



Virginia Center *for* Transportation
INNOVATION
& **RESEARCH**

A Model to Forecast Peak Spreading

http://www.virginiadot.org/vtrc/main/online_reports/pdf/12-r11.pdf

JOHN S. MILLER, Ph.D., P.E.
Associate Principal Research Scientist

Final Report VCTIR 12-R11

Standard Title Page - Report on Federally Funded Project

1. Report No.: FHWA/VCTIR 12-R11	2. Government Accession No.:	3. Recipient's Catalog No.:	
4. Title and Subtitle: A Model to Forecast Peak Spreading		5. Report Date: April 2012	
		6. Performing Organization Code:	
7. Author(s): John S. Miller, Ph.D., P.E.		8. Performing Organization Report No.: VCTIR 12-R11	
9. Performing Organization and Address: Virginia Center for Transportation Innovation and Research 530 Edgemont Road Charlottesville, VA 22903		10. Work Unit No. (TRAIIS):	
		11. Contract or Grant No.: 98770	
12. Sponsoring Agencies' Name and Address: Virginia Department of Transportation Federal Highway Administration 1401 E. Broad Street 400 North 8th Street, Room 750 Richmond, VA 23219 Richmond, VA 23219-4825		13. Type of Report and Period Covered: Final	
		14. Sponsoring Agency Code:	
15. Supplementary Notes:			
16. Abstract: <p>As traffic congestion increases, the K-factor, defined as the proportion of the 24-hour traffic volume that occurs during the peak hour, may decrease. This behavioral response is known as peak spreading: as congestion grows during the peak travel times, motorists may shift their departure time to a non-peak hour. Knowing whether K-factors will remain constant or will change will affect the estimation of travel demand, and the resultant transportation performance, since the traffic volume during a given hour may affect travel speed and vehicle emissions.</p> <p>The purpose of this study was to develop a model for forecasting peak spreading whereby peak spreading is measured as change in the K-factor. Data were collected from 32 continuous count stations in the six Northern Virginia counties of Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford for the period 1997-2010. Because some stations gave two-directional counts and some gave only one-directional counts, there were 52 station-direction combinations, or sites, for analysis purposes. The data collected showed that the average annual K-factor adjusted for months for which data were not available decreased by an average of 0.006 ($p < 0.01$), from 0.103 to 0.097, during the period. The 24-hour volume-to-capacity ratio, which is a surrogate for travel congestion, increased by an average of 0.7 ($p < 0.01$), from 7.3 to 8.0. Both changes were statistically significant.</p> <p>Two models to forecast K-factors were developed in this study. Model 1, for use with an established roadway with an existing K-factor, explained 88% of the variation in K-factors and is based on the previous K-factor, the percentage increase in the jurisdiction's employment, and the roadway functional class. Model 2, for use with a new roadway without an existing K-factor, explained 66% of the variation in K-factors and is based on the percentage change in the jurisdiction's employment; circuitry, i.e., whether the route is radial or circumferential; and for freeways, the 24-hour volume-to-capacity ratio. Use of these variables is advantageous as they are typically available when a 10-year forecast is made. The two models have three implications for forecasting peak spreading. First, site characteristics (e.g., functional class, 24-hour volume-to-capacity ratio) and regional socioeconomic characteristics (e.g., jurisdictional employment growth) affect the K-factor. Second, the 24-hour volume-to-capacity ratio affects the forecasts, even though the effect is evident only after controlling for other variables. Third, the K-factor varies more across sites with the time period held constant than across time periods with the site held constant. The study recommends that VDOT consider the use of the two models when more detailed data are not available; their use would provide an empirically based alternative to assuming the K-factor will remain constant.</p> <p>A potential study limitation is that congestion during the "before" period in Northern Virginia was already so great that any congestion-based effects on peak spreading had already occurred. However, as the large variability in K-factors across sites dampened the overall effect of congestion, it may be the case repeating this study in other locations would yield similar results.</p>			
17 Key Words: planning, peak hour traffic, peak periods, design volume, design hourly volume		18. Distribution Statement: No restrictions. This document is available to the public through NTIS, Springfield, VA 22161.	
19. Security Classif. (of this report): Unclassified	20. Security Classif. (of this page): Unclassified	21. No. of Pages: 54	22. Price:

FINAL REPORT
A MODEL TO FORECAST PEAK SPREADING

John S. Miller, Ph.D., P.E.
Associate Principal Research Scientist

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Virginia Center for Transportation Innovation and Research
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

Charlottesville, Virginia

April 2012
VCTIR 12-R11

DISCLAIMER

The contents of this report reflect the views of the author, who is responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Copyright 2012 by the Commonwealth of Virginia.
All rights reserved.

ABSTRACT

As traffic congestion increases, the *K-factor*, defined as the proportion of the 24-hour traffic volume that occurs during the peak hour, may decrease. This behavioral response is known as peak spreading: as congestion grows during the peak travel times, motorists may shift their departure time to a non-peak hour. Knowing whether K-factors will remain constant or will change will affect the estimation of travel demand, and the resultant transportation performance, since the traffic volume during a given hour may affect travel speed and vehicle emissions.

The purpose of this study was to develop a model for forecasting peak spreading whereby peak spreading is measured as change in the K-factor. Data were collected from 32 continuous count stations in the Northern Virginia counties of Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford for the period 1997-2010. Because some stations gave two-directional counts and some gave only one-directional counts, there were 52 station-direction combinations, or sites, for analysis purposes. The data collected showed that the average annual K-factor adjusted for months for which data were not available decreased by an average of 0.006 ($p < 0.01$), from 0.103 to 0.097, during the period. The 24-hour volume-to-capacity ratio, which is a surrogate for travel congestion, increased by an average of 0.7 ($p < 0.01$), from 7.3 to 8.0. Both changes were statistically significant.

Two models to forecast K-factors were developed in this study. Model 1, for use with an established roadway with an existing K-factor, explained 88% of the variation in K-factors and is based on the previous K-factor, the percentage increase in the jurisdiction's employment, and the roadway functional class. Model 2, for use with a new roadway without an existing K-factor, explained 66% of the variation in K-factors and is based on the percentage change in the jurisdiction's employment; circuitry, i.e., whether the route is radial or circumferential; and for freeways, the 24-hour volume-to-capacity ratio. Use of these variables is advantageous as they are typically available when a 10-year forecast is made.

The two models have three implications for forecasting peak spreading. First, site characteristics (e.g., functional class, 24-hour volume-to-capacity ratio) and regional socioeconomic characteristics (e.g., jurisdictional employment growth) affect the K-factor. Second, the 24-hour volume-to-capacity ratio affects the forecasts, even though the effect is evident only after controlling for other variables. Third, the K-factor varies more across sites with the time period held constant than across time periods with the site held constant.

A limitation of this analysis is that it was based only on six counties in Northern Virginia. This limitation may be substantial if the congestion in these counties during the "before" period was already so great that any congestion-based effects on peak spreading had already occurred. However, as this study suggested, the large variability in K-factors across sites dampened the overall effect of congestion and it may be the case that the results would be similar if the study were repeated in other areas of Virginia.

The study recommends that VDOT consider the use of the two models when more detailed data are not available; their use would provide an empirically based alternative to assuming the K-factor will remain constant.

FINAL REPORT

A MODEL TO FORECAST PEAK SPREADING

John S. Miller, Ph.D., P.E.
Associate Principal Research Scientist

INTRODUCTION

As traffic congestion increases, the duration of the morning, noontime, or afternoon peak period may also increase. This phenomenon is known as *peak spreading*. Peak spreading is a behavioral response: some motorists may shift their travel departure times to slightly before or after the peak period in response to increasing traffic congestion. As a result, the length of the congested period may grow.

For example, Figure 1 shows hourly traffic volumes for SR 7 Westbound in Loudoun County, Virginia, on Wednesday, April 8, in 1998 and 2009. Total volumes increased substantially (by 61%) such that in 2009, peak hour volumes (almost 4,000) approached capacity (i.e., 4,400, based on the Virginia Department of Transportation's [VDOT's] Statewide Planning System [SPS] database). If a *congested period* is arbitrarily defined as a period during which demand is at least half capacity, the length of the congested period grew from 2 hours in 1998 to 5 hours in 1999. Further, the proportion of 24-hour traffic occurring during the hour with the highest volume—i.e., the peak hour—decreased from 0.143 (in 1998) to 0.127 (in 2009). In a future year, if volumes were to continue to grow such that eventually the demanded volume was equal to capacity, the peak should become less sharp relative to traffic volumes during other portions of the day.

Although the general concept of peak spreading is fairly straightforward, several details affect the nature of the spreading; many of these are shown in Figure 1. Morning and evening peak spreading may differ and the peak hour may change. Barnes (1998) suggested that morning peak spreading is generally easier to model explicitly because work trips comprise a greater proportion of morning peak than evening peak trips. Roadway characteristics may change through reconstruction (e.g., adding a lane to increase capacity) or management (e.g., tolling or providing motorist information). The socioeconomic characteristics of a region may change: through increases in population and employment. For example, in the case of a facility that at one time never exceeded capacity and had a high traffic volume during the peak hour, the traffic volume may spread throughout the day if congestion is well in excess of capacity.

Peak spreading matters to VDOT for two reasons. First, the Transportation Planning Research Advisory Committee (2009) of the Virginia Transportation Research Council (now the Virginia Center for Transportation Innovation and Research [VCTIR]) mentioned the need for forecasts of “K and D parameters” on the “Request for Traffic Data” form (Form LD-104) (VDOT, 2008). The “K parameter,” better known as the K-factor (Committee on Highway Capacity and Quality of Service, 2010a), is defined in this report as the proportion of the 24-hour

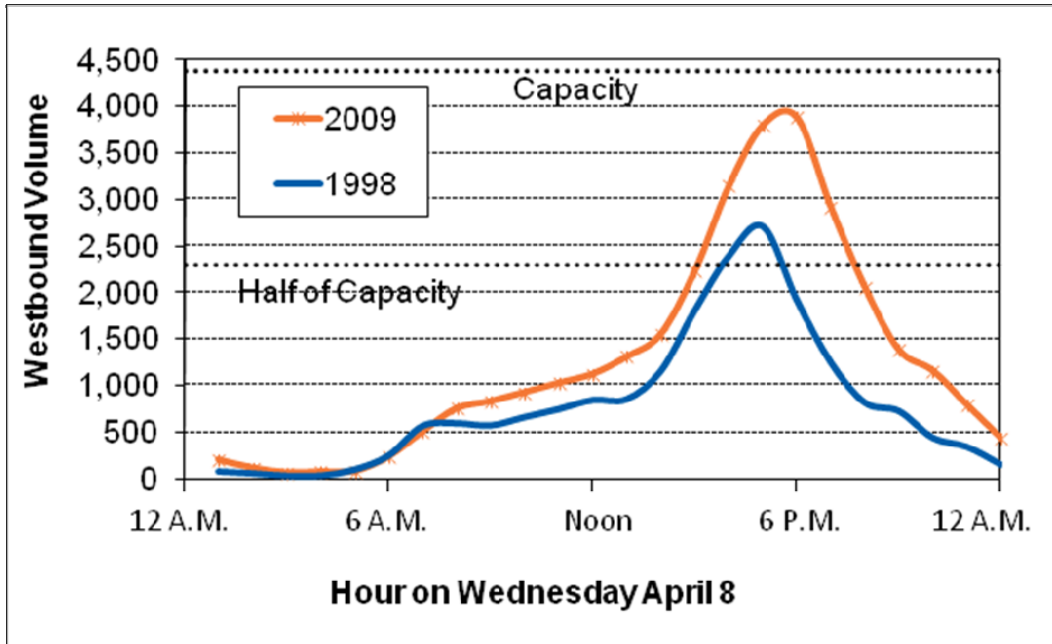


Figure 1. Hourly Volumes on SR 7 Westbound Between SR 9 (Clarkes Gap) and West City Line of Leesburg in Loudoun County (Link 090063)

traffic volume that occurs during the peak hour; it is an indication of peak spreading. Form LD-104 is used by VDOT’s Location & Design Division to request traffic counts 5, 10, 15, and 20 years after a project is completed and to request the design year hourly volume 11 or 22 years after the construction project is advertised, depending on the type of facility (VDOT, 2008). Second, how peak spreading characteristics may change over time affects a number of policy issues including those regarding air quality analysis, evaluation of possible travel demand management strategies, and estimation of required infrastructure investments (Barnes, 1998). For example, a small change in speed may yield a substantive change in emissions. As another example, congested locations can be sensitive to traffic volumes in the sense that at volume-to-capacity ratios that approach 1.0, a modest increase in volume may cause a substantial decrease in travel speed.

PURPOSE AND SCOPE

The purpose of this study was to develop a model to forecast peak spreading in which peak spreading is measured as the proportion of the 24-hour traffic volume that occurs during the peak hour, i.e., the K-factor, and the model includes variables, as appropriate, based on socioeconomic data and characteristics of the roadway facility. The development of this model was recommended by the planning staff of VDOT’s Northern Virginia (NOVA) District (Miller, 2011).

The scope of the study was limited to data collected from continuous count stations in or near VDOT’s NOVA District for the period 1997-2010 and a forecasting horizon of roughly one decade. This period was selected because continuous count stations became operational no

earlier than 1997 and the data collection period ended in December 2010. For forecasting purposes, socioeconomic data that are readily accessible, such as jurisdiction-level estimates of population and employment, were used.

METHODS

A case study methodology was used. After a literature review, K-factors were extracted from various continuous count stations in Northern Virginia over multiple years. Then, models were developed to forecast how these K-factors may change over time based on previous trends. Five tasks defined the methodology:

1. Perform a literature review.
2. Establish a data collection plan for obtaining K-factors.
3. Collect the data.
4. Analyze the data.
5. Develop models to forecast changes in K-factors.

Performance of Literature Review

The purpose of the literature review was twofold: to determine how to define *peak spreading* and to identify empirical approaches for forecasting peak spreading and literature regarding the development of peak spreading models. Relevant literature was identified using the Transportation Research Information Service (TRIS). The empirical approaches considered in the literature review included lookup tables and mathematical techniques.

Establishment of Data Collection Plan for Obtaining K-factors

Data were to be collected from the 50 continuous count stations in the Northern Virginia counties of Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford for the period 1997-2010. The plan was to obtain K-factors at every station for one Tuesday, Wednesday, and Thursday block every month during the period. The goal was to avoid days with snow or ice and federal and state holidays such that the K-factors obtained would be comparable.

Dates for holidays in Virginia were obtained from Time and Date AS (2010), and dates for snow/ice days were obtained from the National Climatic Data Center (undated). For example, for January 2001, holidays occurred on New Year's Day (Monday, January 1), Martin Luther King Jr. Day (Monday, January 15), and Lee-Jackson Day (Friday, January 19); thus, a Tuesday/Wednesday/Thursday combination in early January 2001 that avoided workweeks with holidays was January 9-11. However, for January 2000, holidays occurred on New Year's Day (Saturday, January 1); Lee-Jackson Day (Friday, January 14); and Martin Luther King Jr. Day (Monday, January 17); thus, the best Tuesday/Wednesday/Thursday combination in early January 2000 that avoided holidays during the workweek was January 4-6.

In some cases, there were exceptions to yearly patterns. For example:

- In 2001, with Thanksgiving falling on Thursday, November 22, and Veteran’s Day observed on Monday, November 12, the Tuesday/Wednesday/Thursday combination for November that avoided a holiday was November 27-29, which, unlike in previous years, was after the Thanksgiving holiday.
- In 2006, with New Year’s Day observed falling on Monday, January 2; Lee-Jackson Day on Friday, January 13; and Martin Luther King Jr. Day on Monday, January 16, there was no early January workweek that avoided a holiday. Thus, January 10-12 was chosen: it was recognized that Lee-Jackson Day was a state rather than a national holiday and might have closer-to-normal traffic patterns. A similar situation occurred in 2002: because of snow, January 15-17 was chosen. Although the state holiday on January 18 may also have affected traffic, there were no other periods in early January that avoided snow or a holiday at the end of the workweek.

Table 1 shows the periods for which K-factors were sought.

Table 1. Data Collection Periods for 1997-2010

Month	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Jan.	7-9	6-8	5-7	4-6	9-11	15-17 ^a	7-9	6-8 ^b	4-6	10-12	9-11	8-10	6-8	5-7
Feb.	11-13	10-12	9-11	8-10	13-15	12-14	11-13	10-12	8-10	7-9	13-15	12-14	10-12	9-11
March	11-13	10-12	16-18 ^c	7-9	13-15	12-14	11-13	9-11	15-17 ^d	7-9	13-15	11-13	10-12	9-11
April	15-17	14-16	13-15	11-13	17-19	16-18	15-17	13-15	12-14	11-13	17-19	15-17	14-16	13-15
May	13-15	12-14	11-13	9-11	15-17	14-16	13-15	11-13	10-12	9-11	15-17	13-15	12-14	11-13
June	24-26	23-25	22-24	20-22	26-28	25-27	24-26	22-24	21-23	20-22	26-28	24-26	23-25	22-24
July	15-17	14-16	13-15	11-13	17-19	16-18	15-17	13-15	12-14	11-13	17-19	15-17	14-16	13-15
Aug.	12-14	11-13	10-12	8-10	14-16	13-15	12-14	10-12	9-11	8-10	14-16	12-14	11-13	10-12
Sept.	16-18	15-17	14-16	12-14	18-20	17-19	16-18	14-16	13-15	12-14	18-20	16-18	15-17	14-16
Oct.	21-23	20-22	19-21	17-19	23-25	22-24	21-23	19-21	18-20	17-19	23-25	21-23	20-22	19-21
Nov.	18-20	17-19	16-18	14-16	27-29	19-21	18-20	16-18	15-17	14-16	27-29	18-20	17-19	16-18
Dec.	9-11	8-10	7-9	5-7	11-13	17-19 ^e	9-11	7-9	6-8	5-7	11-13	9-11	8-10	7-9

^aBecause of a January 9 snow/ice event, January 15-17 was used instead of January 8-10 even though January 18 was a state holiday.

^bA snow/ice event occurred on January 8, but not until 5 P.M. As the following week contained a Friday holiday, January 6-8 was still used.

^cBecause of a March 9 snow/ice event, March 16-18 was used instead of March 9-11.

^dBecause of a March 8 snow/ice event, March 15-17 was used rather than March 8-10.

^eBecause of December 5 and 11 snow/ice events, December 17-19 was used instead of December 10-12.

Collection of Data

Three types of data were collected: socioeconomic data, information about the links where the 50 continuous count stations were located, and K-factors at the count stations for the period 1997-2010.

1. *Socioeconomic data.* Population, labor force, and employment data were obtained from the Weldon Cooper Center for Public Service (2003, 2010, 2011) and the Virginia Employment Commission (undated, 2011); data were tabulated by county

(Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford) and year for the locations of the continuous count stations.

2. *Roadway facility data.* Information about links for the locations of the continuous count stations was initially extracted from VDOT's SPS and Traffic Monitoring System (TMS) databases. These data included the jurisdiction in which the link was located; the one-way capacity measured in units of passenger cars per hour (pc/hr); the roadway functional class; and dates when the number of lanes, and hence the capacity, might have changed (see Appendix A). Advice from experts familiar with these data (Dittberner, 2011; Schinkel, 2011b) was essential for verifying changes in the number of lanes at several stations for which information from the database was unclear. A geographic information system was used to classify each link as radial or circumferential; in most cases, this determination of circuitry was straightforward (e.g., I-95 is radial, I-495 is circumferential), but for some routes, examination of the station on a map was necessary. For example, Route 123 has a circumferential function near I-95 in southern Fairfax County but a more radial function near the Arlington County line in northern Fairfax County.
3. *K-factors at the count stations for the period 1997-2010.* Based on data recorded by the continuous count stations, K-factors were extracted through the assistance of VDOT staff who provided .csv files containing all available records for all Northern Virginia stations. Multiple files were necessary because of the large file sizes; for example, the 1997 data consisted of 222,667 records, and the 2010 data consisted of 619,711 records. These data were imported into Microsoft Access, with each year stored as a separate file, and queries that extracted the data and calculated the K-factor were executed for each year in the time periods shown in Table 1; an example of such a query for 1997 is given in Appendix A. Figure 2 shows continuous count stations in and near VDOT's NOVA District. All were initially considered in the analysis, although missing data or changes in the roadway characteristics at some stations meant that not all stations were used to develop forecasting models, as described in Task 4. (Because the K-factor requires obtaining the peak hour volume and the 24-hour volume, these two data elements were also obtained as part of this step in addition to the day, month, and year for which the volumes were recorded.) "Missing data" refers to hourly volume data that for whatever reason were not available at a particular site on a particular day.

Analysis of Data

Two types of data analysis were undertaken.

1. *The variability (i.e., amount of scatter) in the K-factors for the data from all 50 stations was examined using analysis of variance (ANOVA).* This initial screening helped identify which variables were likely to be useful for forecasting the change in K-factors over time. It also showed the feasibility of forecasting a K-factor given the random variation in the dataset. Because some stations provided counts in two

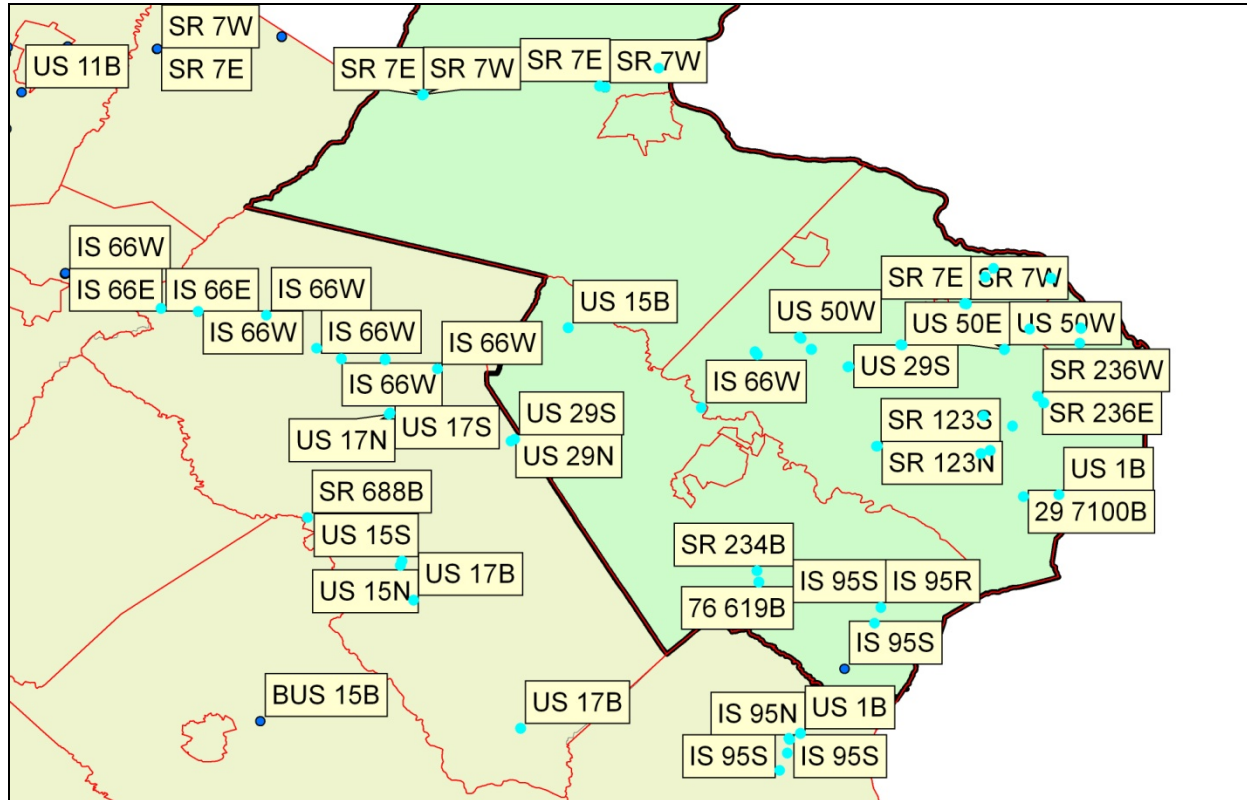


Figure 2. The 50 Continuous Count Stations in Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford Counties. Lighter dots indicate the stations in the six counties; these were used in the study. Darker dots indicate stations outside the six counties; these were not used in the study.

directions (hence 2 sites for analysis purposes) and some in only one direction (hence 1 site for analysis purposes), the 50 stations reflected 80 sites that could be analyzed. For this initial analysis, each daily K-factor at each site was treated as a single observation, with no adjustments made to the dataset. This dataset is defined in this report as the *entire dataset*.

2. *After a smaller dataset was culled from the entire dataset, an annual adjusted K-factor was estimated for each site in the smaller dataset to account for variability that might have resulted from months with missing data. Then, it was determined whether the annual adjusted K-factors in the smaller dataset had changed over time. This smaller dataset was based on 32 continuous count stations, and, since some stations produced two counts (i.e., counts for two directions), it reflected 52 sites. These 52 sites were chosen such that there was usually at least one decade between the first and last years that peak hour factors were available and at least 9 months of data were available for both the first year and the last year. This dataset is defined in this report as the smaller dataset.*

Development of Models to Forecast Changes in K-factors

The development of the models was a two-step process and was based on the smaller dataset:

1. *After it was determined that the change in annual adjusted K-factors was statistically significant, a linear regression model was developed to forecast the change in the K-factor over the period (usually 10 years) as a function of a variety of independent variables (hereinafter called Model 1). These included socioeconomic variables such as employment; facility variables such as roadway functional class; and the K-factor during the before period at the site. Multiple variables were tested, but only those that were statistically significant as determined from linear regression analysis (with a p -value of 0.05 or lower) were retained in the model.*
2. *This same dataset was used to develop another linear regression model to forecast a K-factor in the absence of an existing K-factor (hereinafter called Model 2). Such a model would be appropriate for a site where a new facility was being constructed and for which there were no historical data.*

RESULTS AND DISCUSSION

Literature Review

The concept of peak spreading is widely recognized in the literature. For example, Jin and Chiao (2008) provided an overview of approaches for quantifying peak spreading, and the Institute of Transportation Engineers (2006) provided guidance for using peak spreading to

assess transportation needs relative to land development. The literature review provided in Appendix B covers approaches to forecast peak spreading, applications of these approaches to problems of interest, ways to define *peak spreading*, and data considerations that affect the development of peak spreading models. Table 2 highlights themes found in the literature review.

Three unexpected clarifications resulted when the findings of the literature review were presented to the project's technical review panel, and these clarifications influenced subsequent tasks.

1. *Multiple definitions of peak spreading exist.* VDOT's primary interest concerned how the K-factor changes, and thus this definition was the focus of the study. In addition to the K-factor, dependent variables that characterize peak spreading include the length of the peak period (e.g., Karl and Gaffney, 2008), the proportion of peak hour volume during the peak 3- or 4-hour period (e.g., Cambridge Systematics, Inc., 1997), and the proportion of travel demand during a peak period coupled with a specific mode (e.g., Sall et al., 2010). This finding did not explicitly affect model calibration because VDOT staff clarified their interest in the K-factor, but it did affect the interpretation of the results for the Northern Virginia counties studied.
2. *Either the data collection plan should control for or a model should incorporate key independent variables known to affect peak spreading, such as seasonal factors (Ozbay et al., 2006); congestion-related factors (Margiotta et al., 1999); socioeconomic factors (Replogle, 1990); and roadway-specific attributes (Yang et al., 2009).* This finding led to an emphasis on ensuring that K-factors were adjusted to account for months with missing data and to the consideration of multiple socioeconomic variables, i.e., population, employment, and labor force. Interest was noted in literature that used socioeconomic variables, in particular an approach that considered population and employment (Replogle, 1990).
3. *The model should address challenges known to hamper efforts to forecast peak spreading, such as a lower than ideal number of permanent count locations (Karlström and Franklin, 2009).* This finding led to the inclusion of Northern Virginia sites in addition to those in the counties that comprise VDOT's NOVA District.

Data Collected

The entire dataset was based on 80 sites and contained 23,670 records, each representing a single K-factor collected at a given station in a particular direction for a single day in the period 1997-2010 inclusive. For each record, 50 variables in three categories in addition to the K-factor were available. Table 3 lists and describes some of these variables in alphabetical order.

Socioeconomic variables included Population, Labor Force, and Emp for each year during the period 1997-2010 for the jurisdiction in which each site was located. *Facility variables* included Site, Functional class, Circ, and Capacity. *Variables related to the K-factor* included peak hour Volume, 24-hour Volume, Day, Month, and Year. Some variables spanned two categories; for

example, the 24-hour volume-to-capacity ratio (24VC) is based on a variable related to the K-factor (24-hour Volume) and a facility variable (Capacity).

Table 2. Attributes of Peak Spreading Models Based on the Literature

Theme	Author	Attribute
Dependent variable	Holyoak and Taylor (2006)	The model may forecast two dependent variables: the proportion of travel that occurs during the peak hour and the length of the peak period.
	Barnes (1998)	A dependent variable may be the ratio of the volume in the one-half hour prior to the peak plus the volume in the one-half hour following the peak divided by the peak hour volume.
	Allen and Schultz (1996), Cambridge Systematics, Inc. (1997)	The model may forecast the proportion of peak period volume occurring during the peak hour during the peak period, which may be 3 to 4 hours in length.
	Martland (2005)	There will be considerable variation in the quality of the trip, implying that the model should also describe this variation.
	Martin and McGuckin (1998)	The morning peak may be more critical for air quality analyses (since ground level ozone precursor emissions have a longer time to react in the presence of sunlight); the evening peak may be more critical for congestion analyses (since volumes tend to be higher at that time).
Seasonal factors and tolls	Ozbay et al. (2006), Yang et al. (2009)	Seasonal variation may be significant, so it must be controlled for (through statistical analysis [Ozbay et al.] or the design of the data collection plan with different models for different months [Yang et al.]
	Wolff and Vilain (2007)	The model should control for the effects of tolls and seasonal variation. It may also be appropriate to have link-specific and seasonal factors in the model.
	Allen and Schultz (1996)	Substantially different models may be generated based on trip purpose and length (in terms of distance of the trip).
	Liu and Sharma (2006)	Holiday traffic may merit consideration for inclusion in the model as one study showed holidays provide a substantial portion of the highest hourly volumes per year.
Congestion-related factors	Cottrell (1998), Margiotta et al. (1999), Simons (2006)	The 24-hour volume-to- hour capacity ratio may be used to forecast the degree of peak spreading. This ratio, described in this report as the 24-hour volume-to-capacity ratio, has been used by these authors to represent the degree of congestion.
	Liu et al. (2007)	The availability of a parallel route should be an independent variable in the model (and perhaps the volume-to-capacity ratio on this route should be a factor).
Socioeconomic factors	Habib et al. (2009), Replogle (1990)	Independent variables in the model may be household and employment density, because mixed land uses (e.g., an area containing residences, shopping destinations, and employers) show less peaking (e.g., lower K-factors) than homogenous land uses (Replogle). However, there are some variations by industry type (Habib et al.).
	Yang et al. (2009)	Seasonal factors are also driven by roadway type and characteristics such as number of retirees.
	Ivan and Allaire (2001), Sinha and Thakuriah (2004)	<ul style="list-style-type: none"> Independent variables should include growth industry types, since some industries (e.g., retail) are more likely to have unconventional start times than others (e.g., finance). Independent variables may include income, age, and area type since these also influence work schedules.
	Purvis (2002)	Variables may include travel time, distance, income, and whether the individual worked in the retail employment industry.
Challenges to developing a peak spreading model	Karlström and Franklin (2009)	Variables that one has reason to believe will influence peak spreading, such as income, may not always be shown to be statistically significant with a particular dataset.
	Anderson and Donnelly (2008)	Model development may be hampered by insufficient data in terms of count locations.
	Karl and Gaffney (2008)	Peak spreading can occur within a few years [which may hamper 20-year forecasts].
	Cottrell (1998)	The definition of <i>congestion</i> may vary by region.

Table 3. Definitions of Variables Used in Developing Models to Forecast Peak Spreading

Variable	Definition
Capacity	Maximum vehicle flow rate of facility as obtained from the Statewide Planning System (SPS) based on methodology described in Appendix A
Circ	Circuity (i.e., whether the site is circumferential or radial): 1 if site is circumferential , e.g., I-495; 0 if site is radial , e.g., I-95
Day	The Tuesday, Wednesday, or Thursday during which volume data were obtained
Emp	Percentage change in number of filled jobs during the time period. The number of filled jobs is reported by the Virginia Employment Commission [undated, 2011] as a jurisdiction's employment.
Entire dataset	All data collected from all 50 continuous count stations (reflecting 80 sites) for all time periods
Forecast	A future K-factor that one seeks to determine when applying Models 1 and 2 (e.g., K_{new})
Freeway _{24VC}	24-hour volume divided by hourly capacity during after period for freeway sites only
Free	Freeway or expressway: 1 if facility's functional class is a freeway/expressway, 0 otherwise
Functional class	Facility functional classification, i.e., one of these four: freeway/expressway, urban arterial, rural two-lane, rural multilane
Interaction effect	Effect of the product (cross) of two variables (e.g., 24-hour volume-to-capacity ratio and functional class) on peak hour factor. An example of an interaction effect is the 24-hour volume-to-capacity ratio having a statistically significant impact on the K-factor where functional class equals freeways but not where functional class equals rural multilane.
K-factor	Proportion of 24-hour volume that occurs during peak hour
K_{new}	New K-factor in 2008, 2009, or 2010. Also called a prediction because variable is being predicted (as the dependent variable) by the independent variables in Models 1 and 2
K_{old}	Old K-factor from 1998, 1999, 2000, or 2001
Labor force	Supply (i.e., number) of workers residing in a given jurisdiction as reported by the Virginia Employment Commission (undated)
Link	A longitudinal section of roadway between 2 points
MEF _{Jan1998}	Monthly expansion factor for January 1998
Main effect	In an analysis of variance, the effect of a single variable, such as the year, on the peak hour factor
Month	Calendar month during which volume data were obtained
Peak hour	The single hour of a 24-hour calendar day with the highest traffic volume on a given link
Peak hour volume	The volume of a facility during the peak hour
Population	Number of persons residing in a given jurisdiction as reported by the Weldon Cooper Center for Public Service (2003, 2010, 2011)
Ruralmulti	Rural multilane road: 1 if the facility's functional class is a rural multilane, 0 otherwise.
Site	A combination of a station and a direction; e.g., if 1 station provides both eastbound counts and westbound counts, the station would constitute 2 sites
Smaller dataset	A subset of the entire dataset that reflected data from only 32 continuous count stations (or 52 sites) and only for specific time periods as defined in Table 6
Station	A continuous count station that provides volume data for a particular section of a roadway facility
Two	Rural two-lane road: 1 if facility's functional class is a rural two-lane road, 0 otherwise
Urbanart	Urban arterial facility: 1 if facility's functional class is an urban arterial, 0 otherwise
Volume	The traffic volume for a roadway section over either a 1-hour period (hence an hourly volume) or a 24-hour period (hence a 24-hour volume)
24VC	24-hour volume-to-capacity ratio; the ratio of a link's 24-hour volume divided by its hourly capacity
Year	Calendar year (1997 through 2010) for which traffic volume or a jurisdiction's population, employment, or labor force was obtained

The socioeconomic data (population, employment, and labor force) are available from published sources cited in this report including the Virginia Employment Commission (2011, undated) and the Weldon Cooper Center for Public Service (2003, 2010, 2011). Data concerning the locations of the links are available through the internal TMS and SPS databases operated by VDOT's Traffic Engineering Division and Transportation Mobility Planning Division, respectively. Although hourly volume data are also available through TMS, the large number of records that needed to be processed necessitated acquiring the files from staff of VDOT's Traffic Engineering Division, who provided these files for each year studied (e.g., Schinkel, 2011a). The data used for this study are also available from the author.

Data Analysis

As stated in the "Methods" section, data analyses were performed to determine (1) the variability of the K-factors, and (2) whether the K-factors had changed over time.

Variability in K-factors

As described previously, an initial examination of the entire dataset was undertaken using ANOVA to determine which variables likely had the greatest potential to explain any variability in the K-factors. For this analysis, each daily K-factor at each site was treated as a single observation, with no adjustments made to the dataset.

ANOVA showed that when a site was defined as the station and the direction, the Site variable explained most (61.8%) of all variation. Then, the Site variable was established as a "block" in ANOVA (Montgomery, 2001b), meaning that the Site variable was not crossed with any other variables and the impact of other variables could be detected while controlling for the variation across sites. Keeping the Site variable as a block and including the Month and Year variables (main effects) plus the interaction effect of the Month variable crossed with the Year variable explained 65.2% of the variation; adding the Day variable as a main effect increased only 65.3% of the variation. Thus, the Site variable explained most of the variation in K-factors. Month and year explained some variation. The day explained relatively little of the variation; this was not surprising given that data collection days were Tuesdays, Wednesdays, and Thursdays with the avoidance of holidays and snow/ice days; in such a dataset, the K-factors would not be expected to change substantially because these days were generally workdays with, one would expect, a comparable number of commuters using the facilities. Although there would be variation in the K-factors for such a dataset, when considering only the Day variable and excluding all other variables, one would not expect the variation to have a pattern.

The impacts of year and the month were statistically significant in terms of explaining variation in the K-factor. Figure 3 shows the 95% confidence intervals for the marginal means by year, and Figure 4 shows them by month. As may be seen, the marginal means for the later years (e.g., 2004-2010) were smaller than those for the earlier years (e.g., 1997-2002). Such a result was to be expected to the extent that the later years were periods of heavier congestion when presumably there would be greater peak spreading. However, the practical difference in these K-factors was relatively small, with all average values being between 0.096 and 0.110. Figures 3 and 4 alone do not show the reason for the variation, such as the spike in 1999 in the

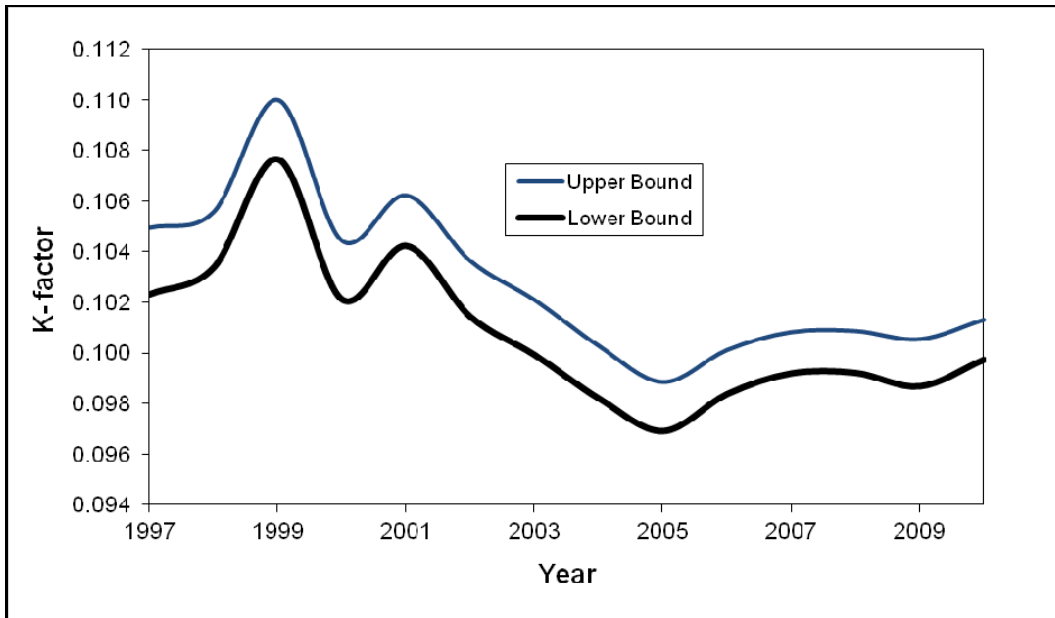


Figure 3. Marginal Means of K-factors for All Sites, All Dates by Year

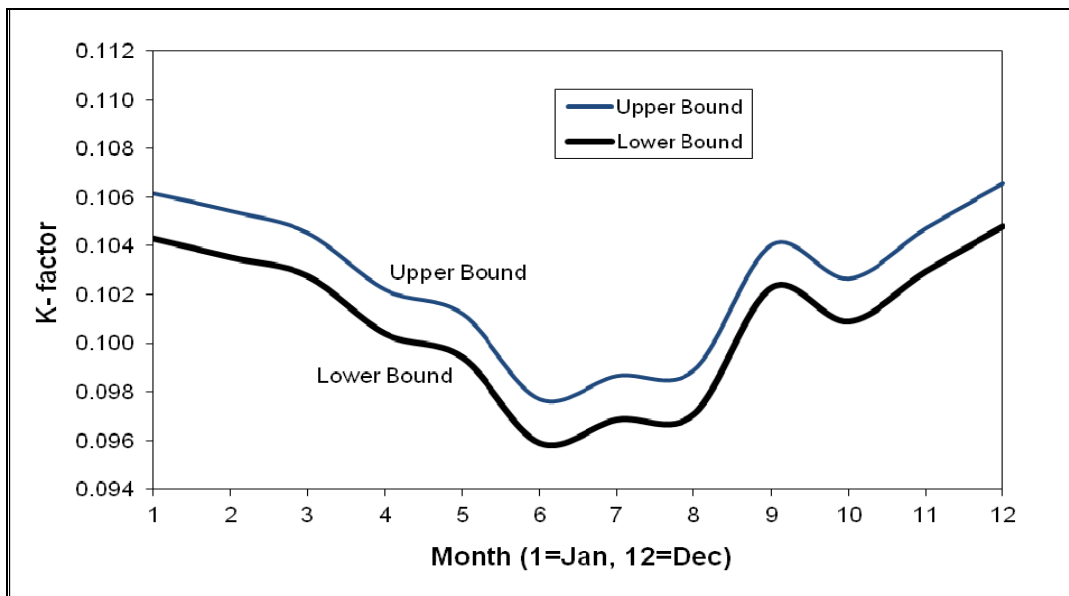


Figure 4. Marginal Means of K-factors for All Sites, All Dates by Month

K-factor. Figure 4 clearly shows that there will be substantial seasonal variation, which has implications for developing an annual K-factor when there are months with missing data.

When the Site variable was removed as an independent variable, it was consistently the case that less variation in K-factors could be explained than was the case when the Site variable was included (65.2%). Table 4 shows that the socioeconomic variables (e.g., Population) coupled with facility variables (e.g., 24VC as an integer, Circ, and Functional class) could explain almost one-half (49.2%) of the variation in the K-factors. (These variables were defined in Table 3.)

Table 4. Variation in K-factors According to Analysis of Variance

Variables^a	% Variation Explained	Discussion
Site, Month, Year	65.2	<ul style="list-style-type: none"> • Most of the variation in K-factors can be explained by the Site variable.
Site, Month, Year, Population, Emp, Labor Force, Circ, Functional class	65.2	
Same as previous entry except without Site	22.9	<ul style="list-style-type: none"> • Socioeconomic variables can partially but not completely replace the Site variable.
Month, Year, Circ, Functional class, 24VC as an integer ^b	36.0	Impact of congestion (defined as 24VC as an integer) varies by facility in terms of explaining K-factor.
Same as previous entry except with a crossed variable added: Functional class crossed with 24VC ^b	49.2	

^aVariables are defined in Table 3. In Table 4, variables include main effects plus an interaction effect: Month crossed with Year.

^bFor Table 4, the 24-hour volume-to-capacity ratio (24VC) as an integer reduced the number of categories required for ANOVA. In later analyses where this was not necessary, this ratio was the full decimal value.

Table 4 also shows that consideration of interaction effects for two variables where the effect on K-factors of Variable 1 varies by the value of Variable 2 might help explain the variation in K-factors. For example, a high 24VC (Variable 1) might have a different effect on peak spreading for a freeway than for an urban arterial facility (Variable 2 being Functional class). A simple correlation analysis confirmed this to be the case in that there was a relatively strong negative correlation between K-factors and the 24VC for freeways (-0.70) but not for other facility types (0.31 or smaller in magnitude).

In summary, these ANOVA results indicated that there was a significant decrease in the K-factor over time and that the single best predictor of a K-factor was the Site variable because only with the Site variable included could a majority of the variation be explained as shown in Table 3.

Table 4 also suggests that when previous K-factors are not available, other variables may explain the variation in K-factors to a lesser degree; however, the variables must be carefully chosen. For example, although congestion matters, its importance is not the same for all facilities, as demonstrated in the last row of Table 4. Table 4 alone does not allow a definitive conclusion that these variables will successfully forecast the K-factor because it does not show the specific models. However, the fact that Table 4 shows that a relationship exists suggests that the development of models for forecasting K-factors may be productive.

Change in Annual K-factors Over Time

As stated in the “Methods” section, (1) a smaller dataset was culled from the entire dataset; (2) an annual adjusted K-factor was estimated for each site in the smaller dataset to account for variability that might have resulted from months with missing data; and (3) then a determination was made with regard to whether the annual adjusted K-factors in the smaller dataset had changed over time.

Completeness of the Entire Dataset

The 23,670 records of data in the entire dataset represented 50 stations. After defining a station that records counts in two directions as 2 sites, 80 sites were available. Given that Table 1 showed 3 days per month, 12 months per year, and an almost 14-year period (continuous count stations became operational no earlier than 1997 and the data collection period ended in December 2010), there was a possibility that there could be almost 500 data points per site. However, the actual number of data points was considerably less as not all sites started collecting data in 1997; some sites were terminated prior to 2010; and data were not available for all periods. Table 5 shows that for 37 of the 50 stations, or 62 of the 80 sites, some data were available for at least a 10-year period.

The data in Figure 3 suggested that for the period 1998-2010, the K-factor decreased overall—with an increase in 1999, a dip in 2000, and an increase in 2005-2010. However, such an initial comparison may have bias because it treats each data point equally; the fact that there were months with missing data for some stations and more stations with data in 2010 than in 1997 may affect the results. Further, examination of the trends in Figure 3 does not indicate whether the changes over time were statistically significant or whether the fluctuations were random.

Table 5. Overview of Data Available From Continuous Count Stations for the Entire Dataset

Period	No. of Stations for Which Data Were Available	No. of Sites^a for Which Data Were Available
1997- 2010	28	47
1998- 2010	7	11
2000- 2010	1	2
1997- 2007	1	2
5 to 9 years of data	8	11
Less than 5 years of data	5	7
Total	50	80

^aA *site* refers to a specific direction at a given station. For example, Station 60148 located on Route 1 reflects 2 sites: northbound Route 1 traffic and southbound Route 1 traffic.

Development of Annual K-factor

Rather than an examination of all K-factors without consideration of the completeness of the entire dataset, a more robust comparison of K-factors was performed by developing monthly expansion factors for each year that were then applied to obtain an estimated K-factor for a given year. Such an approach adjusts for months with missing data and may smooth outliers in the dataset; the latter was a concern to VDOT staff (Jamei, 2011). For example, for year 1998, one site (Station 90027 Direction 1) had no data for the month of June. Accordingly, monthly expansion factors for each of the other 11 months were used to estimate a yearly K-factor for that site. (The monthly expansion factors were based on all sites for each year.) Then, a “before” yearly K-factor from 1998, 1999, 2000, or 2001 was selected and an “after” yearly K-factor from 2008, 2009, or 2010 was selected. The selection criteria were that a given site have at most 3 months with missing data (fewer preferred) and that the time period between the before and after years be as close to 10 years as possible. Figure 3 shows considerable variation by year; the relevant equations were developed for each month and year.

For example, for 1998 only, the average K-factor across all sites was 0.105. However, the average K-factor for January 1998 (across all sites) was 0.109. Accordingly, the monthly expansion factor for January 1998, again for all sites, was $0.105/0.109 = 0.957$. Monthly expansion factors were calculated for the remaining months and then used in Equation 1 along with data from each site to estimate a yearly K-factor for the site. For example, at Station 90027 Direction 1, the January K-factor was 0.112. Thus, this value was multiplied by the corresponding monthly expansion factor of 0.957. A similar multiplication was repeated for each month of 1998 except June, for which no data were available for the site. The sum of these products was divided by 11 to obtain a K-factor for 1998 for Station 90027 Direction 1 as shown in Equation 1.

$$K_{1998} = \frac{K_{\text{Jan}1998} \text{MEF}_{\text{Jan}1998} + \dots + K_{\text{Dec}1998} \text{MEF}_{\text{Dec}1998}}{N} \quad [\text{Eq. 1}]$$

where

K_{1998} = average K-factor for 1998 *only* for Station 90027 Direction 1

$K_{\text{Jan}1998}$ = K-factor for January 1998 *only* for Station 90027 Direction 1; for the same site, K-factors for other months (e.g., $K_{\text{Feb}1998} \dots K_{\text{Dec}1998}$, etc.) were similarly defined

$\text{MEF}_{\text{Jan}1998}$ = monthly expansion factor for January 1998 for all sites; $\text{MEF}_{\text{Feb}1998} \dots \text{MEF}_{\text{Dec}1998}$ and so on were similarly defined

N = number of months for which a K-factor from Station 90027 Direction 1 was available; since June was the only month for which no values were available, $N = 11$ for this site.

A similar approach was used to determine an annual 24-hour volume-to-capacity ratio for each site.

Creation of Smaller Dataset for a Pairwise Comparison of K-Factors at Individual sites

To confirm whether the K-factor changed over time, a smaller dataset was culled from the entire dataset. As described previously, 80 sites were available in the entire dataset. From these, sites were sought for which the available data initially met two criteria:

1. *The site data must include at least two periods of data that are 10 years apart. A before period and an after period with at least 10 years between them were selected. Ten years was selected because it approached one of the horizons on the LD-104 form (VDOT, 2008) and was feasible within the time that continuous count data were available (1997-2010, depending on the site.) Each period was defined as a single year, such that a site had a “before year” and an “after year.”*

2. *The before year and the after year should each contain at least 10 months of data. The before year and the after year should each be based on at least 10 of the 12 months (which minimizes the reliance on monthly expansion factors from Eq. 1).*

Then, to increase the sample size, the criteria were relaxed as follows: 2 additional sites were selected that had 9 rather than 10 years of data between the before and after periods, and 10 sites were selected for which the before periods were based on at least 9 rather than at least 10 months of data. Table 6 summarizes the data completeness of this smaller dataset.

Table 6. Summary of Data Completeness of Smaller Dataset

No. of Months of Data Used to Calculate Annual K-factor	No. of Sites With Before Year of 1998, 1999, 2000, or 2001 ^a	No. of Sites With After Year of 2008, 2009, or 2010 ^b
9	10 sites	0 sites
10	13 sites	1 site
11	17 sites	1 site
12	12 sites	50 sites
Total	52 sites	52 sites

^aOnly 2 sites had a before year of 2001.

^bThe differences between the before and after periods were 9 years (2 sites), 0 years (45 sites), 11 years (4 sites), and 12 years (1 site).

Change in K-Factors at Sites in Smaller Dataset

The average K-factor decreased by 0.006, from 0.103 to 0.097, based on the 52 sites indicated in Table 6. This result was statistically significant ($p < 0.01$). Of the 52 sites, the K-factor decreased for 39 (three fourths) and increased for 13. This increase was detectable only with a site-by-site comparison where the paired version of Student's t -test is used. Without pairing, Student's t -test indicates that the increase was not significant ($p = 0.22$).

The 24-hour volume-to-capacity ratio increased at 40 of the 52 sites, with the increase being statistically significant ($p < 0.01$). Across all sites, the average 24-hour volume-to-capacity ratio increased from 7.3 to 8.0. As with the K-factor, this increase was detectable only with a site-by-site comparison through the use of a paired t -test; in a comparison of the average 24-hour volume-to-capacity ratio, although a nominal increase would have been detected, a t -test without pairing yielded an increase that was not significant ($p = 0.37$). As others (e.g., Margiotta et al., 1999; Simons, 2006) have used the 24-hour volume-to-capacity ratio as a surrogate for congestion, the implication of the increase in the 24-hour volume-to-capacity ratio is that congestion has increased.

Clearly there was variation in the K-factors across sites (when the period was held constant) and across periods (when the site was held constant). ANOVA showed that the period (before or after) and the site (from 1 to 52) were significant in terms of explaining the K-factor ($p < 0.01$). That said, Figure 5 indicates that the variation across sites was greater than the variation across periods. For example, the variance of the 52 site-by-site K-factor differences from the before to the after period (0.0001) was about one-fifth that of the 52 sites during the after period alone (0.0005). (Variance was computed as the sample variance. For example, for the after period, the mean K-factor was calculated, the difference between each of the 52 K-factors and

this mean value was calculated, these differences were squared, and then the sum of these differences was divided by 51.)

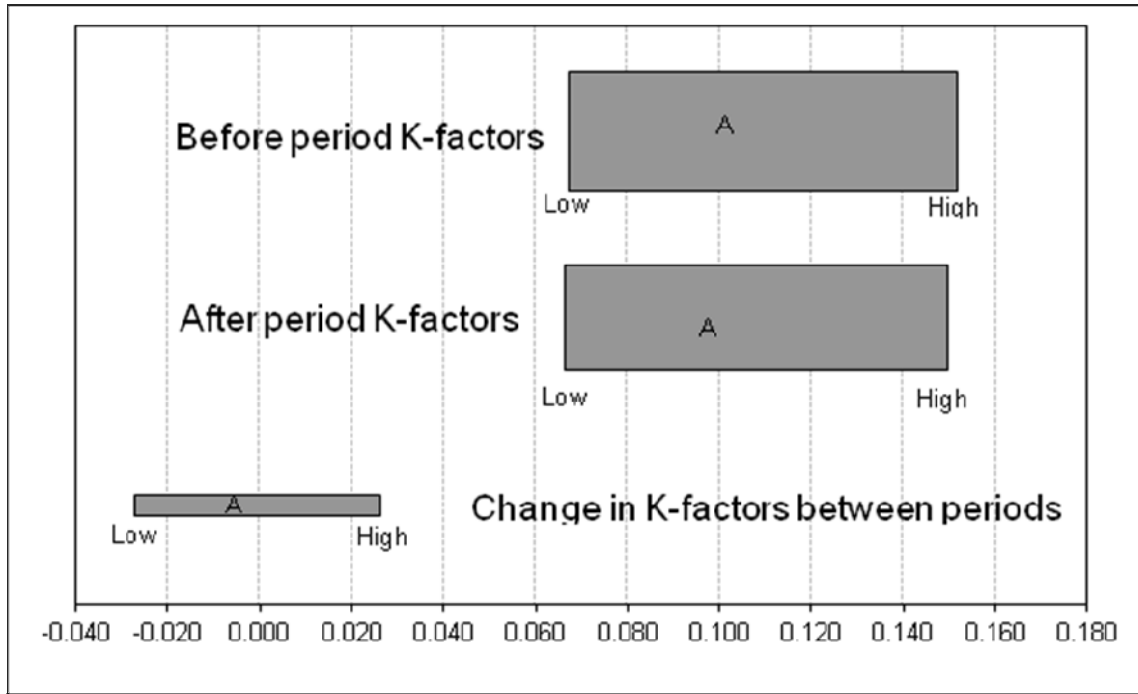


Figure 5. Annual K-factors at the 52 Sites. A = the average value. For example, during the after period, K-factors ranged from 0.066 to 0.150, with an average value of 0.097. The change in K-factors between periods at each site ranged from -0.027 (a decrease) to 0.026 (an increase). The average change was -0.006 (a decrease). The variance for each dataset was proportional to the rectangle’s height: the variance across after period K-factors was 5 times the variance across the change in K-factors between periods.

Models Developed to Forecast a K-factor

Longitudinal Model With Existing K-factor (Model 1)

As described in the “Methods” section, the data from the 52 sites, hence the smaller dataset, were used to develop a linear regression model to predict the K-factor as a function of several variables defined in Table 3: K_{old} ; Population, Labor force, Emp, 24VC; Functional class, and Circ. All variables were entered into the linear regression equation, and then variables that were not statistically significant were removed one at a time until only variables that were statistically significant remained.

The resultant model is given in Equation 2.

$$K_{new} = 0.019 + 0.758K_{old} + 0.022Emp - 0.011Two - 0.007Free - 0.012Ruralmulti [Eq. 2]$$

where

K_{new} = new K-factor in 2008, 2009, or 2010

K_{old} = old K-factor from 1998, 1999, 2000, or 2001

Emp = percentage change in number of filled jobs in a jurisdiction over the time period

Two = Rural two-lane road: 1 if the facility's functional class is a rural two-lane road, 0 otherwise

Free = Freeway or expressway: 1 if the facility's functional class is a freeway/expressway, 0 otherwise

Ruralmulti = Rural multilane road: 1 if the facility's functional class is a rural multilane road, 0 otherwise.

All terms represent variables that were statistically significant; $p < 0.01$ for all variables except the Free variable ($p = 0.024$). The model explained more than 88% of the total variation in the smaller dataset (the adjusted R^2 value was 0.881). (As noted in Table 3, K_{new} as shown in Eq. 2 is called a *prediction* because the value of K_{new} is already known and is being used to develop the model. Later, when the model is applied in a real-life situation in which a future K_{new} is unknown, the future K_{new} is called a *forecast*.)

It was not surprising that Model 1 explained a large portion of the variance given its inclusion of the K-factor from approximately one decade earlier. This was consistent with the ANOVA results that the Site variable, and hence the previous K-factor (K_{old}) (from 1998, 1999, 2000, or 2001), could explain much of the variation in the dataset's new K-factors (K_{new}) (from 2008, 2009, or 2010). It was also not surprising that the Population and Labor force variables explained no additional variation given the inclusion of the Emp variable; at the jurisdiction level, a high correlation among these variables would be expected. It was also not surprising that the Functional classification variable was statistically significant, given that peaking characteristics may vary by facility type. (The travel time impact of a volume increase for a two-lane road and a freeway differ [Committee on Highway Capacity and Quality of Service, 2010b], and generally urban facilities have lower K-factors than rural facilities [Committee on Highway Capacity and Quality of Service, 2010a]).

However, it was surprising to the author that two sets of variables—those denoting whether the road was circumferential or radial (Circ) ($p = 0.25$) and those denoting a measure of the 24-hour volume-to-capacity ratio (24VC) ($p = 0.30$)—were not statistically significant. The author expected radial routes carrying traffic toward the central business district (CBD) to have more severe traffic congestion than circumferential routes carrying traffic around such areas because the radial routes should have heavier traffic volumes in one direction (e.g., heavy flows to the CBD in the morning and heavy flows away from the CBD in the evening) whereas the circumferential routes might have more balanced flows during each time period. By extension, radial routes and circumferential routes would have different K-factors. Further, the author expected that as 24VC increased (meaning more congestion overall), the K-factor would decrease (hence the “spreading of the peak”). The Committee on Highway Capacity and Quality of Service (2010a) reported that for a given highway, the K-factor tends to decrease as the 24-hour volume increases (and since the capacity is constant, this means that the K-factor decreases

as the 24-hour volume-to-capacity ratio increases). Indeed the sign of the coefficient for 24VC was negative, which was consistent with this hypothesis.

Accordingly, several steps were performed to determine if the data could support including the 24-hour volume-to-capacity ratio as a significant variable in the model. For example, a plot of the residuals against the fitted values suggested the possibility of nonconstant variance, which would suggest it would be appropriate to transform the data such that instead of predicting the K-factor directly, the model predicted the square root of the K-factor (Hogg and Ledolter, 1992) or the K-factor raised to the power of 0.2777 (Montgomery, 2001a). However, variability remained in the dataset (Figure 6) and these two transformations did not materially affect the significance of the 24VC variable ($p = 0.22$ and $p = 0.19$, respectively, rather than $p = 0.30$ without the transformation). As these data were indeed proportions, the suggestion by Freund and Wilson (1997) of predicting $\arcsin[K^{0.5}]$ was followed. Although the resulting model had a slightly higher adjusted R^2 of 0.888, the scatter in the residuals did not change substantially. (The significance level of the independent variables also did not change: all variables continued to have significance levels of $p < 0.01$ except the Free variable, for which the previous significance level [$p = 0.024$] changed slightly [to $p = 0.017$].)

The results of Model 1 were confirmed with a simple correlation analysis. There was a strong correlation between the current and previous K-factor (0.91) but virtually no correlation between the change in the K-factor and the later 24-hour volume-to-capacity ratio (-0.08). Stratification by functional class strengthened this correlation for urban arterial sites (-0.52) and freeway sites (0.52) but not substantially for rural two-lane sites or rural multilane sites. Inclusion in the regression model of two products (Product 1 being the cross of the 24VC variable and the Free variable and Product 2 being the cross of 24VC variable and the Urbanart variable), however, showed that these two products were not statistically significant. This does not mean that the 24-hour volume-to-capacity ratio has no effect. It does mean, however, that it has no effect that explains variation that cannot be explained by the other variables already present in Model 1: Functional class, K_{old} , and Emp.

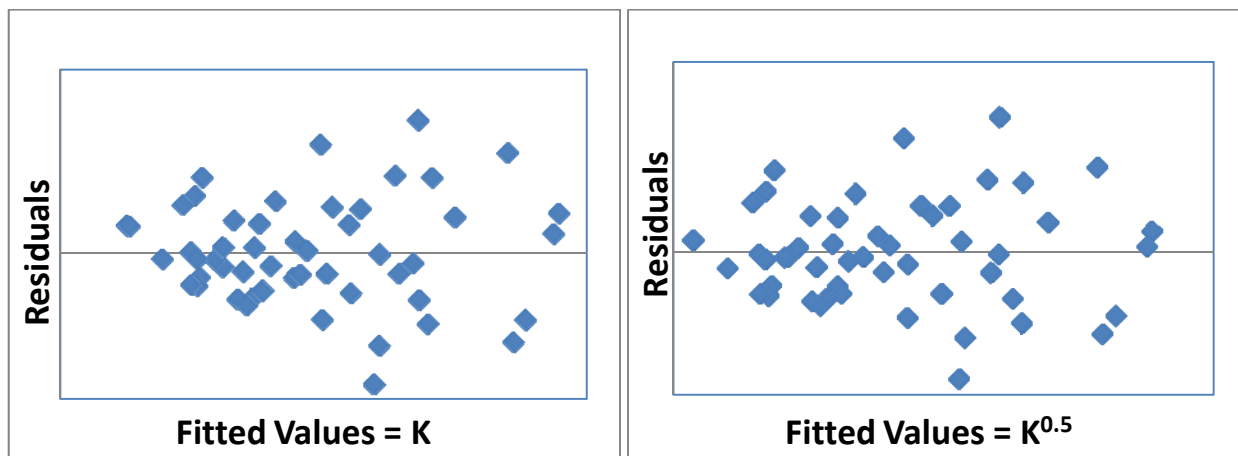


Figure 6. Plots of Residuals From Model 1. *Left*: with no transformation; *right*: with transformation. K = K-factor.

Longitudinal Model Without Existing K-factor (Model 2)

When a new facility is constructed or when historical data are lacking, a future year K-factor cannot be estimated from a previous K-factor. Thus, a second regression model was executed where the K_{old} variable (i.e., the previous K-factor) shown in Equation 2 was excluded. All other variables considered in the development of Equation 2 were initially included, and then variables that were not statistically significant were removed one at a time. The results of this manual approach were also confirmed by executing SPSS with the “forward regression” and “stepwise regression” options and observing that the variables selected by SPSS for inclusion in the model were the same variables selected through the manual approach. The manual approach, however, allows the analyst to consider more carefully the rationale for including or excluding each independent variable.

Equation 3 gives Model 2.

$$K_{new} = 0.080 + 0.059Emp + 0.010Circ - 0.002Freeway_{24VC} \quad [Eq. 3]$$

where

K_{new} = new K-factor in 2008, 2009, or 2010

Emp = percentage change in number of jobs filled over time period

Circ = circuity = 1 if site is circumferential, 0 if site is radial

Freeway_{24VC} = 24-hour volume divided by hourly capacity during after period for freeway sites only.

All terms represent variables that were statistically significant; $p < 0.01$ for all terms except Circ, for which $p = 0.015$. Without an existing K-factor, only 66%, rather than 88%, of the variation was explained. In Model 2, inclusion of the 24-hour volume-to-capacity ratio did add explanatory power for freeways. Model 2 shows an interaction effect: the effect of the 24-hour volume-to-capacity ratio on the K-factor varied according to whether the site is a freeway or not such that the 24-hour volume-to-capacity ratio is multiplied by 1 if the site is a freeway but is multiplied by 0 if the site is not a freeway. This interaction variable, Freeway_{24VC}, is statistically significant. As would be expected, a higher 24-hour volume-to-capacity ratio indicated a lower K-factor. Further, consistent with Equation 2, Emp was a significant variable.

Model 2 (Eq. 3) differs from Model 1 (Eq. 2) in two ways: one expected, and one unexpected. The expected difference was that Circ was a significant variable in Model 2; the model suggested that if the site was circumferential, the K-factor would be higher, all other things being equal, than if the site was radial. The unexpected difference was that in Model 2 Functional class was not a significant predictor of the K-factor unless the site was a freeway. That is, the data used to calibrate Model 2, which were the same data used to calibrate Model 1 except that K_{old} was excluded, suggested that there are no significant differences in the peaking characteristics for urban arterials, rural two-lane facilities, and rural multilane facilities if the

previous K-factor is unknown (and hence cannot be included in the model, in which case other variables with lesser explanatory power than the previous K-factor may then be included in the model) and if changes in Population and Circ have been accounted for. Because Functional class was statistically significant in Model 1 (Eq. 2), one possible explanation for why Functional class was not significant in Model 2 is that the large amount of scatter in the data when a previous K-factor is not available for inclusion in the model, renders Functional class unnecessary (in the sense that Functional class cannot explain any additional variation) except for freeways, which are included in both models. A more general explanation, applicable to any dataset, is that when one or more independent variables are removed from a model, the statistical relationship between the remaining independent variables and the dependent variable may change.

That said, the negative sign for the last variable (i.e., Freeway_{24VC}) in Equation 3 (i.e., Model 2) was as expected. The 24-hour volume-to-capacity ratio does materially affect the K-factor for freeway sites, as demonstrated by its inclusion in the equation ($p < 0.01$). Such a negative correlation was expected to the extent that higher levels of congestion (and hence higher ratios) should be associated with more peak spreading (and hence a lower K-factor). This was also consistent with the strong negative correlation (-0.70) between this ratio and the K-factor for freeway sites only when the entire dataset (based on all 80 sites) was analyzed.

LIMITATIONS OF STUDY ANALYSIS

A limitation of this analysis was that it was based only on counties in Northern Virginia. This limitation may be substantial if the six counties included in the analysis were already experiencing so much congestion during the before period that any congestion-based effects on peak spreading had already occurred. However, as the results of this study suggested, the large variability in the K-factors across sites dampened the overall effect of congestion and it may be the case that similar results would be obtained if the study were repeated in other areas of Virginia. As discussed in the section titled “Suggestions for Further Research,” the collection of data from other regions in Virginia may help determine the extent to which these models are unique to Northern Virginia or transferrable to other locations.

CONCLUSIONS

Except where otherwise noted, the conclusions refer to the smaller dataset collected from the 52 sites in the six Northern Virginia counties of Arlington, Fairfax, Fauquier, Loudoun, Prince William, and Stafford.

- *Limited peak spreading occurred over roughly a decade when congestion increased, suggesting a relationship between congestion and peak spreading.* This conclusion is supported by three findings. First, there was a modest statistically significant decrease in the K-factor over roughly one decade. The average annual K-factor adjusted for months with missing data decreased by an average of 0.006 ($p < 0.01$),

from 0.103 to 0.097, from a before period (1998, 1999, 2000, or 2001) to an after period (2008, 2009, or 2010). This finding is consistent with that of an examination of the K-factors for all 80 sites in the entire dataset (not adjusted for periods with missing data) in which the marginal annual means changed significantly. Second, there was a statistically significant increase in the 24-hour volume-to-capacity ratio. This ratio increased from an average value of 7.3 to 8.0 ($p < 0.01$). Third, the 24-hour volume-to-capacity ratio had a significant impact on the K-factor, but only under certain conditions as shown in Model 2.

- *It is possible to explain most (88.1%) of the variation in K-factors being predicted approximately 10 years into the future if the existing K-factor is known.* Model 1 (Eq. 2) predicts the future K-factor based on the K-factor for the site 10 years prior; the percentage increase in the jurisdiction's employment; and the functional class of the facility. Such a model is appropriate for forecasting K-factors for an existing site for which the existing K-factor is known.
- *It is possible to explain less (65.9%) of the variation in K-factors being predicted approximately 10 years into the future if an existing K-factor is not available.* Model 2 (Eq. 3) predicts the future K-factor based on the percentage change in the jurisdiction's employment; the site's circuitry; and if the site is located on a freeway, its 24-hour volume-to-capacity ratio. Such a model is appropriate for forecasting K-factors when present K-factors are not available, as is the case with new facilities to be constructed.
- *In the forecasting of K-factors, the characteristics of the specific site matter more than the passage of time and/or the increase in congestion.* In this study, there was greater variation in K-factors by site than by time and/or congestion. The variance in annual K-factors for the same time period (whether the before or after period) was approximately 5 times the variance in the change in annual K-factors between the time periods. In fact, had the previous K-factor been the only independent variable in Equation 2 (Model 1), the model would still have been able to explain most (81.9%) of the variance found with the equation in its current form (i.e., 88.1%, as described previously).
- *Congestion, defined here as the 24-hour volume-to-capacity ratio, does affect K-factors even though the effect is evident only after other variables are controlled for.* The 24-hour volume-to-capacity ratio does materially affect the K-factor for freeway sites, as demonstrated by its inclusion in Equation 3 (i.e., Model 2) ($p < 0.01$). As previously discussed with regard to the entire 80-site dataset, this was not surprising given the strong negative correlation (-0.70) between this ratio and the K-factor for freeway sites only.

RECOMMENDATIONS

1. *Planning or engineering staff in VDOT's NOVA District should consider using Model 1 (Eq. 2) or Model 2 (Eq. 3) when future year K-factors are needed and information more detailed than jurisdiction-level forecasts is not available.* Because the models are not explicitly dependent on a single year but rather reflect changes in employment, they are potentially useful for 10-year forecasts. Therefore, the following two factors should be included in any consideration.
 - Site-specific studies may provide better details than the generalized approach provided by the two models, and site-specific knowledge will usually be preferable.
 - As is the case with any model, there is no guarantee that the relationship between the dependent variables (the K-factors) and the independent variables (such as employment), will remain stable. For example, Equation 2 (Model 1) suggests the K-factor will decrease over time if employment remains constant; however, this relationship may or may not hold in the future.
2. *If Model 1 (Eq. 2) or Model 2 (Eq. 3) yields a large value of K, the value should be checked to ensure that capacity is not exceeded by the resultant hourly volume.* For example, if Equation 3 were applied for a jurisdiction that expected a 100% increase in employment for a circumferential non-freeway route, K_{new} would be 0.149. If the facility had a capacity of 2,800 vehicles per hour per direction and a 24-hour volume of 20,000, multiplication of the 24-hour volume (20,000) by K_{new} (0.149) would yield 2,980 vehicles per hour—which is unlikely because capacity (2,800) would be exceeded. In this instance, the analyst, without any additional information, should expect that the K-factor will not likely rise to the extent forecast by the equation but rather would have a maximal value of $(2,800/20,000) = 0.140$. In addition, the 100% increase in employment would exceed the scope of the smaller dataset used to calibrate the models, and thus the use of the models in such situations is limited.

BENEFITS AND IMPLEMENTATION PROSPECTS

The benefit of using one of the two models (i.e., Eq. 2 or Eq. 3) to forecast a change in the K-factor is that the models provide an empirically based alternative to assuming the K-factor will remain constant. The primary benefit is in the short term: VDOT may use such information to address stakeholders' concerns regarding whether the K-factor will change and if so, by how much.

Implementation—defined as execution of Equation 2 (Model 1) or Equation 3 (Model 2) as appropriate—would require approximately 15 minutes for a single case. For example, if a new K-factor for year 2020 for an existing site on Route 1 in Alexandria (an urban arterial) was sought where the year 2010 K-factor was 0.10 and the 2020 forecast for employment in Alexandria suggested a 25% increase relative to 2010, Model 1 (Eq. 2) would be used to compute the new K-factor as follows:

$$K_{\text{new}} = 0.019 + 0.758K_{\text{old}} + 0.022\text{Emp} - 0.011\text{Two} - 0.007\text{Free} - 0.012\text{Ruralmulti}$$

$$K_{\text{old}} = 0.10 \text{ (existing K-factor)}$$

$$\text{Emp} = 0.25 \text{ (reflects a forecast 25\% increase in employment)}$$

$$\text{Two} = 0 \text{ (since facility is not a rural two-lane road)}$$

$$\text{Free} = 0 \text{ (since facility is not a freeway)}$$

$$\text{Ruralmulti} = 0 \text{ (since facility is not a rural multilane road)}$$

$$K_{\text{new}} = 0.019 + 0.758(0.10) + 0.022(0.25) - 0.011(0) - 0.007(0) - 0.012(0)$$

$$K_{\text{new}} = 0.1003, \text{ i.e., virtually no change since this is reported as } 0.10.$$

However, if a K-factor had been sought for a new site along Route 1 in Alexandria, Equation 3 (Model 2) would be applied since Route 1 is a radial route:

$$K_{\text{new}} = 0.080 + 0.059\text{Emp} + 0.010\text{Circ} - 0.002\text{Freeway}_{24\text{VC}}$$

$$K_{\text{new}} = 0.080 + 0.059(0.25) + 0.010(0) - 0.002(0)$$

$$K_{\text{new}} = 0.09475, \text{ which would be reported as } 0.09 \text{ or } 0.095.$$

SUGGESTIONS FOR FURTHER RESEARCH

In concert with Recommendations 1 and 2, a record of how K-factors change over time in other portions of the Commonwealth may be of value, as this study could be repeated in locations where congestion levels differ from those found in the Northern Virginia jurisdictions that comprised this study. The reason is that the study analysis was based only on counties in Northern Virginia and on only one definition of peak spreading. Regarding the conclusion that congestion does affect peak spreading but only to a limited extent, three possible explanations should be considered when these models are applied to other areas of Virginia or when data from those areas are used to develop new models:

1. *The large variability in the K-factors dampens the effect of congestion.* If this is the case, one would expect the models from this study to be applicable in other areas.
2. *The six counties in Northern Virginia were already undergoing so much congestion during the before period of 1998-2000 that any congestion-based effects on peak spreading had already occurred before this study was undertaken.* In this is the case, comparison of these results with longitudinal models developed for other not-so-congested regions would be informative because one could examine whether models

in areas that were already congested differed from models developed for areas that were becoming congested.

3. *Additional peak spreading is indeed occurring in Northern Virginia, but it is manifesting in ways besides a change in the K-factor.* These other ways are given in the literature (see Table 2 and Appendix B) and include a lengthening of the peak morning rush hour; a change in the proportion of peak hour volume during the peak 3- or 4-hour period; and a shift to another mode. Although this study focused on the K-factor because such a focus was the core VDOT objective, further research could look at other indicators of peak spreading through (1) the collection of additional data such as use of other modes and (2) an examination of how these indicators, such as the proportion of peak hour volume during the peak 3-hour period, have changed over time.

With additional data from more sites in other regions one could explicitly account for different congestion levels. It is also possible that nonlinear relationships, rather than the linear relationships that comprised Models 1 and 2, could be examined.

ACKNOWLEDGMENTS

The author thanks the following individuals for their assistance; without them, this research could not have been conducted. Bahram Jamei of VDOT (Chair); Randy Dittberner, Ralph Jones, Tom Schinkel, and Eric Stringfield of VDOT; and Lance Dougald and Amy O'Leary of the Virginia Center for Transportation Innovation and Research served on the technical review panel; Ed Azimi, Bill Mann, Bob McDonald, and JoAnne Sorenson of VDOT provided comments at the inception of the research; Tina Tang of the University of Virginia assisted with obtaining link capacities; and Linda Evans of the Virginia Center for Transportation Innovation and Research edited the report

REFERENCES

- Allen, W.G., and Schultz, G.W. Congestion-Based Peak Spreading Model. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1556. Transportation Research Board of the National Academies, Washington, DC, 1996, pp. 8-15.
- Anderson, R.S., and Donnelly, R.M. Modeling of Peak Hour Spreading with a Disaggregate Tour-Based Model. In *Innovations in Travel Demand Modeling: Summary of a Conference. Volume 2: Papers*. Transportation Research Board of the National Academies, Washington, DC, 2008, pp. 161-164.

- Barnes, J. *Peak Spreading Analysis: Review of Relevant Issues and Synthesis of Current Practice, Phase I*. Washington State Transportation Center, Seattle, 1998.
<http://ntl.bts.gov/lib/21000/21200/21218/PB99104564.pdf>. Accessed July 15, 2010.
- Cambridge Systematics, Inc. *Time-of-Day Modeling Procedures: State-of-the-Practice, State-of-the-Art*. Federal Highway Administration, Washington, DC, 1997.
<http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/time-day/>. Accessed October 1, 2010.
- Committee on Highway Capacity and Quality of Service. *Highway Capacity Manual, Volume I: Concepts*. Transportation Research Board of the National Academies, Washington, DC, 2010a.
- Committee on Highway Capacity and Quality of Service. *Highway Capacity Manual, Volume II: Uninterrupted Flow*. Transportation Research Board of the National Academies, Washington, DC, 2010b.
- Cottrell, W.D. Estimating the Probability of Freeway Congestion Recurrence. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1634. Transportation Research Board of the National Academies, Washington, DC, 1998, pp. 19-27.
- Dittberner, R. Email to John Miller, July 21, 2011.
- Freund, R.J., and Wilson, W.J. *Statistical Methods: Revised Edition*. Academic Press, Inc., San Diego, CA, 1997.
- Habib, K.M.N., Day, N., and Miller, E.J. An Investigation of Commuting Trip Timing and Mode Choice in the Greater Toronto Area: Application of a Joint Discrete-Continuous Model. *Transportation Research Part A*, Vol. 43, No. 7, 2009, pp. 639-653.
- Hogg, R.V., and Ledolter, J. *Applied Statistics for Engineers and Physical Scientists*, 2nd ed. Macmillan Publishing Company, New York, 1992.
- Holyoak, N., and Taylor, M.A.P. Modelling Trip Timing Behaviour and the Influence of Peak Spreading. In *Urban Transport XII. Urban Transport and the Environment in the 21st Century*. WIT Press, Southampton, UK, July 2006, pp. 205-213.
- Institute of Transportation Engineers. *Transportation Impact Analyses for Site Development: An ITE Proposed Recommended Practice*. Washington, DC, 2006.
- Ivan, J.N., and Allaire, S.A. Regional and Area-Type Modeling of Peak Spreading on Connecticut Freeways. *ASCE Journal of Transportation Engineering*, May/June 2001, pp. 223-229.
- Jamei, B. Email to John Miller, April 29, 2011.

- Jin, X., and Chiao, K.-A. A Synthesis of Time-of-Day Modeling Research. In *TRB 87th Annual Meeting Compendium of Papers DVD*. Transportation Research Board of the National Academies, Washington, DC, 2008.
- Karl, C.A., and Gaffney, J. Improving Traffic Network Performance in Australia. In *Australian Road Research Board 23rd Conference CD-ROM*. Australian Road Research Board, Adelaide, 2008.
- Karlström, A., and Franklin J.P. Behavioral Adjustments and Equity Effects of Congestion Pricing: Analysis of Morning Commutes During the Stockholm Trial. *Transportation Research Part A*, Vol. 43, No. 3, 2009, pp. 283-296.
- Liu, Z., and Sharma, S. Predicting Directional Design Hourly Volume from Statutory Holiday Traffic. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1968. Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 30-39.
- Liu, H.X., Jabari, S.E., and Wenteng, M. Peak Spreading Methodology for Future Scenario Evaluations Using Microsimulation. In *TRB 86th Annual Meeting Compendium of Papers CD-ROM*. Transportation Research Board of the National Academies, Washington, DC, 2007.
- Margiotta, R., Cohen, H., and DeCorla-Souza, P. Speed and Delay Prediction Models for Planning Applications. In *Sixth National Conference on Transportation Planning for Small and Medium Sized Communities, September 16-18, 1998, Spokane, Washington*. Transportation Research Board of the National Academies, Washington, DC, 1999. <http://ntl.bts.gov/lib/1000/1100/1195/00780089.pdf>. Accessed November 15, 2010.
- Martin, W.A., and McGuckin, N.A. *Travel Estimation Techniques for Urban Planning*. NCHRP Report No. 365. Transportation Research Board of the National Academies, Washington, DC, 1998. <http://ntl.bts.gov/lib/21000/21500/21563/PB99126724.pdf>. Accessed December 3, 2010.
- Martland, C.D. The Journey to Work: 25 Years on the Jamaicaway. *Journal of the Transportation Research Forum*, Vol. 44, No. 3, 2005, pp. 59-76. http://www.trforum.org/forum/downloads/2005_Jamaicaway_paper.pdf. Accessed September 30, 2010.
- Miller, J.S. Email to Bahram Jamei, Randy Dittberner, Thomas Schinkel, Ralph Jones, Eric Stringfield, and Lance Dougald, April 20, 2011.
- Montgomery, D.C. *Design and Analysis of Experiments*, 5th ed. John Wiley & Sons, Inc., New York, 2001b.
- Montgomery, D.C. *Introduction to Statistical Quality Control*, 4th ed. John Wiley & Sons, Inc., New York, 2001a.

- National Climatic Data Center. Storm Events. <http://www4.ncdc.noaa.gov/cgi-win/wwcgi.dll?wwevent~storms>. Accessed September 28, 2010.
- Ozbay, K., Ozlem, Y.-T., and Holguín-Veras, J. Evaluation of Impacts of Time-of-Day Pricing Initiative on Car and Truck Traffic: Port Authority of New York and New Jersey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1960. Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 48-56.
- Purvis, C. Peak Spreading Models: Promises and Limitations. In *Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods*. Transportation Research Board of the National Academies, Washington, DC, 2002. <http://docs.trb.org/00939750.pdf>. Accessed September 30, 2010.
- Replogle, M. Computer Transportation Models for Land Use Regulation and Master Planning in Montgomery County, Maryland. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1262. Transportation Research Board of the National Academies, Washington, DC, 1990, pp. 91-100.
- Sall, E., Bent, E., Charlton, B., Koehler, J., and Erhardt, G. Evaluating Regional Pricing Strategies in San Francisco: Application of the SFCTA Activity-Based Regional Pricing Model. In *TRB 89th Annual Meeting Compendium of Papers DVD*. Transportation Research Board of the National Academies, Washington, DC, 2010.
- Schinkel, T.O. Email to John Miller, February 24, 2011a.
- Schinkel, T.O. Email to John Miller, July 22, 2011b.
- Simons, C. I-285 Matrix Variegator: Practical Method for Developing Trip Tables for Simulation Modeling from Travel Demand Modeling Inputs. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1981. Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 18-23.
- Sinha, A.K., and Thakuriah, P.V. Relationship Between Occupational, Industrial, and Sociodemographic Characteristics and Job Start Times: Evidence from Current Population Survey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1894. Transportation Research Board of the National Academies, Washington, DC, 2004, pp. 67-74.
- Time and Date AS. Calendar. Stavanger, Norway, 2010. <http://www.timeanddate.com>. Accessed September 28, 2010.
- Transportation Planning Research Advisory Committee. Revised Minutes: November 10, 2009. Virginia Transportation Research Council, Charlottesville, 2009. <http://www.vtrc.net/tprac/pdf/MinutesFall2009Revised.pdf>. Accessed November 2, 2011.

- Virginia Department of Transportation. VDOT Traffic Engineering, undated.
<http://tedweb/tms/jsp/>. Accessed January 15, 2012.
- Virginia Department of Transportation. Department of Transportation, Location and Design, Request for Traffic Data: Form LD-104. Richmond, 2008.
<http://vdotforms.vdot.virginia.gov/SearchResults.aspx?filename=LD-104.doc>. Accessed September 15, 2010.
- Virginia Employment Commission. I Need Labor Market Information. Richmond, undated.
<http://www.vec.virginia.gov/vecportal/index.cfm>. Accessed April 4, 2011.
- Virginia Employment Commission. Labor Market Information. Richmond, 2011.
<http://www.vec.virginia.gov/vecportal/index.cfm>. Accessed July 29, 2011.
- Weldon Cooper Center for Public Service. Demographics and Workforce: Annual Population Estimates for Previous Decades, 2003.
http://www.coopercenter.org/sites/default/files/demographics/intercensal_9000.xls. Accessed April 4, 2011.
- Weldon Cooper Center for Public Service. Demographics and Workforce: 2000-2009 Intercensal Estimates for Localities, 2010.
http://www.coopercenter.org/sites/default/files/node/13/IntercensalEstimatesVA_2000_2009_CooperCenter.xls. Accessed April 4, 2011.
- Weldon Cooper Center for Public Service. July 1, 2010, Population Estimates for Virginia and Its Counties and Cities, April 12, 2011.
http://www.coopercenter.org/sites/default/files/node/13/PopulationEstimates_July2010_VA_CooperCenter.xls. Accessed July 29, 2011.
- Wolff, C., and Vilain, P. Evaluating Congestion Pricing Impacts Under Peak Spreading. In *Proceedings of the 48th Annual Forum of the Transportation Research Forum*, 2007.
http://www.trforum.org/forum/downloads/2007_8A_PeakSpread_paper.pdf. Accessed November 1, 2010.
- Yang, S., Lu, C., Zhao, F., Reel, R., and O'Hara, J.D. Estimation for Seasonal Factors of Similarity-Based Traffic for Urban Roads in Florida. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2121. Transportation Research Board of the National Academies, Washington, DC, 2009, pp. 74-80.

APPENDIX A

PROCEDURE FOR OBTAINING ROADWAY SEGMENT CAPACITIES AND VOLUMES

Introduction

This appendix summarizes how capacities were obtained for each roadway section where a continuous count station is situated. The steps to obtain a capacity included locating the roadway section, calculating a one-way capacity, and determining whether the capacity changed at some point during the period studied (1997-2010). The appendix also provides an example query for extracting volume data.

Steps for Obtaining Capacities for Each Roadway Section

Locating the Roadway Section

The capacities for 50 roadway sections, each monitored by a continuous count station, were obtained using the TMS and SPS databases. The TMS is a VDOT database maintained by VDOT's Traffic Engineering Division that may be accessed within VDOT (VDOT, undated). The SPS database is an application operated by VDOT's Transportation Mobility Planning Division and is accessed via a customized Microsoft Access user interface.

A *link* is a longitudinal section of roadway between two points; for example, a section of interstate between two adjacent exits might be a link. For each link the TMS database contains a five- or six-digit Site ID (e.g., 60148). If the number is only five digits, it is necessary to add a "0" as the first digit to make the number a six-digit number before the number is entered in the TMS database website (VDOT, undated). After the Site ID is entered, the TMS database website displays location information (the route prefix and suffix; the endpoints of the section [as "from" and "to"]; and the jurisdiction). The TMS database also gives the volume and the number of lanes for the roadway section.

This information from the TMS database was used to find the link in the SPS database. The first step for finding the link is to identify the jurisdiction in which the link is located. The second step is to find the correct route. The third step is to use the "from" and "to" designations to find the correct roadway section. In some cases, multiple roadway sections in the SPS database might correspond to a single roadway section as defined in the TMS database. Thus, Google maps and road maps were used to determine the appropriate roadway section in SPS.

Calculating a One-Way Capacity

From the SPS database, the number of lanes in one direction was calculated by dividing the number of lanes given in two directions by 2. Under the "Performance" tab in the SPS

database, the “Capacity Analysis Type” appears in the upper left corner. The number of lanes and the Capacity Analysis Type per link were recorded in a spreadsheet.

The one-way capacity of each link was then calculated using the information given under the Performance tab of the SPS database. Generally, the SPS database provides a capacity in one of three ways, and this capacity can be manipulated to give a one-way capacity as follows:

- *For links for which the SPS database gives a capacity in units of passenger cars per hour per lane (pc/hr/ln), multiply the SPS capacity by the number of lanes in one direction.* For example, for Link 060166 (a section of I-95 in Stafford County), the SPS database gives a capacity of 2,400 pc/hr/ln and there are three lanes in one direction. Thus, the capacity was recorded in the aforementioned spreadsheet as $(2,400 \text{ pc/hr/ln})(3 \text{ ln}) = 7,200 \text{ pc/hr}$. If the TMS database link reflected more than one segment from the SPS database, it was necessary to make sure the capacity was consistent throughout the traffic link.
- *For links for which the SPS database gives only a “Through Capacity,” the units are already in passenger cars per hour so no further computations are needed.* For example, for Link 060148, a four-lane section (two lanes in each direction) of Route 1 in Stafford County, the SPS database gives a Through Capacity of 1,539 pc/hr. (The relatively low capacity reflects the presence of signals along the route.)
- *For rural two-lane traffic links, use the following steps.* First, click the button for “HCS Inputs” at the bottom right of the Performance tab in the SPS database and obtain the “Average Annual Daily Traffic” (AADT) and “Vehicles per Hour” (VPH). Second, execute the Highway Capacity Software (HCS+) to find the two-way flow rate. Third, divide 3,200 [which is the ideal capacity of a two-lane section] by the ratio of the two-way flow rate to the VPH. Dividing that value by 2 gives the one-way capacity in one direction in passenger cars per hour.

For example, for Link 090021 (Route 15 in Loudoun County between the bypass north of Leesburg and Route 662), the capacity was determined as follows:

- Clicking the HCS Inputs tab for this two-lane road (one lane each direction) gave the AADT as 21,692 and the VPH as 2,169.
- Executing HCS+ revealed the two-way flow rate to be 2,562 pc/hr.
- Dividing the two-way flow rate by the VPH revealed the ratio of the flow rate (2,562 pc/hr) to the VPH (2,169 VPH) to be 1.181.
- Dividing the ideal capacity (3,200 pc/hr) by this ratio (1.181) yielded the two-way capacity for the link (2,709.6 pc/hr).
- Dividing this two-way capacity by 2 ($2,709.6/2 = 1,354.7$) yielded the one-way capacity for the link as 1,355 pc/hr when rounded to the nearest integer.

Determining Whether Capacity Changed From 1997 Through 2010

After the capacities for each link were calculated, the TMS database was used to determine if there were updates to the number of lanes of each traffic link during the period reflected by the data. This period began during 1997 and ended December 31, 2010. This was done by clicking “View Raw Data” and then inserting a date in 1997 and then a recent date. Data were not always available for dates that were chosen. If that was the case, the “Catalog of Available Data” was used to determine which days had data. If the number of lanes had changed between the dates, it was necessary to determine approximately when the number of lanes changed and record it in the spreadsheet. Six such links were noted. Although the change in two of the links was straightforward, the change for the other four links was not. The views of expert staff (Dittberner, 2011; Schinkel, 2011b) were sought to determine the number of lanes at the four sites and when, or if, the number of lanes had changed. These conversations revealed two links where capacity had changed: Route 123 (Link 090278), which changed from one to two lanes in 2002, and I-495 N (Link 090138), which changed from four to five lanes in 2005. Ultimately, only one of the changes affected the data: for I-495, data were available prior to October 25, 2001 (when the link had four lanes), and after October 18, 2005 (when the link had five lanes). For Route 123, no data were available prior to November 19, 2002, by which time the link already had two lanes in each direction.

Example Query for Extracting Volume Data

This query was executed after the TMS data for year 1997 had been placed into a single Microsoft Access file. The data in the file included the Site ID for the link, the direction, the date and hour of the day, and the volumes for 15 classes of vehicles such as auto, motorcycle, and heavy vehicle. For each hour of each day shown in Table 1, the query summed the total volume across all 15 classes and computed the ratio of the hourly volume to the peak hour volume of that day. The query provides the following data elements on a daily basis for each link and each direction: the peak hour volume, the 24-hour volume, and the K-factor (i.e., the ratio of the peak hour volume to the 24-hour volume). The queries for 1998 through 2010 were similar to the query shown here except the dates shown in the query were modified to reflect those in Table 1.

```
SELECT
Sum(Data1997!Class1+Data1997!Class2+Data1997!Class3+Data1997!Class4+Data1997!Class5
+Data1997!Class6+Data1997!Class7+Data1997!Class8+Data1997!Class9+Data1997!Class10+
Data1997!Class11+Data1997!Class12+Data1997!Class13+Data1997!Class14+Data1997!Class1
5) AS Total, DateValue(Data1997!StartDate) AS DateOnly,
Max([Class1]+[Class2]+[Class3]+[Class4]+[Class5]+[Class6]+[Class7]+[Class8]+[Class9]+[Cla
ss10]+[Class11]+[Class12]+[Class13]+[Class14]+[Class15]) AS MaxHrVol,
[MaxHrVol]/[Total] AS PHF, Data1997.Direction, Data1997.SiteID INTO Results1997
FROM Data1997
GROUP BY DateValue(Data1997!StartDate), [MaxHrVol]/[Total], Data1997.Direction,
Data1997.SiteID
```

HAVING (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("01/07/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("01/09/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("02/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("02/13/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("03/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("03/13/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("04/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("04/17/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("05/13/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("05/15/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("06/24/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("06/26/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("07/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("07/17/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("08/12/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("08/14/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("09/16/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("09/18/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("10/21/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("10/23/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("11/18/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("11/20/1997"))) OR
 (((Data1997.Direction)=1) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("12/09/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("12/11/1997"))) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("01/07/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("01/09/1997"))) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("02/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("02/13/1997"))) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("03/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("03/13/1997"))) OR
 (((Data1997.Direction)=3) AND

((DateValue([Data1997]![StartDate]))>=DateValue("04/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("04/17/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("05/13/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("05/15/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("06/24/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("06/26/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("07/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("07/17/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("08/12/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("08/14/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("09/16/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("09/18/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("10/21/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("10/23/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("11/18/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("11/20/1997")) OR
 (((Data1997.Direction)=3) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("12/09/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("12/11/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("01/07/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("01/09/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("02/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("02/13/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("03/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("03/13/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("04/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("04/17/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("05/13/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("05/15/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("06/24/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("06/26/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("07/15/1997") And

(DateValue([Data1997]![StartDate]))<=DateValue("07/17/1997")) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("08/12/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("08/14/1997"))) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("09/16/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("09/18/1997"))) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("10/21/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("10/23/1997"))) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("11/18/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("11/20/1997"))) OR
 (((Data1997.Direction)=2) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("12/09/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("12/11/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("01/07/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("01/09/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("02/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("02/13/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("03/11/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("03/13/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("04/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("04/17/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("05/13/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("05/15/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("06/24/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("06/26/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("07/15/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("07/17/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("08/12/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("08/14/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("09/16/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("09/18/1997"))) OR
 (((Data1997.Direction)=4) AND
 ((DateValue([Data1997]![StartDate]))>=DateValue("10/21/1997") And
 (DateValue([Data1997]![StartDate]))<=DateValue("10/23/1997"))) OR

```
((Data1997.Direction)=4) AND
((DateValue([Data1997]![StartDate]))>=DateValue("11/18/1997") And
(DateValue([Data1997]![StartDate]))<=DateValue("11/20/1997")) OR
((Data1997.Direction)=4) AND
((DateValue([Data1997]![StartDate]))>=DateValue("12/09/1997") And
(DateValue([Data1997]![StartDate]))<=DateValue("12/11/1997"))
ORDER BY Data1997.SiteID;
```

References

Dittberner, R. Email to John Miller, July 21, 2011.

Schinkel, T.O. Email to John Miller, July 22, 2011b.

APPENDIX B

LITERATURE REVIEW

Introduction

The topic of land development and site impact analysis, for which the Virginia-specific LD-104 form is completed (VDOT, 2008), has received attention in the literature. For example, the Institute of Transportation Engineers (ITE) (2006) stated: “Care should be taken to consider potential changes in peaking characteristics over time, particularly in growing areas.” ITE (2006) further noted that a change of 100 vehicles per hour “can change the level of service or appreciably increase the volume-to-capacity ratio of an intersection approach.” Despite the mention of the 100 vehicles per hour, ITE (2006) further noted that more urban areas may require an impact analysis for developments generating 30 to 100 vehicles during the peak hour, suggesting that relatively small shifts in peak spreading may have a substantive impact.

The literature also suggests that changes in peak spreading affect policy questions related to air quality analysis, evaluation of possible travel demand management strategies, and estimation of required infrastructure investments (Barnes, 1998). Barnes also suggested that generally morning peak spreading is easier to model explicitly because work trips comprise a greater proportion of morning peak trips than of evening peak trips; this same point was made by Replogle (1990).

A review of Cambridge Systematics, Inc. (1997) indicated that one’s understanding of how peak period travel may change will affect one’s evaluation of transportation policy options that are affected by peak spreading. For example, with regard to a policy option of capacity improvements and operational initiatives to reduce peak period emissions, the extent to which such investments will induce additional peak hour demand will affect the estimates of emissions. Cambridge Systematics, Inc. (1997) discussed the sensitivity of motorists to congestion in the following excerpt:

Peaking and time of travel are critical determinants of level of service, traffic congestion, and concentrations of emissions. For example, the success of strategies to reduce the intensity of highway congestion depends critically on a low elasticity of trip departure time with respect to trip duration, yet common experience on congested facilities suggests otherwise, i.e., peaks narrow but do not decline in intensity very much.

Yet the literature suggests two challenges in particular that affect the ability to forecast peak spreading accurately. First, when peaking characteristics change because of increased congestion, the change can be rapid; Karl and Gaffney (2008) reported that the length of peak periods in Victoria (Australia) increased from 4.5 to 6.0 hours over just a 4-year period (2001 to 2005). Second, there may be insufficient data to forecast peak spreading reliably: although agreeing that available data suggest peak spreading is expected to occur as congestion increases, Anderson and Donnelly (2008) highlighted the difficulty of calibrating an activity-based regional travel demand model with respect to the time of day travel would occur, owing in part to an insufficient number of hourly counts being collected.

Overview of Approaches to Forecast Peak Spreading

Jin and Chiao (2008) developed a taxonomy of approaches for forecasting peak spreading. Three main categories of interest were regional hourly proportion models; link-specific hourly proportion models; and choice models, which are based on the concept of individual utility maximization.

Regional Hourly Proportion Models

An example of a regional hourly proportion model is the use of factors from trip surveys or other sources to assign a proportion of traffic to the peak hour that is not link specific but is based on trip purpose and applies to an entire region. For example, Table B1 lists some of the factors described by Martin and McGuckin (1998) suggesting that in the absence of additional information, one may assume that 14.06% of trips from home to work occur between 7 A.M. and 8 A.M. in metropolitan areas of more than 1 million people. As noted by others (e.g., Liu et al., 2007; Purvis, 2002), however, such factors cannot account for peak spreading resulting from increased congestion during the peak hour.

Table B1. Percentage of Home-Based Work Vehicle Trips for Large Urban Areas^a

Hour Begins	%	Hour Begins	%	Hour Begins	%	Hour Begins	%
Midnight	0.35	6 A.M.	7.90	Noon	2.92	6 P.M.	6.61
1 A.M.	0.21	7 A.M.	14.06	1 P.M.	2.68	7 P.M.	3.26
2 A.M.	0.36	8 A.M.	9.63	2 P.M.	3.80	8 P.M.	2.20
3 A.M.	0.37	9 A.M.	4.30	3 P.M.	6.78	9 P.M.	1.91
4 A.M.	0.88	10 A.M.	2.26	4 P.M.	9.31	10 P.M.	1.75
5 A.M.	2.94	11 A.M.	1.86	5 P.M.	12.04	11 P.M.	1.61

Source: Martin and McGuckin (1998).

^aDefined as areas with population of 1,000,000 or more.

Link-Specific Hourly Proportion Models

An example of a link-specific hourly proportion model was offered by Ivan and Allaire (2001) in which an exponential-based model forecast a peak hour volume's proportion of a 4-hour period for individual Connecticut interstate links. (For example, if a given link had a peak hour volume of 2,000 and a 4-hour volume of 4,000, the proportion would be 0.50.) The model took the form shown in Equation B1

$$\frac{\text{Peak hour volume}}{\text{Peak 4-hour volume}} = \frac{1}{4} + a \left[\frac{\text{4-hour volume}}{\text{4-hour capacity}} \right]^b \quad [\text{Eq. B1}]$$

where a and b are parameters based on the area type (e.g., urban, suburban, exurban, shoreline, or metropolitan area of New York City) and whether the link was in the commute or reverse commute direction (which implicitly captures the effect of different trip purposes, since a link in the commute direction would be expected to have a high proportion of work-related trips). Generally, both sets of variables, i.e., area types and commute direction, were found to be statistically significant, with R^2 values of 0.48 to 0.49 for various models; as a practical matter,

the predicted and actual proportions were relatively close; differences ranged from 0.3% to 5.4%. A model similar to Equation B1 was also cited by Cambridge Systematics, Inc. (1997) for use in the Phoenix area; a 3-hour rather than a 4-hour period was used as the basis for the proportion.

Choice Models

An example of a choice model was offered by Sall et al. (2010). This model jointly considered mode choice and time of day within the regional travel demand planning process for the San Francisco area where the combined utility of five time periods (early morning, morning peak, midday, evening peak, and evening) and mode (e.g., drive alone, a two-person auto trip, transit, and so forth) was estimated using a nested logit model. Sall et al. (2010) reported using the time of day model at two levels: first, the time of day and mode were determined (based on the combined utility of mode choice and time of day); second, the specific half-hour in which auto travel occurred was determined (based on congested auto travel times). Although Sall et al. (2010) did not provide specific equations, presumably the logit model within a given nest might take the form shown in Equation B2, assuming the utilities were scaled (Slavin, 2008).

$$P(D_2 | \text{AutoAM}) = \frac{e^{\frac{U_{\text{alone}}}{\theta_{\text{autoAM}}}}}{e^{\frac{U_{\text{alone}}}{\theta_{\text{autoAM}}}} + e^{\frac{U_2}{\theta_{\text{autoAM}}}} + e^{\frac{U_3}{\theta_{\text{autoAM}}}}} \quad [\text{Eq. B2}]$$

where

$P(D_2|\text{AutoAM})$ = probability of a 2-person auto trip given that an auto trip in the morning peak hour is taken

U = utility of driving alone, a 2-person auto trip, or a 3+-person auto trip

θ_{autoAM} = logsum of the morning peak hour auto trips.

Summary of Approaches

The three categories vary in terms of model sophistication, data requirements, and geographic specificity. For example, choice models require knowledge of individual travel behavior, are based on utility maximization approaches, and concern an entire region; by contrast, the link-specific models use aggregate travel patterns and provide forecasts, as their name suggests, for specific links within the transportation network. These models can also be integrated with the traditional four-step travel demand modeling process as suggested by Purvis (2002). The regional hourly proportion models (e.g., Martin and McGuckin, 1998) offer a fast way to estimate peak spreading but do not explicitly account for changes in congestion effects.

There are other approaches in addition to those presented here. For example, Jin and Chiao (2008) described activity models (which use the concepts of individuals making tours rather than trips). The appeal of such models was articulated by Cambridge Systematics, Inc.

(1997) with the example of five alternatives that individuals could consider in response to increased congestion: change to an alternative route, change the mode of travel, change the time of travel, change the destination, and simply not make the trip. Because such choices are not mutually exclusive and will depend heavily on trip purpose, activity-based models that could simultaneously consider all such purposes are attractive (Cambridge Systematics, Inc., 1997), as would be choice models for which the nesting structure allows consideration of all five alternatives simultaneously.

Definition of Peak Spreading

In peak spreading models, there is some variation in how the dependent variable is established. Often the variable of interest is the proportion of traffic that occurs during the peak hour; for example, a K-factor is the proportion of 24-hour volume that occurs during the peak hour. However, as noted by Equation B2, the denominator is not always a 24-hour period.

Allen and Schultz (1996) explicitly rejected the feasibility of a link-based approach to determining peak spreading, partly because they noted that total trip congestion, rather than congestion on a given link, was responsible for the spreading of the peak. They developed a cross-classification model that predicted a peak hour's share of a 3-hour peak period for six trip purposes and, within those purposes, various trip distances. The general form of the model is shown in Equation B3, where delay is the difference between the free flow time and the congested time.

$$\frac{\text{Peak hour volume}}{\text{Peak 3-hour volume}} = \text{Maximum}[a - b\text{Maximum}(\text{Delay} - c, 0), d] \quad [\text{Eq. B3}]$$

where

Delay = congested travel time minus free flow travel time in minutes

a, b, c, and d = parameters that vary by trip purpose and distance.

Equation B4 illustrates the model for the home-based work trip purpose where the distance is 9 to 14 miles:

$$\frac{\text{Peak hour volume}}{\text{Peak 3-hour volume}} = \text{Maximum}[0.456 - 0.006\text{Maximum}(\text{Delay} - 10, 0), 0.333] \quad [\text{Eq. B4}]$$

For example, for a 10-mile trip for which the congested speed is 20 mph and the free flow speed is 60 mph, the peak hour will have 39.6% of the morning peak; a drop to 15 mph would lower this to 33.6% of the morning peak. Allen and Schultz (1996) reported different sensitivities to congestion based on trip purpose and distance. For example, one category of trip purpose is "home-based work" and another purpose is "home-based other" where the purposes are, respectively, travel to work and travel for some other purpose; in both cases either the origin or the destination is the home end. Allen and Schultz (1996) noted that home-based other trips

are less sensitive to congestion than home-based work trips. The authors also showed that this sensitivity varied by trip distance; for example, home-based work trips exceeding 19 miles were far less sensitive to congestion than home-based work trips with a shorter distance.

Barnes (1998) provided a comprehensive summary of research in peak hour spreading, describing many of the approaches for forecasting peak spreading noted herein (e.g., Allen and Schultz, 1996; Replogle, 1990). Equation B5, which is another approach cited by Barnes (1998), shows that the dependent variable, described as “peakiness,” is the ratio of the sum of the volumes in the one-half hour before and after the peak hour to the peak hour volume.

$$\frac{\text{Volume prior to peak} + \text{Volume after peak}}{\text{Peak hour volume}} = 1 - k(\text{Average peak hour speed})^2 \quad [\text{Eq. B5}]$$

where

k = coefficient for the specific peak period.

Holyoak and Taylor (2006) characterized peak spreading as having two discrete dimensions. One dimension entails travelers’ explicit behavioral change in trip departure time, mode, or destination in response to increased congestion; thus, such a change might be measured as the proportion of 24-hour volume that occurs during the peak hour. The other dimension is the change in the length of the peak period itself, with the implication that this duration does not require an explicit behavioral shift by travelers. Thus, the left side of Equations B2 through B5 might conceivably be replaced with the duration of the peak period.

Although not providing an explicit dependent variable per se, Martland (2005) noted that a weakness of studies of congestion generally is that none of them reveals “any data about how service actually varies over rush hour for any specific commute,” finding that there is considerable variation for transportation performance as experienced by a commuter within a peak, especially for arterial facilities where signal performance may vary considerably.

Use of Congestion to Explain Peak Spreading

It is generally recognized that congestion influences peak spreading and thus should be included in a model if any type of peak spreading is to be forecast. The differences in the literature simply result from multiple ways to account for this congestion and the extent to which variables other than congestion are considered.

Margiotta et al. (1999) reported three different temporal distributions describing the proportion of freeway directional *volume* occurring during a given hour. The appropriate distribution was selected based on the ratio of 24-hour volume (AADT) to hourly capacity (C) where such ratios may be categorized as 7.0 or less, greater than 7.0 through 11.0, and greater than 11.0. (There is some ambiguity about how the boundary conditions are interpreted: the authors reported the middle range as “between 7 and 11” but also provided a table defining this middle range as “7.1 - 11.0” with a subsequent range being “GT 11.0” [where GT is believed to

mean “greater than”) and a preceding range being “LE 7.0” [where LE is presumed to mean “less than or equal to”) .) (For example, if a four-lane freeway has an AADT of 84,000 passenger cars and if each lane has a capacity of 2,100 vehicles, the AADT/C ratio would be 10.0; a lookup table provided by Margiotta et al. (1999) showed that between 7 A.M. and 8 A.M., one should expect 4.59% of the AADT if the peak direction is in the morning and 3.05% of the AADT if the peak direction is in the evening.

Simons (2006) used the concept of these proportions in a travel demand model to determine the proportion of *trips* occurring during the peak hour between each given pair of zones in which the morning and evening directions were summed to accommodate the two-way volumes on links in the model. Figure B1 shows these temporal distributions; for example, during the peak hour of 5 P.M. to 6 P.M., one would expect slightly more than 8% of the 24-hour trips to be on a given link if the 24-hour volume-to-capacity ratio was 7.0 or smaller; if congestion increased such that the 24-hour volume-to-capacity ratio was 11.0 or greater, this proportion would be around only 7%. An interpolation procedure reported by Margiotta et al. (1999) and used by Simons (2006) further distinguished how smaller changes in the AADT/C ratio, such as from 12.0 to 13.0, affected the temporal distribution, but the concept was similar to that illustrated in Figure B1.

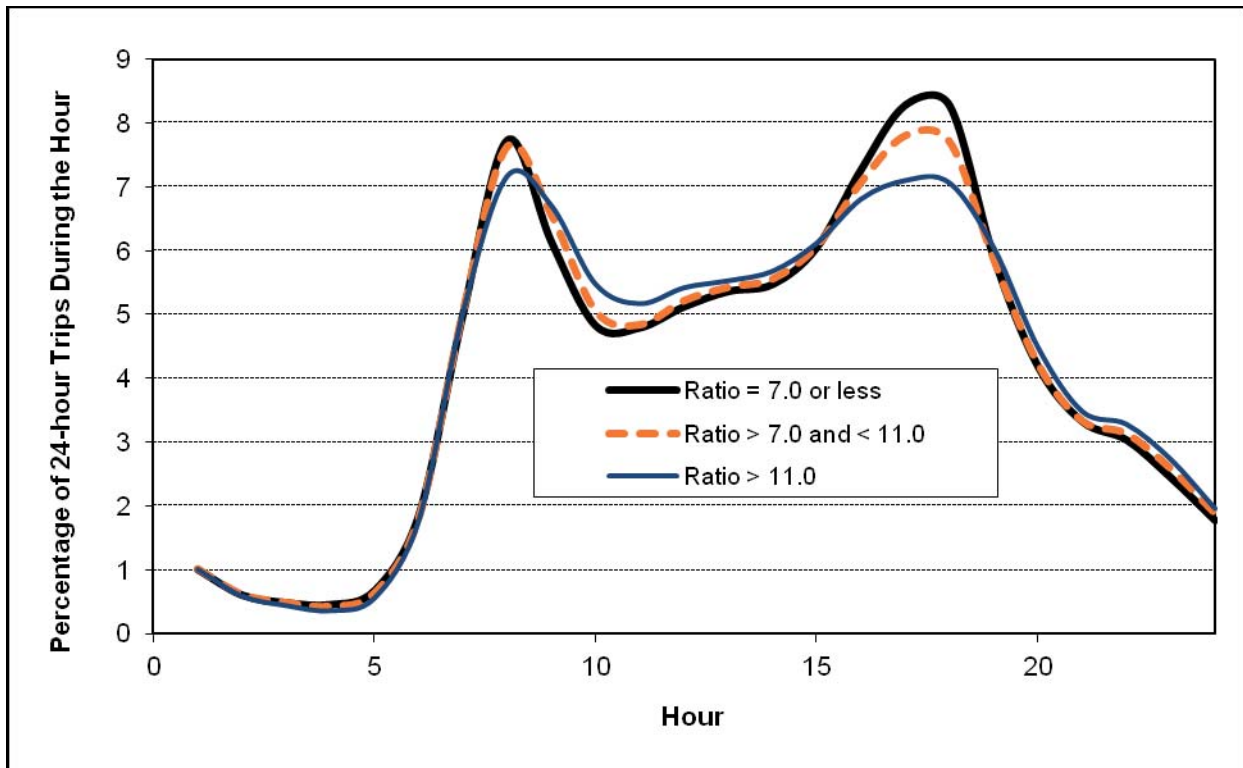


Figure B1. Influence of 24-hour Volume-to-Capacity Ratio on Temporal Distribution of Trips. Drawn from data reported by Simons (2006). Simons reported these categories as “less than or equal to 7, 7 to 11, and greater than 11” and also as “≤7” “7–11,” and “>11.” It is ambiguous whether an exact ratio of 11.0 would be included in the middle or last category.

Cottrell (1998) used a logistic regression model of the form shown in Equation B6 to predict whether based on its 24-hour volume-to-capacity ratio and its K-factor a link would be an uncongested location or a location with recurrent congestion. The implication was that if a given ADT can be accurately forecast, the likelihood of congestion and hence the likelihood that peak spreading will diverge from current values can be forecast. Cottrell recommended a threshold 24-hour volume-to-capacity ratio of 8.5; links with a higher ratio were likely to undergo peak spreading. Because the sites were collected from several geographic locations with different ways of defining congestion, some recalibration of the original dataset would probably be necessary for use in Virginia. For example, congestion at Chicago sites was defined by Cottrell (1998) on the basis of density (percent occupancy time ≥ 30) and congestion at Salt Lake City sites was defined on the basis of speed (≤ 40 mph).

$$P = \frac{1}{1 + \exp\left(7.7966 - 0.2136\left[\frac{24\text{-hour volume}}{\text{capacity}}\right]^{1.2} - 1.2522[K]^{0.6}\right)} \quad [\text{Eq. B6}]$$

where

P = probability of a site having recurrent congestion

K = proportion of traffic occurring during the peak hour.

Liu et al. (2007) formulated a peak spreading approach for use within a corridor-based microsimulation approach in which when forecast peak hour volumes exceeded capacity such that the simulation model could not place all vehicles on the transportation network, the vehicles were first shifted to another route with less congestion and then to a non-peak time period. An implication for peak spreading forecasts for a given link, therefore, is that one independent variable should be the availability of a parallel link with a lower volume-to-capacity ratio. The authors' approach may also be viewed as one way of explicitly considering options travelers have in response to peak spreading other than simply changing the time of departure of their trip.

Use of Socioeconomic Factors to Explain Peak Spreading

In peak spreading models, congestion is not always the only independent variable. Independent variables have included socioeconomic factors—population, employment, income, and land development.

The literature implicitly suggests that although socioeconomic factors may influence peak spreading, detection of these effects remains difficult. For example, Karlström and Franklin (2009) reported that in a survey of users affected by a tolling experiment in Stockholm, users with greater flexibility in their schedule were more able to switch their departure time for work in response to the introduction of tolls during the peak period. However, although income and flexibility were correlated (e.g., 61% of workers in the highest income group had flexible schedules compared to 20% for the second lowest income group), income itself was not a statistically significant variable in explaining time of departure. The authors noted:

... working hours flexibility is strongly correlated with household income and consumption level.
... Finally, we confirm that individuals' own scheduling flexibility and work hour flexibility indeed affect the ability to switch departure time. With respect to equity implications, although income itself was not significant, we suspect that it is a significant determinant of scheduling flexibility, and hence would have an indirect relationship.

A review of Habib et al. (2009) illustrated how employment may influence travel decisions generally and, by extension, peak spreading. Based on a survey sample of Toronto area workers, the authors found that, consistent with Replogle (1990), increasing urban density of the area in which workers live generally led to a lesser likelihood of driving, which the authors attributed to the availability of other transportation options. For individuals in the manufacturing sector in particular, however, Habib et al. (2009) reported the opposite impact: increased density of the area where the workers lived led to *more*, not less, auto use; the authors attributed this finding to the fact that most manufacturing locations are in low-density areas. The larger implication is that the types of employment available in an area may influence mode use. By extension, the amount of peak spreading that will result may be affected, since mode choice and departure time are interrelated.

As an illustration of the interdependency between mode choice and departure time, Purvis (2002) developed a logit model for the San Francisco Bay Area that predicted whether individual commuters would travel during the peak 2-hour morning period or at some other time during the morning based on variables such as travel time, distance, income, and whether the individual worked in the retail employment industry.

A review of Sinha and Thakuria (2004) showed two distinct trends that may accelerate peak spreading: implementation of policies that encourage non-peak travel (e.g., telecommuting, flex-time, and mandates for employers to reduce commute-related auto travel), and growth of industries having non-traditional schedules. For example, metropolitan area data from the 1997 Bureau of Labor Statistics Current Population Survey showed that manufacturing, transportation, trade, and service employees are 3 times more likely to have a non-peak start time (i.e., outside the hours of 6 A.M. to 10 A.M.) than are finance, insurance, and real estate employees. A third factor, which is intertwined with industrial effects, is demographic changes: for example, teenagers were more than 8 times more likely to have a non-peak start time than persons over 50 years of age, and persons with incomes below \$15,000 were almost twice as likely to have a non-peak start time as persons with incomes above \$50,000 (Sinha and Thakuria, 2004).

Replogle (1990) noted that homogenous land uses (whether they are residential only or commercial only) tend to have higher peak hour factors (what this report calls the K-factor) than mixed land uses. Accordingly, Replogle (1990) developed factors for Montgomery County, Maryland (a suburb of Washington, D.C.), that converted daily trips to morning peak hour trips, with these factors based partly on household density in the zone where these trips originated and the employment density in the zone where these trips terminated. For example, regional survey data showed that on average, 19% of trips between home and work occurred during the morning peak. For zones with very low residential densities, defined as 500 households per square mile or less, the percentage of such morning peak hour trips was approximately 23%; for zones with the highest residential densities of 10,000 households per square mile, this percentage decreased to 15.2% (see Figure B2). Replogle (1990) also noted that although a variety of formulations

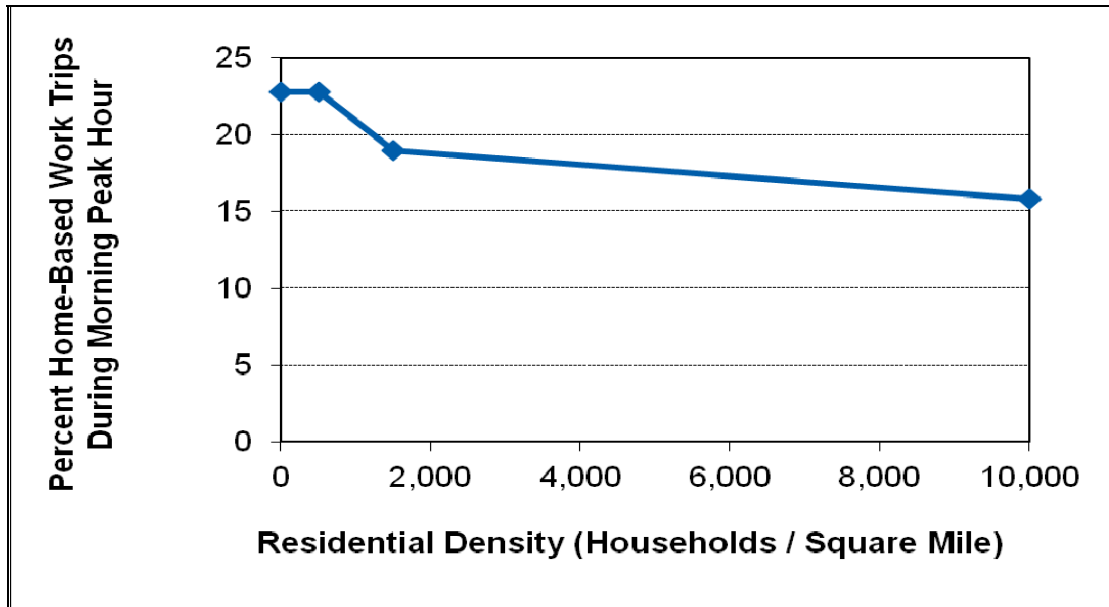


Figure B2. Percentage Home-Based Work Trips During Morning Peak Hour. Values were calculated from origin peak hour adjustment factors provided by Replogle (1990) and do not reflect destination density adjustments.

was used, the model based solely on residential and employment density led to predicted and actual vehicle miles of travel (VMT) being within 5% of each other, with further adjustments being made to the model. For example, recognizing that retail shopping centers open at 10 A.M., the author indicated that a refinement to the model was to “discount” morning work trips related to retail employment (presumably since such trips are outside the morning peak hour).

Thus, returning to Equation B1, although Ivan and Allaire (2000, 2001) used variables such as the area type to calibrate a and b (and in other types of models the authors used variables such as the distance from the site to the central business district and the number of lanes), it is also feasible to use socioeconomic variables to calibrate Equation B1.

Use of Tolls, Seasonal Variation, and Facility-Specific Factors to Explain Peak Spreading

Independent variables may also be selected to capture the effects of tolls, monthly variations, or facility-specific effects. Although an “average day” can be considered when developing a K-factor, it may be less relevant in some situations than holiday traffic. Liu and Sharma (2006) noted that based on rural Alberta routes, holiday traffic can have a substantial impact on peaking characteristics: about two-thirds of the 50 highest hourly volumes per year for regional commuter routes occurred on holidays.

Yang et al. (2009) used regression analysis to predict monthly (seasonal) factors on roadways in Florida as a function of functional class [a roadway could fall into any of four classes: urban freeway, principal arterial, minor arterial, and collector] and socioeconomic data [e.g., population, number of retired households, and total housing units, all obtained at the census tract or transportation analysis zone (TAZ) level of analysis, and employment data such as the

number of employees]. A separate equation was developed for each month and for each of three regions in the state (north, central, and south Florida). For example, for North Florida:

- The January seasonal factor was a function of the percentage of population with an age of 11, 12, or 13; the ratio of seasonal households to permanent households; and the percentages of workers in the fishing and hunting, hotel and camp, and museum sectors.
- The September seasonal factor was a function of the percentage of retired households and the presence of residential universities.

Adjusted R² values for North Florida ranged from 0.26 to 0.75. Yang et al. (2009) concluded that individual models need to be calibrated for each region; for example, in South Florida, winter is the tourist season; in north Florida, summer is the tourist season.

Based on a comparison of 2000 data, Ozbay et al. (2006) reported statistically significant variation among traffic volumes by season for the Goethals Bridge (a facility maintained by the Port Authority of New York and New Jersey) for the evening peak and off-peak hours. Noting also that several other bridges and time periods did not have significant variation by season, the authors recommended analyzing each facility separately.

Wolff and Vilain (2007) used an ordinary least squares regression model to discern the effects of a new toll on travel volumes for bridges and tunnels served by the Port Authority of New York and New Jersey. The toll was \$6 except for motorists who both paid electronically and traveled during off-peak hours. A simplified example of this model is Equation B7; the model was run for a variety of vehicle types, time periods, and payment methods.

$$\ln\left(\frac{\text{Morning peak volume}}{\text{Volume 1 hour prior to the morning peak}}\right) = a + (b)\ln\left(\frac{\text{Morning peak toll}}{\text{Toll 1 hour prior to the morning peak}}\right) + c\left(\frac{\text{Morning peak delay}}{\text{Delay 1 hour prior to the morning peak}}\right) + D(\text{seasonal factors}) + E(\text{link - specific factors}) \quad [\text{Eq. B7}]$$

where

a, b, and c = parameters that vary by period, vehicle type, and time period

ln = the natural logarithm of the expression in parentheses

D = vector that reflects seasonal factors

E = vector that reflects “fixed effects” (which are interpreted here to be factors specific to each of the six facilities).

Although the focus of Equation B7 was eliminating the background noise of peak spreading that would otherwise be occurring (which would mask the true impact of the new tolls), the approach shows a way to capture the degree of peak spreading: if the coefficient of parameter c is negative and statistically significant (apart from parameter b , which captures the impact of the toll increase), peak spreading may be occurring. In their analysis, Wolf and Vilain (2007) found that the parameter b was negative and statistically significant, meaning the tolls contributed to a change in behavior, and parameter c (for drivers who pay electronically and thus are eligible for an off-peak discount) was not significant, meaning peak spreading was not occurring.

References

- Allen, W.G., and Schultz, G.W. Congestion-Based Peak Spreading Model. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1556. Transportation Research Board of the National Academies, Washington, DC, 1996, pp. 8-15.
- Anderson, R.S., and Donnelly, R.M. Modeling of Peak Hour Spreading with a Disaggregate Tour-Based Model. In *Innovations in Travel Demand Modeling: Summary of a Conference. Volume 2: Papers*. Transportation Research Board of the National Academies, Washington, DC, 2008, pp. 161-164.
- Barnes, J. *Peak Spreading Analysis: Review of Relevant Issues and Synthesis of Current Practice, Phase 1*. Washington State Transportation Center, Seattle, 1998.
<http://ntl.bts.gov/lib/21000/21200/21218/PB99104564.pdf>. Accessed July 15, 2010.
- Cambridge Systematics, Inc. Time-of-Day Modeling Procedures: State-of-the-Practice, State-of-the-Art. Federal Highway Administration, Washington, DC, 1997.
<http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/time-day/>. Accessed October 1, 2010.
- Cottrell, W.D. Estimating the Probability of Freeway Congestion Recurrence. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1634. Transportation Research Board of the National Academies, Washington, DC, 1998, pp. 19-27.
- Habib, K.M.N, Day, N., and Miller, E.J. An Investigation of Commuting Trip Timing and Mode Choice in the Greater Toronto Area: Application of a Joint Discrete-Continuous Model. *Transportation Research Part A*, Vol. 43, No. 7, 2009, pp. 639-653.
- Holyoak, N., and Taylor, M.A.P. Modelling Trip Timing Behaviour and the Influence of Peak Spreading. In *Urban Transport XII. Urban Transport and the Environment in the 21st Century*. WIT Press, Southampton, UK, July 2006, pp. 205-213.
- Institute of Transportation Engineers. *Transportation Impact Analyses for Site Development. An ITE Proposed Recommended Practice*. Washington, DC, 2006.

- Ivan, J.N., and Allaire, S.A. *Estimating the Temporal Distribution of Traffic Within the Peak Period*. Connecticut Transportation Institute, Storrs, 2000.
<http://ntl.bts.gov/lib/16000/16100/16143/PB2000102761.pdf>. Accessed July 30, 2010.
- Ivan, J.N., and Allaire, S.A. Regional and Area-Type Modeling of Peak Spreading on Connecticut Freeways. *ASCE Journal of Transportation Engineering*, May/June 2001, pp. 223-229.
- Jin, X., and Chiao, K.-A. A Synthesis of Time-of-Day Modeling Research. In *TRB 87th Annual Meeting Compendium of Papers DVD*. Transportation Research Board of the National Academies, Washington, DC, 2008.
- Karl, C.A., and Gaffney, J. Improving Traffic Network Performance in Australia. In *Australian Road Research Board 23rd Conference CD-ROM*. Australian Road Research Board, Adelaide, 2008.
- Karlström, A., and Franklin J.P. Behavioral Adjustments and Equity Effects of Congestion Pricing: Analysis of Morning Commutes During the Stockholm Trial. *Transportation Research Part A*, Vol. 43, No. 3, 2009, pp. 283-296.
- Liu, Z., and Sharma, S. Predicting Directional Design Hourly Volume from Statutory Holiday Traffic. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1968. Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 30-39.
- Liu, H.X., Jabari, S.E., and Wenteng, M. Peak Spreading Methodology for Future Scenario Evaluations Using Microsimulation. In *TRB 86th Annual Meeting Compendium of Papers CD-ROM*. Transportation Research Board of the National Academies, Washington, DC, 2007.
- Margiotta, R., Cohen, H., and DeCorla-Souza, P. Speed and Delay Prediction Models for Planning Applications. In *Sixth National Conference on Transportation Planning for Small and Medium Sized Communities, September 16-18, 1998, Spokane, Washington*. Transportation Research Board of the National Academies, Washington, DC, 1999.
<http://ntl.bts.gov/lib/1000/1100/1195/00780089.pdf>. Accessed November 15, 2010.
- Martin, W.A., and McGuckin, N.A. *Travel Estimation Techniques for Urban Planning*. NCHRP Report No. 365. Transportation Research Board of the National Academies, Washington, DC, 1998. <http://ntl.bts.gov/lib/21000/21500/21563/PB99126724.pdf>. Accessed December 3, 2010.
- Martland, C.D. The Journey to Work: 25 Years on the Jamaicaway. *Journal of the Transportation Research Forum*, Vol. 44, No. 3, 2005, pp. 59-76.
http://www.trforum.org/forum/downloads/2005_Jamaicaway_paper.pdf. Accessed September 30, 2010.

- Ozbay, K., Ozlem, Y.-T., and Holguín-Veras, J. Evaluation of Impacts of Time-of-Day Pricing Initiative on Car and Truck Traffic: Port Authority of New York and New Jersey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1960. Transportation Research Board of the National Academies, Washington, DC, 2006, pp. 48-56.
- Purvis, C. Peak Spreading Models: Promises and Limitations. In *Proceedings of the Seventh TRB Conference on the Application of Transportation Planning Methods*. Transportation Research Board of the National Academies, Washington, DC, 2002. <http://docs.trb.org/00939750.pdf>. Accessed September 30, 2010.
- Replogle, M. Computer Transportation Models for Land Use Regulation and Master Planning in Montgomery County, Maryland. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1262. Transportation Research Board of the National Academies, Washington, DC, 1990, pp. 91-100.
- Sall, E., Bent, E., Charlton, B., Koehler, J., and Erhardt, G. Evaluating Regional Pricing Strategies in San Francisco: Application of the SFCTA Activity-Based Regional Pricing Model. In *TRB 89th Annual Meeting Compendium of Papers DVD*. Transportation Research Board of the National Academies, Washington, DC, 2010.
- Sinha, A.K., and Thakuriah, P.V. Relationship Between Occupational, Industrial, and Sociodemographic Characteristics and Job Start Times: Evidence from Current Population Survey. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1894. Transportation Research Board of the National Academies, Washington, DC, 2004, pp. 67-74.
- Slavin, H. *Travel Demand Modeling with TransCAD Version 5.0: User's Guide*. Caliper Corporation, Newton, MA, 2008.
- Virginia Department of Transportation. Department of Transportation, Location and Design, Request for Traffic Data: Form LD-104. Richmond, 2008. <http://vdotforms.vdot.virginia.gov/SearchResults.aspx?filename=LD-104.doc>. Accessed September 15, 2010.
- Wolff, C., and Vilain, P. Evaluating Congestion Pricing Impacts Under Peak Spreading. In *Proceedings of the 48th Annual Forum of the Transportation Research Forum*, 2007. http://www.trforum.org/forum/downloads/2007_8A_PeakSpread_paper.pdf. Accessed November 1, 2010.
- Yang, S., Lu, C., Zhao, F., Reel, R., and O'Hara, J.D. Estimation for Seasonal Factors of Similarity-Based Traffic for Urban Roads in Florida. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2121. Transportation Research Board of the National Academies, Washington, DC, 2009, pp. 74-80.