A Retrospective Evaluation of Traffic Forecasting Accuracy: Lessons Learned from Virginia

A Thesis

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

in partial fulfillment

of the requirements for the degree

Master of Science

by

Salwa Anam

December

2016

APPROVAL SHEET

The thesis

is submitted in partial fulfillment of the requirements

for the degree of

Master of Science

AUTHOR

The thesis has been read and approved by the examining committee:

John S. Miller

Advisor

T. Donna Chen

Andrew Mondschein

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, Dean, School of Engineering and Applied Science

December

2016

To everyone who's helped me succeed

Abstract

Understanding the accuracy of techniques for forecasting traffic volumes for a future year, such as extrapolation of previous years traffic volumes, use of regional travel demand models, and use of local trip generation rates, can aid analysts in considering the range of transportation investments for a given location. To determine this accuracy, forecasts from 39 Virginia studies (published from 1967-2010) were compared to observed volumes for the forecast year. The comparison enabled the identification of potential assumptions that might cause variations in the way forecast accuracy is assessed. Some of the assumptions include construction of proposed infrastructure (A new bridge in York river crossing), the appropriate error statistics (Average value or median value of error), chosen observed volume (Volume from continuous count or periodic count), anticipated alignment (relocation of Route 33 where Route 3 and Route 33 are signed together).

Excluding statewide forecasts, the number of roadway segments in each study ranged from 1 to 240 links. For each segment, the difference between the forecast and observed volume divided by the observed volume gives a percent error, such that a segment with a perfect forecast has an error of 0%. The analysis showed that based on 39 Virginia past studies, the median absolute percent error ranged from 1% to 134% with an average value of 40%; forecast volumes tended to be larger than observed volumes. The accuracy of different types of traffic forecasts varies by an order of magnitude: 12% (for a site-specific land development study) to 72% (for statewide forecasts based on historic traffic volumes). The importance of forecast accuracy is determined by whether such errors have any impact on decision taken (i.e. signal warrant, change of road alignment, no remedy for environmental impact).

Slightly more than one-fourth of the variation in such error (29%) was explained by three identifiable factors: the forecast method, the forecast duration (number of years between the base and forecast years), and the number of economic recessions in the same interval (p 0.04). Interaction effects matter: the first two factors have significant and expected impacts on accuracy only if economic changes are explicitly considered. Finally, link-by-link error in a study has sufficient variation ($p = 0.02$) such that if this variation is not controlled, explanatory factors are impossible to detect. Although no forecast is perfect, this study provides an indication of expected forecast error for future studies that might help decision makers to evaluate transportation needs.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

Traffic forecast is a replicable method of forecasting traffic volume on an existing (or not-yet-built) roadway section for some future year. Some of the forecasting methods or techniques include extrapolation of previous years' traffic count, use of a regional travel demand models (TDMs), trip generation rates for specific land use, the FRATAR method for turning movement analyses etc. Traffic forecast can vary by volume type (e.g. daily volume or peak hour volume), duration (e.g. short term or long term forecast), location (e.g. a link or turning movement), method (e.g. statewide or regional forecast) and vehicle type (e.g. passenger car or truck traffic only).

These traffic forecasts are an important input which is needed in transportation planning, policy making, infrastructure design (i.e. thickness or type of pavement), operational changes (signal design or level of service) and improvement of projects (i.e. building bypass or addition of lanes in existing roadway). Traffic forecast helps in making critical decisions regarding project selection. To better manage the scarce resources and ensure sound performance of transportation projects, accurate forecast is needed. Accuracy indicates how close the forecast volume is to the actual volume. Determining the accuracy of Virginia forecasts has been of interest to transportation agencies with an eye towards knowing the expected accuracy over time and the extent to which higher (or lower) accuracy can be explained by other factors (e.g., how forecast was made, changes in economy or traffic pattern, and so on). In order to determine accuracy, past studies of Virginia forecasts were examined where at some point forecast was generated and the year of forecast had been passed. The error is determined by comparing the forecast volume to the actual volume. Some of the examples of respective Virginia studies include a regional travel demand model in the City of Charlottesville, an improvement recommendation of Colonial Heights (i.e. widening of Route 1 and 301), a new upriver crossing over the York River, Statewide Highway needs of Culpeper district for 2010, a mixed use retail development at the quadrant of the primary road (Stonefield development in Albemarle County) etc.

Whether a forecast error is large or small, a related question is–does the error matter? CDM Smith et al. [9] indicated that an error's importance depends on the purpose of the forecast. If a forecast of "2,000 vehicles" denotes an ADT, then error might be ignored as such low-volume facilities "are rarely important to a project decision" [9]. If that 2,000 vehicle forecast denotes a peak hour volume, however, the associated 95% confidence interval (e.g., 1,020 to 2,980) might be reported, as the forecast determines the needed number of lanes [9]. In another case, determination of whether a forecast error affects actions taken depends in part on the magnitude of the error-as one might expect-but also on the location of the error relative to the decision criteria. For instance, in Route 3 Corridor Study [8], the decision criteria was whether a certain performance standard-in this case, level of service C-could be obtained. Although a 61% forecast error was noted, this error would not have affected the decision that the roadway needed to be improved, because both the forecast volume and the observed volume would have indicated that level of service C was not attained. Yet forecast accuracy may not matter: Brett and Snelson [10] reported that despite close alignment of projected and observed bridge counts 15 years after construction, debt was four times higher than forecast, owing to unexpected increases in construction costs that compounded the project debt.

1.1 Background

VDOT's Transportation Mobility Planning Division (TMPD) has commissioned the development of a Traffic Forecasting Guidebook where dozens of techniques for generating traffic forecasts were discussed. One question focuses interest "what is the accuracy of those traffic forecast?" Based on the answer, how should the accuracy influence the performance of projects? and how should the accuracy be addressed in the future revision of that Guidebook? This research is carried out to seek answer to those questions. Moreover, Virginia Department of Transportation (VDOT) which is responsible for project-level, corridor-level, NEPA-level, and statewide-level traffic volumes, does not have a principal set of policies or a guideline related to the development of traffic forecasts. Planners and staff need a document of the methods which they can follow for the level of forecast needed.

Eleven techniques are briefly discussed in Chapter 3. Variations within the techniques depend on the types of data available, time required and reason for the forecast. Idea about the techniques help in understanding how well a technique can perform replicating observed volume and thus influence errors in forecast.

1.2 Contribution

The implication of this research effort is twofold, one concerning how forecasts are performed and one concerning the effect of parameter on forecast accuracy. Considering no forecast is accurate, the study provides an indication of the expected error for future forecasts based on the past projects in Virginia. Such an indication is also reported in other literature, for example, CDM Smith et al. [9]. However, the magnitude of the error alone does not determine whether the error is important. To know the impact of error, the analysts can examine whether a forecast (with expected error) materially affects actions being taken on the basis of the forecast. (For example, one study noted that a certain investment remained "economically feasible" even with an error of 50 %.) Only a portion of the error is explained by the factors identified in this study. However, even with the least successful model the study was able to explain 5.2% of variation in the forecast error which was greater than that reported in the study conducted by Parthasarathi and Levinson [11]. Idea about the factors impacting the forecast error can provide a groundwork for better demand forecasts.

Chapter 2

RELATED WORK

There are various literature out there where the accuracy of travel forecast has been discussed. Some of the literature address accuracy across the country; some compare the accuracy between road project and rail project; some use data for tolled roads or for demand forecasts only; other measure accuracy for forecasting tools. Again, the results of forecast accuracy across literature also vary by number of sample size, study period, geographic location and how they represent error.

Flyvbjerg et al. [12] measured inaccuracy of traffic forecast based on the data from transport projects around the world. The analysis used a sample of 27 rail projects and 183 road projects completed between 1969 and 1998. The projects are located in 14 countries in 5 continents including Brazil, Denmark, Egypt, India, U.K. and U.S projects. The study concluded that the forecasters do a poor job in estimating demand. The results show that the passenger forecasts in 90% rail projects are overestimated. For half of the road projects, difference between actual and forecasted traffic is more than 20% and the actual road traffic on average is 9.5% higher than forecasted traffic. Often it is claimed that forecasts have become better due to improvements in traffic model which is opposed by the study data. For example, the inaccuracy for Danish road projects increased from 3% to 55% over time. The main cause for such inaccuracy, according to the study, is conventional method and political bias.

Buck and Sillence [13] evaluated 131 Wisconsin forecasts completed between 2003 and 2011 and reported that the mean and median absolute differences were 16% and 13%, respectively. The study conducted by Parthasarathi and Levinson [11] investigated 108 Minnesota post construction projects completed about 20 years before. The study did not distinguish between tools or methods used in forecast as did Buck and Sillence [13], but identified some qualitative and quantitative reasons behind inaccuracy. Nicolaisen and Naess [14] focused on 35 projects between 1985 and 2010 that used travel demand forecast for do-nothing alternative and concluded that on average, the demand forecast has been overestimated to 7%. Welde and Odeck [15] investigated the difference of forecast accuracies between 25 Norwegian tolled and toll free road projects. According to the study, traffic forecasts of tolled roads are fairly accurate where toll free roads are generally underestimated; the mean underestimation was 19%.

The large, multinational study conducted by Bain [16] on traffic forecast accuracy showed that half of studied road projects had a difference of over 20% between actual and forecasted traffic. The study also reported that average traffic forecast accuracy on privately financed toll roads, bridges, and tunnels was found to be 130%. The MnDOT study reported average accuracy on existing roadway forecasts around 120% using the same method as Bain study. By utilizing the same method to compare, the WisDOT study found an average accuracy of approximately 115%.

There are also variations among literatures regarding the representation of accuracy since there is no standard guideline on how to measure forecast errors. Some literature measure error as the forecast value minus the actual value where positive value is an over forecast (CDM Smith et al., [9]; Tsai et al., [17]; Buck and Sillence, [13]; Parthasarathi and Levinson, [11]). Some literature does the opposite by estimating error as the actual value minus forecast value where positive value is an under forecast (Flyvbjerg et al., [12]; Welde and Odeck, [15]; Nicolaisen and Naess, [14]). Same with percent error where the error is in numerator and denominator may be actual (Parthasarathi and Levinson, [11]) or forecast (Welde and Odeck, [15]) value. For example, if the forecast value is 1000 and the observed value is 600, the percent error can be measured in terms of observed volume (66.7%) or forecast volume (40%). The advantage of the former is that the percentage is expressed in terms of a real quantity; an advantage of the latter is that uncertainty can be expressed in terms of the forecast value since the observed value is unknown (J. S. Gillespie, personal communication, May 29, 2015). Again, the results can be different whether one uses error or absolute error while determining the average or median value. Because when aggregating, the positive and negative error cancels with each other resulting in lower error. For example, Flyvbjerg et al. [12] examined 183 highway forecasts, reporting an average error of 9.5%. This average value indicates accuracy where over forecast cancels with under forecasts. By contrast, if an absolute value of the errors is computed, the average of the project percentage error would be 32%. The other statistical measure of error includes GEH (Geoffrey E. Havers) statistics (Buck and Sillence, [13]) and RMSE (Root-mean-square error) reported in CDM Smith et al. [9]. The variation in measures of error across different literature in different location made it difficult to compare among themselves; thus hard to evaluate locations with good or bad forecasts. Again, there is no standard range of forecast error that can be used to evaluate a forecast. In comparison to how Virginia forecasts are with those reported elsewhere, the study reported Table 6.1 to present the accuracy of forecast based on 39 Virginia past studies between 1967-2010. On balance, the percent errors shown in Table 6.1 will appear higher than those reported by Buck and Sillence [13], Parthasarathi and Levinson [11], and Flyvbjerg et al. [12], while within the range of errors reported by CDM Smith et al. [9] and Brett and Snelson [10]. For both studies (Buck and Sillence, [13]; Parthasarathi and Levinson, [11]) these reported errors are lower than those shown in the last two rows of Table 6.1. CDM Smith et al. [9] reported on previous work that asked planners what level of accuracy should be expected as a function of the length of the planning horizon. The findings from that piece of literature were that for a five year horizon, one should expect errors of + 20% for an existing facility and + 27.5% for a new facility. For a 20 year horizon, the expected errors for an existing or new facility are + 42.5% and + 47.5%, respectively. No other studies have investigated the accuracy of various forecast techniques available to generate forecasts like the study did here. Only one study [11] investigated the cause of forecast error where the factors can explain 3% variation in forecast.

Chapter 3

FORECAST TECHNIQUES

A variety of forecasting techniques are available that are used to generate forecasts depending on data availability and for what purposes, forecasts are needed. Nine techniques are reported based on a review of the draft *Traffic Forecasting Guidebook*. Two techniques are identified by reviewing additional literature: technique 9 (based on the review of Route 3 corridor study [8]) and technique 11 (Martin and McGuckin [18]). The types of volumes provided in each row differ: techniques 1-4 and 9 give an annual average daily volume, technique 5 gives a seasonal volume, technique 6, 10 and 11 give an hourly volume, technique 7 gives volumes that specifically relate to heavy vehicles and technique 8 gives intersection counts.

In total eleven techniques that are commonly used in generating forecast are listed in Table 3.1. The details of the techniques are discussed in section 5.2.

Table 3.1: Summary of Forecast Techniques

Tech. No	Name	Summary of Techniques[a]
1	Adjusted regional model outputs[b]	Use difference between modeled volume (in 1990) and actual volume (in 1990) to adjust the link forecast (in 2010)
2	Trend analysis of regional model outputs	Determine annual rate of growth from modeled volume (in 1990) and forecasted volume (in 2010); apply to a 1990 volume in order to develop an interim link forecast for year 2000
3	Linear growth based on two traffic counts	Determine annual rate of growth from actual volumes for 1980 and 1990; apply to 1990 volume to get a 2010 forecast
4	Regression based on multiple traffic counts	Determine annual rate of growth from actual volumes for 1975, 1980, 1985, and 1990; apply to 1990 volume to get a 2010 forecast
5	Seasonal adjustment factors	Multiply the 2010 annual forecast from techniques 1, 2, 3, or 4 by the appropriate seasonal adjustment factor to obtain a daily forecast for a given day of the week and month of the year
6	Peak hour link forecasting	Multiply the 2010 ADT from techniques 1, 2, 3, or 4 by the K factor calculated from continuous count stations to forecast the 2010 peak hour volume
7	ESAL (Equivalent Single Axle Load) estimation	Adjust the 2010 annual forecast (techniques 1, 2, 3, or 4) by factors from VDOT Materials Division [19] to estimate 2010 equivalent single axle loads for a given link
8	Fratar Technique	Use current intersection movements (e.g., left, right, and through movements on each of 4 legs in 1990), coupled with 1990 and 2010 link volumes (e.g., two directional volumes on each of 4 legs) to iteratively estimate 2010 turning movements (left, right, and through) on each approach
9	Population-based forecasts	Multiply the 1990 ADT by the ratio of the forecast 2010 population to the actual 1990 population to forecast a 2010 ADT
10	ITE-Based Factoring	Forecast 2010 trips based on land use as noted in ITE Trip Generation [20] and then reduce these trips based on mixed land uses if appropriate as noted by the ITE Handbook [4]
11	Traffic Shift Methodology for Corridors	Based on Martin and McGuckin [18] and an increase in volume (and hence a decrease in travel time) for a given roadway segment, calculate the expected change in volume for a parallel roadway segment

[a] All techniques shown in Table 3.1 presume a base year of 1990 and a horizon year of 2010

[b] All the above techniques excluding 8 and 11 are adopted from chapter 5 of the Virginia Department of Transportations (VDOTs) Traffic Forecasting Guidebook which is in draft form as of March 2014.

Chapter 4

PURPOSE AND SCOPE

The purpose of the study is to respond to Virginia's interest in determining the accuracy of forecasts based on the historical case studies in Virginia and identify potential contributing factors.

The goal of the study is as follows:

1. Determine the forecast accuracy based on the previous case studies.
2. Performance of eleven techniques in terms of replicating observed volumes.
3. Document inherent assumptions in assessing accuracy.
4. Identify factors that may explain variations in forecast error.

The analysis was done based on the data of historical case studies published from 1967-2010 in Virginia.

Chapter 5

METHODOLOGY

In order to meet the goal of the study, four main tasks were conducted. Details of each task are described in the following section.

1. Report the accuracy of traffic forecasts

2. Apply eleven techniques to some of the study

3. Document assumptions while comparing forecast to observed volume

4. Identify explanatory factors

5.1 Report the Accuracy of Traffic Forecasts

To report the accuracy, the following tasks are accomplished.

1. Collect studies in which traffic forecasts were made

2. Obtain forecast volumes from studies

3. Collect observed traffic volumes

4. Summary of Data set

5. Measure accuracy by comparing forecast volumes to observed volumes

5.1.1 Collect Studies in Which Traffic Forecasts were Made

Initially, those studies were sought to obtain where, at some point in the past, traffic forecasts had been generated. Accordingly, the goals and expected methodology of this research effort was presented to VDOT district planners to ask if any studies were available, with the criterion being that the forecast year from any such studies would already have elapsed. Positive responses from that presentation led to one or more members of the research team visiting the Culpeper, Hampton Roads, and Richmond districts in 2014.

Studies were also obtained through follow up communications with district and VDOT TMPD staff and a visit to the VTRC library. For example, during a presentation to the project TRP in March, 2015, attendees suggested that the research team examine the internal VDOT LandTrack database for additional historical studies. A total of 41 studies were obtained in this manner–that is, where some type of forecast was generated–including regional travel demand forecasting models, corridor studies, site impact studies, and statewide regional planning studies. These studies are listed in Table 5.1. While some studies take the form of a bound report with a clear date of publication, other studies take the form of computer files or memoranda that can only be obtained internally. For internal databases, sometimes there are multiple "forecasts" which are then revised as new information becomes available. In other cases, multiple sources are needed to understand how a forecast was generated; for example, previous volume of a study [21] was sought to verify forecast information.

5.1.2 Obtain Forecast Volumes from Studies

The forecast data consists of 39 projects or studies (study 24, 25, 26 in Table 5.1 are basically three sub-projects of one project) where at some point in the past forecasts had been made. The studies included data obtained with different methodology and level of details: for example, the statewide forecast is a trend-based projection for thousands of links throughout Virginia; corridor studies, on the other hand focus on a smaller location and incorporate more detailed analysis. Relative to trend-based forecasts, site impact studies and regional travel demand models use a fundamentally different approach, where some estimates of activity (e.g., changes in population, employment, or developed land) forecast travel.

Each study had four time-based attributes: base year, publication year, forecast year, and observed year. Although the base year and publication year might have been identical, the former refers to the "data" year for the information used to make the forecast. For example, a site study for the University of Virginia Research Park [22] was published in 2008 and used 2006 base year traffic volumes. In some cases, the forecast year (the year for which the forecast was made) and the observed year (the year for which volumes could be obtained) were different. For example, although the Research Park Study had a forecast year of 2015, the most recent year for which volumes were available was 2014, which served as the observed year. The forecast duration, i.e., the number of years between the base year and the observed year, was 8 years. For the 41 studies, forecast durations were from 2 to 28 years, with studies published from 1967-2010 (the published year of some of the studies were unknown, i.e. study 27 in Table 5.1).

5.1.3 Collect Observed Traffic Volumes

Different data sources were used to obtain the observed traffic count. The most common methods used were the sources of annual electronic traffic data publication for various jurisdiction (Virginia Department of Transportation [23]) and VDOT's internal traffic monitoring system (TMS) database (Virginia Department of Transportation [24]). Some special counts tabulated by certain jurisdictions were used under which roads were not state maintained. For long term forecasts, if the forecast year and the observed year were within two years of another, no adjustment was made (e.g., for a forecast made in 2015, it would be ideal to have 2015 observed volumes; however, volumes as far back as 2013 were tolerated.) But if the difference is more than two years, linear interpolation technique was applied between the base year and the forecast year to determine an interim year forecast which was then compared with the available observed year volume. For example, Route 1 Corridor study

[25] showed a forecast for year 2020 with the base year of 1995; thus linear interpolation was applied to generate 2014 forecast and the observed volume of 2014 was used to evaluate the accuracy of the study. When an observed peak hour volume was needed, the ADT was multiplied by the K-factor as recommended by Jones [26]. The observed volumes could not be obtained for all the links forecasted due to lack of information and proper documentation in the study. For example, for Stonefield study [5], there were eight links that were modeled, however, observed data were only available for three of the links.

5.1.4 Summary of Data set

Table 5.1 lists 41 studies and data sources where one could compare a forecast value to an observed value. Column three of Table 5.1 reflects the methodologies used in the study. The total number of links for each of the 41 studies ranged from 1 link for a site passing study to 2,493 links for a statewide travel forecast. For example, George P Coleman bridge study [27] made one forecast for the bridge only. Again, York River Crossing [28] study forecasted for the links of entire area using a travel demand model. The original forecasted link reported in the study is different from the link shown in Table 5.1. Table 5.1 represents the number of links for which the observed data for forecast year was found and ultimately performed in the evaluation. For example, Route 29 Corridor study [6] used a travel demand model for Charlottesville area where forecast was made for 64 links. However, observed volume was found for only 17 links. Duration of forecast also varies by study. For example, Statewide Highway forecasts were made for over 20 years planning horizon. On the other hand, a land development study in Henrico county generated forecast for only 2 yrs. period just after the development was expected to finish. In some cases, when the base year was not available in the study, the published year was used in order to determine the duration of forecast. For example, the forecast duration of Route 360 Improvements to I-295 study [29] is 15 yrs. with a published year of 1990 and forecast year of 2005. For some of the studies, forecasts are iterative. For example, while Stonefield study made a forecast for year 2012 was reported, other earlier forecasts were made in 2001, 2002, 2005 and twice in 2009.

Table 5.1: Studies and data sources supporting a comparison of forecast and observed volumes

No.	**Title (Source)**	**Method**[a]	**Published Year**	**Base Year**	**Forecast Year**	**Unit**[b]	**No of Links**	**No of Recession(s)**[c]
1	Route 3 Corridor Study: Northern Neck and Middle Peninsula [8]	TRE	1988	1986	2010	ADT	19	3
2	George P. Coleman Bridge Financial Alternatives [27]	TRE	1989	1982	2010	ADT	1	4
3	US 15 James Madison Highway Passing Lane Study [30]	TRE	1997	1996	(2014)[d]	ADT	1	2
4	Colonial Heights Thoroughfare Plan [31]	TDM	1970	1966	1985	ADT	21	4
5	York River Crossing Travel Demand Study [28]	TDM	2000	1990	(2014)[e]	ADT	39	3
6	Interstate 66: Fairfax and Prince William Counties [32]	TDM	1986	1985	2010	PHV	11	3

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
7	I-95/Clermont Avenue Interchange & Connector to Duke Street Alexandria, Virginia [33]	TDM	1990	1988	2010	AWT	28	3
8	Route 13 Corridor Study [34]	TRE	1988	1987	2010	ADT	14	3
9	Tappahannock Area: Routes 17 and 360 Corridor Study [35]	TRE	1989	1988	2010	ADT	5	3
10	Route 40 Needs Assessment Study [36]	UK	1999	1996	(2014)[e]	ADT	17	2
11	Routes 20/240 Corridor Study, Albemarle County [37]	UK	1990	1987	2010	ADT	6	3
12	Route 608 Corridor Study: Augusta County [38]	TIA	1996	1994	2014	ADT	25	2
13	Botetourt County Route 220 [39]	TIA	1999	1994	(2014)[d]	ADT	3	2

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
14	Capital Beltway Study I-95/I-495 Northern Virginia [2]	TDM	1989	1988	2010	ADT	19	3
15	Route 221/460 Corridor Study Roanoke and Botetourt Counties [40]	TDM	2002	2000	(2015)[e]	ADT	3	2
16	Route 360 Corridor Study Town Of Warsaw [41]	TRE	1993	1991	2010	ADT	7	3
17	Dulles Toll Road Extension Route 267 Draft Environmental Document [42]	UK	Undated	1986	2010	ADT	8	3
18	Route 240 Corridor Study [43]	UK	1990	1987	2010	ADT	2	3
19	Pulaski Area-Year 2000 Transportation Plan [44]	TDM	1981	1980	2000	ADT	56	3
20	Route 29 Corridor Study [6]	TDM	1990	1987	2010	ADT	17	3

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
21	Peninsula Area Transportation Study [45]	TDM	1967	1967	1985	AWT	174	4
22	Hampton Roads Travel Demand Model [1]	TDM	2004	2000	2011	AWT	42	2
23	2010 Statewide Highway Plan: Culpeper District [46]	TRE	1989	1987	2010	ADT	240	3
24	Statewide Highway Plan: Thomas Jefferson PDC [47]	TRE	1984	1981	2005	ADT	79	3
25	Statewide Highway Plan: Richmond Regional PDC [48]	TRE	1984	1981	2005	ADT	46	3
26	Statewide Highway Plan: 5th Planning District [49] (VDH&T, 1984c)	TRE	1984	1981	2005	ADT	143	3

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
27	Statewide Highway Information Planning System (SHiPS) (an internal database)	TRE	Extracted in 2015	1994	(2013)[d]	ADT	2,493	2
28	University of Virginia Research Park [22]	TRE	2008	2006	(2014)[d]	PHV	9	1
29	Rivanna Village at Glenmore [50]	TIA	2001	2001	2006	PHV	4	1
30	Stonefield at Route 29 [5]	TIA	2010	2010	2012	PHV	3	0
31	Orange County 2020 Small Urban Area Transportation Plan [51]	TRE	2002	2000	2010	ADT	19	2
32	Route 360 Improvements East of I-295 [29]	UK	1990	N/A	2005	ADT	19	2
33	Route 10 in Chesterfield County [52]	UK	2001	2001	2014	ADT	4	1

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
34	Watkins Center Traffic Impact Study	TIA	2007	2005	2024	PHV	NA	NA
35	Trip Generation and Distribution Study, West Creek Parkway, Goochland County, Virginia [53]	TIA	1988	1997	2007	ADT	12	1
36	Route 1: Appomattox River Bridge and Approaches [54]	TDM	1989	1987	2010	PHV	9	3
37	Route 1 Corridor Study: Fairfax and Prince William Counties [25]	TDM	1997	1995	(2014)[e]	ADT	7	2
38	Richmond International Airport Corridor Feasibility Study Report [3]	TDM	1999	1998	(2014)[d]	ADT	23	2

Continued on next page

Table 5.1 – continued from previous page

No.	Title (Source)	Method[a]	Published Year	Base Year	Forecast Year	Unit[b]	No of Links	No of Recession(s)[c]
39	Intersection of Old Keene Mill Road and Rolling Road in Fairfax County [55, 56, 57]	NA	Obtained in 2015	2005	2012	NA	12[f]	NA
40	BJ's Wholesale Club Traffic Impact Assessment [58]	TIA	2007	2007	2009	PHV	10	1
41	Bell Creek Road intersection study [59]	UK	2002	2000	2014	ADT	3	2

[a]TDM = travel demand model, TIA = traffic impact study, TRE = Trend-based study, UK = unknown (e.g.,. methodology is not stated)

[b]ADT = Average Daily Traffic; AWT = Average Weekday Traffic; PHV = Peak Hour Volume

[c]Federal Reserve Bank of Richmond (2016) has a graphic titled Decomposition of Real GDP which shows when economic recessions have transpired. This column is the number of recessions between the base year and the forecast year inclusive.

[d]As the 2015 volumes were not available, volume for the year in parentheses was used to evaluate forecast accuracy.

[e]As noted in Section 6.3.2, the original forecast year was either 2020 or 2018, so linear interpolation was used to determine the forecast for the year in parentheses.

[f]As these were intersections, this field denotes the number of turning movements.

Two out of 41 studies were not used in the analysis of error statistics as shown in Table

5.1; study no 34 (relatively long duration of forecast for which no observed data could be found for the evaluation purposes) and study no 39 (as no forecast was generated by the study itself). Thus, the final dataset excluding statewide forecast (which itself contains 2,493 links) consisted of 1,159 links resulted from 38 studies.

5.1.5 Measure Accuracy by Comparing Forecast Volumes to Observed Volumes

Error is the difference between the forecast volume and the observed volume, such that a positive error means that the forecast was higher than the observed value.

When aggregating multiple forecasts, at least eight measures of accuracy are possible: mean error, mean absolute error, mean percent error, mean absolute percent error, median error, median absolute error, median percent error, and median absolute percent error. The first four measures are defined in Equations 5.1-5.4, with Equations 5.3 and 5.4 based on Wang [60]:

$$MeanError = \frac{1}{n}\sum_{i=1}^{n}(y_{forecast} - y_{observed}) \tag{5.1}$$

$$MeanAbsoluteError = \frac{1}{n}\sum_{i=1}^{n}|y_{forecast} - y_{observed}| \tag{5.2}$$

$$MeanPercentError = \frac{100\%}{n}\sum_{i=1}^{n}(y_{forecast} - y_{observed}) \tag{5.3}$$

$$MeanAbsolutePercentError = \frac{100\%}{n}\sum_{i=1}^{n}|y_{forecast} - y_{observed}| \tag{5.4}$$

where

$y_{forecast}$ = the forecast volume

$y_{observed}$ = the observed volume

n = is the total number of links or intersections in the study.

The last four measures: median error, median absolute error, median percent error, and median absolute percent error, are similar to Equations 5.1-5.4 except "median" (the 50th percentile) replaces "mean". For example, suppose that for a study with four links, the error (e.g., forecast minus observed volume) for each link are, -300, - 100, 100, and 700. Suppose further that each link has an observed volume of 1,000. The mean error, mean absolute error, mean percent error, and mean absolute percent error are, respectively, 100, 300, 10%, and 30%. The median error, median absolute error, median percent error, and median absolute percent error are, respectively, 0, 200, 0%, and 20%.

A set of expected forecast errors based on these results was developed, and these expected forecast errors formed the basis of the recommendation of this report.

5.2 Apply Eleven Techniques to Some of the Studies

In order to determine how other techniques, rather than the one already used by the study perform, one or two techniques from Table 3.1 were applied to some of the study. The applicability was determined by the data requirement to apply those techniques. If the data requirement satisfies studies' data, the technique is applicable to that study. For example, Technique 1 in Table 3.1 determines the difference between modeled ADT and the actual ADT in the base year and uses this difference to adjust the forecasted ADT. Thus, in order to apply Technique 1, modeled volume is required for the base year and forecast year. Hence, the technique can be applied to those studies which are based on a demand model. One of such study in the dataset is Route 29 Corridor study (Study 20) where a travel demand modelling package was used for the City of Charlottesville. Eleven techniques could not be applied to all 39 studies due to data unavailability. Hence, each technique was applied to at least one study. A review of total seven out of eleven techniques from Table 3.1 with their application to the studies are shown below.

5.2.1 Adjusted Regional Model Outputs

The first technique in Table 3.1, also known as "Post-processing model outputs daily link volumes forecast adjustments", corrects local errors in the travel demand model under the premise that, for a specific link, if the model's base year link volume is higher than the actual link volume, then the forecast year link volume will also be higher than the forecast year actual link volume. The method is applied as follows, using an example where, from study 20 [6], a link has an actual base year ADT of 19,400 that exceeds the modeled base year ADT of 23,717. Accordingly, the modeled forecast year link volume of 28,088 needs to be adjusted. The adjustment process requires three main steps: compute the delta factor and ratio factor for the base year, adjust the forecast by each factor, and then compute the average of the adjustments.

Application of technique with an illustration of study 20

1. Compute the delta factor and ratio factor for the base year. For example, the delta factor and ratio factor for one link of study 20 can be calculated as follows.

   The delta factor = modeled ADT - actual ADT = 23,717 - 19,400 = 4,317

   The ratio factor = modeled ADT ÷ actual ADT = 23,717 ÷ 19,400 = 1.22

2. Adjust the forecast by each factor. For example, the delta factor is subtracted from the forecast such that 28,088 - (+ 4,317) = 23,771 and the forecast is divided by the ratio factor such that 28,088 ÷ (1.22) = 22,975.

3. Compute the average of the adjustments. Thus, in this case–

   The average of the adjustments is (23,771 + 22,975) ÷ 2 = 23,373

This technique is also applied if the base year modeled ADT is lower than the base year actual ADT, note that sign retention in these three steps is essential.

5.2.2 Trend Analysis based on Regional Model Outputs

This technique determines annual rate of growth from modeled volume and forecasted volume and then applies the growth rate to existing modeled volume to get the future volume. The same sample data from study 20 is used for the application of technique 2.

Application of technique with an illustration of study 20

1. Compute annual growth rate (linear). The annual growth of the above link is as follows.

   Total growth is (28,088 - 23,717) $\div$ 23,717 = 18.4%

   Annual growth over 23 years is 18.4% $\div$ 23years = 0.8% per year

2. Apply growth rate to existing volumes to compute forecasts. Thus, in this case–

   Interim year forecast: 19,400 $\times$ (1 + (0.008 $\times$ 13)) = 21,417.6

   Design year forecast: 19,400 $\times$ (1 + (0.008 $\times$ 23)) = 22,969.6

Evaluation of Forecast

In order to determine which technique perform well in forecasting traffic, the forecast ADT of 23,373 (Technique 1) or 22,970 (Technique 2) are compared to the observed ADT of 19,000. If it were the case that the regional plan had forecasted AWDT, then the forecast would have been compared to the observed AWDT.

5.2.3 Linear Growth based on Two Traffic Counts

This is the third technique, also known as the “simplistic methodology” for a “traffic count trend analysis” entails the use of past traffic counts to forecast a future traffic count, where a linear change in ADT per year is calculated and then applied to a future year. For example, for a section of Route 3 between Route 301 and Route 205, study 1 [8] showed that the 1965 ADT was 1,810 and the 1985 ADT was 2,885, and a forecast for year 2010 is needed.

Application of technique with an illustration of study 1

The change in ADT per year is thus (2,885 - 1,810)/20 years = 53.75/year. Since the 2010 forecast is 25 years after 1985, the 2010 forecast is computed as 1985 ADT + 25 × (Change in ADT per year) or 2,885 + 25 × 53.75 = 4,229. Because the linear growth rate is based on only two data points, it is sensitive to which two points are selected. For example, for that link it would have been possible to base the calculation on a 1970 ADT (2,540) and a 1986 ADT (3,325), which would yield a forecast of 4,503–which is about a 6% difference from the forecast noted in the paragraph above.

5.2.4 Regression based on Multiple Traffic Counts

As is the case with technique 3, the fourth technique, also known as "least squares regression analysis" for a "traffic count trend analysis," uses historical traffic counts to make a forecast for the future year. However, multiple data points rather than just two are used to make the forecast. For example, for the same segment noted in technique 3–a section of Route 3 between Routes 301 and 205 from study 1 [8]–the following data were available for the period 1965–1986, where a forecast is needed for year 2010.

Table 5.2: Historical ADT for a Route 3 between Route 301 and Route 205

Year	1965	1970	1975	1980	1985	1986
ADT	1,810	2,540	3,160	3,645	2,885	3,325

Application of technique with an illustration of study 1

The technique of least squares linear regression, where the year is the independent variable and the ADT is the dependent variable, yields Equation 5.5. Note that Equation 5.5 can be obtained in three different ways: one can perform linear regression by hand as shown in the Guidebook, one can obtain this from a software package (in Excel, for instance, one uses the steps Data Analysis/Regression), or one can create a graph of the data in Table 5.2 and then add a trend line, with the equation, to the chart.

$$ADT = -115,514 + 59.9 \times (forecastyear) \tag{5.5}$$

For example, for a 2010 forecast year, Equation 5.5 yields a forecast of 4,885. There is some rounding of the coefficients in Equation 5.5 but this does not materially affect the results; without any rounding, the exact forecast from Equation 5.5 is 4,881.

Evaluation of Forecast

Thus the forecast of 4,229 or 4,503 ADT (Technique 3) or 4,881 ADT (Technique 4) is compared to the observed volume of 4,887 ADT.

5.2.5 Seasonal Adjustment Factors

The seasonal adjustment factor can be used to convert a count taken over a given period of time (say a Tuesday and Wednesday in March) to an ADT; one divides the periodic count by the seasonal factor to obtain the ADT (Equation 5.6). The process that VDOT uses to obtain ADTs is similar, where the axle count is multiplied by an annual axle factor (given that different types of vehicles have different numbers of axles) and then this is multiplied by the seasonal factor as shown in Equation 5.7. An example of Technique 5 for determining observed volume is shown using data from study 1 [8]– where the observed volume for a given link is recorded 6,069 ADT in the internal TMS database [24]. Table 5.3 shows how this observed ADT for year 2010 is calculated by the system. For example, the observed counts for the link were taken on Wednesday, August 18 and Thursday August 19 of 2010. These counts were 12,998 axles and 14,497 axles, respectively. Then, the axle counts are multiplied by the axle factor and seasonal factor for year 2010. The observed ADT 6,069 is calculated from the average of 5,876 and 6,262 in Table 5.3. The forecasts of seasonal ADT were also compared to actual seasonal ADT in studies 3 and 38.

$$ADT = Seasonalvolume/Seasonalfactor \tag{5.6}$$

$$ADT = (Seasonalvolume) \times (Annualaxlefactor) \times (Seasonalfactor) \tag{5.7}$$

Table 5.3: Verification of How Seasonal Adjustment Factors Are Used to Estimate ADT for a link of study 1

Day	Axle Count[a]	Axle Factor[b]	Seasonal Factor[c]	ADT
Wednesday August 18	12,998	0.4700982	0.961664	5,876
Thursday August 19	14,497	0.4700982	0.918854	6,262
ADT				6,069

[a]The axle count denotes the raw daily axle counts for that link taken on Wednesday or Thursday.

[b]The axle factor is determined by matching the TMS link id and corresponding factor group number from an internal software application provided by Jones [61]

[c]The seasonal factor is obtained using the factor group number for a particular day, month, and year (e.g., Wednesday in August in 2010) using the internal software application provided by Jones [61].

5.2.6 Peak Hour Link Forecasting

Technique 6 entails multiplying the forecast ADT by the expected proportion of ADT that will occur during the peak hour in order to obtain the peak hour volume. The Guidebook notes that this proportion of ADT will be determined from a "diurnal," which it defines as "the curve of the hourly volume flow of the daily traffic volumes." There are, in fact, multiple ways to determine the proportion of ADT that will occur during the peak hour: (1) generate a diurnal from a continuous count station (as suggested in the Guidebook), (2) assume the proportion will remain the same in the future as what it is at present, and (3) use literature which relates the peak-hour factor to congestion level; see, for example, Simons [62]. This technique is applied to the forecast data of study 1 and 3.

5.2.7 Institute of Transportation Engineers (ITE) based Factoring

Contrary to techniques 3 and 4, ITE based factoring is not based on previous traffic growth at a given location. Rather, technique 10 forecasts travel demand based on anticipated land development to forecast activity. However, a detailed nine-step procedure is available in Chapter 7 of the Trip Generation Handbook (ITE [4]). That procedure is illustrated with study 38 [58] in the subsections that follow.

Application of technique with an illustration of study 38

A proposed development is a wholesale club with 114,576 ft2 and 12 fueling stations, where the concern is additional traffic generated during the evening peak hour.

1. *Identify land use types, their corresponding ITE land use codes, and sizes.* The two land use types are Discount Club (code 857 with 114.6 square feet [in thousands]) and gasoline stations (code 944 and 12 fueling positions).

2. *Pick a time period for analysis.* The analysis time period is evening peak hour of adjacent street traffic.

3. *Compute baseline trip generation for individual land uses.* According to ITE [20] , the discount club will generate $(4.24 \times 114.6) = 486$ (average rate is 4.24 trips/1,000 ft2 for land use code 857) trips and the fueling station will generate $13.87 \times 12 = 166$ (average rate is 13.87 trips/fueling position for land use code 944) trips, for a total of 652 trips. The discount club and fueling station will generate entering trips of 243 and 83 respectively (entering trip is 50% of total trips). The number of exiting trips is identical.

4. *Estimate anticipated internal capture rate between each pair of land uses.* If both land uses are treated as retail (Note that although gasoline stations are considered to be "service" land use category in ITE [20], but it is considered as "retail" category in the

report by DKS Associates [63], then ITE [4] suggests that 20% is an appropriate rate, meaning that the trips in step 3 may be reduced by this amount.

5. *Estimate "Unconstrained Demand" Volume by Direction.* The directional trips from step 3 are multiplied by the percentage in step 4. For example, trips from the discount club to the fueling station are computed as $243 \times 20\% = 49$ exiting trips and also as $83 \times 20\% = 17$ entering trips.

6. *Estimate "Balanced Demand" volume by direction.* For each direction, select the lower value to be recorded as the "balance." In step 5, trips from the discount club to the fueling station were computed as both 49 trips and 17 trips; therefore, the controlling value is 17 internal trips.

7. *Estimate Total Internal Trips to/from Multi-Use Development Land Uses.* The total internal trips are 34, with 17 internal trips to and from the discount club. Overall, 7% of the discount club trips (34 of 486) are internal to the multi-use development.

8. *Estimate the Total External Trips for Each Land Use.* The external trip for the discount club are 243-17 = 226 entering and 226 exiting trips. For the fueling station, there are (83-17) = 66 external entering and exiting trips.

9. *Calculate Internal Capture Rate and Total External Trip Generation for Multi-Use Site.* From step 8, the entering volume estimate of 292 peak hour trips is the sum of the external trips entering the discount club (226 trips) and the fueling station (66 trips). Since the exiting trips have the same value, the net external volume for the multi-use site is $(292 + 292) = 584$ trips. Compared to the 652 trips in step 3, the value of 584 trips represents a trip reduction of 10%. Figure 5.1 summarizes the calculations from steps 3-9 based on ITE [4].

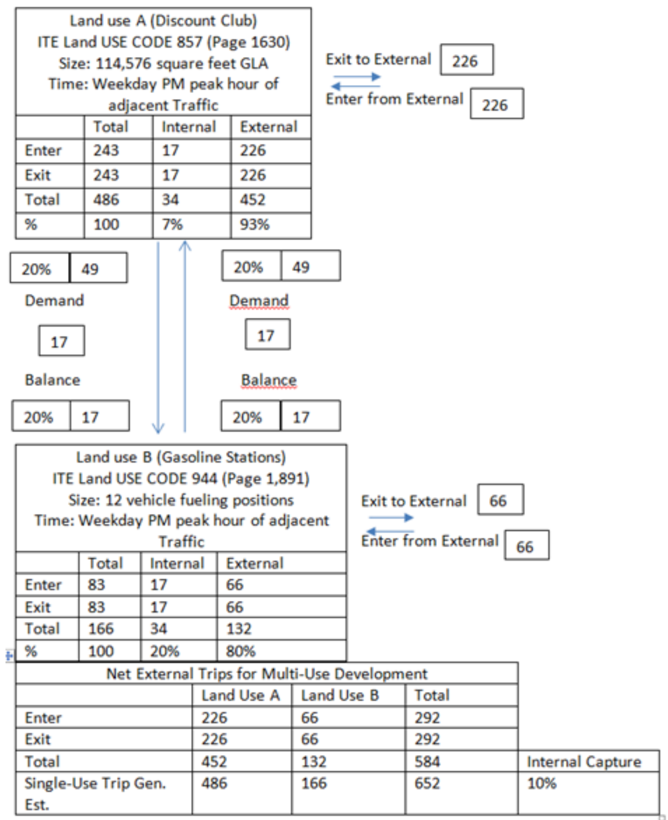

Figure 5.1: Multi-Use-Trip Generation Calculation (adapted from ITE [4])

Evaluation of Forecast

The study (VHB [58]) projected volume for 2009 at the adjacent street and generated site volume considering BJ's wholesale club will be built by 2009. The forecast volume is obtained by taking the summation of the volume if no development is built in 2009 and "site volume" generated by the new development. For example, the study forecasted peak hour volume for Starling Drive, a roadway along the club to be 828 in 2009. Since the study shows turning movement at various intersection, all the Northbound (293), Southbound (487) and turning volumes (0, 12, 12 and 24) to and from the counter location are added to determine the total volume of 828 (Figure 6, VHB [58]). Based on the study, site volumes are generated by some percentage of total entering traffic of 292, derived from the ITE trip generation described above. The total site volume (roughly 264) is determined by adding two directional (40%) and two turning (5%) volume (Figure 10, VHB [58]). Therefore, the total peak hour volume projected by the study is (828+264) = 1,092 for Starling Drive. The volume was taken on weekday of peak hour (5-6 PM) in 2009 and compared to the observed peak volume for 2010, collected from the Henrico County Public Works (Smidler, [64]).

Technique 7 ("ESAL development") and Technique 9("population-based forecasts") were applied to the Route 3 Corridor study (study 1). Technique 11 (the traffic shift methodology for corridors) was applied to a new data set that consisted of two parallel roads which was suitable for that particular methodology (see Appendix A of [65] for details).

5.3 Document Assumptions While Comparing Forecast to Observed Volume

In order to compare forecast to actual counts, the following methodology was used to obtain the desired count in studies where start and end point do not match. This assumption was considered in both cases; (1) accuracy of original forecast study as shown in Table 6.1 and (2) accuracy of other forecast techniques to the same study as shown in Table 6.2. Study-specific assumptions are given in Section 6.3.2.

There is Not a One-to-One Relationship Between Forecast and Observed Volumes

The difference in referencing systems between the roadway network used in the study and the roadway network which held observed counts needed to be understood. For example, in study 1, a 2010 forecast is given for a Route 3 segment between Route 301 and Route 205 [8]; however, this segment reflects two sub-segments in the observed traffic counts (VDOT [23]): a 7.18 mile section with a volume of 5,000 ADT and a 2.84 mile section with a volume of 4,600 ADT. Accordingly, a weighted volume of 4,887 is computed based on these two sub-segments as reported in [66] (see Equation 5.8).

$$\frac{7.18miles}{7.18miles + 2.84miles} \times (5,000ADT) + \frac{2.84miles}{7.18miles + 2.84miles} \times (4,600ADT) = 4,887 \tag{5.8}$$

As noted for several other studies, the link for which a volume is forecast is only a portion of the link for which an observed volume exists. Consider Hydraulic Road in Study 28 and presume that Solomon Road and Commonwealth Drive are sufficiently close (a distance of 0.03 miles) that they can be treated as the same point. Although there is just one observed volume for that section, the study shows three separate forecasts for this section: Route 29 to Swanson Drive, Swanson Drive to Cedar Hill, and then Cedar Hill to Commonwealth Drive. Accordingly, a weighted average of these forecasts was used as per Equation 5.8.

5.4 Identify Explanatory Factors

The accuracy of forecast not only depends on which method of forecast has been applied to the study but also depends on various other factors (i.e. duration, geographic locations, economic factors etc.) In order to determine which factor is the most important to impact accuracy among all other factors, an analysis of variance with a simple hypothesis test was used. In addition to the individual study, four explanatory factors were identified : (1) Forecast Method, (2) Economic Recession, (3) Forecast Duration, and (4) Forecast Unit.

5.4.1 Forecast Method

Forecast method refers to the underlying methodologies or techniques that were used to generate forecasts for some future time period. The dataset of 41 case studies (39 used in the analysis) are based on different forecast methodologies which can be divided into two categories.

Trend Based Forecast: Studies where the travel forecast is generated based on the extrapolation of previous years' traffic count in single or multiple corridor improvement projects are defined as trend based studies. Most of the studies with the word "corridor" in it used some trend based method to generate forecast (exception is Route 29 Corridor Study which was based on a regional travel demand model for Charlottesville city). This categorization also applies to the statewide forecasts shown in studies 23, 24, 25, 26 and 27 in Table 5.1. For study 28 (University of Virginia Research Park), most of the proposed development was not built, and thus the focus of the evaluation was on the growth in background traffic (which was a trend-based forecast as noted in Table 5.1). The forecast methodologies for some of the studies are not clearly stated in the report. Those studies were also treated as trend based studies. 20 out of 41 studies in the database are categorized as trend based forecast studies.

*Activity Based Forecast:*The forecast which is based on some future activity (land development, demographic change etc.) is categorized as activity based forecast. Travel demand model and traffic impact analysis studies fall under this forecast method. Travel demand model studies use four step regional travel demand model and traffic impact studies use the Institute of Transportation Engineers (ITE) trip generation method to generate forecast. 20 out of 41 studies in the database are categorized as activity based forecast.

5.4.2 Economic Recession

Another factor that impacts the traffic forecast is the unanticipated economic or operational changes that may have occurred after the forecast was generated. This is a type of variable which is beyond the forecasters' control. For example, for the Coleman Bridge (study 2), the toll was not precisely what was expected (which could logically affect travel demand), and studies that were based on expected land use changes (such as travel demand modeling efforts) would not have foreseen the impacts of economic recessions that have been observed in Virginia. The number of occurrence of economic recessions between the base year and forecast year "inclusive" was determined based on Federal Reserve Bank of Richmond (National Economic Indicators, [67]). The word "inclusive" means that if an economic recession began or terminated in the base year or forecast year, that recession was considered to have occurred between the base year and the forecast year. For example, three recessions (1991, 2002 and 2009) were recorded between the base year (1987) and the forecast year (2010) inclusive for the Route 29 corridor study as shown in Table 5.1.

5.4.3 Forecast Duration

The duration of the forecast is the difference between the forecast year and the base year. For example, a 2007 traffic impact study made a forecast for year 2015, so the duration is 8 years. For studies where an observed year was earlier than the forecast year, duration was computed as observed year minus base year. Clearly many of these studies support a (roughly) 20 years long range planning horizon; for instance, the I-95 Clermont Avenue Interchange study, published in 1990, provided forecasts for year 2010. However, some studies reflect a shorter time frame, especially some of the land development studies. Notably, studies 30 and 40 (Stonefield at Route 29 and BJ's Wholesale Club) reflect duration less than 5 years; in those cases, immediate land development was expected.

5.4.4 Forecast Unit

Some studies made a forecast of average daily traffic (ADT); but others made a forecast of average weekday traffic (AWT) or the weekday peak hour volume. This is a factor that differentiate among the forecasts for ADT and the forecasts for more detailed variables that build on ADT, such as peak hour volume. As evident by 6.4, the impact on forecast accuracy made by ADT is different than the impact made by peak hour volume. Out of 41 studies, 30 studies made forecast by ADT, 3 studies by AWT and only 7 studies by peak hour volume. For study no 39, forecast volume was not available.

Chapter 6

RESULTS & DISCUSSION

Based on the objective, four sets of results are reported. Each result represents the corresponding task of the methodology.

1. Overall accuracy of Virginia forecasts
2. Accuracy of Forecast Techniques
3. Document assumptions in determining accuracy
4. Contributing factors

6.1 Overall Accuracy of Virginia Forecasts

Table 6.1 summarizes the eight measures of accuracy for each of the 39 Virginia studies. The error statistics are calculated based on Equation 5.1-5.4 where the positive error means over forecast and negative error means under forecast. There are one or more links in a study where forecast volume is higher than the observed volume and one or more links where the forecast volume is lower than the observed volume. The positive and negative error cancels in those links resulting in a lower average or median value than links with all positive errors or all negative errors. For the same reason the absolute error is higher than the mean or median error. This is the case for 26 studies; the remaining 13 studies' mean error and absolute error are identical since the errors in all the links are positive. For example, Colonial Heights study [31] has a mean error of 5,524 which means the average difference between the forecast and observed volume of 21 links is 5,524 vehicles. Out of 21 links, 14 links show over prediction and 7 links show under prediction– which is why the mean error (5,524) is lower than the absolute error (7,690).

The differences in temporal and spatial resolution (see Figure 6.2) indicate difficulty in performing retrospective evaluations over a long period of time. However, a question for further research is how results from other locations compare to those reported here. For example, consider the difference between mean and median error. Table 6.1 suggests that reporting a mean percent error is appropriate only if it does not differ substantially from the median percent error. If these two statistics differ, then there may be a few links with very large errors and many links with smaller errors, such that a confidence interval (e.g., x% of errors are no greater than z%, as done by Welde and Odeck [15]) is more descriptive. It would be helpful to know, based on a larger data set, the types of error distributions that should be expected in forecasts.

For each link, error divided by the observed volume, gives a percent error, such that a link with a perfect forecast has an error of 0%. For Colonial Heights study, each of the error of 21 links is divided by the corresponding observed volume resulting in the average percentage error of 32%. For a larger volume road, the magnitude of error is higher compared to a smaller volume road– in such cases percentage error performs well. For example, the difference of 1,000 vehicles is not large for a road with 10,000 vehicles per day whereas the difference is a concern for a road with 2,000 vehicles per day. The percentage difference of vehicles for the former case is 10% whereas for the latter case it is 50%. Sometimes, there are links with very large errors which influence the mean value– in such cases, the median value can be a useful measure of accuracy. For Colonial Heights study, as only one link has a large error (forecast volume of 5,800 where observed volume is 1,700), the mean (32%) and median

(31%) error for the study is nearly the same. However, for Route 3 Corridor study, two of the 19 links showed large errors (7,100 and 10,125 vehicles per day)– so the median error (41%) is considerably smaller compared to the mean error (61%).

The error varies by study and in some cases, differs up to two orders of magnitude as shown in Table 6.1. The most accurate study is study 18 (Route 240 Corridor study) with the forecast volume and observed volume differing by only 35 vehicles–about 1% of the observed volume. The least accurate study is study 13 (Botetourt County Route 220) which shows an average error of almost 22,000 vehicles per day, such that the error alone was greater than the observed volume (Mean error is 107% and Median error is 134%) in Table 6.1.

For most of the cases with the studies in Table 6.1, forecast volume is higher than the observed volume. About three quarters of the studies' mean error (29 studies) or median error (31 studies) was positive which showed that the studies had been over forecasted. The last two rows of Table 6.1 indicate that the "mean of the means"– that is, the average value of the mean percent errors from each of the 39 studies–is 43%. Since positive and negative errors tend to cancel when not using absolute values, and because a few studies with very large errors (or very small errors) can influence a mean value, the median absolute error can be used to better represent accuracy. That is, Table 6.1 suggests that for the 39 studies, the median absolute percent error ranged from a low of 1% (most accurate) to a high of 134% (least accurate). The average value of the median absolute percent error for all studies was about 40%. That is, for a given study where one has no information about the technique used, forecast duration, or other explanatory variables, on average one would expect the median absolute error for all the links in the study to be about 40%. However, the results in Table 6.1 can be influenced by the following items: (1) the presentation of "observed" volumes are not perfect. The observed volumes used to estimate error in Table 6.1 are, in fact, estimates (e.g., they are obtained from temporary counts and then adjusted based on seasonal factors to provide an annual estimate). Again, in some cases, the observed volume is not for the same year as forecast volume. For example, for study 38, whereas the forecasts were for year

2015, the observed volumes available are for year 2014. (2) some of the studies in Table 6.1 include peak hour volumes rather than ADT or AWT which may affect the magnitude of the error.

Table 6.1: Summary of Error Statistics for 39 Studies

Study No. (1)	Mean Error (2)	Mean Absolute Error (3)	Mean Percent Error (4)	Mean Absolute Percent Error (5)	Median Error (6)	Median Absolute Error (7)	Median Percent Error (8)	Median Absolute Percent Error (9)
1	3,079	3,079	61%	61%	2,430	2,430	41%	41%
2	7,178	7,178	24%	24%	7,178	7,178	24%	24%
3	4,375	4,375	73%	73%	4,375	4,375	73%	73%
4	5,524	7,690	32%	53%	5,830	6,170	31%	45%
5	4,992	6,108	102%	118%	1,431	2,128	56%	58%
6[a]	-73	934	-5%	22%	-325	854	-13%	20%
7	10,861	11,844	69%	72%	8,577	8,865	50%	50%
8	1,964	2,735	12%	16%	1,829	2,404	16%	16%
9	6,388	6,388	112%	112%	4,521	4,521	57%	57%
10	756	1,439	39%	53%	812	972	32%	39%
11	-1,110	1,227	-18%	20%	-430	570	-6%	8%
12	6,737	7,017	118%	126%	3,749	3,749	71%	71%
13	21,959	21,959	107%	107%	18,971	18,971	134%	134%
14	41,726	41,726	24%	24%	40,471	40,471	22%	22%
15	7,503	7,503	30%	30%	6,708	6,708	31%	31%
16	3,804	3,804	66%	66%	4,266	4,266	37%	37%

Continued on next page

Table 6.1 – continued from previous page

Study No. (1)	Mean Error (2)	Mean Absolute Error(2)	Mean Percent Error (4)	Mean Absolute Percent Error (5)	Median Error (6)	Median Absolute Error (7)	Median Percent Error (8)	Median Absolute Percent Error (9)
17	-7,225	9,775	-16%	23%	-5200	6,900	-14%	19%
18	-34	35	-1%	1%	-34	35	-1%	1%
19	2,118	2,492	61%	68%	2,488	2,570	49%	51%
20	3,508	4,055	52%	58%	2,585	2,585	15%	22%
21	-5,887	8,100	2%	56%	-1175	3,605	50%	38%
22	1,036	4,689	14%	47%	570	3,136	4%	35%
23	1,153	2,419	33%	48%	549	987	19%	31%
24	-539	1,449	-3%	36%	-276	664	-13%	26%
25	-972	2,425	12%	48%	-93	864	-6%	37%
26	-216	1,590	58%	70%	210	439	25%	36%
27	4,258	4,368	112%	113%	1,794	1,854	72%	72%
28[a]	383	383	35%	35%	141	141	29%	29%
29[a]	-6	86	5%	10%	40	61	8%	10%
30[a]	1,016	1,016	40%	40%	509	509	35%	35%
31	1,358	1,697	22%	27%	1,064	1,330	16%	17%
32	2,304	3,157	58%	64%	810	1,000	59%	59%
33	4,450	4,450	40%	40%	4,665	4,665	34%	34%
35	-7,689	8,683	-20%	33%	-5,219	5,219	-19%	24%
36[a]	776	776	103%	103%	721	721	86%	86%
37	22,006	22,006	78%	78%	21,502	21,502	65%	65%

Continued on next page

Table 6.1 – continued from previous page

Study No. (1)	Mean Error (2)	Mean Absolute Error(2)	Mean Percent Error (4)	Mean Absolute Percent Error (5)	Median Error (6)	Median Absolute Error (7)	Median Percent Error (8)	Median Absolute Percent Error (9)
38	3,452	9,217	8%	43%	1,405	8,439	7%	36%
40[a]	142	251	13%	17%	131	234	10%	12%
41	11,847	11,847	114%	114%	11,740	11,740	40%	40%
Min	-7,689	35	-20%	1%	-5,219	35	-19%	1%
Max	41,726	41,726	118%	126%	40,471	40,471	134%	134%
Mean	4,177	6,153	43%	55%	3,829	4,970	31%	40%
Median	2,118	4,055	35%	48%	1,405	2,570	31%	36%

[a] For studies 6, 28, 29, 30, 36, and 40, error statistics reflect the peak hour volume rather than the 24 hour volume.

6.1.1 Two Caveats in Assessing Accuracy

While determining the accuracy of Virginia studies, two key points were noted that might cause variations in the way forecast accuracy is assessed.

1. There Are Multiple Ways to Summarize Forecast Error

2. Accuracy may vary depending on chosen observed volume

There are Multiple Ways to Summarize Forecast Error

Route 3 corridor study forecasted ADT volumes for 19 segments. There are several ways to tabulate the 19 link errors.

1. To sum the difference between forecast and observed values, such that positive and negative errors cancel. This forecast is about 108,000 ADT–about 7% under the observed corridor ADT (116,488).

2. If the average percent error is computed by segment, the average error would be reported as only 3%; another low percentage results because the presence of positive and negative errors reduces aggregate error.

3. If an absolute value of the errors is computed, the average of the link-by-line errors is 30% as shown in Figure 6.1.

A related complication is whether to report the median or mean error. For example, in 1994 a trend-based statewide system forecast ADTs for 2,493 links for a 2015 horizon year for which the misalignment between forecast and observed links was avoided. When these 2015 forecasts were compared to observed 2013 volumes (as a surrogate for 2015 volumes), the MAE was 4,368. However, the median absolute error for these links was only 1,854–less than one-half the mean value–because several links had large forecast errors. The choice of whether to exclude links where unforeseen changes occur can influence the reported accuracy. For example, one link for Route 3 study (Link 9 in Figure 6.1) shows a 290% error. When the forecast was made in 1985, Route 3 shared this link with another large facility (Route 360). By the horizon year (2010), Route 360 had been reconstructed and no longer shared the link with Route 3-a change that was not anticipated. When it is removed from the data set, the MAE falls by one half, from 30% to 15%.

Accuracy may vary depending on chosen observed volume

The imperfections in observed volume may change the reported accuracy since observed volume is the basis on which the accuracy of forecast is evaluated. Observed volume may come from–(1) a continuous count station which gives a true estimate of annual average daily traffic (AADT), (2) special count over a short period of time and adjusted to estimate AADT

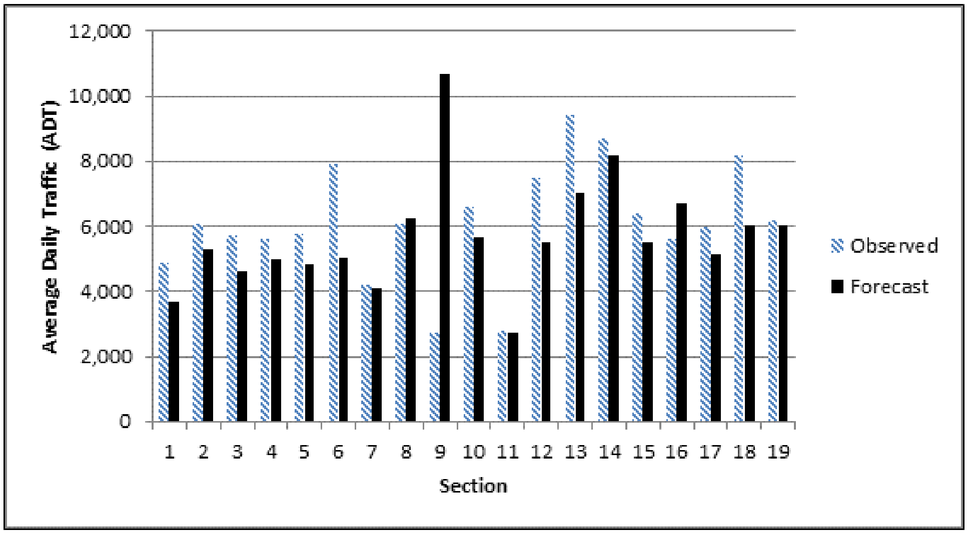

Figure 6.1: Difference between observed and forecast volumes for the Route 3 corridor study.

and (3) not an observation but rather an historical estimate. For example, the observed volume of Route 360 for the year 1989 was 30,135 ADT. However, a special count taken for the same year was 29,250 ADT. For whichever volume one chooses, the difference of 1,000 ADT (or 3%) is a rough indication of noise which will be present in the data.

6.2 Accuracy of Forecast Techniques

As described in Section 5.2, in order to verify how other techniques would have performed in terms of forecasting traffic volume, each of the techniques was applied to at least one study. The verification was made with determining the accuracy of those techniques in forecasting traffic as opposed to the original one already applied in the study. For example, Technique 1 and 2 was applied to study 20 as shown in Table 6.2; technique 0 is the original forecast. Technique 1 and 2 showed an error (median percent error) within the range between 25% and 35% when applied to study 20. However, the original methodology in this case, had done a good job in forecasting traffic with an error of 15%. Table 6.2 shows the performance of technique 1 and 2 along with the original method in terms of eight error statistics.

Table 6.2: Summary Error Statistics for Technique 1 and 2 to Study 20

No.	Forecasting Technique	Mean Error	Mean Absolute Error	Mean Percent Error	Mean Absolute Percent Error	Median Error	Median Absolute Error	Median Percent Error	Median Absolute Percent Error
20	0	3,508	4,055	52%	58%	2,585	2,585	15%	22%
	1	5,074	5,139	54%	56%	3,900	3,900	26%	26%
	2	5,538	5,615	57%	59%	3,783	3,783	34%	34%

Technique 1 when applied to the 17 links of the study shows that the average difference between the adjusted forecast volume and the observed volume is 5,074 ADT, where one link has a higher-than-predicted value and the remaining 16 links have a lower than predicted value. The median absolute percent error for this technique is 26%. Accordingly, the average difference between the forecasted ADT and the observed ADT for technique 2 is 5,538 ADT with median absolute error of 34%.

All three techniques have positive errors, which mean those techniques made an over forecast of traffic volume. Also, they appeared to show relatively similar range of errors. Thus, in this particular case, the two adjustment techniques did not materially improve forecast accuracy relative to original forecast–in fact, they nominally lowered the accuracy.

Similarly, technique 3 (which uses only two years of traffic volumes) and technique 4 (which uses multiple years of traffic volumes) were applied to four of the corridor studies (study 1, 8, 9 and 10) as shown in Table 6.3. Technique 3 and 4 are commonly used forecast techniques and can be applied to several studies since they do not require much data.

The last row of Table 6.3 indicates the p-value for the particular error statistic based on the two techniques. For example, a paired t-test between technique 3 and technique 4 shows that the difference in the mean error for these two tests has a p-value of 0.25, meaning there is not a statistically significant difference between technique 3 and technique 4 in terms of how they affect the mean error. The most accurate technique used in this study was Technique 4, which was extremely accurate when applied to this particular case study. Technique 4 is also a fairly reliable Technique when forecasting because of its simplicity and straight forward method of using past ADT values to predict future ADT values, rather than looking at other attributes.

Table 6.3: Summary Error Statistics for Technique 3 and 4

No.	Forecasting Technique	Mean Error	Mean Absolute Error	Mean Percent Error	Mean Absolute Percent Error	Median Error	Median Absolute Error	Median Percent Error	Median Absolute Percent Error
1	3	1,071	1,600	27%	35%	515	7%	1,080	18%
	4	1,436	1,791	34%	39%	1,102	20%	1,336	22%
8	3	3,486	3,357	21%	21%	2,662	18%	2,551	15%
	4	43	1,837	0%	11%	-452	-3%	1,596	8%
9	3	4,412	4,412	54%	54%	2,615	33%	2,615	33%
	4	1,632	4,294	29%	45%	3,457	43%	4,205	43%
10	3	-871	1,702	-17%	58%	-694	-37%	1,650	59%
	4	-825	2,066	-6%	65%	-439	-14%	1,518	51%
p-value		0.23	0.57	0.51	0.68	0.73	0.56	0.74	0.96

There is a great deal of variation in Table 6.3: the mean absolute percent error for the four studies, for example, ranges from a low of 11% (based on technique 4 for study 8) to a high of 65% (the same technique but for study 10). However, there is not a statistically significant difference between the accuracy of these two techniques based on any of the eight error measures shown in Table 6.3. That is, all of the p-values shown in Table 6.3 exceed 0.05.

6.3 Document Assumptions in Determining Accuracy

Two type of assumptions were reported while determining the accuracy of eleven techniques as well as original forecasts.

1. Assumptions in applying techniques

2. Assumptions reported in each data source (39 studies)

6.3.1 Assumptions in Applying Techniques

When determining the accuracy in Table 6.2 and Table 6.3, planners have to make several assumptions whether implicit or explicit in order to apply techniques to the data from previous studies. Five assumptions are listed while applying those techniques to the studies:

1. The technique may require input data which themselves must be forecast.
2. Technique selection entails an implicit judgment about whether historical trends or expected future activity is a better predictor of future travel demand.
3. Additional data may not necessarily improve forecast accuracy.
4. Data requirements is not the same for all the techniques.

The Technique May Require Input Data That Must Be Forecast

One observation is that errors can compound when techniques that build on other techniques are applied. For example, consider technique 3 (Table 3.1) (trend line forecasting) and technique 6 (computation of a peak hour volume given a K-factor). Since technique 6 requires two forecasting elements– first an ADT from technique 3 and then a K-factor to obtain a peak hour volume– it is possible for errors to compound. This was observed in evaluating the third study, where a forecast ADT of 10,360 exceeded the observed ADT of 5,985 by 73% (technique 3) yet the peak hour forecast volume of 1,036 exceeded the observed hourly volume of 548 by 89% (technique 6). A contributing factor to the greater error of technique 6 was that the K-factor had decreased from a forecast value of 0.10 to an observed value of 0.092. Table 6.4 shows results for the three studies that forecast both an ADT and peak hour volume: in all three cases, the use of the extra data element– the K factor–led to a larger median percent error than was the case with ADT. (The difference is significant; $p = 0.04$).

As a second example, consider the traffic shift methodology for corridors (Martin and McGuckin [18]), which can be derived from the theory of utility maximization and which

Table 6.4: Accuracy for Three Studies That Forecast Both an ADT and a Peak Hour Volume

Study	Forecast year	Average Daily Traffic		Peak Hour	
		Median Erroe	Median Percent Error	Median Error	Median Percent Error
2	2010	7,178	24%	712	34%
31	2006	3,125	19%	1,131	85%
31	2014	4,665	34%	1,211	100%
39	2003	8,200	30%	771	57%
39	2014	11,740	40%	1,152	52%

forecasts how an improvement to a given route will attract traffic to that route from alternative routes. An example is two roughly parallel routes in Virginia (I-64 and U.S. 60) where a lane was added to I-64, with construction completed in 2006. Based on the volumes in 2001 (prior to the improvement and a year for which volumes are available) and a forecast of how the improvement would affect speed on I-64, the change in volume for each route in 2007 (after the improvement has been made) can be forecast. If the total volume (from both routes) is known for the forecast year of 2007, the method is strikingly accurate, with an average absolute percent error of about 2%. However, because the total volume decreased–which might not have been expected in this urban area–the average absolute percent error is about 27%. To be clear, although the purpose of the traffic shift methodology for corridors is to forecast diversion based on an improvement, the method's accuracy is affected by the accuracy of its inputs.

As a third example, consider techniques 1 and 2 which adjust regional travel demand model outputs based on differences between the observed and forecast volume. Technique 1 and 2 needs input data from Technique 0 which themselves must be forecasts. For example, in order to apply technique 1 in one link of study 20, original forecast of 28,088 ADT is required which results in an output forecast of 23,373 ADT as shown in Section 5.2. Table 6.2 shows that the output forecast errors from technique 1 and 2 are higher than the input forecast errors from technique 0.

Technique Selection Entails Implicit Judgment About Whether Historical Trends or Expected Future Activity Is Better Predictor of Future Travel Demand

Application of the techniques as shown in Section 5.2 suggests that most techniques can be split into two categories based on the data elements that are required.

1. *Some of the techniques are based on future expectations of land and demographic activity.* Techniques 1 and 2 (which adjust outputs from the regional travel demand model), technique 9 (which uses an estimate of future population growth), and technique 10 (ITE-based factoring) are largely based on one's expectation of how land development or population will change. Thus, if one believes that a change in land development–whether through population growth or estimates of what will be built–will drive travel demand, then one of these techniques is appropriate.

2. *Some of the techniques are based on historical trends of travel demand.* Techniques 3, 4, 5, and 8, while they all encompass different methodologies, essentially use existing patterns of travel to forecast future travel patterns. Techniques 3 and 4 extrapolate observed traffic counts for an annual average 24-hour day, technique 5 uses previous seasonal trends to convert a forecast average to a count one would expect on a specific day of the week for a given month, and technique 8 estimates intersection turning movements based on historical movements. Thus if one believes that the past is a good predictor of the future, these techniques can be appropriate.

Two techniques do not fall neatly into the above two categories because they can build on forecasts provided by the above techniques. Technique 7 estimates equivalent single axle loads based on forecast vehicle types for a given section of roadway. Technique 11 determines how a given travel time improvement for one facility will affect the quantity of traffic attracted to it from other parallel facilities. Both techniques require an input forecast volume: for technique 7, this forecast volume must be decomposable into heavy trucks, medium trucks, and autos and for technique 11, this forecast volume represents the total volume on two or more parallel routes.

Additional Data May Not Necessarily Improve Forecast

Table 6.2 showed that the inclusion of additional information–the difference between the base year modeled and observed volume, as illustrated by Technique 1–did not materially improve forecast accuracy relative to not applying the technique at all (and simply using the travel demand model forecast alone). As another comparison, technique 3 and technique 4 were applied to four of the corridor studies shown in Table 6.3. The accuracy of the studies did not change substantially by applying those techniques rather than those originally used in the studies. Again, a statistical test conducted with the four corridor studies shows that technique 3 and technique 4 are not statistically significant at 95% confidence interval in terms of impacting the forecast error. Thus, applying other forecast technique does not necessarily improve the accuracies.

Data Requirements is Not the Same for All Techniques

It is somewhat difficult to directly compare the number of data elements required for each technique because not all data elements require the same degree of effort. For example, techniques 1 and 4 each require, at a bare minimum, just three data elements. Technique 4 requires a set of historical traffic counts in order to use a linear trend line to forecast future traffic counts, and at a minimum, a trend line could be established with three observations. Technique 1 requires an observed volume from a base year, a forecast volume for both the base year and forecast year from a regional travel demand model. However, execution of a regional travel demand model (associated with technique 1) requires considerably more effort than obtaining traffic volumes from historical data sources. Because a travel demand model requires additional information (i.e. demographic and employment information) along with traffic data. Accordingly, Table 6.5 compares the data required for each of the techniques in matrix form, recognizing that some techniques require more data elements than others.

Table 6.5: Data Elements Required for the 11 Techniques Given in Table 3.1

Data Element	**Technique**										
	1	2	3	4	5	6	7	8	9	10	11
Base year ADT forecast from regional demand model	X	X									
Future year ADT forecast from regional demand model	X	X									
Base year ADT	X		X	X				X	X	X	
Observed ADT from any two past years			X								
Observed ADT from any three or more past years				X							
Annual axle factor					X						
Seasonal factor					X						
Hourly volumes from a comparable route or forecast year volume/capacity ratio						X					
Number of lanes for the facility							X				
Percent cars							X				
Percent single unit trucks							X				
Percent tractor trailer trucks							X				
Base year intersection turning movements								X			
Base year population									X		
Forecast year population									X		
Forecast year land use types										X	
Free flow speed, capacity for base year parallel routes											X
Free flow speed, capacity for forecast year parallel routes											X
K-factor to convert ADT to peak hour volume						X					X
A forecast year ADT from techniques 1, 2, 3, 4, or 9					X	X	X	X			X

Table 6.5 shows that five of the techniques (5 [seasonal adjustment factors], 6 [peak hour link forecasting], 7 [ESAL estimation], 8 (Fratar) and 11 (traffic shift methodology) themselves require a forecast year volume generated from techniques 1 or 2 (along with a travel demand model), 3 or 4 (extrapolation of volume), or 9 (population based forecast). Technique 11 is a special case where the forecast for each parallel route can again be used as

the input for techniques 5 through 8. Table 6.5 also shows that except for a base year estimate of ADT, several of the techniques do not have overlapping data elements. For example, a set of historical traffic volumes is needed for techniques 3 and 4 in order to develop a trend line, but not for ITE-based factoring which requires land use codes for use with Trip Generation (ITE, [20]). Estimation of equivalent single axle loads requires some way of determining the portion of ADT that is heavy trucks, medium trucks, and passenger automobiles–but such vehicle classifications are not needed for the other techniques. In terms of data availability, two caveats are noted. First, travel demand model is applied to Virginia's MPO areas. Based on a review of VDOT [68], about half (61) of Virginia's 134 independent cities and counties are entirely or partially located within an MPO area. Techniques 1 and 2 can be used to the data available for locations under MPO areas. Second, the road which is city maintained, does not have continuous count rather periodic counts and no routine approach to convert annual average volume (Smidler [64]).

6.3.2 Assumptions Reported in Each Data Source (39 Studies)

A review of 39 Virginia studies enabled the identification of the following study specific assumptions that was encountered by the research team. Assumptions varied by each study (see Appendix B of [65]) which included topics such as classifying the forecast methodology for the study, determining what was built (for studies that have multiple forecasts depending on which improvements are constructed), connecting the forecast year (from the study) to the observed year (based on available data), and other factors that might have contributed to forecast error. The summary of those assumptions reported in the study are as follows.

1. The definition of a traffic forecasting study may vary
2. The number of links that used in the analysis and for which forecast was made vary
3. Background methodology is not clear
4. How are temporal and spatial ambiguities addressed?

5. Forecast volume for AWT, yet observed volume for ADT
6. The Year of forecast volume differs from the year of observed volume
7. Which alternative to choose?
8. Seasonality may or may not be Addressed in forecast or observed volume
9. Methodology that fall under both categories
10. Alternative that reflects the present scenario
11. Observed volume could not be obtained for forecast year; hence interim forecast is used
12. Is proposed alternative built or not?
13. Turning movement was converted into link volumes

The Definition of a Traffic Forecasting Study may Vary

The 38 studies included statewide forecasts, regional travel demand models, corridor studies, and site impact studies and, hence, different methods and level of details. For example, the statewide forecast is a trend-based projection for thousands of links throughout Virginia; corridor studies, focus on a smaller location and incorporate more details. Relative to trend-based forecasts, site impact studies and regional travel demand models use a fundamentally different approach, where estimates of activity (e.g., changes in population, employment, and developed land) forecast travel. Although many factors influence forecast accuracy, the median absolute percent error varies by the subset of studies where accuracy was determined, from relatively low values of 12% (for a traffic impact study [58]) or 28% in average for the two regional travel demand models (19,20) to a higher value of 72% (for an aggregate statewide study).

The Number of Links That Used in the Analysis and for Which Forecast was Made Vary

Links that could not be matched to locations in the VDOT TMS database (VDOT [24]) or existing VDOT count locations were excluded. For example, out of 64 links shown in the Route 29 Corridor study, 47 were excluded because the start or end location of the link could not be determined. As another example, the Hampton Roads Travel Demand Model required the merging of several different data sources into one application. In order to determine the best links for the study, an application was developed in ArcGIS using the Model builder tool supplemented with a script in Python. The application assigned a buffer to the roadway segment based on the number of lanes and the facility type, and associated with a roadway segment to the proper traffic counter. These links were manually verified using the VDOT TMS web application (VDOT [24]). This resulted in 42 forecasted links whose location and distance matched the TMS traffic counter locations exactly (VDOT [24]). However, there were 163 links that had an imperfect match; those links (as well as links that had no match) were excluded from the analysis.

Background Methodology is not Clear

The methodology that underlies the forecasts in the study was documented–but sometimes this methodology was not clear. For example, for study 1, the report notes that the forecasts are "based upon the historical traffic growth trends for the past twenty years"– which seems similar to the method of trend analysis based on traffic counts (technique 3). However, the exact method used to generate these forecasts could not be determined: That is, although the study reported, for each link, ADTs from 1965, 1970, 1975, 1980, 1985, and 1986, the research team could not find a way to use these volumes to generate the 2010 forecasts reported in the study. Again, Route 608 Corridor Study [38] is similar to a TIA that uses trip generation rates, as the study states "Through data supplied by the Augusta County Department of Community Development, a forecast of land use for the year 2014 was developed along the corridor. Traffic volumes generated from this land use were added to the 1994 traffic volumes to represent future traffic volumes on the corridor.” Given that the forecast traffic volumes are for the PM peak hour in a rural area in 1996, it appears unlikely that a travel demand model was used; at that time, travel demand models were not used in rural areas.

How are Temporal and Spatial Ambiguities Addressed?

A link’s temporal resolution and spatial resolution may differ. For example, Stonefield development (study 30) in Figure 6.2 (left) forecast a volume for each of four subsections; however, there was only one observed volume available for the entire segment: a temporary counter could have been placed anywhere on this segment. The temporal resolutions differ: each forecast is a single point, whereas the observed volume is measured on three different days. Based on the three observed volumes (Tuesday, Wednesday, or Thursday) and the four forecast volumes (one for each subsection), there are 12 possible comparisons between an observed and forecast volume. The magnitude of the absolute error thus ranges from 2 to 271 vehicles during the peak hour or 0.2% and 36% of the observed volume. Another example is study 20; where Figure 6.2 (right) shows that the forecast and observed segments do not align perfectly. The length of forecast link as reflected in the travel demand model (0.05 mile) differs from the length as reflected in the count database (0.70 mile).

Forecast volume for AWT, yet observed volume for ADT

The Peninsula Area Transportation Study [21] forecasts average weekday traffic (AWT) for the year 1985; however, AWT volumes are not available for year 1985. As a result, Table 6.1 compares forecast AWT to the observed ADT for this study. In another case, where it was difficult to identify whether the forecast was for average weekday traffic or average daily traffic, both the accuracies were determined by comparing with observed AWT and ADT. For example, Cube Model application for the Hampton Roads Regional Travel Demand Model generated forecast for the year 2011. Initially, it was not clear about the model’s forecast unit.

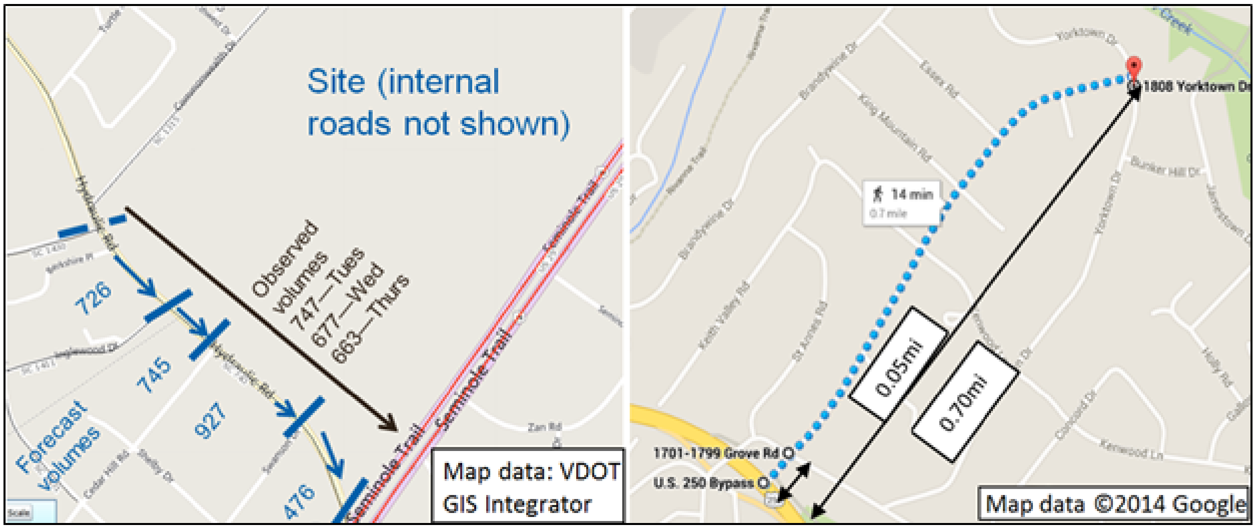

Figure 6.2: Comparison of temporal and spatial resolution for forecast vs. observed volumes in Charlottesville. Left, a site impact study [5]); right, a regional travel demand model [6] .

Additional inquiries suggested that the study was likely forecasting AWT, but the difference in error rates depending on whether one uses AWT or ADT for the observed volume are shown in Table 6.6. The mean and median absolute error would have each changed by two percentage points.

Table 6.6: Performance of Forecasts for Study 22 [1]

Compare Forecast from Study 22 to observed	Mean Error	Mean Absolute Error	Mean Percent Error	Mean Absolute Percent Error	Median Error	Median Absolute Error	Median Percent Error	Median Absolute Percent Error
ADT	1,957	4,776	21%	49%	1,579	3,610	11%	33%
AWT	1,036	4,689	14%	47%	570	3,136	4%	35%

The Year of Forecast Volume Differs from the Year of Observed Volume

For I-95 Clermont Avenue Interchange & Connector to Duke Street study, weekday observed volumes for 2010 were used to evaluate the forecasts. However, the weekday ramp volumes could not be obtained for year 2010 as they were not routinely collected until 2012 (Dunnavant [69]). Thus, the ramp volume of 2012 was compared with the 2010 volume. Accordingly, for the University of Virginia Research Park and Richmond International Airport study, 2014 traffic counts were used as a surrogate of 2015 since the most recent traffic count available in the database was for 2014.

Which alternative to choose?

All the four model alternatives in Richmond International Airport study [3] showed a connector to the airport (between Route 895 and I-64) forecast to be built at the time of study with some additional improvements to I-64. Modified base condition reported in the study refers to neither the Airport Connector nor the changes in I-64 interchange are made. However, in practice, only the airport connector is built (thus modified base condition cannot be used), yet none of the I-64 improvements were made (thus alternatives 1, 2, 3, and 4 cannot be used). The only visible change that had been made since then was the connector to I-895 (verified by VDOT GIS integrator and aerial image of Google Map, 2015), which makes Model Alternative 3 as the closest alternative. Another example is study 6 where no alternatives made a forecast for the HOV-2 lane which is the current scenario west of I-495. The two closest alternatives that reflect the current I-66 scenario are (1) full improvement to I-66, Separate, Reversible HOV-3 with access to I-495; and (2) full improvement to I-66, Concurrent flow HOV-3. The second option appears closest to the reality, although it shows a 4-foot buffer between the HOV lane and the general purpose lane. That said, the mean absolute percent error is similar for the two alternatives: 21.2% ("Reversible HOV-3" alternative) and 22.3% ("Concurrent flow HOV-3" alternative).

Seasonality may or may not be Addressed in Forecast or Observed Volume

Monthly volume variation affects forecast evaluation. For example, BJ's wholesale club land development study forecast 1,084 vehicles during the peak hour for a particular location (Starling Drive). During the forecast year, no continuous counts were available; rather, a single count was collected once in August. Although seasonal adjustment factors for Starling

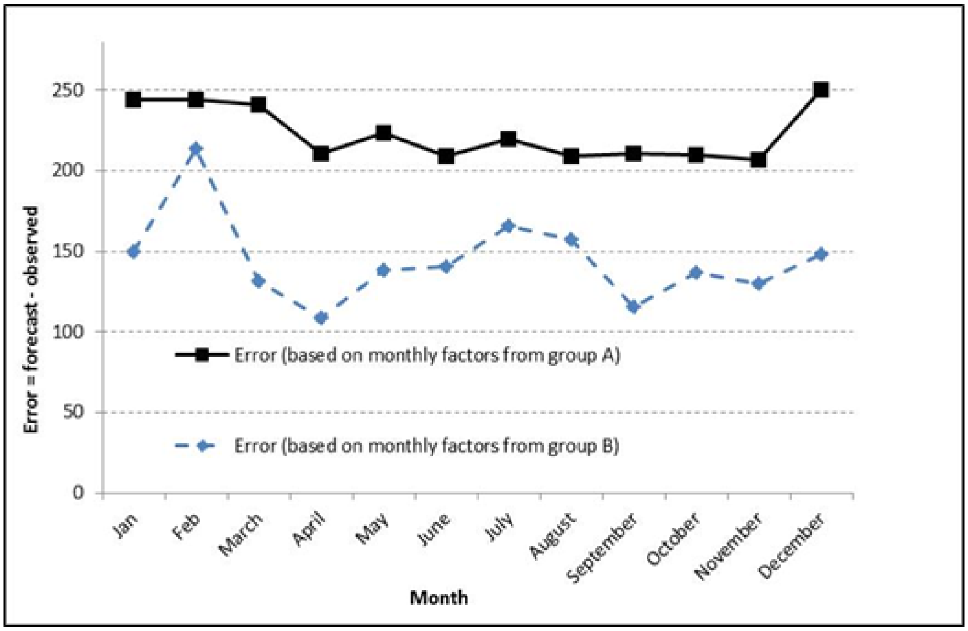

Figure 6.3: Forecast errors for Starling Drive. A single count was observed in August 2010 at Starling Drive, and this count was adjusted to estimate the "observed" volumes for other months using seasonal factors from two candidate route groups. The difference between each resultant monthly "observed" volume and the single forecast for the study is shown in the figure

Drive are not available, factors from two candidate route groups can be used to obtain an "observed volume" for Starling Drive for each month. Because these monthly volumes will vary and the site impact study provides just a single forecast, the forecast error will vary as shown in Figure 6.3. Depending on which month and set of seasonal factors are used to evaluate the forecast of 1,084, the monthly error varies from 10% to 23%. Another subset of data (Figure 6.2, left) shows modestly greater variation among weekdays: in February, a count taken on a Friday tends to be about 9% higher than a count taken on a Monday.

Methodology That Fall Under Both Categories

Rivanna Village at Glenmore [50] is a land development study where the subject (Rivanna Village) was not built. However, a portion of the development was built (which entails the use of ITE trip generation rates) yet a portion of the traffic was forecast using a trend-based analysis. Thus there is some debate as to whether the study's forecast method should be a trend analysis or traffic impact analysis. The study was evaluated as a trend based study as shown in Table 5.1 since the background forecast is trend based. A similar category of study is University of Virginia Research Park which is also categorized as a trend-based study. A third example is Botetourt County Route 220 where the background traffic was generated using a regional travel demand model; then, additional trips were generated with what the study referred to as "the Institute of Transportation Engineers' (ITE) Trip Generation Manual". Because of the ITE emphasis, the study is categorized as a traffic impact analysis.

Alternative That Reflects the Present Scenario

The George P. Coleman Bridge Study [27] forecasts the 2010 ADT for the Coleman Bridge under different build alternatives. However, Alternative 12A, which is the reconstruction of the Coleman Bridge, represents the actual scenario based on a review of the study and Kemper [70]. The methodology for developing this forecast is not stated in the study, except that tolls of $1.00 would reduce traffic by 5% whereas tolls of $0.75 or less would not affect traffic volumes (VDOT [27]).

Observed Volume could not be Obtained for Forecast Year; Hence Interim Forecast is Used

For some studies, the length between the observed year for which volume is available and the forecast year is so long that an interim year forecast is generated. Most cases, interim forecast is generated by linear interpolation method assuming that the change of traffic between those years remains the same. For example, Route 1 Corridor Study identifies current and future transportation needs for year 2020. Given the 2020 forecast year, a linear interpolation was used with the observed 1995 ADT and the forecast 2020 ADT to obtain a 2014 forecast which could be compared to a 2014 observed value. One such example is segment 1, which had 11,000 ADT in 1995 and a forecast of 46,000 ADT for 2020. Assuming linear growth, the segment would add 1,400 ADT annually during that 25-year period. Accordingly, for the 19-year period from 1995 to 2014, one would expect the segment to add (191,400) to the 11,000 to generate an interim forecast of 37,600 ADT.

Is Proposed Alternative Built or Not?

The extent to which the forecast depends on certain improvements was reported. For example, the Route 3 Corridor study provides a list of approximately 50 improvements, but the forecasts given are independent of these improvements. Logically, however, one would expect that the extent to which the improvements were made would affect the forecast. For example, that study suggested that, at a point where the road is signed as both Route 3 and Route 33, that in the future Route 33 could be relocated and Route 3 could be widened from two to four lanes. However, this relocation and widening had not happened as of 2015. In another case, for capital beltway study, partial improvements were built. A detailed review of ten out of the 45 segments showed that five segments were widened to 4 or 6 lanes while other five were not based on the use of VDOT's Statewide Planning System and Google Maps street view application in 2015 (see Table 6.7). Since, five of the ten proposed improvements were made and all four new facilities were built by 2010, a build scenario named "Scenario A" was chosen for the study. Finally, 2010 forecasts from "Scenario A" were compared to the observed 2010 volume.

Turning Movement was Converted into Link Volumes

The I-95 Clermont Avenue study [33] forecasted 24-hour intersection turning movements for 2010 based on a travel demand model. For comparison purposes, turning movements had to be converted to link volumes. For example, the link between Edsall Road and South Pickett Road has the intersection of S Van Dorn/Edsall at one end and S Van Dorn/S. Pickett at the other end. The link has two southbound forecast volume; 32,000 (=800+29600+1600) AWT for S Van Dorn/S. Pickett intersection and 31,400 (=21500+2800+7100) AWT for S Van Dorn/Edsall intersection. Accordingly, the two potential northbound volumes are

Table 6.7: Comparison of Proposed and Actual Improvements for the I-95/I-495 Capital Beltway Study [2]

Road	**From**	**To**	**Proposed Improvements for 2010 (Both Direction)**	**Existing Facility in 2010 (SPS & Google Street View)**
Route 1	Route 233	15th St	Widen to 6 Lanes	Built
Route 1	Occoquan River	Route 623	Widen to 6 Lanes	4 lanes (not built)
Route 7	DATR	Fairfax Co. Line	Widen to 6 Lanes	4-5 lanes (not built)
Route 7	Falls Church	I-495	Widen to 6 Lanes	4 lanes (not built)
Route 7	Columbia Pike	7 Corners	Widen to 6 Lanes	5 lanes (not built)
Route 7	Loudoun Co.		Widen to 6 Lanes	Built
Route 29	Graham Rd	Fairfax Co. Line	Widen to 4 Lanes	Built
Route 28	I-66	Route 7	Widen to 6 Lanes	Built
Route 123	Occoquan	Fairfax Co. Parkway	Widen to 4 Lanes	Built
Route 236	I-395	I-495	Widen to 6 Lanes	4 lanes (not built)

29,600 AWT and 27,800 AWT for S Van Dorn/S. Pickett and S Van Dorn/Edsall intersection respectively. As is done with all other studies, the higher of the two volumes is selected; which is 29,600 AWT. Given a 32,000 AWT southbound volume, the two-way link volume is 61,600 AWT which may be compared to the 2010 observed volume of 54,017 AWT.

6.4 Contributing Factors

In order to know which method has been performed well in forecasting traffic and to investigate whether any external attributes (i.e., duration, economy) have impact on error, an Analysis of Variance (ANOVA) model was conducted with the 39 studies. ANOVA can explain the data better where the distribution of data deviates from normality compared to regression model. The model helps determining how much the external variables associated with the forecast can explain the variation in accuracy. Among various factors, the reasons behind choosing these 3 factors (Forecast method, Duration and Economic recession) as shown in Table 6.8 are– First, the agency was interested in the accuracy of all eleven techniques– which led to focus on the method used to generate forecasts for 39 studies. Second, pick a few variables that appears critical to the analysis to avoid having misleading results with too many factors. For example, 39 studies were split into two major groups: those that were based on an extrapolation of previous years' traffic counts and those that were based on some activity. The analysis was conducted to see whether forecast method along with duration could lead to a useful conclusion about errors and how long term (or short term) forecast performs in error perspective. Third, identify factors that are beyond the forecasters' control, such as changes in the economy. However, forecast unit (e.g., 24-hour volume or a peak hour volume) was not considered as an independent variable because it would yield an unbalanced study with 33 studies having a 24-hour volume and six studies having a peak hour volume. Further, an analysis with forecast unit as an independent variable showed that forecast unit alone was not significant (p=0.59) in terms of explaining forecast accuracy. Thus, in order to differentiate among the forecasts for ADT and the forecasts for peak hour volume, separate ANOVA models with the above 3 factors were analyzed with the ADT volumes only. Table 6.8 lists three explanatory factors, each with two or more levels and their sample sizes based on 39 studies.

Table 6.8: Categories for Each Factor

Factor	**Levels**	**Sample Size for Model**	
		MAPE (Number of Study)[a]	**APE (Number of Links)**[b]
Forecast Method	Trend Based Forecast	20	517
	Activity Based Forecast	19	642
Economic Recession	Smaller (0-1)	6	38
	Medium (2)	12	166
	Large (3-4)	21	955
Forecast Duration	Long Term (>20 yrs)	20	781
	Short-to-Medium Term (≤19 yrs)	19	378

[a] For full factorial model 39 studies were used for the analysis (excluding study no. 34 and 39)
[b] For nested model 1,159 links from 38 studies were analyzed (excluding study no. 27)

6.4.1 Models that Explain the Absolute Percent Error for Each Link

A conventional ANOVA assumes that observations are independent. This assumption is violated when multiple links from the same study are considered. (For example, the links in a single study will have the same forecast method and forecast duration; they may also share other characteristics not explicitly considered by the analyst, such as the expected seasonal adjustment factors.) Accordingly, a four-stage nested ANOVA, also known as a hierarchical ANOVA [71], was used for the link-by-link analysis. The three upper levels of the nest-forecast method, number of economic recessions in the forecast period, and forecast duration-were treated as fixed factors as they represented the full number of levels shown in Table 6.9: for example, for forecast duration, no levels other than long term and short-to-medium term were considered. However, the lowest level of the nest-study number-was itself a random factor, as the 39 levels of this factor (i.e., Study 1, Study 2, etc.) were used to make inferences about the whole population of forecast studies

Table 6.9 shows that only one factor-the study itself-was statistically significant ($p = 0.02$); forecast method ($p = 0.65$), number of economic recessions ($p = 0.22$), and forecast duration ($p = 0.20$) were not. However, the model for Table 6.9 explains only a small percentage (5.2%) of the variance in link errors. That is, if the analysis ceased with Table 6.9, the conclusion to be drawn would be that the study itself was a substantial contributor to the variation in forecast accuracy. However, with regard to the factors stated above, only a small amount of the variation in error, about one-twentieth, could be explained by the study itself.

Certainly, one possible additional line of inquiry would be to look for link-by-link causal factors (e.g., perhaps one or two different links in the same study had substantially different types of improvements made between the base year and the forecast year). However, given that many of the studies would not have supported such a level of analysis, a different approach was pursued in an effort to reduce within-study variability and then determine the effects of potential explanatory factors. One way to reduce this within-study variable is

Table 6.9: Results of Applying the Nested Analysis of Variance[a]

Level	Factor	Degrees of Freedom	Mean Square	p value
1 (highest)	Forecast method	1	0.17	0.65
2	Number of economic recessions	4	1.18	0.22
3	Forecast duration	3	1.63	0.20
4 (lowest)	Study number	26	1.02	0.02

[a]Because the number of links for Study 27 (2,493) was so much larger than the number of links for other studies, the analysis in Table 6.9 was done by excluding study 27.

through use of the median percent error by each study.

6.4.2 Models that Explain the Median Absolute Percent Error for Each Study

Although it is possible to develop a full factorial ANOVA with all three independent variables (forecast method, number of economic recessions, and forecast duration) explaining 41% of the variation in accuracy, such a model appears inappropriate because two of the factors (number of economic recessions and forecast duration) are highly correlated (0.85). Two highly correlated variables interact with each other could lead to the false belief that spurious impacts were significant. Hence, two different full factorial models were tested: Model 1 (forecast method and number of economic recessions) and Model 2 (forecast method and forecast duration). However, the two correlated variables (forecast duration and number of economic recessions) had different forms: in Model 1, number of economic recessions was significant (p=0.03) as a main effect only (and forecast method is not), and in Model 2, forecast duration was significant but only when interacting with forecast method (p=0.01). These two models explain less than one-fifth (15%-18%) of the variation in forecast accuracy– and significance levels were similar whether all 39 studies or the 33 studies that forecast a ADT volume only were used.

Models 1 and 2 generated two additional concerns. First, the interpretation of the effects

Table 6.10: Results of Applying the Traditional Analysis of Variance that Explain Variation in Forecast Accuracy

Model[a]	Key Factors	p-value[b]	Variance Explained	Interpretation of Model Effects
1	Forecast method	0.48 (0.11)	1.82% (17.2%)	A trend-based forecast is more accurate than an activity-based forecast when the number of economic recessions between the base and forecast years is 2 or more.
	No. of economic recession(s)	0.03 (0.04)		
2	Forecast method × Forecast duration	0.01 (0.02)	15% (15.4%)	A trend-based forecast is more accurate than an activity-based forecast for long-term duration but not for short-term duration. Further, long-term trend-based studies are more accurate than short-term trend-based studies.
3	Forecast method × Forecast duration	0.04 (0.07)	29.1% (26.8%)	For both activity-based forecasts and trend-based forecasts, accuracy increases when forecast duration decreases. For both cases, accuracy is lowest under the case of exactly 2 economic recessions.
	No. of economic recession(s)	0.02 (0.05)		For long-term duration, trend-based analysis is more accurate than activity-based analysis, but for short-term duration, activity-based analysis is slightly more accurate than trend-based analysis.

[a]Models 1 and 2 were run as full factorials, but the interaction effect in Model 1 (Forecast method × No. of economic recessions) and the main effects in Model 2 (forecast method, forecast duration) are not significant. All terms for Model 3 are shown except the intercept, which was significant in all three models.
[b]The value in parentheses is the result when the six studies that forecast a peak hour volume, rather than a 24-hour volume, were excluded.

of these three factors was counterintuitive; notably, for Model 2, when a trend-based analysis was performed, studies with a forecast duration of 19 years or less appeared to be less accurate than studies with a forecast duration of 20 years or more. In addition, the hypothesis of normality (desirable for ANOVA) can be rejected based on the Kolmogorov-Smirnov test ($p = 0.02$ and 0.03, respectively), which would suggest the need for some type of transformation or another model formulation.

Accordingly, a third model was considered where the number of economic recessions had been treated as a block [71]. (A block can be described generally as "a set of relatively homogeneous experimental conditions [Montgomery, 2001]"; in practice, a block may be a factor that should affect the results but which is not controlled by the forecaster.) While the forecaster can choose forecast method and duration, the number of recessions is in fact a nuisance variable: it is an effect that must be controlled for when analyzing results, but it is not a decision that the modeler can choose at the outset of the study. Model 3 explicitly accounts for the impacts of economic recessions but allows for one to consider the impact of forecast duration (addressing a concern with Models 1 and 2). Model 3 also shows that the residuals are normally distributed ($p = 0.75$), thereby addressing a second concern with Models 1 and 2 (see Table 6.10).

Table 6.11: Estimated Marginal Means for Model 3[a]

Type of Study	Forecast Duration (No. of Years Between Base and Forecast Years)	No. of Economic Recessions Between Base and Forecast Years Inclusive		
		0 or 1	2	3 or 4
Activity	Short term (≤19 years)	22%	49%	23%
	Long term (>20 years)	50%	78%	51%
Trend	Short term (≤19 years)	23%	50%	24%
	Long term (>20 years)	26%	53%	27%

[a]For example, one would expect a long-term forecast that used an activity method where there were 2 economic recessions between the base and forecast years to have a median percent error of 78%.

Model 3 has several implications. First, the number of economic recessions has a nonlinear impact; that is, errors were highest when the number of economic recessions was 2, rather than at a smaller (0 or 1) or a larger (3 or 4) number of recessions. (Although it is expected that the increase in the number of economic recessions from 0 or 1 to 2 would increase error, the decrease that results when it is increased to 3 or 4 is surprising and shows that the number of recessions have nonlinear impact on accuracy.) Repeating Model 3 with two levels of ANOVA (e.g., 0-1 recessions and 2-4 recessions) did not affect the significance level. The length of recession along with the number might help in explaining such conditions. For example, the impact of one-time 3 years long recession is not the same compared to three recessions each with 1-year duration that occurred between base and forecast year. Second, as one might expect, a shorter forecast duration increases accuracy, but this is evident only when controlling for the confounding effect of the number of economic recessions. That is, had economic recessions not been controlled for, one would have believed, based on Model 2, that in some cases having a longer horizon increases accuracy. Third, as shown in Table 6.11, activity-based approaches are slightly more accurate than trend-based approaches, but only for shorter term studies; for longer term studies, Model 3 shows that trend-based forecasts are more accurate. This third result is surprising: one would have expected the activity-based approaches, which incorporate more detail (e.g., impacts of land development, for instance), to be more accurate than the extrapolation of past trends. Although this explanation cannot be proven, one possible reason is that in the short term, behavioral assumptions (e.g., attitudes toward driving or using public transportation) are more likely to remain constant, such that additional model detail (which comes from activity-based approaches) increases accuracy. However, in the longer term, it is possible that as changes in behavior become more likely, such additional detail is not helpful.

6.4.3 Discussion of Explanatory Factors of Variation in Forecast Accuracy

As expected, forecast method and forecast duration have significant impacts on forecast accuracy but only if confounding effects are controlled for in at least three ways. First, one must consider the interaction effect between forecast method and forecast duration: in the short term, activity-based forecasts are slightly more accurate than trend-based forecasts, but in the long term, the accuracy of these activity-based forecasts degrades substantially.

A less dramatic decrease in accuracy is noted for trend-based forecast as one changes from short to long term. In order to improve the accuracy of activity based method for longer term forecasts, the underlying assumptions regarding travel behavior has to remain valid over the years between the forecast preparation and the horizon year. To do this, we need to develop ways to consider the uncertain circumstances (i.e. economic recession) as also suggested by Hartgen [72]. Further research regarding this area can be conducted to better quantify uncertainty in land use, economic development and employment near the proposed projects. Idea about the explanatory factors contributing to the forecast inaccuracies can provide groundwork for better demand forecasts. Second, one must control for changes in economic condition, which in this report is the number of economic recessions between the base year and the forecast year. Third, given that the studies have different facilities (e.g., eight-lane interstate highways versus two-lane local roads), it is appropriate to control also for differences in volumes by using the median absolute percent error instead of the median absolute error as the dependent variable.

It is acknowledged that other factors besides the four noted may have affected forecast accuracy. For example, more than a decade before this research was undertaken, FHWA issued guidance to Virginia stating that the Transportation Improvement Program (TIP) should only include those projects that could realistically be built with the funds available (Debruhl [73]). It is possible that in urban areas which use a travel demand model, one would expect to see forecasts become more accurate after this period, given that the models would have a more realistic assessment of which projects would be built. However, of the studies shown in Table 5.1, only eight were published in 2003 or laterand of those, only one was based on a travel demand model. Accordingly, this particular factor was not used in the analysis of variance.

The errors determined are somewhat higher, but not unreasonably so, than those in some related literature. By contrast, the 32% for the average of the absolute project percentage errors noted by Flyvbjerg et al. [12] is nominally lower than the 55% reported in Table 6.1.

Further, Buck and Sillence [13] evaluated 131 Wisconsin forecasts and reported that the mean and median absolute differences were 16% and 13%, respectively. For both studies, however, one could argue that comparison with Virginia is not appropriate: Table 6.1 does include forecasts for peak hour volumes, whereas Flyvbjerg et al. [12] appears to be focused on average daily traffic, and for the Wisconsin study, none of the forecasts spanned a 20-year horizon, meaning they had shorter time periods, generally, than those shown in Table 6.1.

Finally, although some errors in Table 6.1 may seem large at a glance, they are not always meaningful in terms of influencing decisions. For example, the aforementioned Richmond Airport Study [3] performed a sensitivity analysis and concluded that a certain alternative would remain "economically feasible" provided the forecast error (as defined in this report) was no greater than 50% (for a connector from one interstate noted in the study) and no greater than 400% (for a connector from another interstate noted in the study). Such a practice appears to implement a suggestion noted by Hartgen [72], which was to understand what magnitude of forecast error is tolerable based on the purpose of the study.

6.5 The Impact of Forecast Error on Chosen Decision

Detailed analysis of one study shows that how the impact of forecast error on a decision is driven by situation for which forecast is needed. For example, the decision criteria for study 1 was whether a certain performance measure (which is level of service C) could be obtained. Another example of study 38 [3] shows that some performance measures do not explicitly use the forecast.

6.5.1 Performance Measure That Uses the Forecast

The ADT for one section of the two-lane aforementioned Route 3 [8] was forecast to be 9,800 in 2010. With an observed ADT of 6,100, the error was 3,700 (61% relative to the observed value). The question arises as to what error would cause decision makers to change their course of action.

The study [8] identified improvements to achieve level of service (LOS) C: having channelization at two intersections, widening the substandard pavement to 24 feet, improving the shoulder, and widening the segment from two to four lanes–a major capacity recommendation. LOS was the determinant of action, therefore, how the forecast error affected LOS can be examined. The report does not provide LOS computational details; however, when the forecast was generated in 1988, the methodology in the 1985 Highway Capacity Manual (described by Garber and Hoel [7]) was standard practice.

Figure 6.4 illustrates how the forecast error influences decisions. Assuming no change to the roadway geometry, 726 vehicles per hour was calculated as the maximum volume that could support LOS D. The observed peak hour volume in 2010 was 537 (LOS D). The forecast ADT was 9,800, yielding LOS E. Therefore, if the forecasters had assumed that the K-factor–the peak hour volume divided by ADT–would remain constant for the period 1986-2010, the forecast hourly volume would be 998 (86% error). If the forecasters had instead anticipated some peak spreading where the K-factor dropped to its observed value, the forecast hourly volume would have been 862 (61% error). Regardless of whether peak spreading was assumed, the forecast error results in an LOS E, rather than the correct value of LOS D, being computed. Yet because both the forecast and observed volumes are below LOS C, the error does not alter the judgment that the segment is deficient.

However, for a low-cost alternative where the pavement is widened to 24 feet, 6–foot shoulders are added, and no lanes are added, an hourly volume of 689 vehicles accommodates LOS C. The observed hourly volume (537) meets this level of service, but the forecast–with or without peak spreading–does not. Hence, the forecast error means that decision makers would have rejected (incorrectly) this low-cost alternative as they would have thought that widening to four lanes was essential to achieve LOS C. Figure 6.4 suggests two distinct factors affect the impact of error on decision making: (1) the magnitude of the error, and (2) the

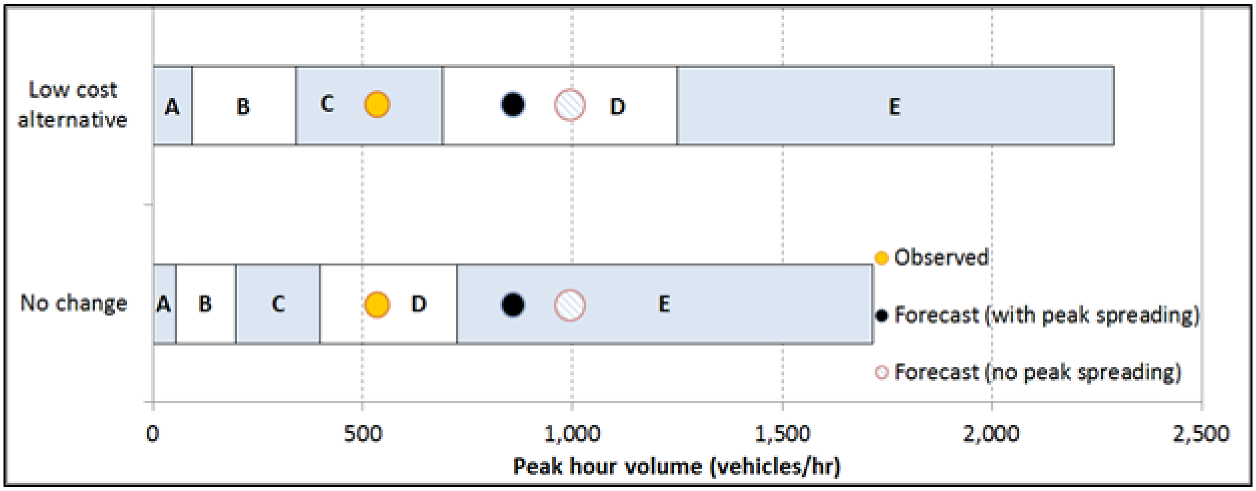

Figure 6.4: Level of service (LOS) standards for Section of Route 3. Criteria for LOS were calculated by the research team based on the methodology of Garber and Hoel [7] and data available in VDOT [8].

location of the error relative to the performance criterion. In Figure 6.4, if the magnitude of the error had remained the same but its location had been shifted rightward by 152 or more vehicles, the decision maker would have assumed widening was necessary–regardless of whether the forecast value or the observed value was used in the decision.

Assuming no change to the roadway geometry, the researcher team calculated that 726 vehicles per hour was the maximum volume that could support LOS D. The observed peak hour volume in 2010 was 537, which thus yields LOS D. The forecast ADT was 9,800, yielding LOS E. Therefore, if the forecasters had assumed that the K-factor–the peak hour volume divided by ADT–would remain constant for the period 1986-2010, the forecast hourly volume would be 998 (86% error). If the forecasters had instead anticipated some peak spreading where the K-factor dropped to its observed value, the forecast hourly volume would have been 862 (61% error). Regardless of whether peak spreading was assumed, the forecast error results in an LOS E, rather than the correct value of LOS D, being computed. Yet because both the forecast and observed volumes are below LOS C, the error does not alter the judgment that the segment is deficient.

6.5.2 Performance Measure That does not Explicitly Use the Forecast

Generally, when evaluating alignments, a variety of environmental, social and economic impacts are considered (including the no-build and other alternatives: commercial and residential relocation, right of way required, construction and engineering costs, safety impacts, air quality or noise impacts). For example, relative to the no-build scenario, the preferred Alternative 2 for Richmond International Airport Study [3] would cost 25.9 million dollars which include construction, right of way, relocation and engineering cost. In addition to that an estimated cost ranges between $3,400 and $10,000 per lane-mile would require for maintenance, operation and administrative purposes. Thus the above metrics are not dependent on traffic volume. However, there are two elements of the study that do rely

Table 6.12: Ranking of various alternatives for I-895 Corridor based on total possible score of 100% (Richmond International Airport Feasibility Study [3])

<table>
<tr><td colspan="10">I-895 Alternative Evaluation Matrix</td></tr>
<tr><td rowspan="2">Points Earned</td><td rowspan="2">30%</td><td colspan="4">40%</td><td rowspan="2">10%</td><td rowspan="2">10%</td><td rowspan="2">10%</td><td rowspan="4">Total</td></tr>
<tr><td>20%</td><td>10%</td><td>5%</td><td>5%</td></tr>
<tr><td rowspan="2">Alternative</td><td rowspan="2">Cost</td><td colspan="4">Environmental</td><td rowspan="2">Future Development Potential</td><td rowspan="2">Travel Efficiency/ Safety</td><td rowspan="2">Expansion Potential</td></tr>
<tr><td>W</td><td>C</td><td>RR</td><td>RC</td></tr>
<tr><td>1[a]</td><td>$24.3M 30%</td><td>10%</td><td>6%</td><td>5%</td><td>4%</td><td>5%</td><td>10%</td><td>0%</td><td>70%</td></tr>
<tr><td>2</td><td>$25.9M 29%</td><td>20%</td><td>3%</td><td>2%</td><td>4%</td><td>10%</td><td>7%</td><td>5%</td><td>80%</td></tr>
<tr><td>3</td><td>$26.5M 28%</td><td>10%</td><td>6%</td><td>3%</td><td>4%</td><td>5%</td><td>7%</td><td>5%</td><td>68%</td></tr>
<tr><td>4</td><td>$30.4M 20%</td><td>8%</td><td>6%</td><td>4%</td><td>4%</td><td>3%</td><td>7%</td><td>10%</td><td>62%</td></tr>
<tr><td>5</td><td>$30.6M 20%</td><td>20%</td><td>10%</td><td>0%</td><td>4%</td><td>10%</td><td>7%</td><td>5%</td><td>76%</td></tr>
</table>

[a] Example. For alternative 1, the low cost meant the alternative earned 30 of the total points for cost. Alternative 1 also earned 25% of the maximum of 40% of points for the environmental categories based on wetlands takings, cultural resources, residential relocations, and commercial relocations. It earned 5%, 10%, and 0% of the maximum points in the categories of future development, travel efficiency/safety, and expansion potential

on the forecast. First, a sensitivity analysis conducted within study mentioned that the Alternative I-895 connectors would be beneficial if the observed traffic volume was at least 20% of the projected value. (Benefits are computed as monetized delay savings, monetized crash reductions, monetized carbon monoxide reductions, and savings from reduced vehicle operating costs and reduced fuel consumption.) For example, the sensitivity analysis suggested that for the I-895 connector, the benefit-cost ratio would have increased from 2.981 to 4.205 if observed traffic volumes increased from 90% of the forecast value to 125% of the forecast value. Thus because the forecast volume for the connector was generally higher than the observed volume, the forecast error did not affect the former project level decision. Second, a comparison of the four alternatives suggested that a grand score of 80% (for Alternative 2) rendered it preferable to the remaining alternatives for a preliminary screening. Because traffic safety and efficiency (which are based on the forecast) are only a small portion of this grand score (see Table 6.12), however, it does not appear likely that the forecast errors altered the decision that was taken. However, additional information regarding calculation of these percentages would be needed to confirm this assessment.

6.6 Limitation

As noted in the methodology, a limitation of any retrospective evaluation is– it is not always possible to align a forecast value and the observed value. Two examples of this limitation were noted in the methodology: (1) the forecast may have presumed certain improvements (which were never built) and (2) the forecast may be for a portion of a road segment, whereas an observed volume may be for a much longer portion of that segment. An additional example is the unanticipated economic or operational changes that may have transpired after the forecast was generated. For example, for the Coleman Bridge (study 2), the toll was not precisely what was expected (which could logically affect travel demand), and studies that were based on expected land use changes (such as travel demand modeling efforts) would not have foreseen the impacts of economic recessions that have been observed in Virginia. Thus, there may be a variety of reasons for why forecast and observed volumes differ.

Chapter 7

CONCLUSIONS

Several conclusions regarding the accuracy of retrospective forecasts may be drawn.

1. *Based on 39 case studies in Virginia, roadway volumes have been over forecasted for the past two decades.* The magnitude (or percent) value of average and median error for all the 39 studies as shown in Table 6.1 are positive which indicates that the forecast volume is greater than the observed volume.

2. *On average about three fifth of the Virginia forecasts are accurate. The median absolute percent error for 39 studies are, on average 40%.* For each of the 39 studies analyzed in this report, one could (for each link in the study) divide the absolute error (the magnitude of the difference between the forecast volume and observed volume) by the observed volume. Then, for each study, one could then select the median absolute percent error from these ratios. The average of these median absolute percent errors was 40%, as shown in Table 6.1.

3. *On a percentage basis, the accuracy of different types of traffic forecast studies may vary by almost an order of magnitude.* Certainly, several factors may influence a study's forecast accuracy, such as the length of the horizon year or unanticipated developments. That disclaimer aside, a subset of the Virginia studies showed that the median absolute

percent error was as small as 12% (for the site impact study in Figure 6.3), slightly larger at 22% or 35% (from two travel demand modeling studies), or as large as 72% (for a single statewide study).

4. *Three factors affect this median absolute percent error in a statistically significant manner: forecast duration (number of years between the base and forecast years), forecast method, and number of economic recessions between the base and forecast years.* The ANOVA showed that forecast duration and forecast method significantly affect median percent error but only when both factors are considered together ($p = 0.04$). In the short term, a trend-based forecast is slightly less accurate than an activity-based forecast (by about 1 percentage point). However, although both the trend- and activity-based approaches show a decrease in accuracy when one shifts from a short-term forecast to a long-term forecast, the degradation is greater for activity-based forecasts (which see a change of about 30 percentage points in the median percent error) than for trend-based forecasts (which see a change of about 3 percentage points). These effects are not evident, however, without controlling for the number of economic recessions ($p = 0.02$).

5. *Most of the variations in accuracy are random. The three external variables identified (forecast method, duration and recession) can explain more than one fourth of the variation in forecast accuracy.* A model based on three of the above factors can explain about 29% of the variation in median percent error. That is, most of the variation in forecast accuracy is either random or based on factors not identified in this report.

6. *Random link-by-link variation must be controlled in order to detect factors that influence forecast accuracy.* A hierarchical ANOVA, which considers the accuracy of individual links, indicated that the only statistically significant causal factor was the study itself ($p = 0.02$). That is, there is a large amount of variation in individual links in the same study such that other potential causal factors do not explain a significant amount of

this variation. Only when this link-by-link variation in a study is somehow controlled, as is done when a single study median percent error is used, can other causal factors be identified. A practical implication of this finding is that even in a given study that is performed under the best circumstances, there may continue to be links that have a low forecast error and links that have a high forecast error.

7. *There is not necessarily a linear relationship between the magnitude of the error and the impact such an error would have on decision making.* Figure 6.4 illustrates that a chosen decision criterion may or may not be sensitive to forecast error. In that instance, although the error was seemingly large at 61%, or about 370 vehicles during the peak hour, the error mattered only to the extent that it caused a shift from LOS C to LOS D. Had volumes been higher, even with error a decision maker would have reached the same decision that a widening was necessary.

Chapter 8

RECOMMENDATION AND FUTURE STUDIES

Considering no forecast is likely to be perfect, analysts should consider including some indication of expected forecast error. That indication may be based on prior studies that have used similar techniques where one compared the forecast volume to the observed volume. Alternatively, if no such studies are available, one should expect a median forecast error of about 40% for studies where an overforecast matters as much as an underforecast. Interestingly, this error is within the same range of errors as found by CDM Smith et al. [9]. The use of ranges supports a practice noted by Welde and Odeck [15]. For studies where an overforecast on one link and an underforecast on another link tend to cancel (which might be appropriate, for instance, if one was using forecasts to estimate total vehicle miles traveled), the median percent error for all links within the study may be expected to be 31% While these two values were selected from Table 8.1, for some forecasts, it may be appropriate to assume a larger error, and Table 8.1 may help one make that determination. For example, Table 8.1 suggests that if one were planning for an almost-worst case scenario, one might anticipate that, when one determined the errors for all links within the study, the median of these errors was 74%. Such a study would be in the 95th percentilethat is, for most studies, the median percent error would be less than 74%.

Table 8.1: Expected Percent Errors

Descriptor	Mean Percent Error	Mean Absolute Percent Error	Median Percent Error	Median Absolute Percent Error
Minimum	-20%	1%	-19%	1%
Mean	43%	55%	31%	40%
Median	35%	48%	31%	36%
95th percentile	112%	114%	74%	74%
Maximum	118%	126%	134%	134%

If it is not desirable to use expected errors based on the studies that provided the basis for Table 8.1, then an alternative is to present expected errors for individual links based on the literature. For example, CDM Smith et al. [9] indicates expected errors of 20% for a five year forecast and 42.5% for a 20 year forecast (for an existing road), with slightly higher errors of 27.5% and 47.5% (for a new road.) One benefit from this indication would be that decision makers could be better positioned to evaluate transportation needs if the forecast should diverge from the future observed value. (For example, suppose a given study has forecasts for several links during the peak hour, and suppose further that it appears that changing signal timing could accommodate these future volumes. Decision makers could use Table 8.1 to increase the volumes such that the increased volumes differed by the original forecast volumes by a median value of 40% and then determine whether changing signal timing could still accommodate such future volumes.)

These study results raise one tangible long-term research need: to quantify how changes in underlying assumptions (such as expected population growth for travel demand models, expected annual percentage increases in volumes for trend-based models, or expected trip generation rates for land development studies) affect forecast error. Such research is potentially infeasible with some older studies as critical assumptions may not be fully documented [65]; not surprisingly, others have called for better archiving to remedy such studies' "lack of data availability" [74]. That said, a longer term effort to separate the impacts of assumptions from the forecasting technique itself is a worthwhile endeavor. Philosophically, Table 6.9 may portend a theme for future validation efforts: the link-by link variation in accuracy within a single study is inherently large. For this reason, it is quite possible that even larger scale efforts to develop explanatory models may show similar percentages of variation explained as documented herein.

Bibliography

[1] Inc Michael Baker Jr. 2000 hampton roads model users guide. Virginia Department of Transportation, Richmond, study 22, 2004.

[2] Dewberry & Davis JHK & Associates, Rhodeside, and Harwell. Capital beltway phase ii: Short term and mid-term recommendations report. Virginia Department of Transportation, Richmond, Study 14, 1989.

[3] The LPA Group and David Volkert & Associates. Richmond international airport corridor feasibility study report, capital region airport commission (crac). Richmond, study 36, February 1999.

[4] Institute of Transportation Engineers. *Trip Generation Handbook, Second Edition: An ITE Recommended Practice*. Washington, DC, 2004.

[5] E. Memorandum to Chuck Proctor Strohhacker. The memo states reference: Albemarle place phase i, albemarle, virginia. the subject states: Access management exceptions. study 28, June 1, 2010.

[6] Federal Highway Administration U.S. Department of Transportation and Virginia Department of Transportation. Route 29 corridor study: City of charlottesville and albemarle county. Virginia Department of Transportation, Richmond, Study 20, March, 1990.

[7] Nicholas J Garber and Lester A Hoel. *Traffic and highway engineering*. West Publishing Company, St. Paul, MN, 1988.

[8] Virginia Department of Transportation. Route 3 corridor study: Northern neck and middle peninsula. Richmond, study 01, July, 1988a.

[9] CDM Smith, A Horowitz, T Creasey, R Pendyala, and M Chen. Analytical travel forecasting approaches for project-level planning and design, 2014.

[10] A Brett and P Snelson. Traffic and revenue forecasting for privately funded transport projects. In *Financing Infrastructure. Proceedings of Seminar C, European Transport Conference, 27-29 September 1999, Cambridge, UK. Volume P431*, 1999.

[11] Pavithra Parthasarathi and David Levinson. Post-construction evaluation of traffic forecast accuracy. *Transport Policy*, 17(6):428–443, 2010.

[12] Bent Flyvbjerg, Mette K Skamris Holm, and Søren L Buhl. How (in) accurate are demand forecasts in public works projects?: The case of transportation. *Journal of the American Planning Association*, 71(2):131–146, 2005.

[13] Karl Buck and Mike Sillence. A review of the accuracy of wisconsins traffic forecasting tools. In *Transportation Research Board 93rd Annual Meeting*, number 14-1717, 2014.

[14] Morten Skou Nicolaisen and Petter Næss. Roads to nowhere: The accuracy of travel demand forecasts for do-nothing alternatives. *Transport Policy*, 37:57–63, 2015.

[15] Morten Welde and James Odeck. Do planners get it right? the accuracy of travel demand forecasting in norway. *EJTIR*, 1(11):80–95, 2011.

[16] Robert Bain. Error and optimism bias in toll road traffic forecasts. *Transportation*, 36(5):469–482, 2009.

[17] Chi-Hong Patrick Tsai, Corinne Mulley, Geoffrey Clifton, et al. Forecasting public transport demand for the sydney greater metropolitan area: a comparison of univariate and multivariate methods. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*, 23(1):51, 2014.

[18] William A Martin and Nancy A McGuckin. *Travel estimation techniques for urban planning*, volume 365. National Academy Press Washington, DC, 1998.

[19] Virginia Department of Transportation. Materials division: Manual of instructions, chapter vi: Pavement design and evaluation. Richmond, http://www.virginiadot.org/business/materials-download-docs.asp, 2014. [online; Accessed 24-April-2015].

[20] Institute of Transportation Engineers. Trip generation: An ite informational report. Washington, DC, 2008.

[21] Wilbur Smith and Associates. Southeastern virginia regional transportation study, current and projected travel patterns, vol. i. Richmond, study 21, 1965.

[22] The Cox Company. Traffic impact analysis: University of virginia research park, albemarle county, virginia. University of Virginia Foundation, Charlottesville, Study 26, 2008.

[23] Virginia Department of Transportation. Traffic data. Richmond, http://www.virginiadot.org/info/ct-TrafficCounts.asp, 2014a. [online; Accessed 13-April-2015].

[24] Virginia Department of Transportation. Traffic monitoring system. Richmond, http://tedweb/tms/jsp/, undated a. [online; Accessed 13-April-2015].

[25] A. Morton Thomas TransCore, Dewberry & Davis, Lardner/Klein Landscape Architects Associates, Hunter Interests, and Mary Means & Associates. Route 1 corridor study: Fairfax and prince william counties. Virginia Department of Transportation, Study 35, Richmond, 1997.

[26] Jones R. email to J.S. Miller. Private Communication, October 13, 2015.

[27] Virginia Department of Transportation. George p. coleman bridge financial alternatives. Richmond, Study 02, 1989a.

[28] TransCore and Inc. Intermodal Engineering Buchart-Horn. York river crossing travel demand study, final report. Virginia Department of Transportation, Richmond, Study05, August, 2000.

[29] Virginia Department of Transportation. Route 360 corridor study, hanover county, from chickahominy river bridge (henrico county) to pamukey river bridge (king william county). Richmond, study 30, 1990b.

[30] Vols. 1 US 15 James Madison Highway Passing Lane Study RMP 0.0 to RMP 15.5 and 2. Uma engineering, inc. Virginia Department of Transportation, Study 03, Richmond, 1997.

[31] Virginia Department of Highways. Colonial heights thoroughfare plan. Richmond, Study 04, 1970.

[32] Schuh & Jernigan Post, Buckley, Jack E. Leisch & Associates, and Inc.B & E Engineers Barton-Aschman Asssociates. Interstate 66: Fairfax and prince william counties, commonwealth of virginia, hov feasibility study. Virginia Department of Highways & Transportation, Richmond, study 06, 1986.

[33] Inc Louis Berger & Associates. I-95 clermont avenue interchange & connector to duke street, alexandria, virginia. Virginia Department of Transportation, Richmond, study 07, 1990.

[34] Virginia Department of Transportation. Route 13 corridor study, northampton county and accomack county. Richmond, study 08, 1988b.

[35] Virginia Department of Transportation. Routes 17 and 360 corridor study, town of tappahannock. Richmond, Study 09, 1989b.

[36] Virginia Department of Transportation. Route 40 needs assessment study, campbell county, charlotte county, lunenburg county, and nottoway county. Richmond, study 10, 1999a.

[37] Virginia Department of Transportation. Routes 20/240 corridor study, albemarle county. Richmond, study 11, 1990a.

[38] Virginia Department of Transportation. Route 608 corridor study. Richmond, study 12, 1996.

[39] Virginia Department of Transportation. Botetourt county route 220 (draft). Richmond, study 13, 1999b.

[40] Virginia Department of Transportation. Route 221/460 corridor study: Roanoke and botetourt counties. Richmond, study 15, 2002.

[41] Virginia Department of Transportation. Route 360 corridor study: Town of warsaw. Richmond, study 16, 1993.

[42] Virginia Department of Transportation. Dulles toll road extension route 267: Draft environmental document. Richmond, study 17, undated b.

[43] Virginia Department of Transportation. Route 240 corridor study, albemarle county. Richmond, study 18, 1990c.

[44] Virginia Department of Highways and Transportation. Pulaski area-year 2000 transportation plan. Richmond, study 19), 1981.

[45] Cather & Associates Deleuw. Peninsula area transportation study, recommended transportation plan, vol. ii. Commonwealth of Virginia, Richmond, Study 21, 1967.

[46] Virginia Department of Highways and Transportation. Virginia 2010 statewide highway plan: Culpeper district. Richmond, study 23, October, 1989c.

[47] Virginia Department of Highways and Transportation. Statewide highway plan, thomas jefferson planning district, pd-10. Richmond, study 24A, April, 1984a.

[48] Virginia Department of Highways and Transportation. Statewide highway plan, richmond regional planning district, pd-15. Richmond, study 24B, April, 1984b.

[49] Virginia Department of Highways and Transportation. Statewide highway plan, 5th planning district, pd-5. Richmond, study 24C, April, 1984c.

[50] The Cox Company. Traffic impact analysis: Rivanna village at glenmore, charlottesville. Charlottesville, Study 27, 2001.

[51] Inc Michael Baker Jr. Town of orange, 2020 transportation plan, technical report. Virginia Department of Transportation, Richmond, study 29, 2002.

[52] Jr. Johnson, W.L. Traffic data for design purposes route 10 from route 10 (cmaq intersection improvement) to route 746 (enon church road). VDOT Inter-Departmental Memorandum. Richmond, study 31, April 24, 2001.

[53] Frank Coleman & Associates and Inc WGM Development Consultants. Trip generation and distribution study, west creek parkway, goochland county, virginia. Study 33, 1988.

[54] Mattern & Mattern Inc Hayes, Seay. Route 1 appomattox river bridge and approaches, location study report, petersburg to colonial heights. Virginia Department of Transportation, Richmond, Study 34, April 1989.

[55] Azimi E. Email to J.S. Miller. Private Communication, February 13, 2015.

[56] Azimi E. Email to J.S. Miller. Private Communication, March 31, 2015.

[57] Azimi E. Email to J.S. Miller. Private Communication, October 9, 2015.

[58] Inc Vanasse Hangen Brustlin. Bjs wholesale club traffic impact assessment. Richmond, VA, study 38, October 16, 2007.

[59] Jr. Johnson, W.L. Traffic data for design purposes (route 360 from bell creek road relocated to closing crossover @ rte. 360). VDOT Inter-Departmental Memorandum. Virginia Department of Transportation, Richmond, study 39, February 14, 2002.

[60] Yinhai Wang, Bahar Namaki Araghi, Yegor Malinovskiy, Jonathan Corey, and Tianxing Cheng. Error assessment for emerging traffic data collection devices. Technical Report 2012-S-UW-0018, University of Washington, Seattle, 2014. [http://depts.washington.edu/pactrans/wp-content/uploads/2012/12/PacTrans-18-624909-McCormack-Edward-Small-Project.pdf; Accessed 2-June-2015].

[61] Jones R. Email to C. Lee and J.S. Miller. Private Communication, October 14, 2014.

[62] Chris Simons. I-285 matrix variegator: Practical method for developing trip tables for simulation modeling from travel demand modeling inputs. *Transportation Research Record: Journal of the Transportation Research Board*, (1981):18–23, 2006.

[63] DKS Associates. Curtis park village fuel center transportation analysis, technical report. Sacramento, CA, http://sierra2.org/LinkClick.aspx?fileticket=lOz60wfMvfg%3d&tabid=80, April 10, 2015. [online; Accessed 3-June-2015].

[64] Smidler S. Email to J.S. Miller. Private Communication, May 27, 2015.

[65] J.S. Miller, Salwa Anam, and Raleigh A Matteo. A retrospective evaluation of traffic forecasting techniques. Technical report, Virginia Transportation Research Council, Charlottesville, VA, 2016. [http://www.virginiadot.org/vtrc/main/online_reports/pdf/17-r1.pdf, accessed 10-Nov-2016].

[66] John S Miller, Salwa Anam, Jasmine Amanin, Raleigh Matteo, and Deven Barkley. Assumptions inherent in assessing traffic forecast accuracy. In *Transportation Research Board 95th Annual Meeting*, number 16-1555, 2016.

[67] Federal Reserve Bank of Richmond. National economic indicators. Richmond, May 9, 2016, http://ntl.bts.gov/lib/23000/23100/23121/09RoadFunction.pdf, November, 2000. [Online; Accessed 02-February-2016].

[68] Virginia Department of Transportation. Transportation enhancement program: Listing of mpo areas. http://www.virginiadot.org/projects/resources/mpo_listing.pdf, 2007. [online; Accessed 30-March-2016.].

[69] Dunnavant D. Email to S. Anam. study 07, Private Communication, March 3, 2016.

[70] B. Kemper. Coleman bridge to be widened: $1 toll charge could begin as early as july 1. The Daily Press, http://articles.dailypress.com/1989-03-17/news/8903170125_1_toll-money-coleman-bridge-board-members., study 02, March 17, 1989. [Online; Accessed 12-September-2014].

[71] Douglas C Montgomery. *Design and analysis of experiments, 5th edition*. John Wiley & Sons, Inc., 2001.

[72] David T Hartgen. Hubris or humility? accuracy issues for the next 50 years of travel demand modeling. *Transportation*, 40(6):1133–1157, 2013.

[73] J. DeBruhl. Development and financial constraint of virginias stip: Joint fhwa/fta/vdot/drpt process review. Federal Highway AdministrationVirginia Division, Richmond, http://www.virginia.edu/crmes/stip/STIP.pdf, 2002. [Online; accessed 18-Feb-2016].

[74] Morten Skou Nicolaisen and Patrick Arthur Driscoll. Ex-post evaluations of demand forecast accuracy: A literature review. *Transport Reviews*, 34(4):540–557, 2014.

[75] Bowman Consulting. Traffic impact analysis: Stonefield, albemarle county, virginia. Richmond, VA, May 27, 2011.

[76] Federal Highway Administration. Road function classifications. Federal Highway AdministrationVirginia Division, Richmond, http://ntl.bts.gov/lib/23000/23100/23121/09RoadFunction.pdf, November, 2000. [Online; Accessed 10-June-2015].

[77] Nicholas J Garber and Lester A Hoel. *Traffic and highway engineering*. Cengage Learning, Stamford, CA, 2015.

[78] Xiangdong Xu, Anthony Chen, SC Wong, and Lin Cheng. Selection bias in build-operate-transfer transportation project appraisals. *Transportation Research Part A: Policy and Practice*, 75:245–251, 2015.

[79] Inc Ramey Kemp & Associates. Technical appendix: Volume 2 revised traffic impact analysis report for the proposed albemarle place town center located in albemarle county, virginia. Raleigh, NC, January, 2002a.

[80] Inc Ramey Kemp & Associates. Technical appendix: Volume revised traffic impact analysis report for the proposed albemarle place town center located in albemarle county, virginia. Raleigh, NC, January, 2002b.

[81] Roadway Design Manual. Texas department of transportation. *Austin, Texas*, February, 2004. [http://onlinemanuals.txdot.gov/txdotmanuals/rdw/rdw.pdf; Accessed 11-June-2015].

[82] Virginia Center for Transportation Innovation and Research. Project detail: Assessment of trip generation estimation methods for mixed-use, infill, transit-oriented, or smart growth (mitus) developments. http://vtrc.virginiadot.org/ProjDetails.aspx?ID=587, 2013. [online; Accessed 18-March-2016].

[83] Virginia Department of Highways. Average daily traffic volumes on interstate and rural primary routes: 1960-1961. Richmond, 1961.

[84] Virginia Department of Transportation. Average daily traffic volumes with vehicle classification data on interstate, arterial and primary routes: 2004. Richmond, http://www.virginiadot.org/info/resources/AADT_PrimaryInterstate_2004.pdf, 2005. [online; Accessed 17-May-2016.].