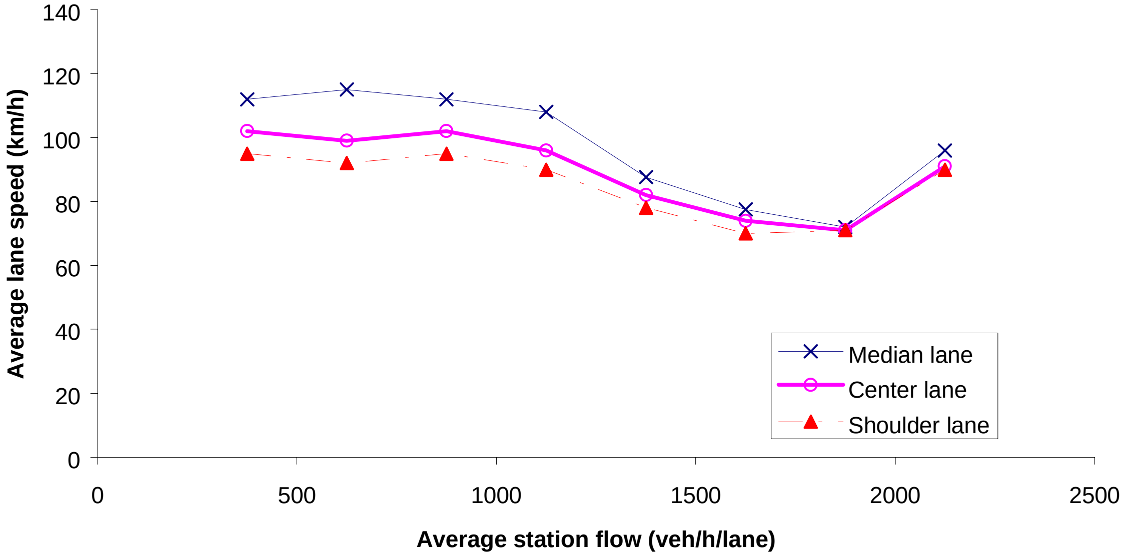

Fig. 4. Average speeds per lane as a function of average station flow

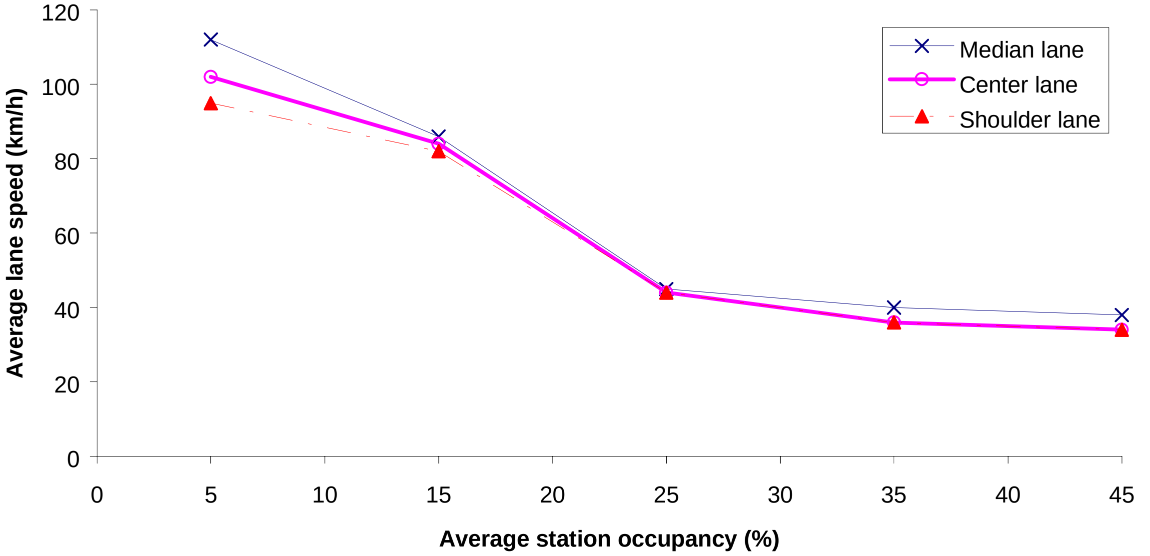

Fig. 5. Average speeds per lane as a function of average station occupancy

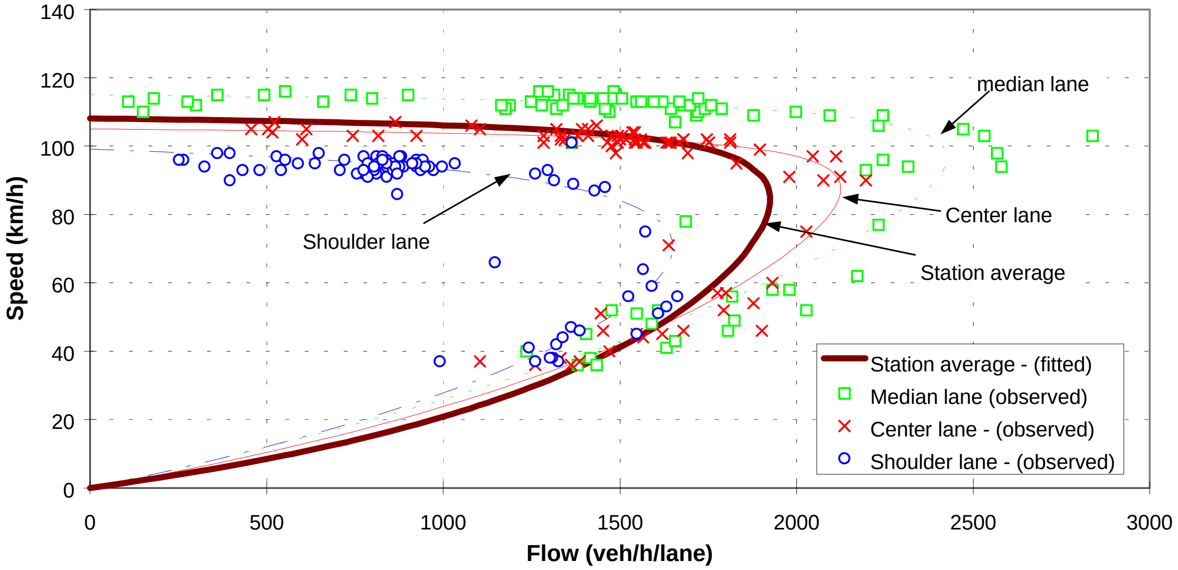

Fig. 6. Speed-flow relationship for all lanes (station 12 - 04/05/94)

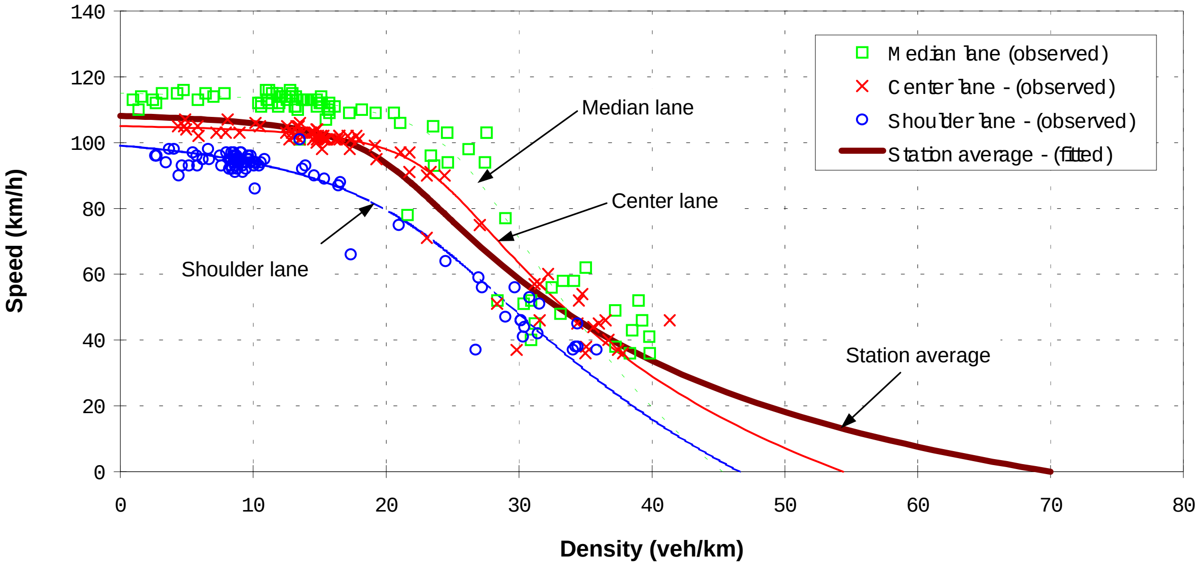

Fig. 7. Speed-density relationship for all lanes (station 12 - 04/05/94)

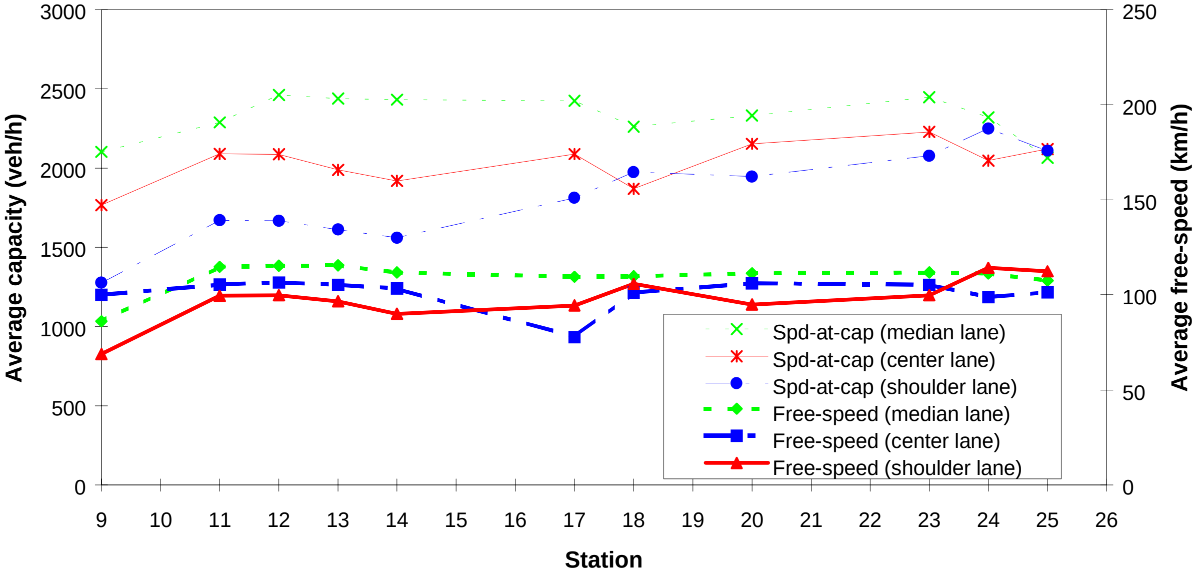

Fig. 8. Variation in estimated free-speed and capacity as a function of location and time