

1. Report No. SWUTC/06/167556-1		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle Robust Design and Evaluation of Transportation Networks with Equilibrium Under Demand Uncertainty				5. Report Date October 2006	
				6. Performing Organization Code	
7. Author(s) Jennifer Duthie and S. Travis Waller				8. Performing Organization Report No. Report 167556-1	
9. Performing Organization Name and Address Center for Transportation Research University of Texas at Austin 3208 Red River, Suite 200 Austin, Texas 78705-2650				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No. 10727	
12. Sponsoring Agency Name and Address Southwest Region University Transportation Center Texas Transportation Institute Texas A&M University System College Station, Texas 77843-3135				13. Type of Report and Period Covered	
				14. Sponsoring Agency Code	
15. Supplementary Notes Supported by general revenues from the State of Texas.					
16. Abstract As long-term forecasts of origin-destination (O-D) travel demands are inherently uncertain, future network performance cannot be predicted with certainty. Neglecting the uncertainties present can result in inaccurate measures of network performance that may lead to incorrect policy decisions. This work relaxes the assumptions of determinism and independence of O-D demand in solving user equilibrium assignment. Insight is presented into the impacts of these assumptions through numerical analyses, where demand takes on various types of normal and lognormal multivariate distributions. The results indicate that incorrect assumptions of independence can, depending on the actual relationship between demands, lead to significant over- or under-estimation of network performance and incorrect network improvement decisions.					
17. Key Words Origin-Destination, Travel Demand, Travel Patterns, Transportation Network Performance			18. Distribution Statement No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161		
19. Security Classif.(of this report) Unclassified		20. Security Classif.(of this page) Unclassified		21. No. of Pages 57	22. Price

**Robust Design and Evaluation of Transportation Networks
with Equilibrium Under Demand Uncertainty**

By

Jennifer Duthie
S. Travis Waller

Research Report SWUTC/06/167556-1

Southwest Region University Transportation Center
Center for Transportation Research
University of Texas at Austin
Austin, Texas 78712

October 2006

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGEMENTS

Appreciation is expressed to The Southwest Region University Transportation Center for sponsoring this research study.

Our sincere appreciation is also extended to Avinash Unnikrishnan at the University of Texas at Austin's Center for Transportation Research for his assistance, particularly with implementing various sampling methods.

This publication was developed as part of the University Transportation Centers Program, U.S. Department of Transportation, which is funded 50% with general revenue funds from the State of Texas.

ABSTRACT

As long-term forecasts of origin-destination (O-D) travel demands are inherently uncertain, future network performance cannot be predicted with certainty. Neglecting the uncertainties present can result in inaccurate measures of network performance that may lead to incorrect policy decisions. This work relaxes the assumptions of determinism and independence of O-D demand in solving user equilibrium assignment. Insight is presented into the impacts of these assumptions through numerical analyses, where demand takes on various types of normal and lognormal multivariate distributions. The results indicate that incorrect assumptions of independence can, depending on the actual relationship between demands, lead to significant over- or under-estimation of network performance and incorrect network improvement decisions.

EXECUTIVE SUMMARY

How does a region know what its future transportation needs will be? Planning a transportation system to meet future needs requires information on future demographic and travel patterns. Many techniques exist to predict travel demand, but regardless of how fundamentally sound a technique is, it will never be able to predict the future with complete accuracy.

Once it is accepted that the future is unlikely to be realized exactly as forecast, it no longer makes sense to base decisions on a single point prediction. Instead of designing a system that will succeed under the forecasted conditions, it is more prudent to design a system that is robust, meaning it functions well under a wide variety of feasible future scenarios. Thinking of transportation systems in this manner is analogous to a school of thought already prevalent in the business community. Companies such as Smith and Hawken, and Royal Dutch/Shell, have achieved successes by consciously planning for feasible future scenarios, both unlikely and likely (Schwartz, 1991). By considering the implications of various future scenarios on their business plan, organizations will have a greater chance at success than those that pretend current trends will continue or those that neglect planning and hope for the best.

In order to achieve accurate results for transportation network performance, accurate predictions of travel demand are needed. In practice, a single value is used for demand between each origin-destination pair. Scenario-based analysis, a tool increasingly used by planners to identify several possible futures, can be an improvement over single point estimation when scenarios are used to represent possible outcomes. Instead of looking at each scenario as possible, many planning exercises allow the public to choose their most desired scenario. Choosing one scenario to use as a planning guide, however, is equivalent to choosing a single point estimate of the future.

To gain insights into the problem of uncertainty in long-term forecasting, this paper assigns multivariate distributions to the origin-destination demands. Especially in scenario analysis, treating demand as correlated rather than independent is more intuitive because the demand patterns within each scenario are related to one another. This paper compares measures of network performance from user equilibrium assignment between independent and correlated demand scenarios. Since the exact relationships between origin-destination demands are likely to be network dependent, a variety of correlation structures and probability distributions are considered.

The method for assessing the effect of origin-destination correlations on network performance is as follows. Origin-destination correlations are checked for positive semi-definiteness and modified if necessary using Hypersphere Decomposition. The matrix correlations, along with expected demands and variances, are used to randomly generate a finite number of origin-destination demand matrices, each matrix representing one possible scenario, using the Box-Muller and Cholesky Decomposition methods. Finally, deterministic user equilibrium is applied for each demand realization. Statistics, such as mean and variance of the results, are then collected and analyzed.

This research is the first to explore the effects of considering long-term origin-destination demand as correlated in traffic assignment. It motivates research into determining if network design decisions will change based on forecasts of correlations, and into achieving these correlations. In particular, it shows how network robustness relies heavily on the type of correlations that exist between origin-destination pairs. Although no exact prescription is given

for when correlations should be considered, general trends may help guide this decision. These trends include:

- 1.) As correlations become more positive, the error in TSTT caused by using a deterministic model may increase.
- 2.) As the magnitude of correlations increase, the error in mean TSTT caused by using a stochastic and independent model may increase.
- 3.) As the magnitude of correlations increase, the error in variance of TSTT caused by using a stochastic and independent model may increase.

The first trend suggests that current models using a deterministic value of demand may be inadequate if origin-destination demands have positive correlations. TSTT, a common measure of network performance, may in some cases be grossly underestimated. Future research is needed to test if this trend becomes less significant as network size grows.

The second trend suggests that just considering demand uncertainty may not be sufficient to achieve accurate TSTT results if the correlations between origin-destination demands are neglected. The more the correlations differ from zero, the worse the predictions of mean TSTT. Future research is needed to test if this trend becomes less significant as network size grows.

The third trend is especially important. As the importance of network robustness is more widely recognized and applied to practice, the consideration of correlations becomes vital. If large correlations exist between future origin-destination demand predictions and they are not considered, the potential errors are enormous.

TABLE OF CONTENTS

ABSTRACT	v
ACKNOWLEDGEMENTS.....	iv
EXECUTIVE SUMMARY	vii
TABLE OF CONTENTS	ix
LIST OF FIGURES	xi
LIST OF TABLES.....	xii
CHAPTER ONE - INTRODUCTION	1
1.1 BACKGROUND	1
1.2 MOTIVATION.....	2
1.3 PROBLEM STATEMENT.....	3
1.4 CONTRIBUTIONS	5
1.5 ORGANIZATION.....	5
CHAPTER TWO - LITERATURE REVIEW	7
2.1 INTRODUCTION	7
2.2 TRAFFIC EQUILIBRIUM	7
2.3 UNCERTAINTY IN TRANSPORTATION MODELING	8
2.4 ROBUSTNESS.....	9
2.5 CORRELATED DEMAND	10
CHAPTER THREE - METHODOLOGY.....	13
3.1 INTRODUCTION	13
3.2 USER EQUILIBRIUM.....	13
3.4 METHODOLOGY FOR CONSIDERING CORRELATED DEMAND	14
3.4.1 Scaling Factors	16
3.4.2 Correlation Matrices.....	17
CHAPTER FOUR – NUMERICAL ANALYSIS.....	19
4.1 SMALL EXAMPLE.....	19
4.1.1 Deterministic Demands	19
4.1.2 Independent Demands	20
4.1.3 Positively Correlated Demand.....	21
4.1.4 Negatively Correlated Demands.....	22
4.2 NETWORKS AND INPUTS	22
4.2.1 Test Network	22
4.2.2 Nguyen-Dupuis Network.....	25
4.2.3 Sioux Falls Network	26
4.3 Results	30
4.3.1 Test Network	30
4.3.2 Nguyen-Dupuis Network.....	33
4.3.3 Sioux Falls Network	36
4.3.4 Conclusions	38

CHAPTER FIVE - CONCLUSIONS.....	39
5.1 IMPLICATIONS OF WORK.....	39
5.2 FUTURE WORK	39
5.2.1 Achieving Correlations.....	40
5.2.1.1 Four Step Process	40
5.2.1.2 Activity-Based	41
REFERENCES	43

LIST OF FIGURES

Figure 1.1a Four Step Planning Process	1
Figure 1.2a Example	3
Figure 1.3a Robust vs. Deterministic Planning	4
Figure 3.4a. Numerical method for assessing uncertain and correlated long-term O-D demands	15
Figure 4.1a Example Network	19
Figure 4.1.1a Example Network: Arc Flows in Deterministic Case	19
Figure 4.1.2a Example Network: Arc Flows in Scenario 1	20
Figure 4.1.2b Example Network: Arc Flows in Scenario 2	20
Figure 4.1.2c Example Network: Arc Flows in Scenario 3	21
Figure 4.1.2d Example Network: Arc Flows in Scenario 4	21
Figure 4.2.1a Test Network	23
Figure 4.2.1b Test Network: Correlation Matrices	24
Figure 4.2.1c Test Network: Congestion Levels	24
Figure 4.2.2a Nguyen-Dupuis Network	25
Figure 4.2.2b N-D Network: Correlation Matrices	26
Figure 4.2.2c N-D Network: Congestion Levels	26
Figure 4.2.3a S.F. Network:	27
Figure 4.2.3b S.F. Network: Correlation Matrices	29
Figure 4.2.3c S.F. Network: Congestion Levels	30
Figure 4.3.1a Test Network: TSTT v. Total System Demand, (given C1)	31
Figure 4.3.1b Test Network (VSF ¹): % Diff from Deterministic v. Congestion Level	31
Figure 4.3.1c Test Network (VSF ¹): % Diff from Mean of Independent Case v. Congestion Level	32
Figure 4.3.1d Test Network (VSF ⁶): % Diff from StdDev of Independent Case v. Congestion Level	33
Figure 4.3.2a Test Network: TSTT vs. Total System Demand, (given C3)	33
Figure 4.3.2b N-D Network (VSF ¹): % Diff from Deterministic v. Congestion Level	34
Figure 4.3.2c N-D Network (VSF ¹): % Diff from Mean of Independent Case v. Congestion Level	35
Figure 4.3.2d N-D Network (VSF ⁶): % Diff from StdDev of Independent Case v. Congestion Level	35
Figure 4.3.2a S.F. Network: TSTT vs. Total System Demand, (given C1)	36
Figure 4.3.3b S.F. Network (VSF ¹): % Diff from Deterministic v. Congestion Level	36
Figure 4.3.3c S.F. Network (VSF ²): % Diff from Mean of Independent Case v. Congestion Level	37
Figure 4.3.3d S.F. Network (VSF ²): % Diff from StdDev of Independent Case v. Congestion Level	38
Figure 5.2.1.1a Four-Step Process	40

LIST OF TABLES

Table 4.1a Example Network: O-D Demands.....	19
Table 4.1.2a Example Network: O-D demands assuming independence.....	20
Table 4.1.2b Example Network: Results Assuming Independence.....	21
Table 4.1.3a Example Network: O-D Demands assuming positive correlations	22
Table 4.1.4a Example Network: O-D demands assuming negative correlations	22
Table 4.2.1a Test Network: Expected Demand.....	23
Table 4.2.1b Test Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k	23
Table 4.2.2a N-D Network: Expected Demand.....	25
Table 4.2.2b N-D Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k	25
Table 4.2.3a S.F. Network: Expected Demand	27
Table 4.2.3b S.F. Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k	28

CHAPTER ONE - INTRODUCTION

1.1 BACKGROUND

Since the Federal Surface Transportation Assistance Act of 1973, urbanized areas with 50,000 or more people have been required to form metropolitan planning organizations (MPOs) to ensure current and future projects use the 3-C (continual, cooperative, and comprehensive) planning process. One of the main functions of an MPO is to maintain a long-range transportation plan (LRTP). “Long-range” is defined as at least twenty years, though longer planning horizons are often used. A second important function is to evaluate transportation alternatives, considering realistic available options.

The development of a LRTP involves determining the transportation demand over the course of the plan, identifying future projects and policies, and identifying goals for future land use and demographics. To determine demand for transportation infrastructure, a four step process is typically employed. This process is shown in Figure 1.1a.

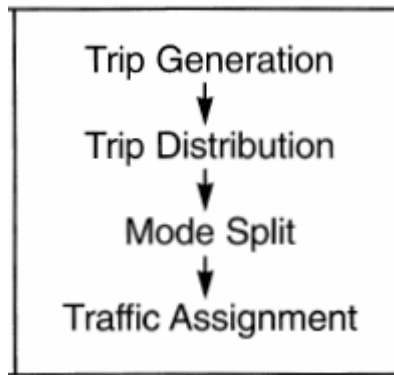


Figure 1.1a Four Step Planning Process

In the first step, “Trip Generation,” the number of trips from each origin and the number of trips to each destination are predicted. The second step, “Trip Distribution,” assigns the results of the first step into O-D demands. The third step, “Mode Split,” determines which modes will be used for each trip. The fourth step, “Traffic Assignment,” assigns each trip to a specific route within the network.

This research focuses on the traffic assignment step. At this step, trip demands are taken as an input and flows on each roadway link are output. This output is then often used to inform policy decisions regarding capacity improvements. Total system travel time (TSTT), the total time spent in the system for all vehicles, is another output of traffic assignment and is the most common measure of performance for a transportation network. Network designs are often chosen from a set of options such that minimize TSTT.

For LRTPs, the trip demands that are input into traffic assignment are predictions of the future and are therefore not known with certainty. This research explores the results of traffic assignment when trip demands are considered to be random variables instead of deterministic quantities.

1.2 MOTIVATION

This research is motivated by the fact that the transportation planning process relies on current data and predictive models to forecast an uncertain future. It is widely accepted that transportation forecasts are intimately connected to factors including land use decisions, demographic changes, economic conditions, and technological innovations. Instead of trying to perfectly predict the future, an impossible task, this research discusses ways to recognize the uncertainties associated with travel demand and accommodate them to evaluate networks to determine which ones perform the best overall given a host of future scenarios.

In order to achieve accurate results for network performance, accurate predictions of demand are needed. In practice, planners typically use a single value for each expected demand or in some cases use multiple qualitatively-based scenarios. In this research, not only are future origin-destination (O-D) demands considered to be uncertain, but both independent and correlated cases are explored. For example, if the demand for one O-D pair is higher than expected, it is foreseeable that all O-D demands are higher than expected. The reverse could also be true. Exact relationships between O-D demands are network dependent and may be difficult to capture. Considering long-term O-D demand as correlated yields different results for system cost and robustness than the independent and deterministic cases, and may lead to different network design decisions.

As an example, consider a town with three zones (Figure 1.2a): A, B, and C. Let zone A be a residential zone, B a commercial zone, and C an industrial zone. In this town, there is one origin, A, and two destinations, B and C. The town planners decide that depending on the number of people who move to and from zone A and the number of businesses located in zone B, the future demand from A to B will be in the range of 100 to 400 trips per day. Also, depending on the number of people who move to and from zone A and the number of industries located in zone C, the future demand from A to C will be in the range of 50 to 500 trips per day. In addition to the future number of residents, different economic scenarios and possible future demographics should also be taken into account in computing these ranges. Intuitively, the demand from A to B is unlikely to be independent of the demand from A to C. If, however, future O-D demands are treated as uncertain and independent when analyzing future traffic conditions, it is just as likely to have a scenario where demand A to B is 100 and A to C is 500 as it is to have the demand from A to B as 100 and B to C as 50. If the demands are instead taken to be correlated, both of these scenarios may still be possible, but their probability of occurrence may change drastically.

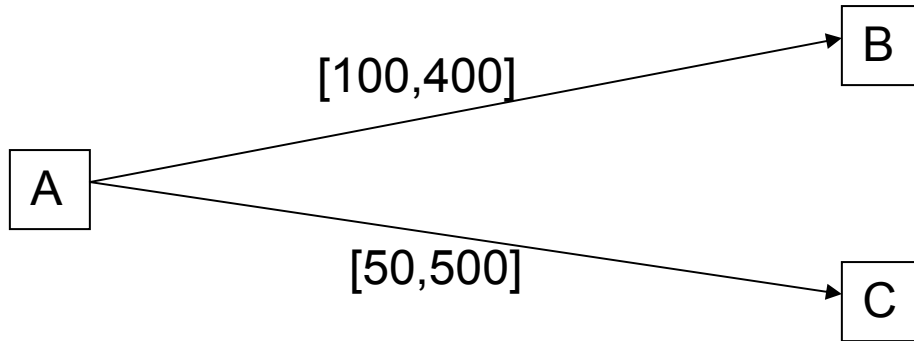


Figure 1.2a Example

For simplicity, only one origin is considered for the example network in Figure 1.2a. However, it is possible that the demands for O-D pairs with different origins and different destinations would also be correlated if their future conditions are likely to be similar or systematically dissimilar. Intuitively, positive correlations between O-D demands seem more likely, but negative correlations may also be possible if the O-D pairs are somehow competing.

Not only is it important to design a network that performs well if the realized future demand is exactly the expected demand, but it is also important that the network behaves well if the realized future demand differs from what was expected. This behavior can be observed by testing the network under a variety of future scenarios and seeing how the results vary. A robust network, as defined in this research, is one which varies little given unexpected conditions. Here, system robustness is measured as the variance in TSTT. It seems intuitive that treating demand as correlated will affect the variance of TSTT. As correlations become more positive, it seems intuitive that the variance of TSTT will increase because, as one O-D demand strays from the mean, the likelihood that other demands stray from the mean in the same direction is increased. This will lead to more frequent extreme low and extreme high TSTT values as compared to the independent case. As correlations become more negative, it seems intuitive that the variance of TSTT will decrease because as one O-D demand strays from the mean, the likelihood that other demands stray from the mean in the opposite direction is increased. This leads to a “balancing effect” where each higher than average demand is likely to be countered by a lower than average demand, causing a smaller variance in TSTT.

The next section describes in greater detail the problem to be solved, and the approach taken in this research to solving it.

1.3 PROBLEM STATEMENT

The goal of this work is to determine whether considering long-term O-D demand as correlated will impact predictions of network performance. Long-term demand refers to demand many years into the future, as opposed to short-term demand which focuses on day-to-day fluctuations in the number of trips. Treating long-term demand as uncertain is important for network design decisions such as capacity improvements that are typically made in the present to meet predicted future demand. This research considers only long-term uncertainty.

Not only may long-term O-D demand be uncertain, it may also be correlated. This research seeks to determine if considering these correlations will influence the results of traffic assignment. If uncertainty or correlations impact results, then neglecting these may result in misled network design and policy decisions.

Two measures of performance are considered: mean TSTT and variance of TSTT. Variance of TSTT is one possible measure of system robustness. Higher moments are other possible measures. A robust network is one that is least affected by the uncertainty of future demand. Given a set of future demand scenarios, the most robust network may not be the best choice in each case, but its performance varies the least. This concept is illustrated in Figure 1.3a.

In this figure, “Deterministic Planning” means a network design is chosen such that network performance is optimized for a pre-specified level of demand. “Robust Planning” means a network design is chosen such that network performance varies the least over the range of uncertainty. The “Deterministic Planning” and “Robust Planning” curves show the resulting network performance for all possible values of the stochastic parameter. “Optimal Solution” refers to the TSTT predicted by deterministic planning. If deterministic planning is used and the stochastic parameter is realized at its pre-specified value, then the optimal solution will be achieved. Deterministic planning performs better than robust planning if the stochastic parameter is realized at or near its pre-specified level. If the stochastic parameter varies more than a small amount from its pre-specified level, robust planning gives a better solution. In general, robust planning yields a more stable, though in some cases worse, solution than deterministic planning.

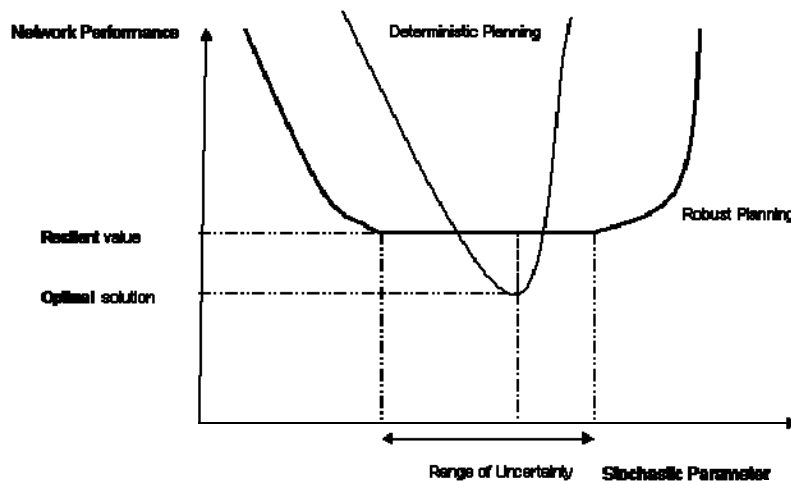


Figure 1.3a Robust vs. Deterministic Planning

A small network using discrete distributions to model demand is solved analytically in section 4.1. Larger networks with continuous demand distributions are solved using Monte Carlo simulation. Each model is run for a variety of correlations, expected demands, and variances. Varying the expected demands gives insight into how TSTT results differ depending on network congestion. The demand variances indicate the uncertainty of the expected demand value.

Section 1.4 gives a succinct description of the contributions of this research to the field of transportation network analysis and planning.

1.4 CONTRIBUTIONS

To our knowledge, this work is the first to explore in detail how uncertain and correlated long-term O-D demands affect network performance. Existing literature has mentioned the possible importance of correlations, but this research, we believe, is the first to use numerical analyses to test the degree of this importance. The results of this work have the potential to change how transportation networks are analyzed and designed. O-D demand uncertainty and correlations are shown in most cases to have a significant impact on both mean TSTT and network robustness; therefore ignoring them may yield inaccurate traffic assignment results and poor network design decisions. This work may motivate research in creating a methodology for determining realistic long-term O-D demand distributions and correlations.

1.5 ORGANIZATION

Chapter 2 will describe relevant existing literature in the areas of traffic equilibrium, uncertainty in transportation modeling, robustness, and correlated demands. Chapter 3 describes user equilibrium formulations that treat demands as either deterministic or stochastic. A methodology is described for considering O-D demand correlations. Chapter 4 shows analytically how the methodology presented in chapter 3 can be applied to a small network, and gives the numerical results for three networks and general conclusions. Chapter 5 concludes the research and discusses opportunities for future work.

CHAPTER TWO - LITERATURE REVIEW

2.1 INTRODUCTION

The focus of this research is to determine the effect of uncertain and correlated long-term O-D demands on network performance. This section will review relevant literature used to motivate the problem in the following order: Traffic Equilibrium, Uncertainty in Transportation Modeling, Robustness, and Correlated Demands.

2.2 TRAFFIC EQUILIBRIUM

It is widely accepted in transportation literature that drivers route themselves to minimize their personal travel costs, without regard for others. This was first stated by Wardrop in what is commonly referred to as Wardrop's (1952) first principle: the journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route. A system with drivers exhibiting this behavior is said to be in User Equilibrium (UE), meaning no user can unilaterally choose a different path and improve his travel time. The work presented in this research assumes users behave in this manner.

A second type of equilibrium, system optimal (SO), is described in Wardrop's second principle: At equilibrium the average journey time is at a minimum. Here, a user may choose a costlier path if it benefits the overall system. Solving for SO outputs the total time spent in the system by all users. This output, TSTT, is one measure of evaluation used in this research.

In both the UE and SO formulations, there must be a method for describing link performance, or how the time to traverse a link depends on its current flow. One typical method is using link performance functions such as The Bureau of Public Records (BPR) formulation (U.S. Department of Commerce, 1964). BPR cost functions are expressed as follows:

$$T = T_0[1 + B(V/C)^\gamma] \quad \text{where,}$$

T is link travel time,

T_0 is free-flow travel time,

V is hourly volume,

C is hourly capacity, and

B and γ are parameters that depend on link geometry.

The mathematical programming formulations of static UE and SO problems are described in detail in section 3.2. Since both formulations are non-linear, solving the problem is typically done iteratively. Method of Successive Averages (MSA) and Frank-Wolfe (F-W) are two common approaches. Both methods solve the shortest path problem given free-flow travel times, then update travel costs, solve the new problem, update travel costs for a weighted average of the current and previous solutions, and repeat until convergence is achieved. All iterations of MSA use predetermined, fixed weights. MSA finds each new solution by adding $(1-1/n)$ multiplied by the previous solution and $1/n$ multiplied by the current solution, where n is the number of iterations

performed. F-W is similar to MSA except it uses a line-search to determine optimal weights, λ , to use within a given iteration. The F-W algorithm is as follows (Sheffi 1985):

Step 0: Initialization. Perform all-or-nothing assignment based on $t_a = t_a(0), \forall a$. This yields $\{x_a^1\}$. Set counter $n:=1$.

Step 1: Update. Set $t_a^n = t_a^n(0), \forall a$.

Step 2: Direction finding. Perform all-or-nothing assignment based on $\{x_a^n\}$. This yields a set of (auxiliary) flows $\{y_a^n\}$.

Step 3: Line Search. Find α_n that solves $\min_{0 \leq \alpha \leq 1} \sum_a \int_0^{x_a^n + \alpha(y_a^n - x_a^n)} t_a(\omega) d\omega$

Step 4: Move. Set $x_a^{n+1} = x_a^n + \alpha(y_a^n - x_a^n), \forall a$

Step 5: Convergence test. If a convergence criterion is met, stop (the current solution, $\{x_a^{n+1}\}$, is the set of equilibrium link flows); otherwise, set $n := n+1$ and go to step 1.

This line-search adds complexity to the problem; however, F-W is used in the numerical analyses presented in this research because it often converges to a solution much more quickly than MSA.

If future demand is assumed to be known, than only one run of the user equilibrium algorithm is required. If the uncertainty in demand is considered, however, the algorithm may have to be repeated many times.

2.3 UNCERTAINTY IN TRANSPORTATION MODELING

Uncertainty is inherent in all aspects of the transportation modeling process. Work has been done in considering error propagation and modeling uncertainties in link capacities, user's perceptions of travel time, and short-term demand and long-term demand. Long-term demand uncertainty is the focus of this research.

One type of uncertainty that exists in transportation analysis is not true uncertainty, but rather errors that propagate through the analysis. These errors may have a variety of origins including imprecise calculations, systematic biases, and data gathered from non-representative or small samples. Frey (1992) and Stopher and Meyburg (1975) discuss these sources of error. Harvey and Deakin (1995) postulate that socioeconomic projections yield more errors than any other part of transportation planning. Zhao and Kockelman (2002) state that errors compound and magnify over the course of the planning process, however, the traffic assignment step has a centralizing effect that reduces errors back to their input levels. These types of errors are not dealt with in this research, only true uncertainty in future demand.

True uncertainty in transportation has been studied primarily in terms of capacity reliability (Iida and Wakabayashi, 1989). A comprehensive review of the various definitions of capacity reliability is given in Bell and Cassir (2000). Chen et al (1999) defined capacity reliability as the probability that the network can accommodate a certain demand at a given service level. Chen et al (2002) introduced capacity reliability as a network performance index. Lo and Tung (2002) dealt with the problem of link

capacities subject to stochastic degradations day-to-day, and define capacity reliability as the maximum flow that the network can carry given link capacity and travel time reliability constraints. Du and Nicholson (1997) proposed a conventional equilibrium approach with variable demand to describe flows in a network with degradable link capacities.

A third type of uncertainty in transportation modeling is in user perception of travel time. This is typically dealt with using stochastic user equilibrium (SUE) methods. SUE allows for different user classes with perfect or imperfect perception of travel time. The probability of choosing each route is determined by a stochastic loading model such as a logit model, the STOCH algorithm developed by Dial (1971), or the SAM probit model (Maher(1992) and Maher and Hughes (1997a)). A mathematical program for SUE is given in Sheffi and Powell (1982).

A fourth type of uncertainty is in the day-to-day, or short-term, fluctuations in O-D demand. Short-term demand uncertainty is important when analyzing the present day network under present day conditions. Instead of assuming that the number of trips is the same every day and using a constant O-D trip matrix, it may be more accurate to assume that demand varies according to some probability distribution.

In long-term planning where demand is forecast many years into the future, there are more opportunities to plan ahead and alter or add to the network infrastructure. Waller et al (2001) describe how to evaluate traffic assignment under demand uncertainty. The researchers employ Jensen's Inequality (1906) and numerical analysis to show that using expected demand levels systematically underestimates the actual expected future demand costs. Jensen's Inequality states that for convex functions, $f(x)$, the following condition holds:

$$E[f(x)] \geq f(E[x])$$

The SO objective function is convex, and numerical analyses of the UE objective function show that it tends to be convex.

Waller and Ziliaskopoulos (2001) propose a dynamic network design model as a two stage stochastic linear programming problem where the Cell Transmission Model (Daganzo, 1994, 1995) is the embedded traffic flow model in the second stage and O-D demand was modeled as a random variable. Long-term demand uncertainty can be accounted for using stochastic optimization methods with either a recourse or a chance-constrained formulation as demonstrated by Waller and Ziliaskopoulos (2001). All of the aforementioned research assumes O-D demands are independent. The next section describes how to use this uncertainty to capture another measure of network performance.

2.4 ROBUSTNESS

While some literature exists on accounting for system robustness in transportation, it is more prevalent in other fields such as finance. Markowitz (1952) introduced the idea of a trade-off when investing, between risk and expected return. He developed the "critical line algorithm" that found all feasible portfolios that use variance to minimize risk for a given expected return, and also minimize expected return for a given level of risk. Markowitz's model is applicable to static, single-period situations, but may not be appropriate for long-run decisions.

Mulvey et al (1995) present a general robust optimization formulation that balances solution robustness (close to optimal for all scenarios) with model robustness (almost feasible for all scenarios). Highly volatile solutions are discouraged by minimizing variance in the objective function. This approach is appropriate if the input parameters are uncertain with known, symmetric distributions or if there exist multiple bounded random input parameters with unknown, symmetric distributions.

A model for asymmetric random variables is presented in Keeney and Raiffa (1976). This approach is based on the von Neumann-Morganstern expected utility functions. Paraskevopoulos et al (1991) define robust capacity planning under certainty for the plastics industry. Sengupta (1991) describes robustness in stochastic programming models. Escudero et al (1993) proposed a robust optimization formulation for the problem of outsourcing in manufacturing.

All of the above literature considers only independent variables. Correlated variables, described in the next section, may lead to different results and require different methodologies.

2.5 CORRELATED DEMAND

A correlation coefficient is a measure of the degree of linear relationship between two variables. It can take on any value between -1 and +1 inclusive. A positive correlation implies that as the value of the first variable increases, the value of the second value is likely to increase; and as the value of the first variable decreases, the value of the second value is likely to decrease. A negative correlation implies that as the value of the first variable increases, the value of the second variable is likely to decrease, and vice versa. A correlation coefficient of zero implies no linear relationship exists, and correlations of +1 and -1 imply perfect linear relationships.

Correlations, ρ , can be calculated as follows:

$$\rho_{x,y} = \frac{\sum_{i=1}^N z_x z_y}{N-1}$$

where,

$$z_x = \frac{X - \bar{X}}{\sigma_x}, \quad z_y = \frac{Y - \bar{Y}}{\sigma_y},$$

X is a sample from population 1

Y is a sample from population 2

\bar{X} is the mean of the sample taken from population 1

\bar{Y} is the mean of the sample taken from population 2

N is the sample size

σ_x is the standard deviation of the sample taken from population 1

σ_y is the standard deviation of the sample taken from population 2

Certain properties must hold for a matrix to be a valid correlation matrix. It must be symmetric, positive-semidefinite, and have all ones on the main diagonal. Rebonato and Peter (1999) describe a method for using Hypersphere Decomposition to transform

any symmetric matrix into a valid correlation matrix. As a correlation matrix must be positive semi-definite by definition, it is not trivial to obtain one as even small changes are likely to alter the structure of the correlation matrix in such a way that it is no longer valid and cannot be sampled from for system evaluation. The algorithm can be implemented as follows:

- 1.) Choose an $n \times n$ target matrix, C
- 2.) Let $\{\theta_{ij}\}$ be an $n \times (n-1)$ matrix of arbitrary angles
- 3.) Create a $n \times n$ matrix, B , as follows:

$$b_{ij} = \cos \theta_{ij} \cdot \prod_{k=1}^{j-1} \sin \theta_{ik} \text{ for } j = 1..n-1$$

$$b_{ij} = \prod_{k=1}^{j-1} \sin \theta_{ik} \text{ for } j = n$$

- 4.) Find the $n \times n$ matrix, \hat{c} such that χ^2 is minimized where

$$\chi_{Elements}^2 := \sum_{ij} (c_{ij} - \hat{c}_{ij})^2$$

Use of correlated demands appears sparingly in other domains. Liu and Yuan (2000) present a Markovian model for a two-item inventory system with correlated demands and show how correlations influenced replenishment policy.

Little work exists in correlated demands in transportation network analysis. Waller and Ziliaskopolous (2001) mention the need to consider correlations, but no analysis has been done until this research. The following chapter presents a methodology for this analysis.

CHAPTER THREE - METHODOLOGY

3.1 INTRODUCTION

In this section, a methodology is presented to consider uncertain and possibly correlated demand in traffic assignment, where expected system cost and robustness are the measures of performance. To motivate this methodology, the traditional UE model is presented first followed by a discussion of extensions to account for demand uncertainty and robustness.

3.2 USER EQUILIBRIUM

The traditional user equilibrium formulation models Wardrop's first principle: the journey times in all routes actually used are equal and less than those which would be experienced by a single vehicle on any unused route. Beckmann (1956) introduced the following optimality conditions for UE:

$$\begin{aligned} h_{wr} (c_{wr} - \pi_w) &= 0, r \in R_w, w \in W, \\ c_{wr} - \pi_w &\geq 0, r \in R_w, w \in W, \\ \sum_{r \in R_w} h_{wr} &= d_w, w \in W \end{aligned}$$

where,

W is the set of O-D node-node pairs

R_w is the set of routes joining node pair $w \in W$

h_{wr} is the flow on route $r \in R_w$

π_w is the minimum cost on any route joining node pair w

d_w is the demand for service between the node pair w

c_{wr} cost or delay on that route $r \in R_w$

Adding nonnegativity restrictions $h_{wr} \geq 0$ and $\pi_w \geq 0$, the resulting system of equalities and inequalities become the Karush-Kuhn-Tucker (KKT) optimality conditions of the following optimization problem, known as the Beckmann transformation (Beckmann, 1956).

$$\begin{aligned} \min f(x) &= \sum_{l \in A} t_l(x_l) dx = \sum_{l \in A} \int_0^{\sum_{i \in N} x_{il}} t_l(x) dx \\ \text{subject to} \quad & \sum_{r \in R_w} h_{wr} = d_w, w \in W, \\ & \sum_{w \in W} \sum_{r \in R_w} h_{wr} \delta_{wr}^l = x_l, l \in A, \\ & x_l \geq 0, l \in A, \end{aligned}$$

where,

A is the set of links

N is the set of nodes

x_{il} is the flow of users from node i on link l

x_l is the total flow on link $l \in A$

$t_l(x_l)$ is the cost on link l given link flow x_l , $l \in A$

δ_{wr}^l is a 0-1 indicator function that takes the value 1 when link l is present on route $r \in R_w$

Assume that c_{wr} is the sum of link costs $t_l(x_l)$ of the links l along route r between node pair w , and that the cost t_l is a function only of the total flow x_l over each link l . Some of the pitfalls of this formulation include that it assumes static conditions, cannot adequately model congestion effects (i.e. queuing and shockwaves), cannot model traffic signals, and all parameters are assumed to be known with certainty.

The model above can be extended to account for demand uncertainty by making d_w a random variable. Instead of minimizing for TSTT, the objective function then works to minimize the expected value of TSTT. The problem increases in difficulty as the number of demand scenarios increases. To account for system robustness, a variance (or higher order moment) term can be added to the objective function. Section 3.4 describes a method for considering demand correlations within the UE framework.

3.4 METHODOLOGY FOR CONSIDERING CORRELATED DEMAND

The method for assessing the effect of O-D correlations on network performance is as follows and is summarized in Figure 3.4a. As stated previously, O-D correlations are checked for positive semi-definiteness and modified if necessary using Hypersphere Decomposition (see Section 2.5). The matrix correlations, along with expected demands and variances, are used to randomly generate a finite number of O-D demand matrices, each matrix representing one possible scenario, using the Box-Muller and Cholesky Decomposition methods. In Figure 3.4a, the arrow from expected O-D demand to variance of O-D demands is present because in the numerical analyses demonstrated in Section 4, variances are obtained by scaling the expected demands. This is done for the sake of simplicity. In an actual network, variance would be separately estimated. Finally, deterministic user equilibrium is applied for each demand realization. Statistics, such as mean and variance of the results, are then collected and analyzed.

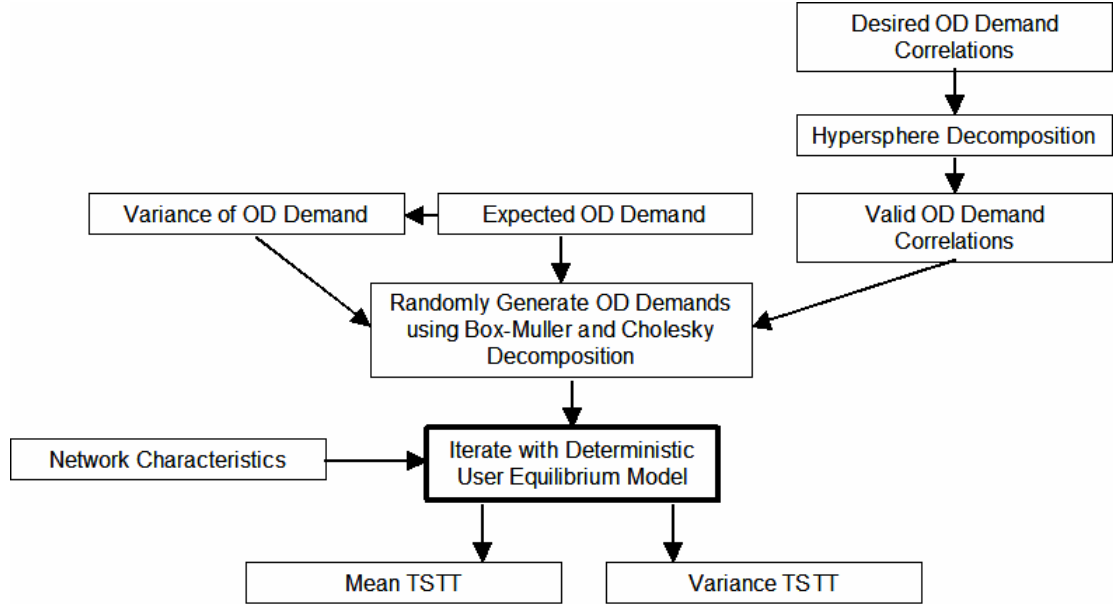


Figure 3.4a. Numerical method for assessing uncertain and correlated long-term O-D demands

The Box-Muller Method inputs uniform[0,1] variables and transforms them into standard normal random variables. Let x_1 and x_2 be numbers generated from a uniform[0,1] distribution. Solve for two standard normal random variables, y_1 and y_2 , as follows:

$$y_1 = \sqrt{-2 \ln x_1} \cos 2\pi x_2$$

$$y_2 = \sqrt{-2 \ln x_1} \sin 2\pi x_2$$

To transform y_1 and y_2 into normal variables, z_1 and z_2 , with a mean μ and variance σ^2 :

$$z_1 = \sigma^2 * y_1 + \mu \quad \text{and} \quad z_2 = \sigma^2 * y_2 + \mu$$

Only one of the normal variables is needed and the other may be discarded.

Cholesky Decomposition inputs a positive semi-definite matrix, A , and outputs a lower triangular matrix, L , such that $A = L * L^T$. The elements of L are defined as:

$$l_{ii} = \sqrt{(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2)}$$

$$l_{ji} = (a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik}) / l_{ii}$$

Since A is symmetric and positive semi-definite, all l_{ij} are real, and the expression under the square root is always positive.

In this research, OD demand is assumed to be normally distributed with given mean, variance, and correlations. The distribution was always truncated at zero to eliminate the possibility of negative demands. Alternate distributions such as uniform and truncated exponential should be considered in the future. The normal distribution was chosen because it is common and generating correlated random normal variables is

relatively simple. Although vehicles are discrete units, a continuous distribution is still likely appropriate due to the large number of vehicles on a network.

The following measures were computed to determine the impact of varying the congestion level, O-D variance, and correlations.

- 1.) Percent difference in mean Total System Travel Time (TSTT) using correlated vs. uncorrelated O-D demands
- 2.) Percent difference in mean TSTT between the deterministic case vs. correlated O-D demands
- 3.) Percent difference in the standard deviation of TSTT using correlated vs. uncorrelated O-D demands

If the first measure is determined to be substantial, then treating demand as uncertain but neglecting correlations will result in a poor estimate of average system performance. If the second measure is determined to be substantial, then using evaluating network performance by setting O-D demands to their expected value is a poor estimate of mean TSTT. If the third measure is substantial, then treating demand as uncertain but neglecting correlations will lead to a poor estimate of system robustness.

3.4.1 Scaling Factors

To obtain a wide variety of network scenarios, scaling factors are applied to the original demand matrices to test the impact of congestion and demand uncertainty on the expected value and variance of TSTT.

$$E[D_{ij}^k] = E[D_{ij}^0]/DSF^k, \text{ for all } (i,j) \text{ in the set of O-D pairs, and } k=1..K_d,$$

$$\text{Var}[D_{ij}^k] = E[D_{ij}^k]/VSF^k, \text{ for all } (i,j) \text{ in the set of OD pairs, and } k=1..K_v,$$

where,

D_{ij}^k := demand of OD pair (i,j) using scaling factor k,

D_{ij}^0 := original demand of OD pair (i,j) ,

DSF_k is the demand scaling factor for scenario k

VSF_k is the variance scaling factor for scenario k

K_d is the total number of demand scaling factors,

K_v is the total number of variance scaling factors,

$E[-]$ represents the expected value, and

$\text{Var}[-]$ represents variance.

These scaling factors are presented for each network in Section 4.2 . Scaling the expected value of OD demands is important to mitigate for congestion effects. Variations in demand of an uncongested network will have less of an impact on travel time than if the network were congested. A variance of zero is equivalent to the deterministic case. Increasing variances means the future long-term OD demands are less certain.

3.4.2 Correlation Matrices

An infinite number of correlation matrices could be created for the O-D demands. A finite number of correlation matrices were chosen in an attempt to get a representative sample. For each network, the following types of correlations were tested:

Perfectly Correlated – This is represented by a correlation matrix of all ones. If one O-D demand is known, all O-D demands are known.

No Correlations – This is represented by a correlation matrix with ones on the diagonal and zeroes elsewhere. In this case, the O-D demands are random and independent.

Identically Negatively Correlated – This is represented by a correlation matrix with ones on the diagonal and identical negative numbers elsewhere. As the network size increases, the magnitude of the negative correlations decreases. It was found that the maximum negative correlation possible, such that all O-D demands have the same correlation, is equal to $1/(n-1)$ for an $n \times n$ matrix.

To see trends, other correlation matrices were tested with each network, such as smaller positive correlations, smaller negative correlations, and matrices with both positive and negative correlations.

CHAPTER FOUR – NUMERICAL ANALYSIS

4.1 SMALL EXAMPLE

To motivate the problem, a small network is used with discrete random variables that can be solved by hand. The example network, shown in Figure 4.1a, has four nodes, five links, and two origin nodes and one destination node. The arc costs are specified where x_{ij} indicates the flow on link (i,j) . The O-D demands and associated probabilities are shown in Table 4.1a.

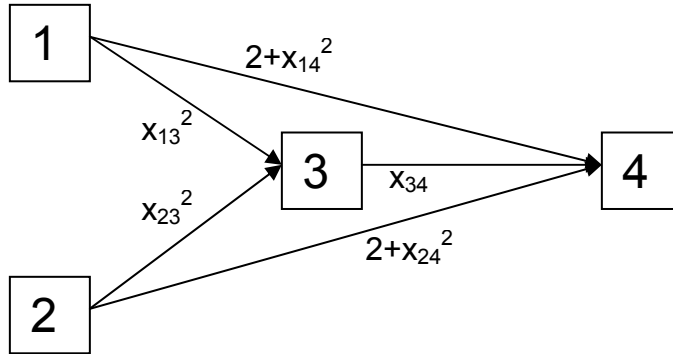


Figure 4.1a Example Network

Table 4.1a Example Network: O-D Demands

d_{14}	5	11
d_{24}	5	11
prob(scenario)	$\frac{1}{2}$	$\frac{1}{2}$

No assumptions are made yet about demand dependence or independence.

4.1.1 Deterministic Demands

Ignoring the randomness of O-D demands, the deterministic case equilibrates traffic assuming O-D demand is at its expected value, $E[d_{14}] = E[d_{24}] = (1/2)*5 + (1/2)*11 = 8$. The resulting link flows are shown in Figure 4.1.1a.

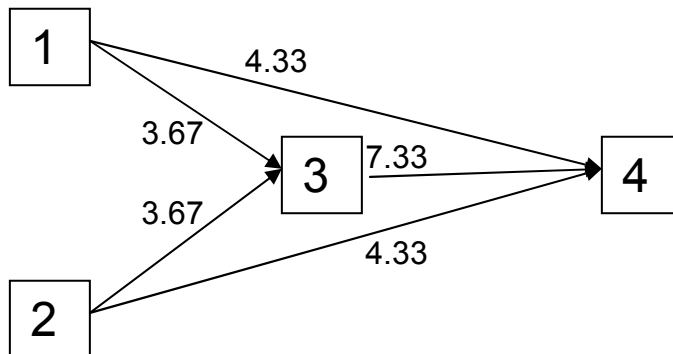


Figure 4.1.1a Example Network: Arc Flows in Deterministic Case

Summing the costs of flow on each arc, multiplied by its flow, gives the total system travel time, 332.4.

4.1.2 Independent Demands

Neglecting any correlation between d_{14} and d_{24} , and assuming they are independent and identically distributed, there are four possible realizations:

Table 4.1.2a Example Network: O-D demands assuming independence

	Scenario			
	1	2	3	4
d_{14}	5	11	5	11
d_{24}	5	11	11	5
prob(scenario)	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Figures 4.1.2a-4.1.2d show the resulting flows. Table 4.1.2a shows the resulting total system travel times.

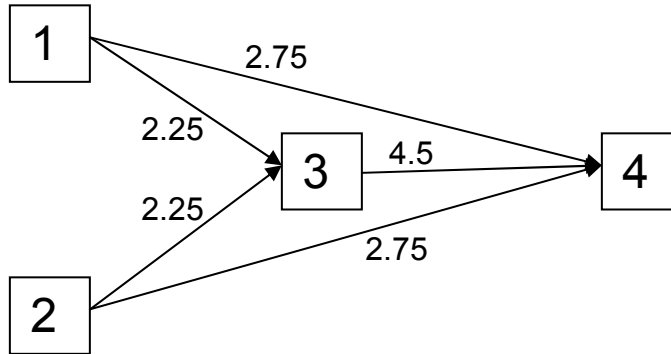


Figure 4.1.2a Example Network: Arc Flows in Scenario 1

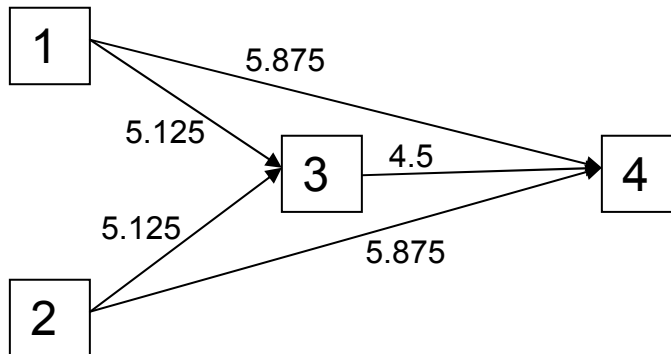


Figure 4.1.2b Example Network: Arc Flows in Scenario 2

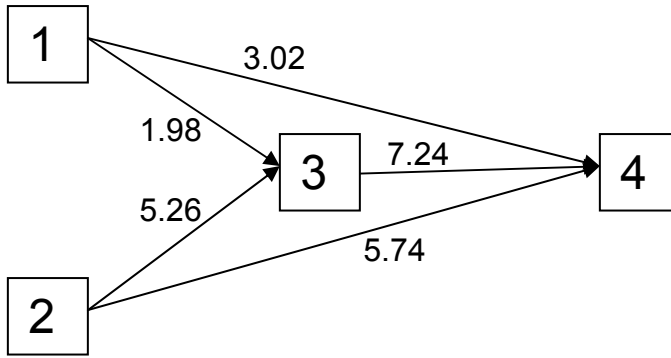


Figure 4.1.2c Example Network: Arc Flows in Scenario 3

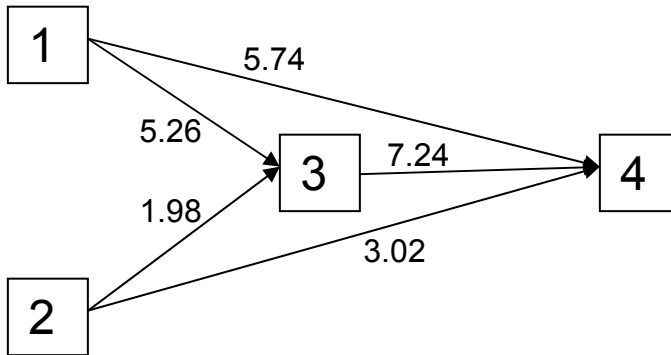


Figure 4.1.2d Example Network: Arc Flows in Scenario 4

The total system travel time for each scenario is listed in Table 4.1.2b. Expected total system travel time, $E[TSTT]$, was calculated by taking a weighted average of the results, using the scenario probabilities as weights. The variance of total system travel time, $V[TSTT]$, is equal to $E[TSTT^2] - (E[TSTT])^2$. Notice that $E[TSTT]$ is greater than the TSTT from the deterministic case. In this case, ignoring the randomness of O-D demand underestimates $E[TSTT]$.

Table 4.1.2b Example Network: Results Assuming Independence

	Scenario				E[TSTT]	V[TSTT]
	1	2	3	4		
TSTT	95.63	803.34	439.90	439.90	444.69	83508.26
prob(scenario)	1/4	1/4	1/4	1/4		

4.1.3 Positively Correlated Demand

When the demands are considered to be perfectly positively correlated, there are two possible scenarios shown in Table 4.1.3a. These scenarios are equivalent to the first two scenarios considered in section 4.1.2 (Table 4.1.2a).

Table 4.1.3a Example Network: O-D Demands assuming positive correlations

	Scenario	
	1	2
d_{14}	5	11
d_{24}	5	11
prob(scenario)	$\frac{1}{2}$	$\frac{1}{2}$

The resulting flows are shown in Figure 4.1.2a-4.1.2b. $E[TSTT] = 449.48$ and $V[TSTT] = 250432.9$. Notice that both $E[TSTT]$ and $V[TSTT]$ are higher than in the independent case. Here, if demand is modeled as random but uncorrelated, the true expected congestion and system reliability would both be underestimated.

4.1.4 Negatively Correlated Demands

When the demands are considered to be perfectly negatively correlated, there are two possible scenarios shown in Table 4.1.4a.

Table 4.1.4a Example Network: O-D demands assuming negative correlations

	Scenario	
	1	2
d_{14}	5	11
d_{24}	11	5
prob(scenario)	$\frac{1}{2}$	$\frac{1}{2}$

The resulting flows are shown in Figure 4.1.2c-4.1.2d. $E[TSTT] = 439.9$ and $V[TSTT] = 0$. Note that perfect negative correlations cannot occur if more than two items are being compared. Zero variance will be achieved only if each scenario yields the exact same total system travel time. This will happen only if the problem is deterministic, or demands are perfectly negatively correlated such that the total network demand is the same for each scenario. Perfect negative correlations do not imply zero variance. If the costs in the example network are changed to be asymmetric, variance would likely be positive. Here, if demand is modeled as deterministic, the true expected congestion would be underestimated.

This small example shows that results can differ substantially if demand is treated as deterministic, stochastic and independent, or stochastic and correlated; thereby motivating further study of the problem. The following section describes the larger networks and inputs used to examine the impact of demand stochasticity and correlations through numerical analysis.

4.2 NETWORKS AND INPUTS

4.2.1 Test Network

The Test network, shown in Figure 4.2.1a, has four nodes, five links, and three O-D pairs. Table 4.2.1a shows the demand for each O-D pair. Table 4.2.1b shows each of the demand and variance scaling factors. Seven different levels of congestion and six different variance levels were examined.

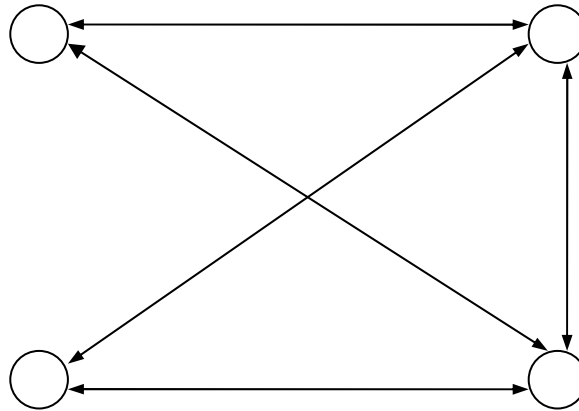


Figure 4.2.1a Test Network

Table 4.2.1a Test Network: Expected Demand

Origin/Dest. Nodes	2	4
1	800	200
2	0	100

Table 4.2.1b Test Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k

k	1	2	3	4	5	6	7
DSF^k	0.25	0.5	0.75	1	1.25	1.5	2
VSF^k	0.01	0.05	0.1	0.5	1	2	-

Nine correlation matrices (shown in Figure 4.2.1b), labeled C1 through C9, were tested. C1 represents perfect correlations. C2, C3, and C4 have positive correlations. C5 is the independent case. C6 and C7 are mixes of positive and negative correlations. C8 and C9 are all negative correlations.

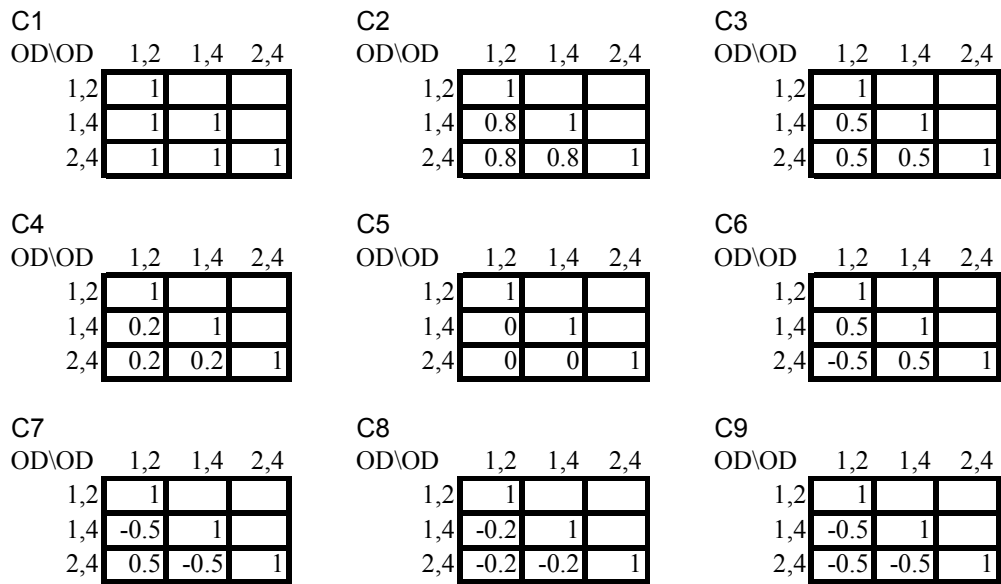


Figure 4.2.1b Test Network: Correlation Matrices

Figure 4.2.1c shows the congestion level in the Test network associated with each demand scaling factor, given VSF^1 and C2. The congestion levels given other variance matrices and scaling factors are similar. Congestion was measured as the ratio of total system travel time with actual costs to total system travel time with free flow costs (i.e. link costs do not increase as vehicles are added). For example, a congestion level of two indicates that TSTT is twice what it would be under free flow conditions. As the inverse of the demand scaling factor increases, congestion increases exponentially. DSF^1 - DSF^3 are very near free flow conditions. DSF^4 - DSF^6 are moderately congested relative to the others. DSF^7 is the most severely congested.

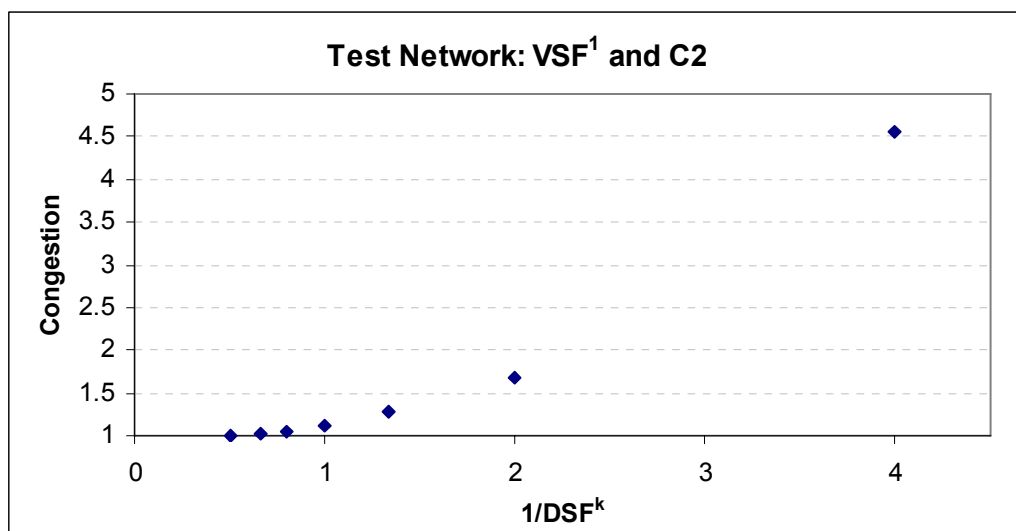


Figure 4.2.1c Test Network: Congestion Levels

4.2.2 Nguyen-Dupuis Network

The Nguyen-Dupuis (N-D) network has 13 nodes, 19 links, and four O-D pairs. Table 4.2.2a shows the demand for each O-D pair. Table 4.2.2a shows each of the demand and variance scaling factors. Six different levels of congestion and variance levels were examined.

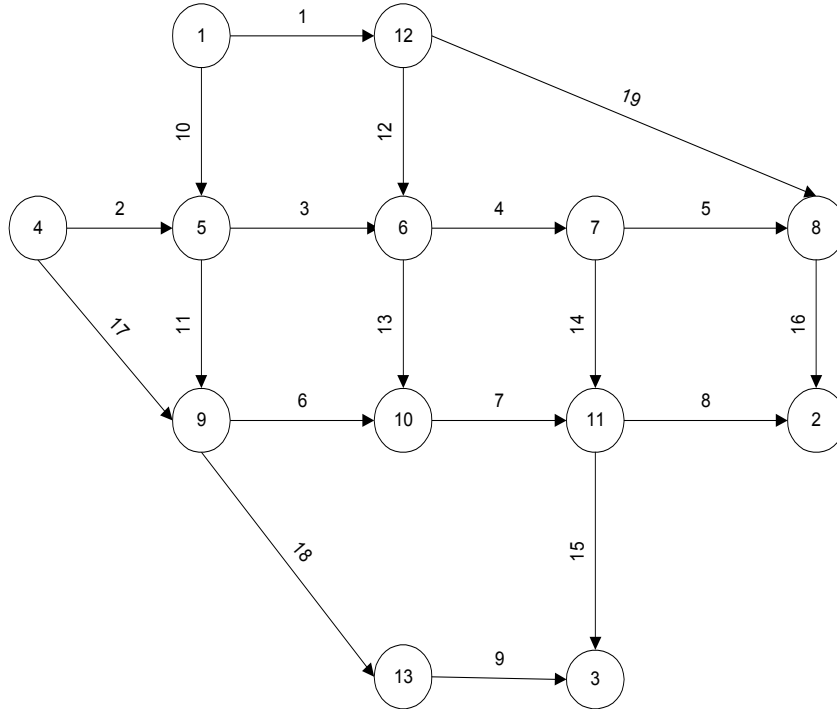


Figure 4.2.2a Nguyen-Dupuis Network

Table 4.2.2a N-D Network: Expected Demand

Origin/Dest. Nodes	1	4	2	3
1	0	0	400	800
4	0	0	600	200
2	0	0	0	0
3	0	0	0	0

Table 4.2.2b N-D Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k

k	1	2	3	4	5	6
DSF^k	0.5	0.75	1	1.25	1.5	2
VSF^k	0.01	0.05	0.1	0.5	1	2

The nine correlation matrices used, C1 through C9, are shown in Table 4.2.2b. C1 represents perfect correlations. C2, C3, and C4 have positive correlations. C5 is the

independent case. C6 and C7 are mixes of positive and negative correlations. C8 and C9 are all negative correlations.

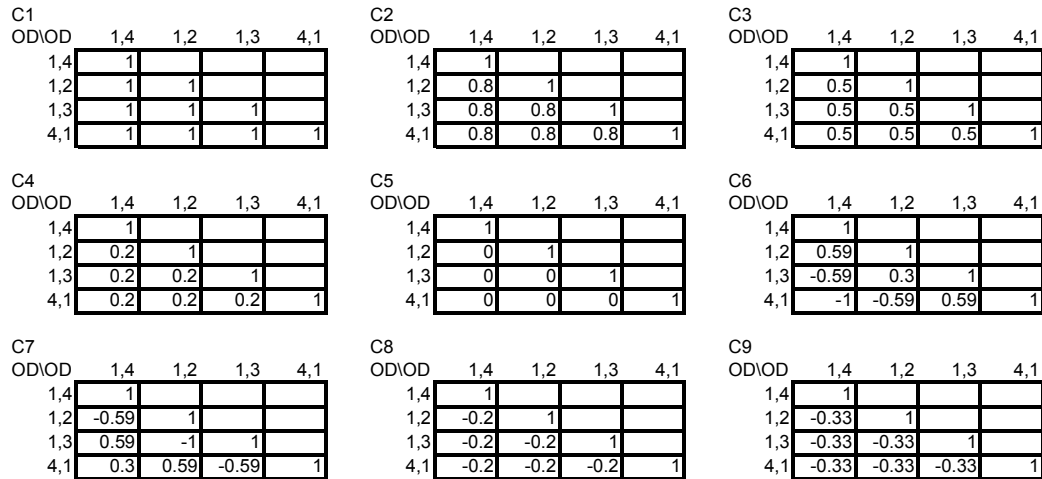


Figure 4.2.2b N-D Network: Correlation Matrices

Figure 4.2.2c shows the congestion level in the test network associated with each demand scaling factor, given VSF^1 and C1. The congestion levels given other variance matrices and scaling factors are similar. As the inverse of the demand scaling factor increases, congestion increases. DSF^1 - DSF^8 are all congested with DSF^8 representing very extreme congestion.

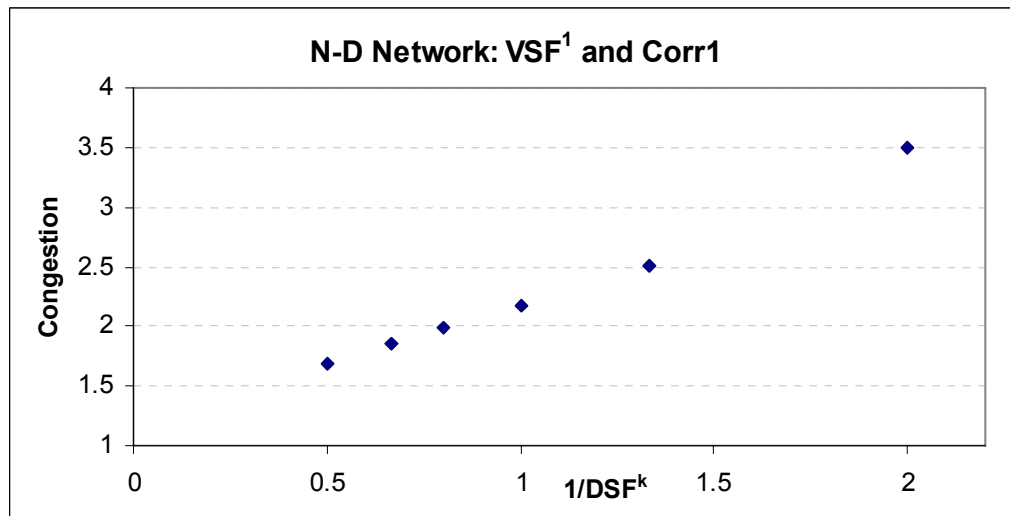


Figure 4.2.2c N-D Network: Congestion Levels

4.2.3 Sioux Falls Network

The Sioux Falls (S.F.) network, shown in Figure 4.2.3a, has 24 nodes, 76 links, and 12 O-D pairs. Table 4.2.3a shows the demand for each O-D pair. Table 4.2.3b shows each of the demand and variance scaling factors. Six different levels of congestion and five variance levels were examined.

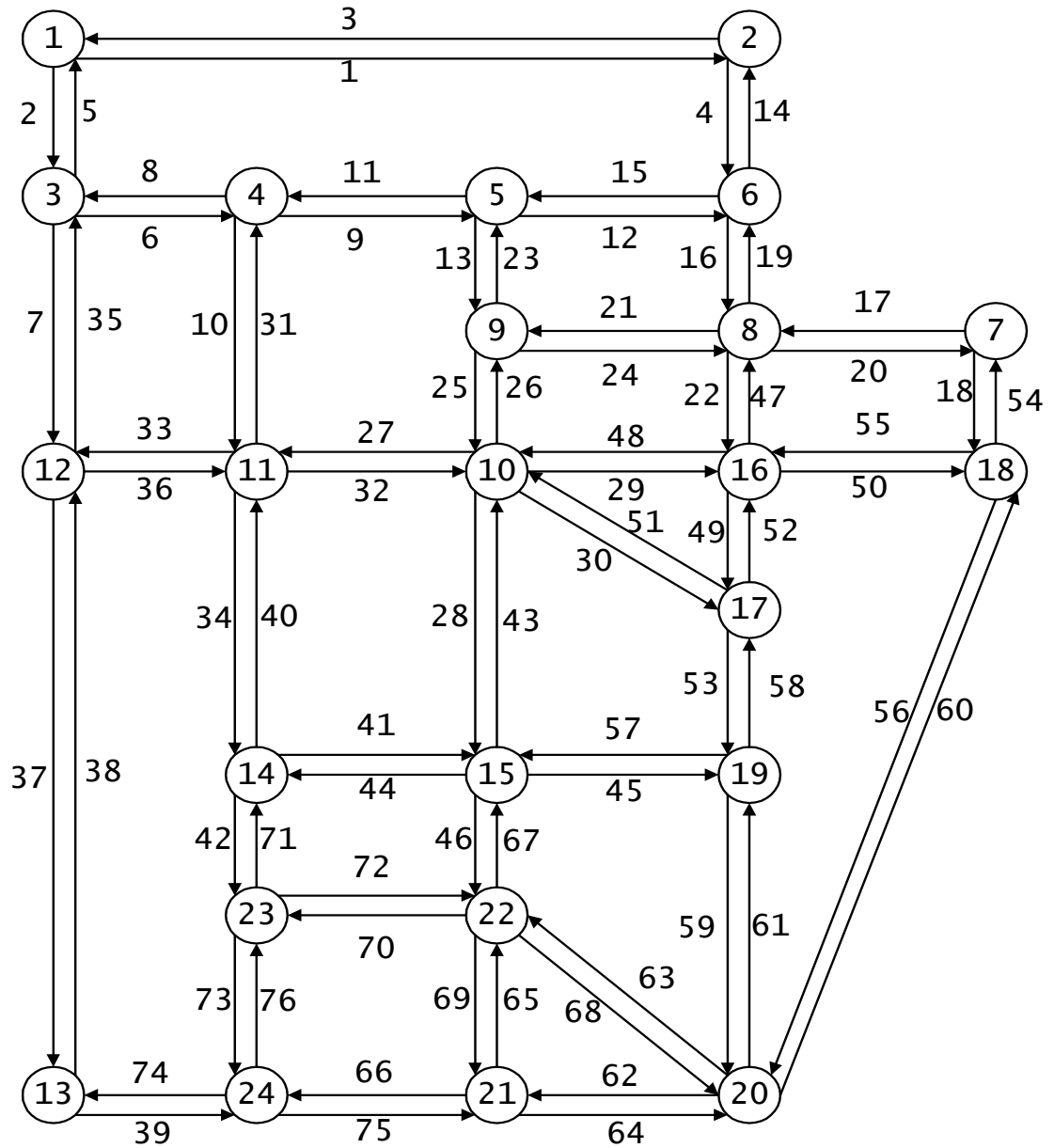


Figure 4.2.3a S.F. Network

Table 4.2.3a S.F. Network: Expected Demand

Origin/Dest. Nodes	20	18	7	6
1	0	800	900	600
2	900	0	1200	500
3	1230	500	0	1860
13	1540	600	400	0

Table 4.2.3b S.F. Network: Demand Scaling Factors, DSF^k , and Variance Scaling Factors, VSF^k

k	1	2	3	4	5	6
DSF^k	0.5	0.75	1	1.25	1.5	2
VSF^k	0.05	0.1	0.5	1	2	-

The five correlation matrices used, C1 through C5, are shown in Figure 4.2.3b. C1 represents perfect correlations. C2 is the independent case. C3 and C4 are mixes of positive and negative correlations. C5 is all negative correlations.

Corr1 OD\OD	1,18	1,7	1,6	2,20	2,7	2,6	3,20	3,18	3,6	13,20	13,18	13,7
1,18	1											
1,7	1	1										
1,6	1	1	1									
2,20	1	1	1	1								
2,7	1	1	1	1	1							
2,6	1	1	1	1	1	1						
3,20	1	1	1	1	1	1	1					
3,18	1	1	1	1	1	1	1	1				
3,6	1	1	1	1	1	1	1	1	1			
13,20	1	1	1	1	1	1	1	1	1	1		
13,18	1	1	1	1	1	1	1	1	1	1	1	
13,7	1	1	1	1	1	1	1	1	1	1	1	1

Corr2 OD\OD	1,18	1,7	1,6	2,20	2,7	2,6	3,20	3,18	3,6	13,20	13,18	13,7
1,18	1											
1,7	0	1										
1,6	0	0	1									
2,20	0	0	0	1								
2,7	0	0	0	0	1							
2,6	0	0	0	0	0	1						
3,20	0	0	0	0	0	0	1					
3,18	0	0	0	0	0	0	0	1				
3,6	0	0	0	0	0	0	0	0	1			
13,20	0	0	0	0	0	0	0	0	0	1		
13,18	0	0	0	0	0	0	0	0	0	0	1	
13,7	0	0	0	0	0	0	0	0	0	0	0	1

Corr3 OD\OD	1,18	1,7	1,6	2,20	2,7	2,6	3,20	3,18	3,6	13,20	13,18	13,7
1,18	1											
1,7	-0.09	1										
1,6	-0.09	-0.09	1									
2,20	-0.09	-0.09	-0.09	1								
2,7	-0.09	-0.09	-0.09	-0.09	1							
2,6	-0.09	-0.09	-0.09	-0.09	-0.09	1						
3,20	0.09	0.09	0.09	0.09	0.09	0.09	1					
3,18	0.09	0.09	0.09	0.09	0.09	0.09	-0.09	1				
3,6	0.09	0.09	0.09	0.09	0.09	0.09	-0.09	-0.09	1			
13,20	0.09	0.09	0.09	0.09	0.09	0.09	-0.09	-0.09	-0.09	1		
13,18	0.09	0.09	0.09	0.09	0.09	0.09	-0.09	-0.09	-0.09	-0.09	1	
13,7	0.09	0.09	0.09	0.09	0.09	0.09	-0.09	-0.09	-0.09	-0.09	-0.09	1

Corr4 OD\OD	1,18	1,7	1,6	2,20	2,7	2,6	3,20	3,18	3,6	13,20	13,18	13,7
1,18	1											
1,7	0.52	1										
1,6	0.47	0.52	1									
2,20	0.47	0.52	0.46	1								
2,7	0.47	0.52	0.46	0.47	1							
2,6	0.49	0.51	0.49	0.49	0.49	1						
3,20	-0.48	-0.51	-0.47	-0.48	-0.48	-0.49	1					
3,18	-0.55	-0.46	-0.56	-0.55	-0.55	-0.52	0.54	1				
3,6	-0.53	-0.48	-0.54	-0.53	-0.53	-0.51	0.53	0.44	1			
13,20	-0.45	-0.54	-0.43	-0.44	-0.44	-0.48	0.46	0.59	0.56	1		
13,18	-0.54	-0.47	-0.56	-0.55	-0.55	-0.52	0.54	0.41	0.45	0.60	1	
13,7	-0.55	-0.46	-0.56	-0.55	-0.55	-0.52	0.54	0.41	0.44	0.60	0.41	1

Corr5 OD\OD	1,18	1,7	1,6	2,20	2,7	2,6	3,20	3,18	3,6	13,20	13,18	13,7
1,18	1											
1,7	-0.09	1										
1,6	-0.09	-0.10	1									
2,20	-0.09	-0.09	-0.09	1								
2,7	-0.09	-0.10	-0.09	-0.09	1							
2,6	-0.09	-0.09	-0.09	-0.09	-0.09	1						
3,20	-0.09	-0.09	-0.09	-0.09	-0.09	-0.10	1					
3,18	-0.10	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	1				
3,6	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	1			
13,20	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	1		
13,18	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.10	-0.09	-0.09	1	
13,7	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09	1

Figure 4.2.3b S.F. Network: Correlation Matrices

Figure 4.2.3c shows the congestion level in the S.F. network associated with each demand scaling factor, given VSF^2 and C1. The congestion levels given other variance matrices and scaling factors are similar. As the inverse of the demand scaling factor

increases, congestion increases. DSF^1 - DSF^3 are fairly uncongested, with DSF^1 exhibiting near free-flow travel times. DSF^4 and DSF^5 are more moderately congested. DSF^6 leads to the highest levels of congestion tested for this network, with travel times almost eighty percent above free-flow. The congestion levels tested for the S.F. network are in a similar range as those tested in the Test network with the first six demand scaling factors.

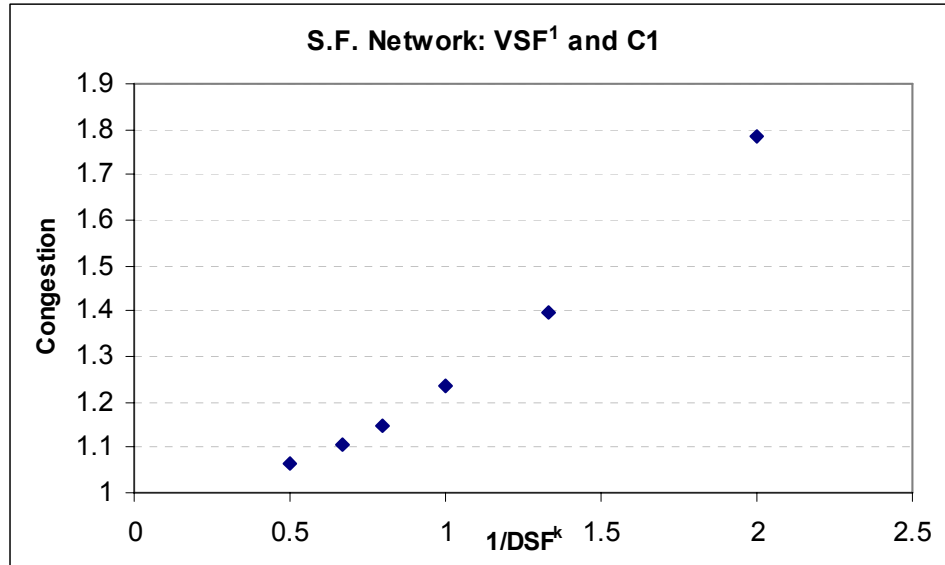


Figure 4.2.3c S.F. Network: Congestion Levels

4.3 Results

As described in section 3.4, the following three parameters are used in evaluating the results of numerical analysis.

- 1.) Percent difference in mean Total System Travel Time (TSTT) using correlated vs. uncorrelated O-D demands
- 2.) Percent difference in mean TSTT between the deterministic case vs. correlated O-D demands
- 3.) Percent difference in the standard deviation of TSTT using correlated vs. uncorrelated O-D demands

4.3.1 Test Network

Figure 4.3.1a shows how TSTT varies with demand as congestion increased using the correlation matrix, C1. Although it cannot be proven that TSTT v. Demand is a convex function, the results show that the curve is non-decreasing.

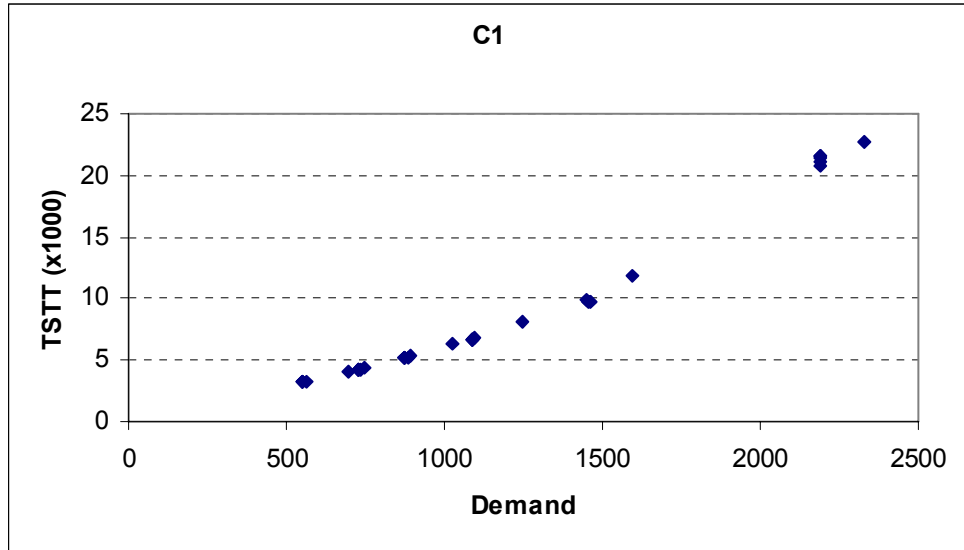


Figure 4.3.1a Test Network: TSTT v. Total System Demand, (given C1)

Figure 4.3.1b shows, for each congestion level (given VSF^1), the difference in TSTT between using the expected value of demand (deterministic case) and assuming O-D demand is variable and potentially correlated. The results show that, for every correlation matrix tested, if O-D demand is actually variable, but is assumed to be deterministic, the resulting TSTT will be inaccurate. Extreme positive correlations show the greatest percent difference. In almost all cases, the percent difference is positive, meaning that using the expected value of demand gave an underestimate of TSTT. This does not hold true in all cases because there is not a one-to-one matching between the user equilibrium objective function (value minimized) and the system optimal objective function (value reported). In general, the smaller the magnitudes of the correlations, the smaller the percent difference in results from the deterministic case.

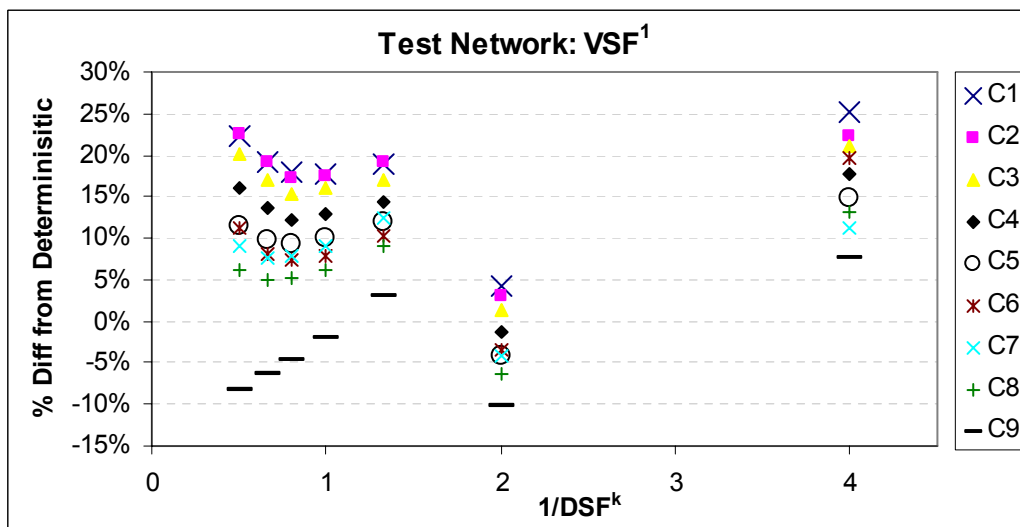


Figure 4.3.1b Test Network (VSF^1): % Diff from Deterministic v. Congestion Level

Figure 4.3.1c shows the percent difference between mean TSTT when O-D demand correlations are considered and mean TSTT when O-D demand is assumed to be independent. The most extreme differences occur when the magnitude of O-D correlations are large, positive or negative. C6 and C7, both mixes of positive and negative correlations, had resulting mean TSTTs very close to the mean of the independent case. If correlations are positive, mean TSTT is greater than the mean TSTT in the independent case. If correlations are negative, the opposite result occurs. This suggests that if the network being examined has future O-D demands with strongly positive or strongly negative correlations, demand independence is not a valid assumption. The curve in the data may be due to the congestion level. When the network is experiencing near free flow conditions, it may be very sensitive to changes in demand. As the network becomes more congested, additional demand may have less of an impact on TSTT.

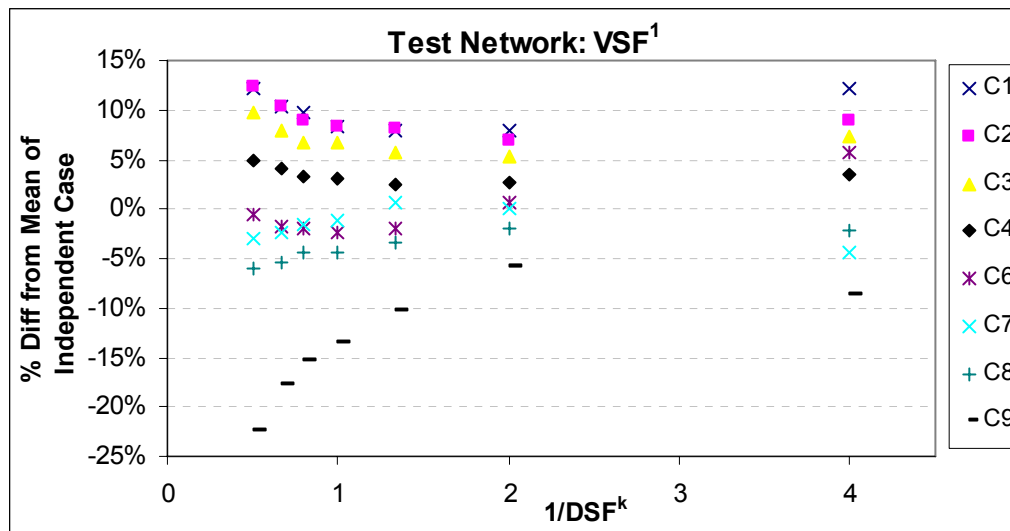


Figure 4.3.1c Test Network (VSF¹): % Diff from Mean of Independent Case v. Congestion Level

Figure 4.3.1d shows the percent difference in standard deviation between the cases where O-D demands are considered to be independent, and when correlations are considered. This measure is critical when planning for a robust transportation network. The more positive the correlations, the more assuming demand independence will underestimate the standard deviation of TSTT. This is intuitive since positive correlations imply that if one O-D demand is higher than expected, other O-D demands are likely to be higher than expected; leading to more extreme (higher variance) TSTT results. Conversely, the more negative the correlations, the more assuming demand independence will overestimate the standard deviation of TSTT. This result is intuitive since negative correlations imply that the O-D demands have a balancing effect on one another; if the demand for one O-D pair is higher than expected, demands for other O-D pairs are more likely to be lower than expected. This “balancing” decreases the standard deviation of TSTT.

As congestion increases through DSF^6 , the percent difference in standard deviation decreases. When the network is uncongested the marginal impact of each additional vehicle on the network is relatively large. As the network congestion increases, it may be less sensitive to the addition of vehicles.

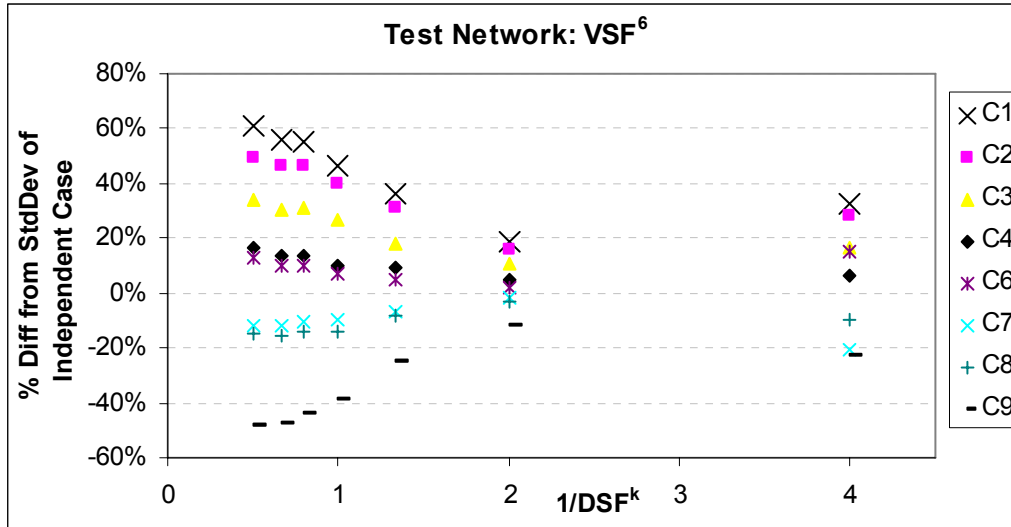


Figure 4.3.1d Test Network (VSF⁶): % Diff from StdDev of Independent Case v. Congestion Level

4.3.2 Nguyen-Dupuis Network

Figure 4.3.2a shows how TSTT varies with demand as congestion increases, given correlation matrix, C3. The curve formed by the data appears piecewise convex. The other correlation matrices give similar results.

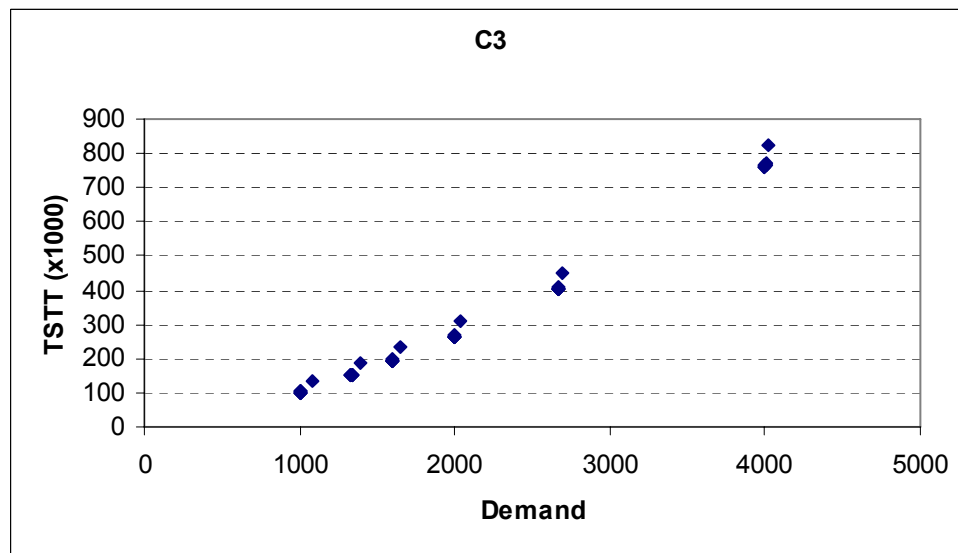


Figure 4.3.2a Test Network: TSTT vs. Total System Demand, (given C3)

Positively correlated demand leads to mean TSTT values that differ more from the deterministic result, than the mean TSTT values found using mixed or negatively correlated demands. The perfectly correlated case caused the greatest percent difference between its TSTT and TSTT of the deterministic case. As the degree of positive correlation decreased, so did the percent difference from deterministic results. Mean TSTT results for negatively correlated demands are the most similar to the deterministic TSTT. As congestion increases, the percent difference from deterministic decreases for all correlation matrices tested. This may be due to the decreasing marginal impact on the system per vehicle as the network becomes more congested. Figure 4.3.2b shows the results given VSF^1 . Similar results were achieved given other variance scaling factors.

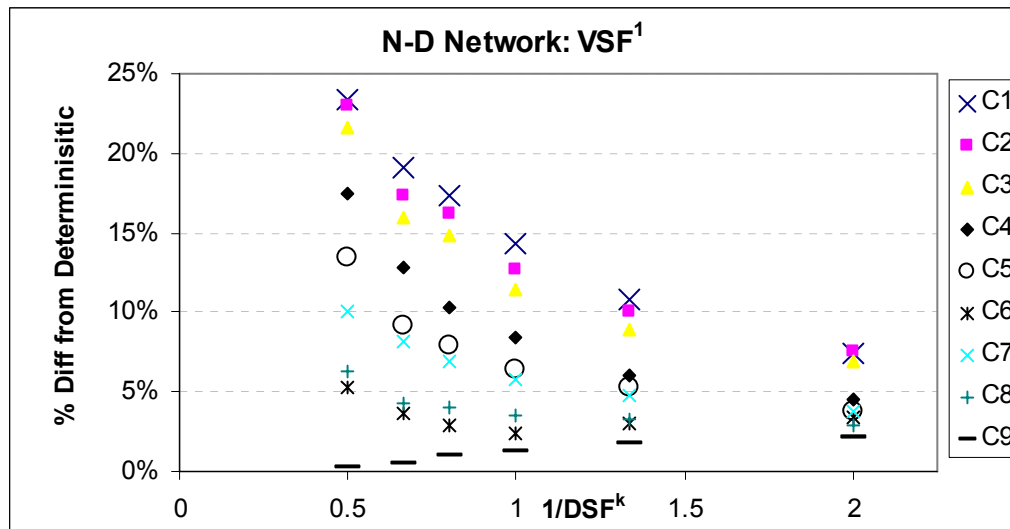


Figure 4.3.2b N-D Network (VSF^1): % Diff from Deterministic v. Congestion Level

As congestion increases, the percent difference in mean TSTT between the correlated demand scenario and the independent scenario decreases. The magnitude of this difference increases as the correlations become more negatively or more positively correlated. The smallest percent difference was seen for mixed correlation matrices. Figure 4.3.2c shows the resulting data given VSF^1 . Although the magnitude of the percent difference decreases as the demand variances decrease, the results appear the same.

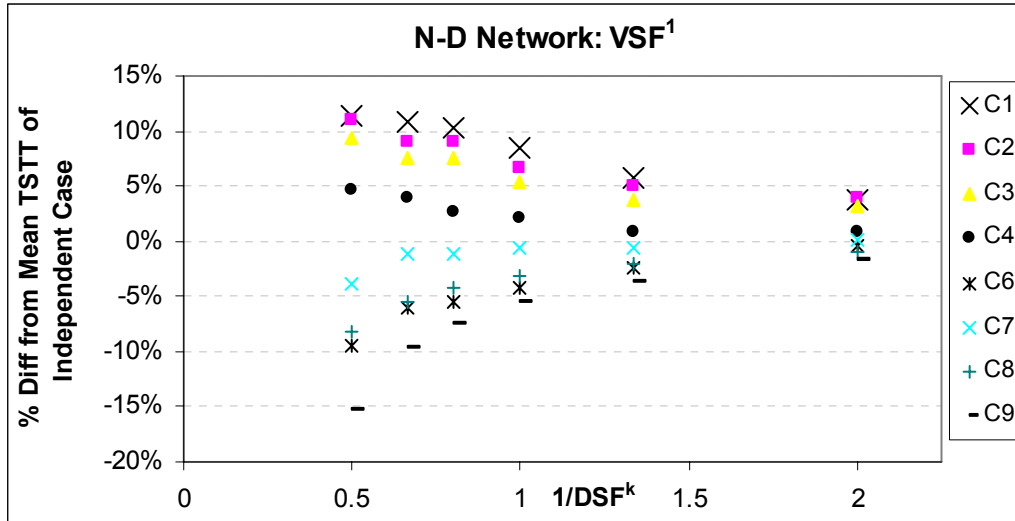


Figure 4.3.2c N-D Network (VSF¹): % Diff from Mean of Independent Case v. Congestion Level

Figure 4.3.2d shows the percent difference in standard deviation between the cases where O-D demands are considered to be independent, and when demand correlations are considered. As with the Test network, the more positive the correlations, the more likely it is that assuming demand independence will underestimate the standard deviation of TSTT. This is intuitive since positive correlations imply that if one O-D demand is higher than expected, other O-D demands are likely to be higher than expected; leading to more extreme (higher variance) TSTT results.

Unlike the Test network, the percent difference in standard deviation is relatively stable for all congestion levels. This is likely the case because the N-D network is congested for all levels of DSF, whereas the Test network is congested only for larger values of DSF.

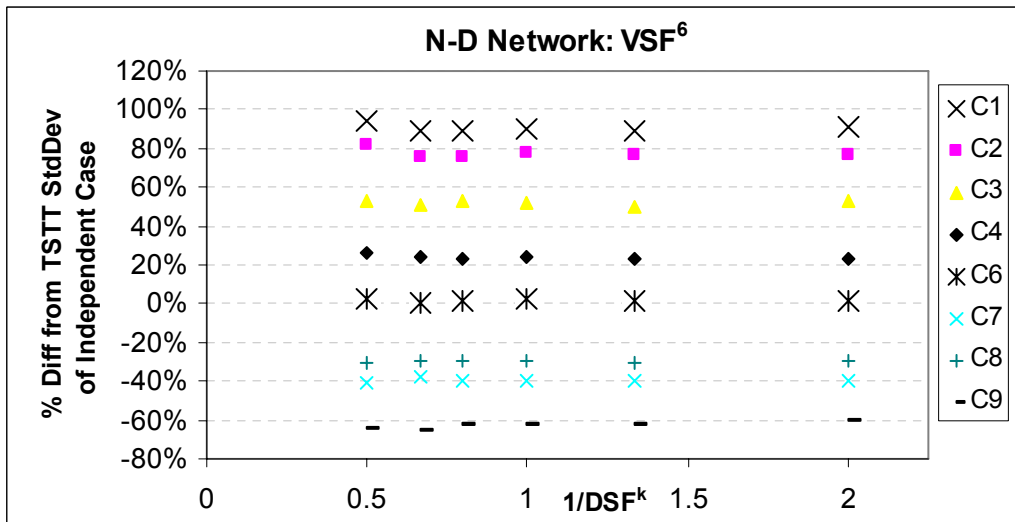


Figure 4.3.2d N-D Network (VSF⁶): % Diff from StdDev of Independent Case v. Congestion Level

4.3.3 Sioux Falls Network

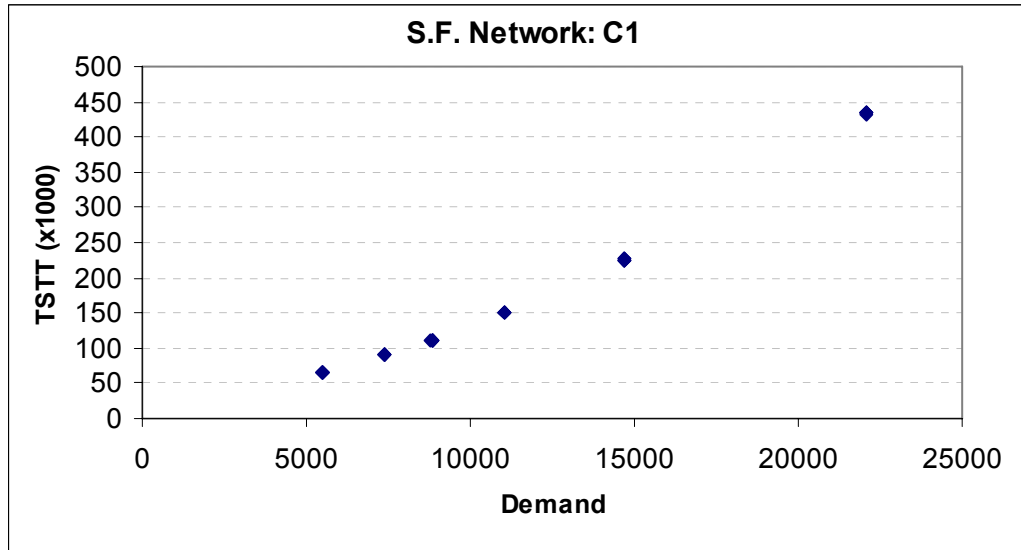


Figure 4.3.2a S.F. Network: TSTT vs. Total System Demand, (given C1)

Figure 4.3.3b shows, for each congestion level (given VSF^2), the difference in TSTT between using the expected value of demand (deterministic case) and assuming O-D demand is variable and potentially correlated. The results show that, for every correlation matrix tested, if O-D demand is actually variable, but is assumed to be deterministic, the resulting TSTT will be inaccurate. These inaccuracies, however, are much less than the inaccuracies seen in the previous two networks. This may be due to the larger size of the S.F. network. Further testing is needed to determine the exact cause of this result, but perhaps as the network size grows, the impact of correlations on $E[TSTT]$ decreases. As in previous tests, larger magnitudes of correlations lead to greater percent differences from the deterministic TSTT result.

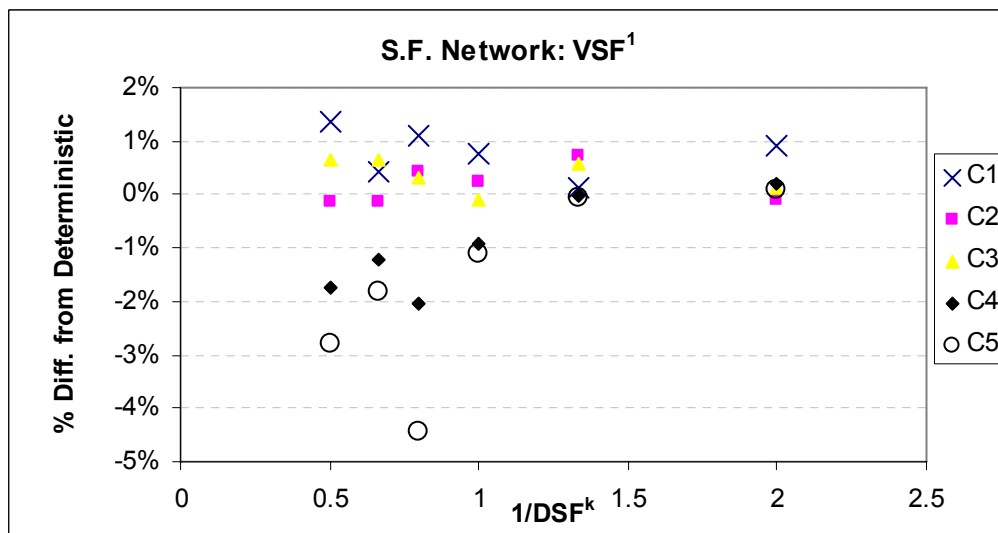


Figure 4.3.3b S.F. Network (VSF^1): % Diff from Deterministic v. Congestion Level

Figure 4.3.1c shows the percent difference between mean TSTT when O-D correlations are considered and mean TSTT when O-D demand is assumed to be independent. The most extreme differences occur when the magnitude of O-D correlations are large, positive or negative; however the differences are much less than those seen for the previous two networks examined. This result coincides with the hyporesearch in the previous paragraph: the impact of correlations on mean TSTT tends to decrease as network size grows. As congestion increases, the percent difference in TSTT from the mean of the independent case becomes negligible, especially for the mixed and negative correlation matrices. This decreasing trend was seen for the other networks as well and it is hypothesized that as the network becomes more congested, additional demand may have less of an impact on TSTT.

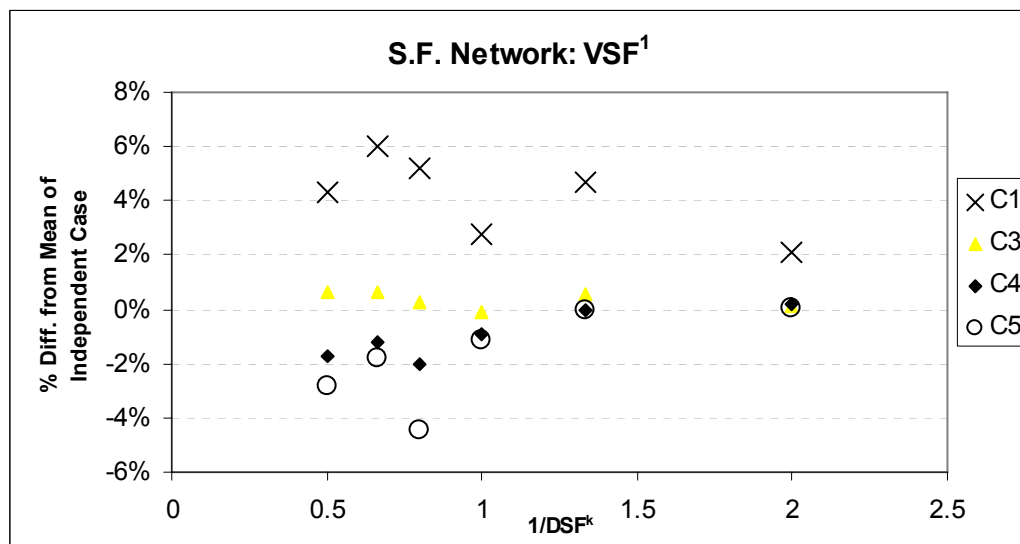


Figure 4.3.3c S.F. Network (VSF²): % Diff from Mean of Independent Case v. Congestion Level

Figure 4.3.3d shows the percent difference in standard deviation between the cases where O-D demands are considered to be independent, and when demand correlations are considered. As with the Test and N-D networks, the more positive the correlations, the more assuming demand independence will underestimate the standard deviation of TSTT; and the more negative the correlations, the more assuming demand independence will overestimate the standard deviation of TSTT.

For all levels of congestion, the results are relatively stable, except for fluctuations in the results for the perfectly correlated case. Comparing the results for the Test and S.F. networks, for similar levels of congestion, they behave differently. The S.F. network results are more stable, and the Test network results decrease as congestion increases. For both networks, however, the magnitudes of the percent differences for extreme correlations are very great and present strong evidence for the need to consider correlations when attempting to determine a system's robustness.

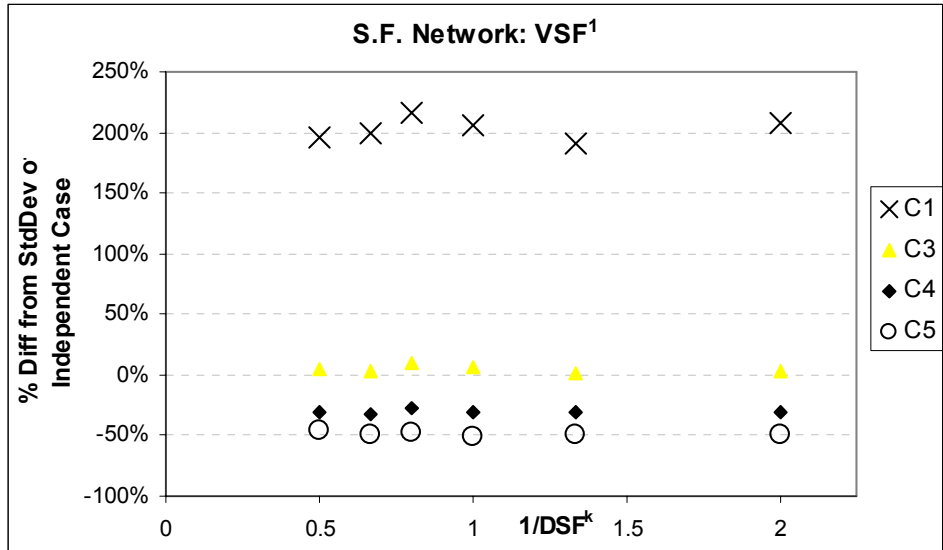


Figure 4.3.3d S.F. Network (VSF²): % Diff from StdDev of Independent Case v. Congestion Level

4.3.4 Conclusions

The results from all three of the examined networks show that correlations can be very influential in determining system performance. The most prominent result is that if correlations between O-D demands are very positive or very negative, and are not considered, then measures of system robustness will likely be very inaccurate. Mean TSTT may also be inaccurate if correlations are ignored, especially if congestion is relatively low. Treating demand as uncertain is shown to be very important in achieving accurate results, but independence is rarely a good assumption.

CHAPTER FIVE - CONCLUSIONS

5.1 IMPLICATIONS OF WORK

This research is the first to explore the effects of considering long-term O-D demand as correlated in traffic assignment. It motivates research into determining if network design decisions will change based on forecasts of correlations, and into achieving these correlations. In particular, it shows how network robustness relies heavily on the type of correlations that exist between O-D pairs. Although no exact prescription is given for when correlations should be considered, general trends may help guide this decision. These trends include:

- 1.) As correlations become more positive, the error in TSTT caused by using a deterministic model may increase.
- 2.) As the magnitude of correlations increase, the error in mean TSTT caused by using a stochastic and independent model may increase.
- 3.) As the magnitude of correlations increase, the error in variance of TSTT caused by using a stochastic and independent model may increase.

The first trend suggests that current models using a deterministic value of demand may be inadequate if O-D demands have positive correlations. TSTT, a common measure of network performance, may in some cases be grossly underestimated. Future research is needed to test if this trend becomes less significant as network size grows.

The second trend suggests that just considering demand uncertainty may not be sufficient to achieve accurate TSTT results if the correlations between O-D demands are neglected. The more the correlations differ from zero, the worse the predictions of mean TSTT. Future research is needed to test if this trend becomes less significant as network size grows.

The third trend is especially important. As the importance of network robustness is more widely recognized and applied to practice, the consideration of correlations becomes vital. If large correlations exist between future O-D demand predictions and they are not considered, the potential errors are enormous.

The following section describes the work that remains to be done on this subject.

5.2 FUTURE WORK

Since exploring the effects of uncertain and correlated O-D demand is a new area, there are many opportunities for future research. The most obvious opportunity is developing a method to determine these correlations. Some ideas are discussed in Section 5.2.1. Research is also needed to develop an analytical formulation for including correlated O-D demand within the traffic assignment models and network design decision process. Although this research does show the significant impact of correlations on network performance, the impact on network design decisions remains to be examined. More important than measures of network performance is project rankings. It remains to be seen if considering O-D demand correlations causes policy decisions to be made differently. Examining the impact of correlations on the network design problem will answer this question.

This research attempts to examine a representative sample of congestion levels, variance matrices, and correlation matrices, however testing more scenarios can only help

strengthen the understanding of the problem. Investigating larger networks would also be useful, but will require the implementation of a more advanced sampling procedure. As seen with the S.F. network, larger networks may be less sensitive to whether or not O-D demands are considered to be correlated, especially when determining the mean TSTT.

5.2.1 Achieving Correlations

The focus of this research so far has been on determining the impact of O-D demand correlations on network performance, but the issue of how to achieve these correlations has not yet been addressed. The method for achieving actual long-term O-D demand correlations depends on the demand forecasting tool employed. The two most popular forecasting tools, the four-step process and activity-based, will be addressed briefly here.

5.2.1.1 Four Step Process

The four step process for transportation planning is shown in Figure 5.2.1.1a. In the first step, “Trip Generation,” the number of trips from each origin and the number of trips to each destination are predicted. The second step, “Trip Distribution,” assigns the results of the first step into O-D demands. The third step, “Mode Split,” determines which modes will be used for each trip. The fourth step, “Traffic Assignment,” assigns each trip to a specific route within the network.

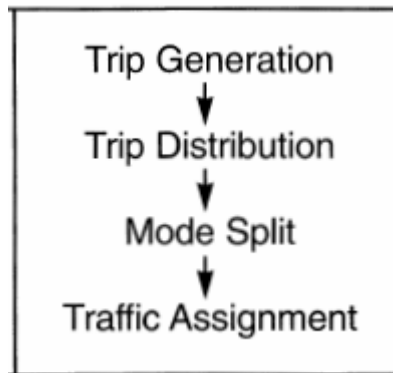


Figure 5.2.1.1a Four-Step Process

Most of the methodology presented in this research is concerned with the fourth step; however, the correlations will be determined during the first two steps. Trip Generation outputs trips produced by each origin, i , and trips attracted to each destination, j . Trip Distribution uses a “gravity model” to predict the number of trips between each origin and destination.

$$T_{ij} = \frac{P_i(A_j F_{ij})}{\sum A_j F_{ij}}$$

where,

T_{ij} is the number of trips from origin i to destination j ;

P_i is the number of trips produced by origin i ;

A_j is the number of trips attracted to destination j ; and

F_{ij} represents the effect of travel time and distance between origin i and destination j .

To determine a distribution of O-D demands, the following methodology is proposed.

1. Create a list of possible future scenarios with differing growth rates, household size, employment rate, etc...
2. Assign each scenario a probability of occurrence
3. Run Trip Generation for each scenario to obtain productions and attractions
4. Run Trip Distribution using the gravity model for each scenario
5. Fit a probability distribution to each T_{ij} and solve for the correlations using the following equation

$$Corr(T_{ij}, T_{i'j'}) = \frac{E[T_{ij}T_{i'j'}] - E[T_{ij}]E[T_{i'j'}]}{\sqrt{Var(T_{ij}) * Var(T_{i'j'})}}$$

where $E[-]$ represents an expected value and $Var(-)$ represents a variance; the subscripts i and i' are origins, and j and j' are destinations.

5.2.1.2 Activity-Based

Achieving correlations for O-D demands for the activity based approach could be done similarly to the method for the four step process (section 5.2.1.1) except demand for O-D tours would be correlated instead of simply the demand between each O-D pair. Again, the input parameters could be varied to come up with feasible future scenarios. Each scenario would lead to a different set of future O-D tours. Pooling the results together, an expected value and variance of each O-D tour could be found, as well as correlations amongst the tours. This approach may be very difficult since there are a near infinite amount of possible O-D tours. However, limiting the number of stops along tours could be helpful. No method currently exists for accurately predicting traffic assignment given the O-D tours. Future research is essential to fill this gap.

REFERENCES

- Beckmann, M., C.B. McGuire, and C.B. Winsten. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven, Connecticut.
- Bell, M.G.H and C. Cassir. (2000). *Reliability of Transport Networks*. Research Studies Press Limited.
- Bureau of Public Roads. (1964). *Traffic Assignment Manual*, Urban Planning Division, U.S. Department of Commerce, Washington, D.C.
- Chen, A., H. Yang, H. Lo and W.H.Tang. (1999). A Capacity Related Reliability for Transportation Networks. *Journal of Advanced Transportation* **33(2)**, 183-200.
- Chen, A., H. Yang, H.K. Lo, and W. Tang. (2002). Capacity Reliability of a Road Network: An Assessment Methodology and Numerical Results, *Transportation Research, Part B: Methodological*, **36B**, 225-252.
- Daganzo, C.F. 1994. The Cell Transmission Model: a Simple Dynamic Representation of Highway Traffic. *Transportation Research* **28B**, 269-287.
- Daganzo, C.F. 1995. The Cell Transmission Model, Part II: Network Traffic. *Transportation Research* **29B**, 79-93.
- Dial R.B. (1971). A Probabilistic Multipath Traffic Assignment Model which Obviates Path Enumeration. *Transportation Research*, **5**, 83-111.
- Du, Z.P. and A. Nicholson. (1997). Degradable Transportation Systems: Sensitivity and Reliability Analysis. *Transportation Research*, **31B**, 225-237.
- Escudero, L.F., Kamesam, P.V., King, A., Wets, R.J.-B. (1993). Production Planning Via Scenario Modeling. *Annals of Operations Research*, **43**, 311-335.
- Frey, H.C. (1992). *Quantitative Analysis of Uncertainty and Variability in Environmental Policy Making*. AAAS/EPA Environmental Science and Engineering Fellows Program, American Association for the Advancement of Science, Washington, DC.
- Harvey, G. and E. Deakin. (1995). *Description of the STEP Analysis Package*. Berkeley, California.
- Iida, Y. and H. Wakabayashi. (1989). An Approximate Method of Terminal Reliability of a Road Network Using Partial Minimal Path and Cut Set. *Proceedings of the 5th WCTR, Yokohama*, 367-380.

Jensen, J.L. (1906). Sur les Fonctions Convexes et les Inégalités Entre les Valeurs Moyennes. *Acta. Math*, **30**, 175-193.

Keeney, R.L. and H. Raiffa. (1976). Decisions with Multiple Objectives : Preferences and Value tradeoffs. John Wiley, New York.

Liu, L. and X. Yuan. (2000). Coordinated Replenishments in Inventory Systems with Correlated Demands. *European Journal of Operational Research*. **123**, 490-503.

Lo, H.K. and Y.K. Tung. (2003). Network with Degradable Links: Capacity Analysis and Design. *Transportation Research B*, **37**, 345-363.

Maher, M.J. (1992). SAM - A Stochastic Assignment Model. In: *Mathematics in Transport Planning and Control* (ed. J.D. Griffiths), Oxford University Press.

Maher, M.J. and P.C. Hughes (1997a). A Probit-based Stochastic User Equilibrium Assignment Model. *Transportation Research*, **31B**(4), 341-355.

Markowitz, H. M. (1952). Portfolio Selection. *Journal of Finance*. **7**(1).

Mulvey, J.M., R.J. Vanderbei, and S.A. Zenios. (1995). Robust Optimization of Large-Scale Systems. *Operations Research*. **43**(2), 264-281.

Paraskevopoulos, D., E. Karakitsos, and B. Rustem. (1991). Robust Capacity Planning Under Uncertainty. *Management Science*, **37**, 787-800.

Rebonato, R. and J. Peter. (1999). The Most General Methodology to Create a Valid Correlation Matrix for Risk Management and Option Pricing Purposes. *The Journal of Risk*, **2** (2), 1-12.

Sengupta, J.K. (1991). Robust Solutions in Stochastic Linear Programming. *JORSA*, **42**(10), 857-870.

Sheffi, Y. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods, Prentice-Hall, Englewood Cliffs, N.J, 1985.

Sheffi, Y. and W. Powell. (1982). An Algorithm for the Equilibrium Assignment Problem with Random Link Times. *Networks*, **12**, 191-207.

Stopher, P.R. and A. H. Meyburg *Urban Transportation Modeling and Planning*. D.C. Health and Company, 1975.

Waller, S.T. and A.K. Ziliaskopoulos. (2001). A Dynamic and Stochastic Approach to Network Design. *Journal of the Transportation Research Board No. 1771*, 106-114.

Waller, S.T., J.L. Schofer, and A.K. Ziliaskopoulos. (2001). Evaluation with Traffic Assignment Under Demand Uncertainty. 1. *Journal of the Transportation Research Board*, **1771**, 69-75.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. In *Proceedings of the Institution of Civil Engineers*, **1**, 325-378.

Zhao, Y. and K. Kockelman. (2002). The Propagation of Uncertainty Through Travel Demand Models: An Exploratory Analysis. *Annals of Regional Science*, **36(1)**.