A STATISTICAL ANALYSIS OF DAY-TO-DAY VARIATIONS IN REAL-TIME TRAFFIC FLOW DATA

H. Rakha and M. Van Aerde
Department of Civil Engineering,
Queen's University, Kingston, Ontario
Canada K7L 3N6
PHONE: (613) 545-6000 (ext. 5259)
FAX: (613) 545-2128
rakhah@jeff-lab.queensu.ca

ABSTRACT

In the absence of IVHS technologies, commuters tend to select their routes through a congested network based primarily on the expected average link travel times. In order for this average to be representative of the current day, it is essential that the traffic conditions be relatively similar each day. However, if the traffic conditions vary considerably from one day to the next the historical information will be insufficient for commuters to find the optimum routes through the network, and the provision of real-time traffic information could provide major benefits.

Furthermore, simulation is becoming an important tool in evaluating different traffic control strategies. As a result it has become more and more important that not only the average typical traffic conditions be established but that the upper and lower bounds of these average conditions also be estimated.

Consequently, this paper examines two related issues: the spatial and temporal magnitude of the variability in traffic conditions during typical non-incident conditions and the magnitude of this variability during incident conditions.

The study showed that, in the absence of incidents, the temporal and spatial variations in traffic conditions were very similar for weekdays but varied considerably relative to the typical conditions during weekends. Major incidents, however, were found to drastically alter the average recurring conditions, thus creating a window of opportunity for achieving travel benefits by utilizing dynamic data in real-time.

(Keywords: Simulation, Traffic flow, ATIS, Statistical analysis of traffic)

INTRODUCTION

ATIS Travel Time Estimation

The main objective of most Advanced Traveler Information Systems (ATIS) is to provide drivers with accurate and timely real-time information on traffic conditions. Drivers can select optimum routes to their intended destinations based on this information. Various studies have investigated the potential benefits of ATIS systems (*1*,*2*).

In general, the benefits of such ATIS systems have been shown to depend on the level of market penetration and on the relative accuracy of the real-time information provided to the equipped vehicles when compared to the accuracy of the historical data that is available to the non-equipped vehicles.

Furthermore, as simulation becomes an important evaluation tool, it is important that one calibrates these simulation models to the existing traffic conditions.

Objectives of Paper

Therefore, various questions remain to be addressed. For example: How large must typical day-to-day variations in weekday traffic conditions be before they provide a sufficient window of opportunity for benefits to be accrued through the provision of real-time data to equipped vehicles? By how much do traffic conditions typically vary from day to day? By how much do incidents increase the window of opportunity for achieving benefits through the provision of real-time data?

This paper attempts to address most of these questions through a qualitative and quantitative analysis of 75 days of Freeway Management Center (FMC) data along the Interstate Freeway, I-4 in Orlando, Florida.

The specific objectives of this paper are two-fold:

- To attempt to investigate the variability in traffic conditions during typical non-incident conditions.
- To attempt to investigate the variability of traffic conditions during incident conditions.

It is anticipated that the findings will be of assistance to both IVHS system designers and to those who simulate such systems, as they will be able to perform their analysis based on tangible traffic network statistics rather than hypothetical ones.

Paper Layout

Initially, the background to this study will be presented followed by a description of the network configuration that was utilized in the study conducted in this paper. This description provides an overview of the layout of the I-4 section in Orlando and the spacing of the FMC detector stations. Subsequently, the study analysis period and the number of days for which FMC data were available are presented. In the following section, typical weekday flow, speed and occupancy estimates are presented and compared to the average weekday surface data in order to demonstrate the similar trend in weekday traffic conditions. Subsequently, traffic data for different non-incident days are compared in order to investigate the variability of average traffic conditions between the different days of the week. Incident traffic conditions are also presented in order to demonstrate the potential window for providing real-time traffic information during incidents. In addition, the traffic conditions for the various days are quantitatively compared. Finally, the conclusions of the paper are presented.

BACKGROUND

As part of the Advanced Driver and Vehicle Advisory Navigation Concept (ADVANCE) developed static prediction models which could be applied to a series of traffic flow data: travel time, volume and occupancy (*3*). In their model they studied the effect of link type, time of day, day of week, and season on the flow and occupancy measurements for arterial and freeway links. This work was an extenuation to previous work on travel time analysis on links (*4*).

Shbaklo *et al*. (*3*), using 5-min loop detector data, for 72 days conducted Analysis of Variance (ANOVA) tests on freeway data in Chicago. They found the season to be an insignificant factor, and the day of the week (2.5% to 9.7% of squared error) and time period (50% to 77% of squared error) to both be significant factors on the flow and occupancy measurements. In their analysis, Shbaklo *et al*. (*3*) did not examine whether Fridays or Mondays were statistically different from midweek weekdays (Tuesdays, Wednesdays and Thursdays). Furthermore, they did not study the effect of incidents on these typical traffic conditions.

This paper extends the work conducted by Shbaklo *et al*. (*3*) to investigate variability within weekdays, spatial variability, and to investigate the effect of incidents on typical traffic conditions. Furthermore, this paper investigates the temporal and spatial variability in flow, speed and occupancy measurements about a typical average temporal and spatial surface in an attempt to estimate statistical bounds to identify non typical weekday traffic conditions.

STUDY DESCRIPTION

Network Configuration

A 16 kilometer (10 miles) portion of the I-4 freeway in Orlando, Florida was considered in this study. I-4 serves as a major route that travels across the center of Florida from the south-west (Tampa) to the north-east (Daytona) passing by Disney World. The detectorized portion of the I-4 freeway is located in the vicinity of the downtown of Orlando, extending from 33rd street to the south-west and ending downstream of Maitland Boulevard to the north-east.

There was a total of 24 loop detector stations located along I-4 numbered from 1 to 25, with no data being provided for station 10. The spacing of the detector stations ranged from approximately 0.40 to 0.90 kilometers (0.25-0.54 miles). There were no major terrain variations along the detectorized section of the I-4 freeway, as Orlando is rather flat. However, at many interchanges with arterials the freeway was elevated. The entire detectorized section of I-4 was composed of three lanes in each direction.

Data Collection Time Frame

The analysis period included traffic data for portions of a four month time period during the winter of 1992-93. The data included 11 days in November 1992, 29 days in January 1993, 26 days in February 1993, and 11 days in March 1993. This amounted to a total of 75 days of 30-s data, yielding approximately ten different days of data for each day of the week.

The FMC dual loop detectors measured and logged the flow, occupancy and space mean speed for each of the three lanes at 30-s intervals. These data were aggregated into 5-min data summaries in order to reduce the level of data to be processed, while still capturing most of the trends in the varying traffic conditions. An average lane flow, occupancy and space mean speed estimate was generated from the individual loop detector measurements for each station. In estimating the average lane speed at a specific station, the loop speeds were weighted by the volume on each set of dual loops.

INITIAL ANALYSIS OF FMC DATA

This section of the paper first presents an analysis of the FMC traffic data in order to attempt to assess the variability in traffic conditions within weekdays. Subsequently a comparison of different weekdays and an assessment of the effect of incidents on the average typical traffic conditions is presented later. The analysis presented in this paper defines Saturdays and Sundays to constitute the weekend days.

Generation of Typical Weekday Surfaces

Based on the FMC data available during the above noted four month period, it was possible to generate a surface which represented the average for all the days at a particular station of all the speed, flow and occupancy measurements at a particular time of day. Equations 1 and 2 demonstrate how an estimate of each observation for the flow and occupancy was generated. In the case of the speed surface, a volume weighted average was utilized. Core weekdays were considered to be Tuesday through Thursday, as it was initially not clear if Mondays or Fridays would be consistently similar to Tuesdays, Wednesdays and Thursdays. There was a total of 33 pure core weekdays during the analysis period. These core weekdays were checked for any abnormal traffic conditions such as vehicle detector failures (indicated as -1) or major incidents, as indicated in the incident database that was provided by the FMC. These suspected days were removed from the estimated average.

The above selection process resulted in 22 weekdays being considered in developing the average Eastbound weekday surfaces (nd = 22). The entire 33 weekdays were utilized to generate the average Westbound weekday surfaces (*nd* = 33). The resulting average flow surface for only the Eastbound direction is presented in figures 1, as the results for the Westbound direction were very similar.

$$x_{i,j}^{n} = \sum_{k=1}^{10} x_{i,j,k}^{n} \quad \forall\ x_{i,j,k}^{n} \geq 0 \tag{1}$$

$$\overline{x}_{i,j} = \frac{\sum_{n=1}^{nd} x_{i,j}^{n}}{nday} \quad \forall\ x_{i,j}^{n} \geq 0 \tag{2}$$

Where:

nd = total number of non-incident weekdays.

nday = number of good observations ($x_{i,j}^{n} \geq 0$)

$x_{i,j,k}^{n}$ = 30-s observation on day *n* at station *i*, at 5-min time interval *j*, at 30-s period *k* during the 5-min interval.

$x_{i,j}^{n}$ = 5-min observation on day *n* at station *i* at time interval *j*.

$\overline{x}_{i,j}$ = average weekday 5-min observation at station *i* at time interval *j* (flow or occupancy; speed was generated as a volume weighted average).

The typical average spatial and temporal flow variation in the Eastbound direction for an entire 24 hour period along the detectorized I-4 section is presented in figure 1. The x-axis represents the time of day that ranges from 0, at midnight at the start of the day, to 24 at midnight at the conclusion of the day, while the y-axis represents the station numbers traversed. The eastbound flow proceeds in the upward direction from station 1 to station 25. For each cell combination of time-of-day and station the z-axis represents the average hourly lane flow measured.

It can be noted from figure 1 that the flow gradually increased at 6:00 a.m. along all stations until it reached a flow of approximately 2000 vph/lane at 7:30 a.m. along most of the detector stations. The flow increased again during the p.m. peak at approximately 3:00 p.m. until 6:30 p.m. at stations 12 through 22. It appears from figure 1 that the flow from 5:00 to 7:00 p.m. at stations 7 through 12 was lower (ranging from 1000-1500 vph/lane). However, after examining figure 2 it appears that the speed was also low, ranging from 20 to 40 km/hr. Thus the lower flow measurements were most likely due to the presence of congestion, rather than a reduction in demand. It appears from figures 1 and 2 that a strict analysis of flow contours can be deceiving, as it is not clear if a reduction in flow is caused as a result of congestion or as a result of a simple reduction in demand.

Single-Factor ANOVA of Weekday Data

In order to investigate whether the variability in traffic conditions between the different days of the core of the week (Tuesday, Wednesday and Thursday) was statistically significant, a single-factor ANOVA was conducted using the SYSTAT model (*5*). The ANOVA tested if the Root Mean Square Error (RMSE) associated with the different day surfaces about the typical average weekday surface was greater than the variation within the samples for each specific day of the week using equation 3. Table 1 demonstrates the ANOVA results for flow variations in the Eastbound direction. These results, based on the 22 observations, indicate that the different days were not found to be statistically different at a level of significance of 95%. Similar results were obtained when comparing the speed in the Eastbound direction, as demonstrated in table 1, and the occupancy in the Eastbound direction, as demonstrated in table 1. Consequently, the observations in the Eastbound direction for Tuesdays, Wednesdays and Thursdays were all grouped together as weekdays.

$$RMSE = \sqrt{\frac{\sum_i \sum_j (x_{i,j}^n - \bar{x}_{i,j})^2}{nobs}} \quad \forall x_{i,j}^n, \bar{x}_{i,j} \geq 0 \tag{3}$$

Where:

nobs = number of good observations ($x_{i,j}^n, \bar{x}_{i,j} \geq 0$)

A similar single-factor ANOVA on the different weekdays in the Westbound direction was conducted as presented in table 1. Again, the ANOVA results demonstrated that there was no statistical difference between the observations for Tuesdays, Wednesdays and Thursdays at the 95% confidence level. Consequently, the data for these days were grouped together as core weekdays.

In order to examine the ANOVA assumption of homogeneity of variance, the variation in residuals as a function of the estimated values (day mean) is plotted in figure 4. The studentized residuals were used because it is convenient to reference them against a *t* distribution. In figure 4 the residuals for the typical weekdays (Tuesdays, Wednesdays and Thursdays) were all within two standard deviations. It appears from figure 4 that the residuals are homogeneous as there appears to be no trend to the residuals. Similar trends were found for the residual plots generated for the Eastbound speed and occupancy surfaces. In addition, similar trends were also found for the Westbound flow, speed and occupancy surfaces, however, due to the limited space are not presented.

COMPARISON OF MEAN SURFACES

In this section, a typical average core weekday is compared to a typical Monday, a typical Friday, a typical Saturday and a typical Sunday to determine if the traffic conditions are qualitatively and statistically different. An incident scenario is also compared to the typical average weekday conditions in order to demonstrate qualitatively the relative difference in flow conditions from one day to the next, versus an incident day to a non-incident day.

Average Monday Surface

The average Monday flow, speed and occupancy Eastbound and Westbound surfaces were generated in a similar fashion to the average core weekday surfaces. The Eastbound average Monday surfaces were estimated by averaging over nine Mondays, while the Westbound average surfaces were estimated by averaging over 10 days.

The average Monday flow surface was found to be quite similar to the core weekday surface and thus a typical Monday may qualitatively be considered to be similar to a core weekday. The same trends were found in comparing the occupancy and speed surfaces. However, due to the limited space in this paper these surfaces are not presented.

In order to quantitatively verify the similarity or variability between the Monday traffic conditions and the typical core weekday conditions, a single-factor ANOVA was conducted. The results of the ANOVA analysis for the Eastbound direction, that are presented in table 1, demonstrate that the Monday flow conditions were statistically different from the typical weekday conditions at the 95% confidence level. However, the speed and occupancy measurements were not statistically different from the typical core weekday measurements (at the 95% confidence level) as illustrated in table 1. The same trend of results was obtained in conducting an ANOVA analysis for the Westbound direction as indicated in table 1.

It appears that Mondays are different from core weekdays in terms of flow, but not in terms of speed or occupancy. Mondays were therefore not included in the data sample to create an average core weekday. These results were found to be consistent with the homogeneity assumption of ANOVA as illustrated by the residuals in figure 4.

Average Friday Surface

The Eastbound and Westbound average Friday flow, speed and occupancy surfaces were generated by averaging over 10 Fridays. By comparing the weekday and Friday surfaces it was found that the p.m. peak on Friday started earlier (11:00 a.m. versus 12:00 p.m.) and extended over an extra hour (until 8:00 p.m. versus 7:00 p.m.).

The statistical results were found to verify the above qualitative comparison as illustrated in table 1. Specifically, the ANOVA results for the Eastbound direction indicated that the flows, speeds and occupancies on a typical Friday were statistically different from the traffic conditions of typical core weekdays at the 95% confidence level. The results for the Westbound direction were similar, as illustrated in table 1. These results, again, were found to be consistent with the homogeneity assumption of ANOVA as illustrated by the residuals in figure 4.

Average Saturday Surface

The Eastbound and Westbound average Saturday flow, speed and occupancy surfaces were generated by averaging over 10 Saturdays, however, due to the limited space in this paper the surface plots are not presented. For the average Saturday flow surface, the traffic flows increased gradually from 7:00 a.m. until they reached a maximum flow of approximately 1800 vph/lane at noon at station 15. The flow characteristics for a typical Saturday were very different from the traffic characteristics of a typical core weekday, as might be expected. The ANOVA results for the Eastbound direction, presented in table 1, demonstrate the Saturday traffic conditions were statistically different from the typical weekday conditions. The results for the Westbound direction, that are presented in table 1, also demonstrate this trend.

It is noteworthy that, in terms of Eastbound flow and speed, Saturdays were much more distinct from core weekdays than Fridays. However, in terms of occupancy, Saturdays were different from core weekdays by only as much as did Fridays. In the Westbound direction, flow and occupancy were much different, but speeds were not quite as different.

These results, again, were found to be consistent with the homogeneity assumption of ANOVA as illustrated by the residuals in figure 4.

Average Sunday Surface

An ANOVA of the Eastbound Sunday traffic conditions and the weekday conditions, presented in table 1, demonstrate that traffic conditions on Sundays were also statistically different from typical weekday conditions. Similar results were found for the Westbound direction as illustrated in table 1. As in the case of Saturdays, the results presented in these tables indicate that the flow and speed on a typical Sunday were very different from a typical core weekday for the Eastbound direction. Also the flow and occupancy in the Westbound direction were very different from the core weekday. These results, again, were found to be consistent with the homogeneity assumption of ANOVA as illustrated by the residuals in figure 4. However, there appeared to be an outlier point as illustrated in figure 4.

Incident Effects

During the analysis of the core weekday data, a severe incident, that resulted in the total closure of the Eastbound direction of I-4, occurred on Thursday 5/11/1993, as illustrated by the speed surface plot presented in figure 3. The incident started at approximately 3:20 p.m. and lasted until approximately 5:00 p.m. The incident site was located between detector stations 9 and 11 at Robinson Street, as indicated by the stationary frontal shock wave.

Following the clearance of the incident it can be noted in figure 3 that the traffic proceeded downstream as a continuous platoon, and thus one can observe a surge of low speeds proceeding downstream up to station 21. The forward forming shock wave appears to be steeply sloped because the vehicles proceeded to station 21 within one 5 minute analysis period. The above incident resulted in a queue that extended as far back as station 1. An important point to also note is that it can be observed that a vehicle entering the system at 6:00 p.m. would experience delay at a location downstream of the incident at a point in a time period after the incident was actually cleared.

Summary

In summary, based on statistical comparison of the traffic conditions for various days, the following conclusions can be made:

- Traffic flow conditions within core weekdays appear to be highly similar and consistent.
- Some traffic flow parameters on Mondays are similar to traffic conditions on core weekdays (Tuesday, Wednesday and Thursday).
- Traffic conditions on Fridays differ from core weekday conditions in each of the three measures. Specifically, it appears that the p.m. peak on Fridays extends further in the day.
- Traffic conditions on weekends differ from traffic conditions on weekdays, and Saturdays differ in flow from Sundays.
- Major incidents can cause significant disruptions to the typical weekday traffic conditions.

OVERALL COMPARISON OF TRAFFIC CONDITIONS

The traffic conditions for each day, were compared to the average weekday flow, speed and occupancy surfaces. Two measures of comparison were estimated. The first measure was an estimate of the Coefficient of Determination (R^2) and will be labeled the Regression Measure. The second measure was an estimate of the number of observations within 2 standard deviations of the average weekday observation assuming a normal distribution and will be labeled the Success Measure. The findings for each of these two evaluation measures are discussed in this Section.

The Regression Measure

A Regression Measure similar to R^2 was utilized to compare the traffic conditions for each day. For each day three matrices of flow, speed and occupancy observations were generated. These matrices were 288 rows (number of 5 minute intervals in the day) by 24 columns (number of loop detector stations). A separate overall mean for the average weekday flow, speed and occupancy measurements was also estimated as demonstrated in equation 4 (mean over all stations and all time periods $\overline{x}$).

For each of these surfaces, an estimate of the squared error about the average core weekday surface was estimated as the difference for each station and time-of-day combination from the average core weekday surface using Equation 5 (Sum of squared errors about the average surface S_1). The sum of squared errors for the flow, speed and occupancy measurements of each day about their respective overall mean was also estimated using equation 6 (S_t). The sum of squared error, explained by each of the flow, speed and occupancy average weekday surfaces, S_2 was estimated as the difference between S_t and S_1 based on equation 7. The R^2 measure for each of the three surfaces for each day was calculated as the ratio of S_2 to S_t (S_2/S_t). Thus, R^2 was a measure of the amount of error captured by the average weekday surface. An R^2 of 1 would mean that the average surface explained 100% of the squared error while an R^2 of 0 would mean that the average surface did not explain any of the error.

$$\overline{x} = \frac{\sum_{i=1}^{24}\sum_{j=1}^{288}\overline{x}_{i,j}}{nobs} \quad \forall\ \overline{x}_{i,j} \geq 0 \tag{4}$$

$$S_1 = \sum_{i=1}^{24}\sum_{j=1}^{288}(x_{i,j}^n - \overline{x}_{i,j})^2 \quad \forall\ x_{i,j}^n, \overline{x}_{i,j} \geq 0 \tag{5}$$

$$S_t = \sum_{i=1}^{24}\sum_{j=1}^{288}(x_{i,j}^n - \overline{x})^2 \quad \forall\ x_{i,j}^n \geq 0 \tag{6}$$

$$S_t = S_1 + S_2 \tag{7}$$

Where:

nd, $x_{i,j,k}^n$, $x_{i,j}^n$, and $\overline{x}_{i,j}$ are as defined earlier.

nobs = number of good observations ($\overline{x}_{i,j}$≥0; max.= 6912)

$\overline{x}$ = overall average observation (flow, speed, and occupancy).

S_t = total sum of squared errors about overall mean (flow, speed, and occupancy).

S_1 = sum of squared errors about average surface (flow, speed, and occupancy).

S_2 = sum of squared errors explained by average surface (flow, speed, and occupancy).

The variation of R^2 over the 75 day analysis period from the average core weekday flow surface in the Eastbound direction is presented in figure 5. It appears from figure 5 that the R^2 for weekdays exceeded 90% and that an R^2 of 30% was estimated for the major incident day (11/5/1992, day 24). This low R^2 indicated that this incident had a substantial effect on the average traffic conditions. The Mondays also had a relatively high R^2 (exceeded 90%) except for a single Monday that had an incident in addition to a failure in some loop detectors. Fridays had a lower R^2, ranging from 75% to 90%. The Saturday and Sunday flow surfaces differed considerably from the weekday average surface (R^2 ranged from 0% to 60%). The same trend was found for the Westbound direction, however, due to the limited space these results are not presented here.

The variation, from the average weekday speed surface in the Eastbound direction, in R^2 during the 75 day analysis period was also analyzed by is not presented in this paper due to lack of sufficient space. Unlike the flow surface comparisons illustrated in figure 5, the speed variation appeared to be much more scattered. The scatter in the speed variation was most probably a result of shock waves proceeding along the detectorized section at different rates, even though the overall flow remained very similar. Interestingly, the major incident did not result in an R^2 worse than non-incident weekdays (day 24).

The variation, from the average weekday occupancy surface in the Eastbound direction, in R^2 during the 75 day analysis period was less scattered than the speed variation. Specifically, the R^2 ranged from 65% to 95% for the core weekdays, 45% to 90% for Mondays, 60% to 90% for Fridays and 0% for Saturdays and Sundays. As was the case for the flow, the R^2 for the major incident day (day 24) was much lower than the typical weekday R^2 (38%).

The Success Measure

The original loop detector measurements, that were made at 30 30-s intervals, were aggregated into 5-min observations for purposes of the analysis in this paper. Each 5-min observation was the sum of ten measurements. Based on the central limit theorem, it can therefore be assumed that each of these 5-min observations may become normally distributed because the 5-min observation on one day should not be correlated with the same observation on another day. In order to verify this assumption, a 5-min estimate of flow for the 22 core days in the Eastbound direction were estimated and stratified into bins. The observed probabilities were then tested using a Chi squared goodness of fit test in order to establish whether the normal distribution assumption was valid as illustrated in figure 6. The Chi squared type of analysis showed the observed 5-min flows to not be statistically different from the expected outcome of a normal distribution at the 95% confidence level. The test was also repeated for higher average flows in the range of 1800 vph/lane and similar results were found. Similar tests were conducted for speed and occupancy 5-min observations with similar outcomes. Thus, it appears that the normal distribution assumption is valid.

The three average weekday surfaces were obtained by averaging each cell of the matrix over the non-incident weekdays using equations 1 and 2. For each cell of these matrices the standard deviation of the mean observation was estimated using equation 8 and an upper bound and lower bound was estimated assuming a normal distribution using equations 9 and 10, respectively. An estimate of the proportion of similar observations was estimated as the ratio of observations within the upper and lower bounds to the total number of good observations using equation 11. An average proportion of cells within the average weekday confidence limits was subsequently estimated for the weekdays using equation 12. Using this proportion of successful observations, a lower confidence limit was estimated using equation 13 (*6*).

$$\sigma_{i,j} = \sqrt{\frac{\sum_{n=1}^{nd} (\overline{x}_{i,j} - x_{i,j}^n)^2}{(nday-1)}} \quad \forall \overline{x}_{i,j}, x_{i,j}^n \geq 0 \tag{8}$$

$$x_{i,j}^u = \overline{x}_{i,j} + 1.96 \times \sigma_{i,j} \quad \forall\, \overline{x}_{i,j} \geq 0 \tag{9}$$

$$x_{i,j}^l = \overline{x}_{i,j} - 1.96 \times \sigma_{i,j} \quad \forall\, \overline{x}_{i,j} \geq 0 \tag{10}$$

$$p^n = \frac{n_0^n}{n^n} \tag{11}$$

$$\overline{p} = \frac{\sum p^n}{nd} \tag{12}$$

$$p^l = \overline{p} - 3\sqrt{\frac{\overline{p}(1-\overline{p})}{nobs}} \tag{13}$$

Where:

nd, $x_{i,j,k}^n, x_{i,j}^n$, and $\overline{x}_{i,j}$ are defined as before.

nobs = 24×288=6912

$\sigma_{i,j}$ = standard deviation of 5-minute observation distribution at station *i* at time interval *j*.

$x_{i,j}^{u}$ = upper 95% confidence limit of 5-minute observation at station *i* at time interval *j*.

$x_{i,j}^{l}$ = lower 95% confidence limit of 5-minute observation at station *i* at time interval *j*.

n_0^n = number of observations for day *n* within the confidence limits of the average weekday surface.

n^n = number of good observations for day *n* (observation≥0).

p^n = proportion of observations for day *n* within the confidence limits of the average weekday surface.

$\overline{p}$ = average weekday proportion of observations within confidence limits.

p^l = lower bound of proportion of observations within confidence limits.

Figure 7 illustrates how flow p^n varied for the different days of the analysis period in the Eastbound direction. It appears that most of the non-incident weekdays were within the confidence limit (16 of 22 observations). The high number of observations outside the range was caused because the number of good observations (non-negative) for these extreme non-incident weekdays was less than *nobs* (used in estimating the confidence limits) and thus the lower confidence limit should have decreased to reflect the smaller number of observations, however, this was not done. The major incident (day 24) did not have a major influence on p^n, which was 78%, indicating that traffic conditions were similar, based on this index, to typical core weekday conditions prior and after the effects of the incident were removed. This high p^n resulted because this measure is not affected by how much the observation is outside the confidence limits, and thus the fact that the incident had an extreme effect on traffic flow was not reflected. It is important to note that, except for a single incident day, all the incident days fell outside the above confidence range.

The Monday flows appeared to be near the borderline of the weekday flows (20% of the observations fell within the confidence range). Fridays differed from the weekday conditions and so did the Saturdays and Sundays (as 0% of the days fell within the confidence range). The Westbound direction experienced a similar trend in variation of the flow p^n.

Summary

In this Section two methods for distinguishing typical traffic conditions from non-typical traffic conditions were investigated. The Regression Method, that utilizes the flow and occupancy surface, could distinguish typical weekday traffic conditions from non-typical conditions. However, the noise in the speed surface was too large to enable the identification of any systematic underlying variations. In the Regression Method it was not possible to determine any statistical confidence limits, which limits its practical usage.

The Success Measure of the flow had the advantage of yielding confidence limits in order to statistically distinguish between significant and insignificant variations from the typical traffic conditions. This method could potentially be developed further as an on-line incident detection routine by decreasing the averaging process from 5 minutes to 2 minutes and estimating a *p* value on-line for each station. A value outside the confidence limits would indicate some suspicious observation and a second *p* value outside the confidence range could set off an alarm. Such an incident detection approach differs from other techniques that detect incidents on the basis of the traffic state at an upstream and downstream detector station (*7*), rather than the deviation of the current observation from some time-of-day based bounds.

CONCLUSIONS AND RECOMMENDATIONS

The premise of most equilibrium traffic assignments is that drivers select efficient routes for the current day based on the assumption that, in the absence of incidents, temporal traffic patterns are very similar from one day to the next. Many IVHS technologies attempt to explore that fact that, even in the absence of incidents, traffic conditions on one day may be quite different from a similar previous day. This paper attempted to quantify the above similarities and differences, both for incident and non incident days.

It is recommended that the quantification of these similarities and differences should be directly incorporated in any IVHS system designs and benefit simulations. The present frequent use of hypothesized similarities or differences of day-to-day traffic may lead to designs or benefit estimates which are not consistent with the actual behavior of traffic. Based on this paper this behavior has at least been quantified for a one location and a potential step towards a standardized procedure for analyzing others in a comparable fashion can be adopted from this initial study.

ACKNOWLEDGMENTS

The authors of this paper would like to acknowledge the comments, critiques and suggestions made by Hana Suleiman at Queen's University.

REFERENCES

1. Rakha H., M. Van Aerde M., E.R. Case E.R. and A. Ugge (1989), "Evaluating the Benefits and Interactions of Route Guidance and Traffic Control Strategies Using Simulation," Proceedings of the IEEE Vehicle Navigation and Information Systems Conference, Toronto, September, pp. 296-303.

2. Rilett L.R. (1992), *Modelling of TravTek's Dynamic Route Guidance Logic Using the INTEGRATION Model*, Ph.D. Thesis, Queen's University.

3. Shbaklo S., F. Koppelman, and C. Bhat (1993), "Static Prediction Models of Flow and Occupancy," ADVANCE Program Report, TRF-TT-05.

4. Bhat C., F. Koppelman, A. Sen, P. Thakuriah, P. Li, and N. Rouphail (1992), "Short-Term Travel Time Prediction," ADVANCE Project Report, TRF-TT-02.

5. SYSTAT (1992), *SYSTAT for Windows: Statistics*, *Version 5 Edition*, Evanston, IL.

6. Crow E.L., F.A. Davis and M.W. Maxfield (1960), *Statistics Manual*, Dover 0-486-60599-X.

7. Gall A.I. and F.L. Hall (1989), "Distinguishing Between Incident Congestion and Recurrent Congestion: A Proposed Logic," Transportation Research Board 68th Annual Meeting, Paper No. 880225.

Table 1: Single-Factor ANOVA Results

Description	ANOVA groups	DF (within groups)	DF (total)	F	F_{crit}	Sig (95%)
Flow (EB)	Tue. vs. Wed. vs. Thur.	19	21	1.16	3.52	No
	weekday vs. Mon.	29	30	5.32	4.18	Yes
	weekday vs. Fri.	30	31	101.87	4.17	Yes
	weekday vs. Sat.	30	31	682.84	4.17	Yes
	weekday vs. Sun.	32	33	384.79	4.15	Yes
Speed (EB)	Tue. vs. Wed. vs. Thur.	19	21	2.76	3.52	No
	weekday vs. Mon.	29	30	2.40	4.18	No
	weekday vs. Fri.	30	31	101.87	4.17	Yes
	weekday vs. Sat.	30	31	682.84	4.17	Yes
	weekday vs. Sun.	32	33	384.79	4.15	Yes
Occ. (EB)	Tue. vs. Wed. vs. Thur.	19	21	1.88	3.52	No
	weekday vs. Mon.	29	30	1.20	4.18	No
	weekday vs. Fri.	30	31	17.13	4.17	Yes
	weekday vs. Sat.	30	31	16.01	4.17	Yes
	weekday vs. Sun.	32	33	47.25	4.15	Yes
Flow (WB)	Tue. vs. Wed. vs. Thur.	30	32	0.85	3.32	No
	weekday vs. Mon.	41	42	7.03	4.08	Yes
	weekday vs. Fri.	41	42	66.39	4.07	Yes
	weekday vs. Sat.	41	42	1678.67	4.08	Yes
	weekday vs. Sun.	43	44	1668.55	4.07	Yes
Speed (WB)	Tue. vs. Wed. vs. Thur.	30	32	0.55	3.32	No
	weekday vs. Mon.	41	42	0.11	4.08	No
	weekday vs. Fri.	41	42	12.15	4.07	Yes
	weekday vs. Sat.	41	42	22.34	4.08	Yes
	weekday vs. Sun.	43	44	23.54	4.07	Yes
Occ. (WB)	Tue. vs. Wed. vs. Thur.	30	32	0.62	3.32	No
	weekday vs. Mon.	41	42	0.30	4.08	No
	weekday vs. Fri.	41	42	15.98	4.07	Yes
	weekday vs. Sat.	41	42	113.03	4.08	Yes
	weekday vs. Sun.	43	44	208.02	4.07	Yes

List of Tables

List of Figures

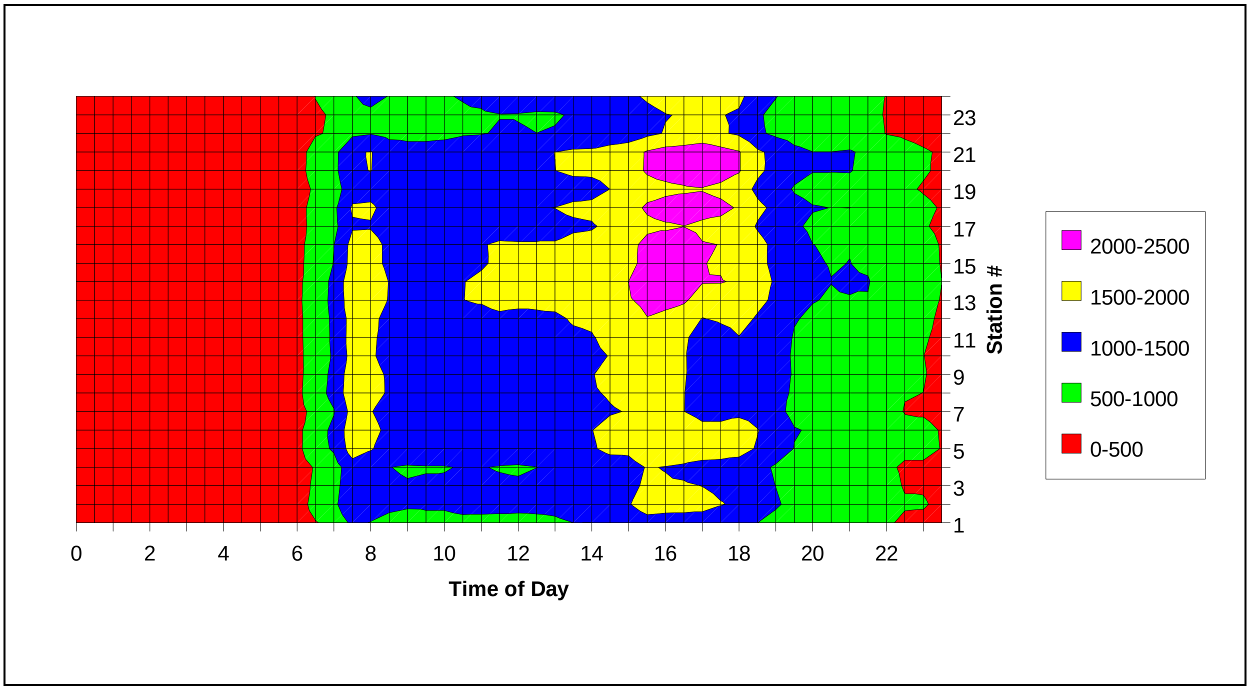

Figure 1: Spatial and temporal Eastbound flow variation for an average weekday (vph)

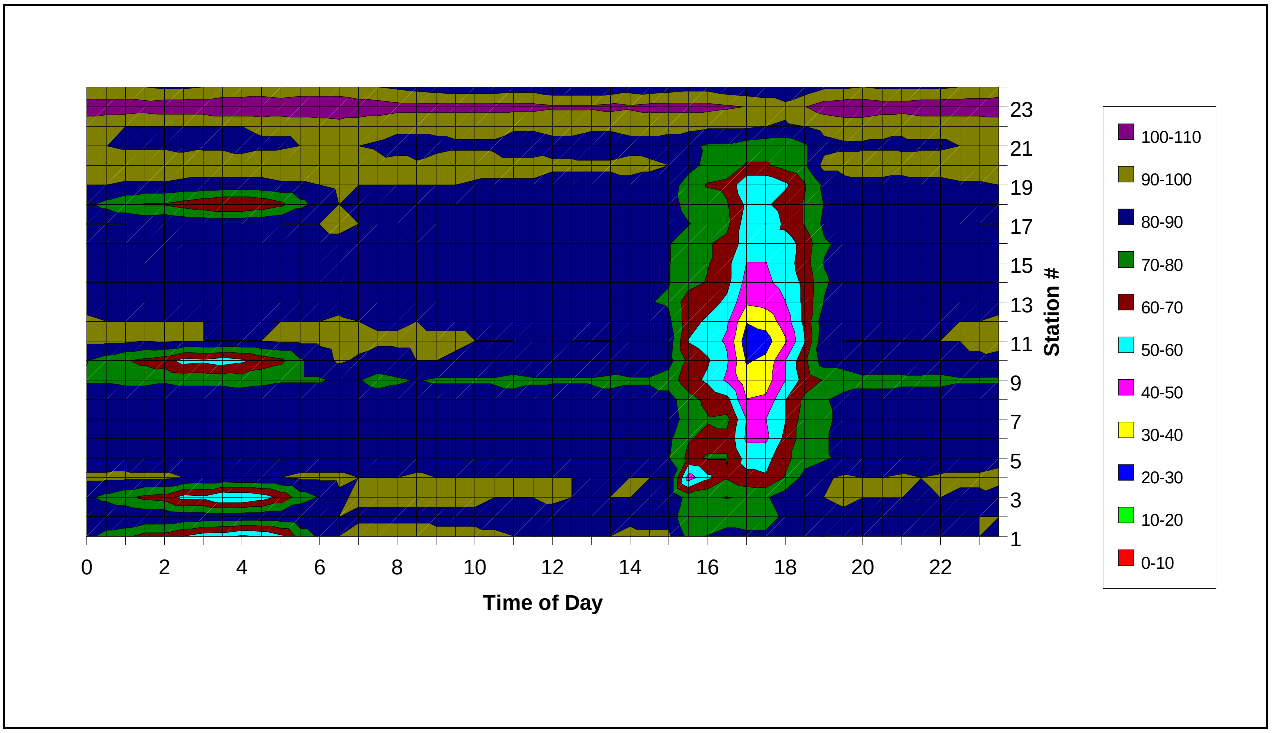

Figure 2: Spatial and temporal Eastbound speed variation for an average weekday (km/hr)

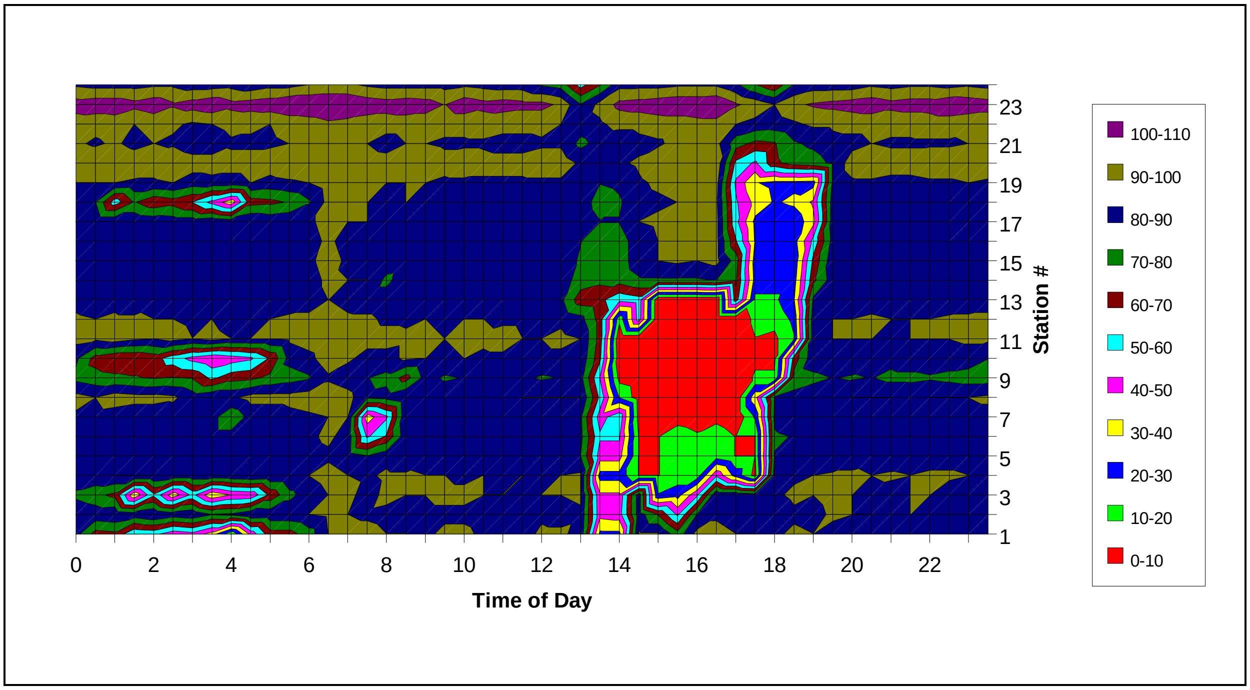

Figure 3: Spatial and temporal Eastbound speed variation for during an incident (km/hr)

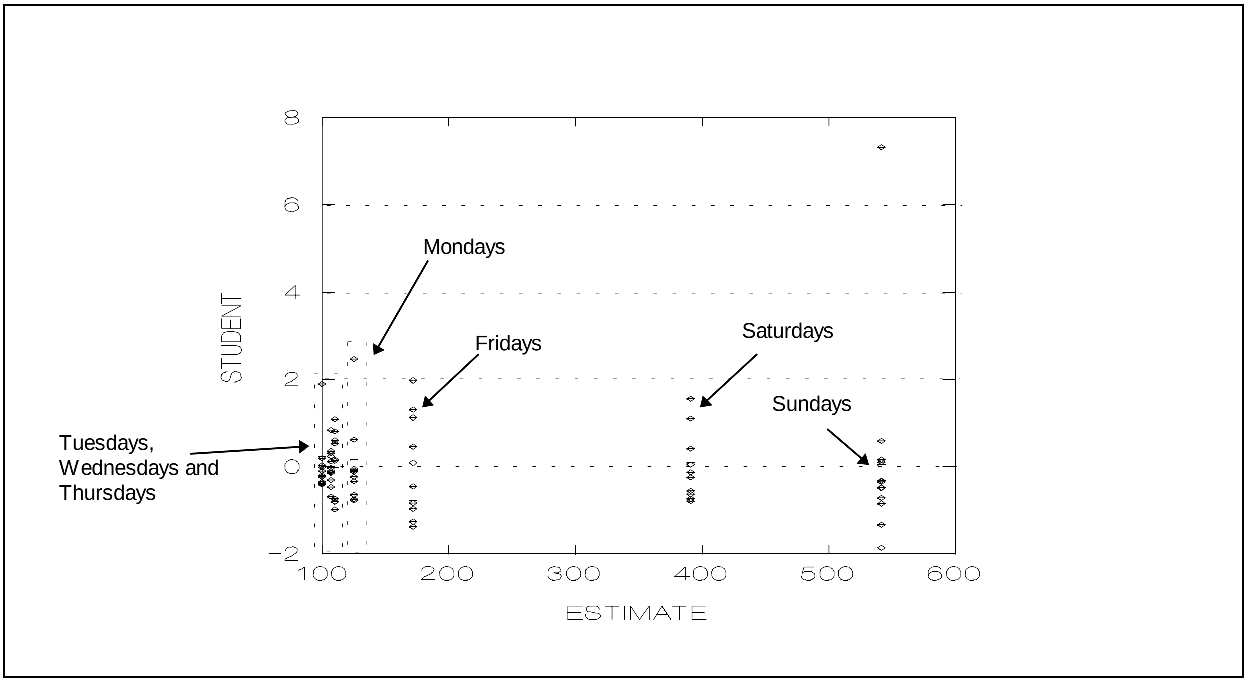

Figure 4: Variation in residual error as a function of the RMSE estimate for Easbound flows

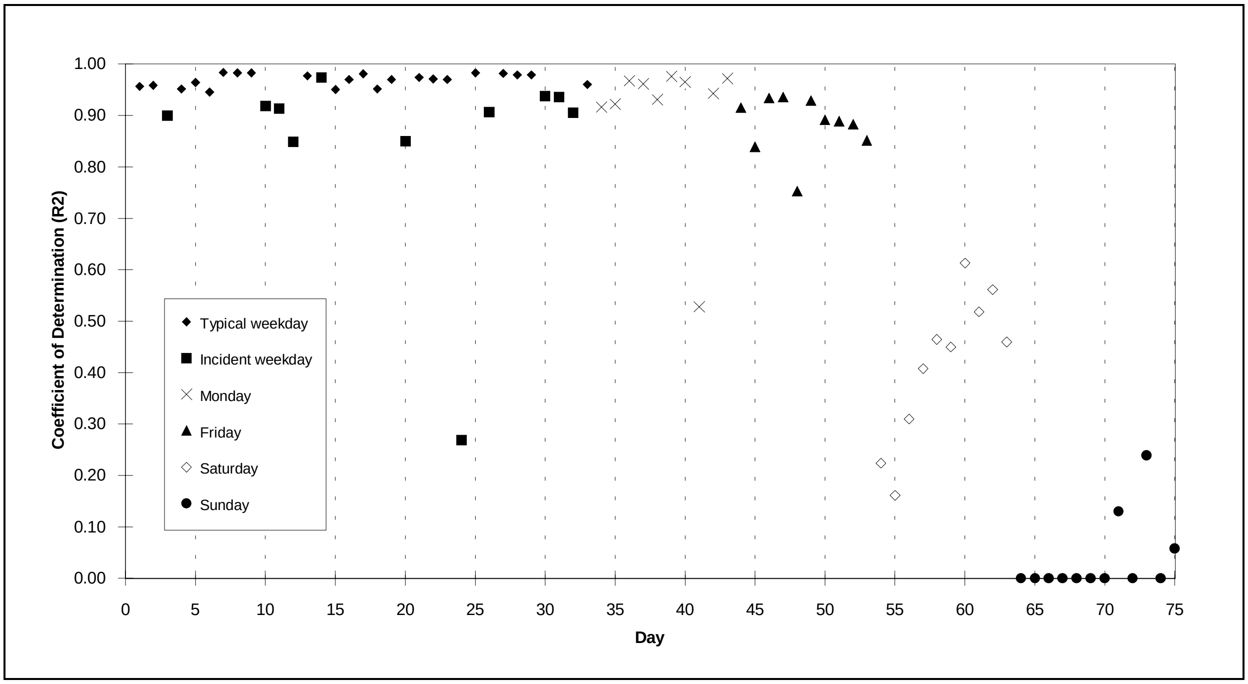

Figure 5: R^2 variation for flow in Eastbound direction

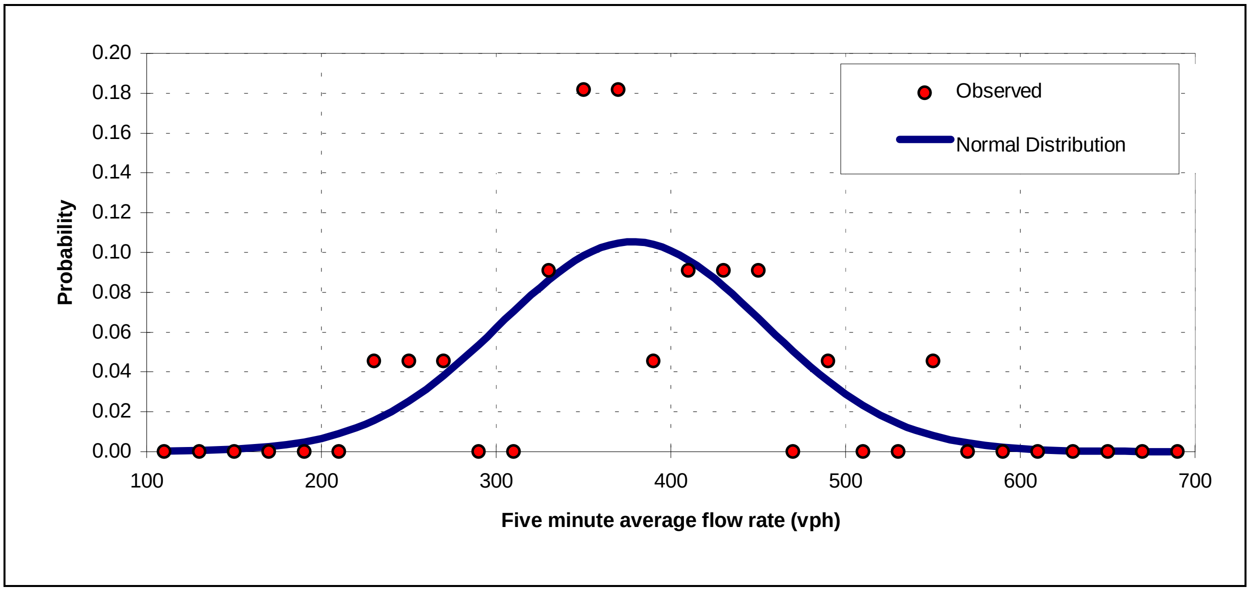

Figure 6: Observed and normal distribution 5-min flow rate estimates for 22 weekdays

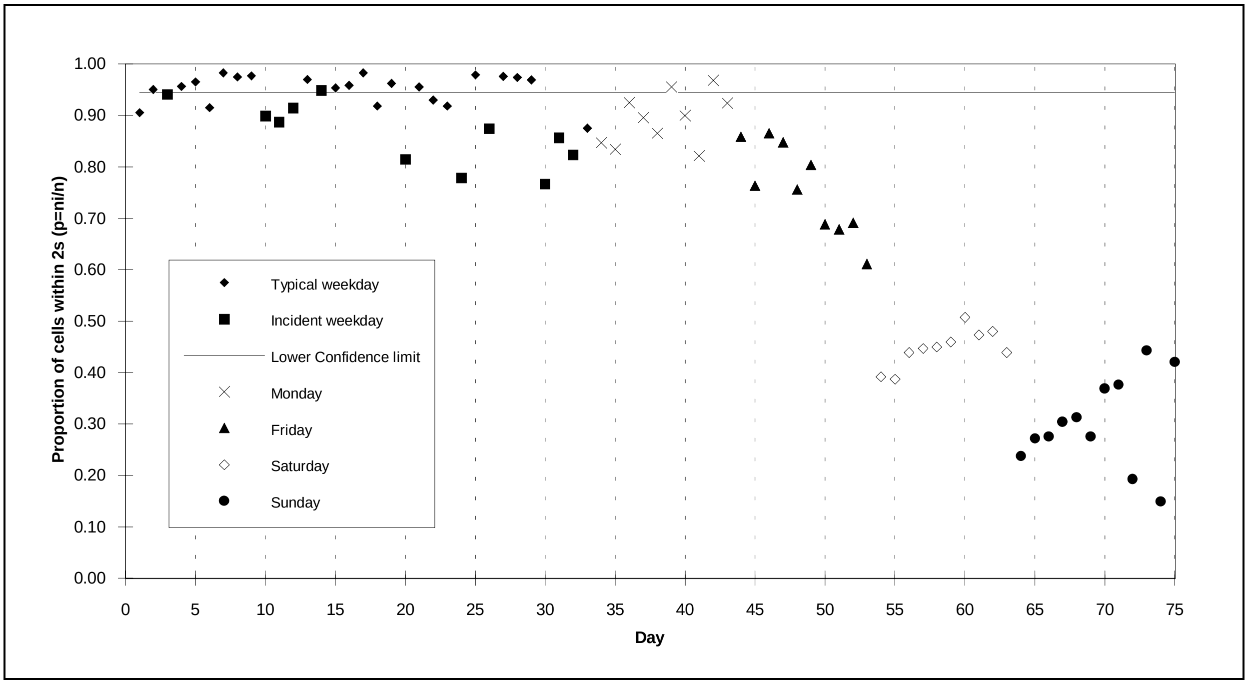

Figure 7: Variation of *p* for flows in the Eastbound direction