An Integrated Decision-Making Framework for Transportation Architectures: Application to Aviation Systems Design

A Dissertation
Presented to
The Academic Faculty

by

Jung-Ho Lewe

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

School of Aerospace Engineering
Georgia Institute of Technology
April 2005

An Integrated Decision-Making Framework for Transportation Architectures: Application to Aviation Systems Design

Approved by:

Daniel P. Schrage, Ph.D.
Professor, Aerospace Engineering, Committee Chair

Dimitri N. Mavris, Ph.D.
Professor, Aerospace Engineering, Co-advisor

Amy R. Pritchett, Sci.D.
Associate Professor, Aerospace Engineering

Daniel A. DeLaurentis, Ph.D.
Assistant Professor, Aeronautics and Astronautics, Purdue University

Alan Wilhite, Ph.D.
Langley Professor, National Institute of Aerospace

Mark D. Moore
Personal Air Vehicle Sector Manager, NASA Vehicle Systems Program

Date Approved: April 18, 2005

To my family

"There's nothing new under the sun, but there are lots of old things we don't know."

ACKNOWLEDGEMENTS

I would like to thank my committee for their guidance and support in the course of my research. It was their vision, intelligence, expertise, dedication, academic rigor and integrity that saw me through. Thank you again Drs. Schrage, Mavris, Pritchett, DeLaurentis, Wilhite and Mr. Moore. I truly consider this dissertation a culmination of your efforts with only my two cents.

There are many smart and friendly students and colleagues whose contributions I also appreciate; each member of the SATS competition team, the PAVE project team and the MIDAS squad. I have enjoyed the friendship and collaboration which I wish to continue, and I thank you all.

I also want to hug my family wholeheartedly. My precious son Eric who still believes every daddy goes to school and earns a living by studying, my beloved wife Katie—what can I say about you? You're my all in all—and my loving parents across the sea who have always prayed for me, I love you for ever, ever and ever.

Jungho Lewe

Atlanta, GA, Spring 2005

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS OR ABBREVIATIONS

$A \circ B$	Hadamard product of matrices A and B.
α	Choice model parameter.
B	Mobility budget.
$\square(a,b)$	Uniform distribution.
C_{ij}	Generalized impedance matrix for O-D matrix calculation.
D	Great circle, one-way trip distance.
D_R	Route distance.
D_T	Total trajectory distance.
e	Arc elasticity.
ε	elasticity.
f_G	Gumberl distribution function.
F_G	Cumulative Gumbel distribution function.
f_n	Normal distribution function.
F_N	Cumulative normal distribution function.
γ	Gravity model parameter.
I_k	Inclusive value of nest N_k.
PI	Productivity Index.
$\pi(m)$	choice probability of alternative m.
px	Price index.
$R(m)$	On-time reliability of mode m.
S	Circuitry factor.
σ	standard deviation.
τ_p	Perceived delay.
τ_r	Real delay.
t_{ij}	Number of trips from locations i to j.

$\triangle(a,b,c)$ Triangular distribution.

U_m Utility of alternative m.

ε_m Random utility of alternative m.

V_c Cruise speed.

V_m Deterministic utility of alternative m.

ABM Agent Based Model(ing).

ABM/S Agent Based Modeling and Simulation.

AIA Aerospace Industries Association.

AIAA American Institute of Aeronautics and Astronautics.

ALN Commercial air transport.

AP:HUB Large- or medium-hub airport.

AP:SML Small- or non-hub airport.

ATS Americal Travel Survey.

BTS Bureau of Transportation Statistics.

CAR Personal automobile.

CONUS Continental United States.

DOT Department of Transportation.

FAA Federal Aviation Administration.

GA General Aviation.

GAJ General aviation aircraft with jet engine.

GAMA General Aviation Manufacturers Association.

GAP General aviation aircraft with piston engine.

GDP Gross Domestic Product.

GUI Graphic User Interface.

IFR Instrument Flight Rules.

IIA Independence of Irrelevant Alternatives.

LAR Large-metro area.

MADM Multi-Attribute Decision Making.

MAS Multi-Agent System.

MCDM Multi-Criteria Decision Making.

MED Medium-metro area.

Mi Simulation code ***Mi*** .

MIDAS Multi-platform Integrated Development Aid System.

MNL Multinomial Logit.

M/S Market Share.

M&S Modeling and Simulation.

NAS National Airspace System.

NASA National Aeronautics and Space Administration.

NLM Nested Logit Model.

NOM Non-metro area.

NPIAS National Plan of Integrated Airport Systems.

NTS National Transportation System.

O-D Origin-Destination.

PAV Personl Air Vehicle.

PAVE Personl Air Vehicle Exploration.

R&D Research and Development.

RE Requirements Engineering.

SATS Small Aircraft Transportation System.

SML Small-metro area.

SoS System-of-Systems.

TAF Transportation Architecture Field.

TLM Tournament Logit Model.

TRB Transportation Research Board.

UC Urban Cluster.

UZA Urbanized Area.

VFR Visual Flight Rules.

VMT Vehicle-Mile Traveled.

VV&A Verification, Validation and Accreditation.

WSM Weighted Sum Model.

SUMMARY

The National Transportation System (NTS) is undoubtedly a complex system-of-systems—a collection of diverse 'things' that evolve over time, organized at multiple levels, to achieve a range of possibly conflicting objectives, and never quite behaving as planned. The purpose of this research is to develop a virtual transportation architecture for the ultimate goal of formulating an integrated decision-making framework. The foundational endeavor begins with creating an abstraction of the NTS with the belief that a holistic frame of reference is required to properly study such a multi-disciplinary, trans-domain system. The culmination of the effort produces the Transportation Architecture Field (TAF) as a mental model of the NTS, in which the relationships between four basic entity groups are identified and articulated. This entity-centric abstraction framework underpins the construction of a virtual NTS couched in the form of an agent-based model. The transportation consumers and the service providers are identified as adaptive agents that apply a set of preprogrammed behavioral rules to achieve their respective goals. The transportation infrastructure and multitude of exogenous entities—disruptors and drivers—in the whole system can also be represented without resorting to an extremely complicated structure. The outcome is a flexible, scalable, computational model that allows for examination of numerous scenarios which involve the cascade of interrelated effects of aviation technology, infrastructure, and socioeconomic changes throughout the entire system.

CHAPTER I

INTRODUCTION

Contents

"Heavier-than-air flying machines are impossible."

– Lord Kelvin, President of the Royal Society, 1895.

"Mark my word: A combination airplane and motor car is coming. You may smile. But it will come."

– Henry Ford, Ford Motor Company Founder, 1940.

"Prediction is always easy; either keeping it accurate or making it possible is the only difficult part."

– An anonymous aerospace engineer, 2003.

1.1 *Motivation*

1.1.1 Need for a New System

The National Transportation System (NTS)—composed of transportation vehicles, their supporting infrastructure and the people who use, operate and build the system itself—is one of the largest and most complicated systems of modern civilization. The NTS has continuously expanded its capability through technology revolutions. This historic progress is conceptually portrayed in Figure 1, where the nation's increasing mobility—the capability to move passengers and cargo—is depicted as a superposition of the various mode of transportation.

Before the Industrial Revolution, people depended on bio-mechanic systems and nature for transportation. In 1796, the locomotive engine was invented as a precursor of revolutionary change. For the first time in human history, a small crew could carry almost a limitless amount of passengers and cargo crossing the continent on the railroad. From the 1910s onward, it has been mass production of automobiles that has allowed the general public to travel freely. Cars are inexpensive, easy-to-use, on-demand vehicles completely suitable for the diverse needs of diverse individuals. The last evolution since the 1950s has

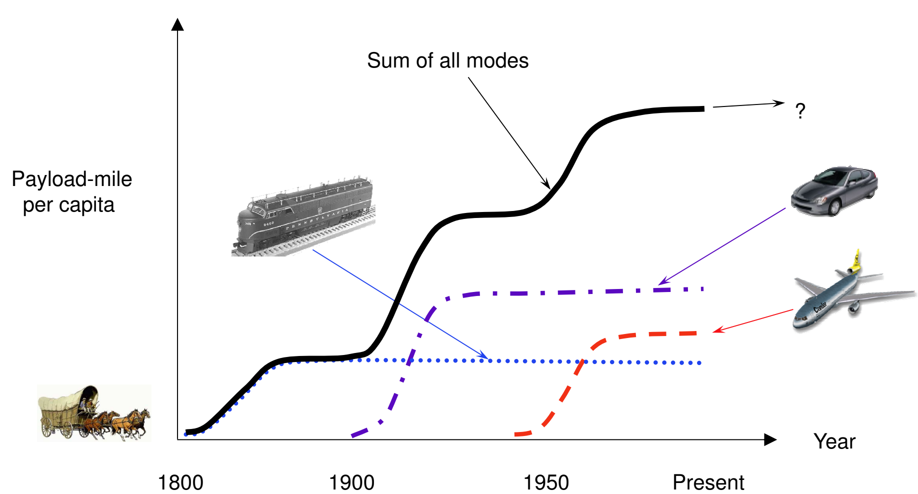


Figure 1: Mobility Progress in the Transportation System—the ordinate indicates the mobility capability notionally obtained from the aggregated product of weight and distance for all transportation activities per capita.

occurred through the expansion of commercial air transportation. The nation's travelers have been able to enjoy long-range, safe travel with considerably improved mobility, at an affordable cost after airline deregulation. This eventually reshaped the pattern of life for many individuals. A question that may immediately arise is: When will the next revolutionary leap occur, and what form will it take? It is not the easiest question to answer. Nevertheless, this can be said at least—there is a problem in the current transportation system.

At present, the nation's air travelers are predominantly served by large airlines using the traditional hub-and-spoke system. While this system is an efficient method in many ways and has increased its capability in recent years, air travelers are increasingly dissatisfied with the current air transportation system as it gradually becomes plagued by delays, long waits, and built-in inefficiencies both in the air and on the terminal areas. (See Figure 2.) It is projected that even with the implementation of planned enhancements, aircraft delays are anticipated to grow exponentially from less than 4% of flights in 2000 to more than 38% in 2015. (AIA 2001) Even though the September 11, 2001 incident mitigated the situation, gridlock is still approaching. For someone who belongs to the middle class income level or below, alternatives to the commercial transport are limited to personal cars, trains or buses, but these options have a downside. Additional use of cars yields more road congestion and air pollution, which is already a serious concern. While trains or buses are very convenient in some cities, they are basically for scheduled trips and are not an efficient method to cover door-to-door trips. Other alternatives that can be considered are high-end

	1995	2000
Operations (M)	62.0	68.7
Enplanements (M)	598.0	706.0
Delays (K)	237.0	450.0

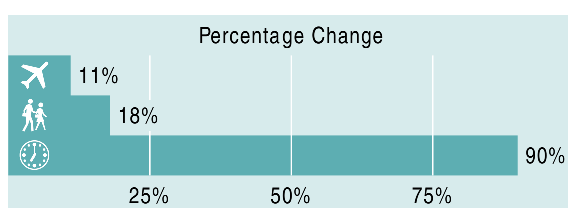


Figure 2: Percentage Changes in Operations, Enplanements and Delays [Source: FAA (2001, p. 2)]

transportation methods such as rotorcraft, business jets, etc. However, these options exist only for licensed pilots or the wealthy; the majority of the public simply cannot afford to use these high-speed, on-demand modes.

Potential solutions to this could lie in many areas such as intelligent ground transportation systems or high-speed rail transportation. (BTS 2000, Ch. 6) In the present research, however, the scope of possible alternatives will remain within the aerospace field. Increases in airspace system capacity would obviously help, and much energy is being expended in trying to upgrade the airports and enroute systems. For example, NASA's Aviation Systems Capacity (ASC) program addresses this directly. NASA is carrying out other initiatives as well: namely, the Small Aircraft Transportation System (SATS) project and the Personal Air Vehicle Exploration (PAVE) project.[1]

NASA and FAA found that smaller airports outnumber large-hub airports almost 200:1 but sit nearly idle while general aviation manufacturers struggle to maintain even a fraction of their former production levels. (See Figures 3 and 4.) The SATS vision is one where personal-use or shared, low-cost general aviation aircraft exercise new technologies to fly into thousands of under-utilized airports throughout the nation in practically any weather. On the other hand, the PAVE project, as the title implies, aims to develop / identify / evaluate Personal Air Vehicle (PAV) concepts and the associated technologies required for the future. Although SATS and PAVE have different foci, viewpoints and ground rules, they share the same spirit—massive operations of small vehicles operating within a distributed air traffic control architecture. Many in the aerospace community feel that this is the most promising solution to the public's on-demand air travel needs.[2]

To sum up, the present research is motivated by accepting the following assumption: The National Transportation System (NTS) is under pressure to improve, and evolutionary improvements to existing elements in the system alone are not sufficient to achieve the de-

[1]Detailed information about each project is given in Appendix A.2.
[2]Some criticism can be found, and will be discussed later in §4.2.2.2.

sired level of improvement. Therefore, a "new mobility system"—a new product resulting from innovative ideas and revolutionary technologies—has to be developed and introduced. Easy-to-use PAV systems still belong to the uncharted territory of the second centennial of human flight and they represent one of the many conceivable new mobility systems needed to improve the overall efficiency of the NTS. Accordingly, the need to design the most viable PAV system is established, aiming at a bold new era in which the aerospace field—at least the general aviation segment of the industry—will be rejuvenated and will regain the leadership of technology development.

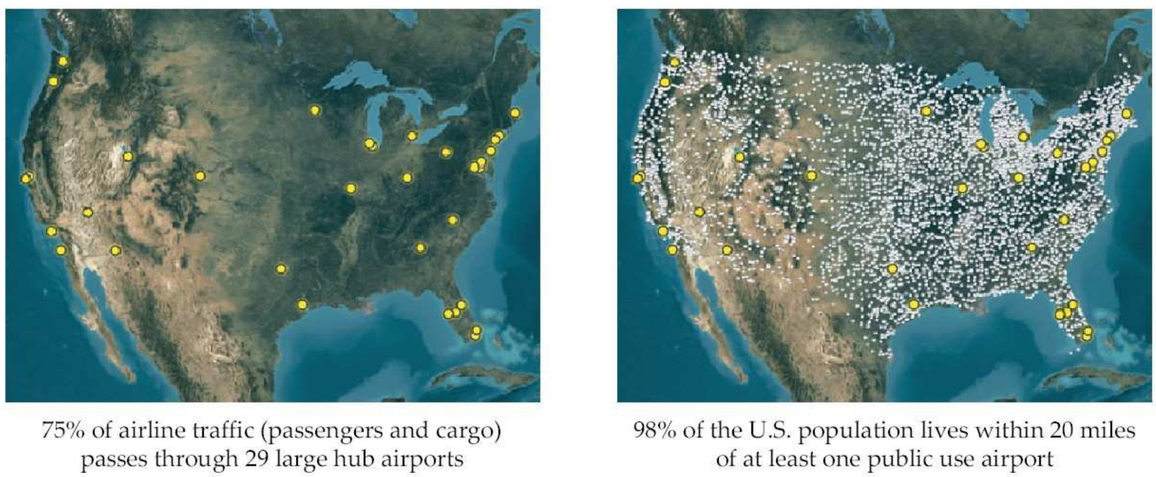


Figure 3: The Status of Airports in the U.S. [Source: NASA LaRC (2001)]

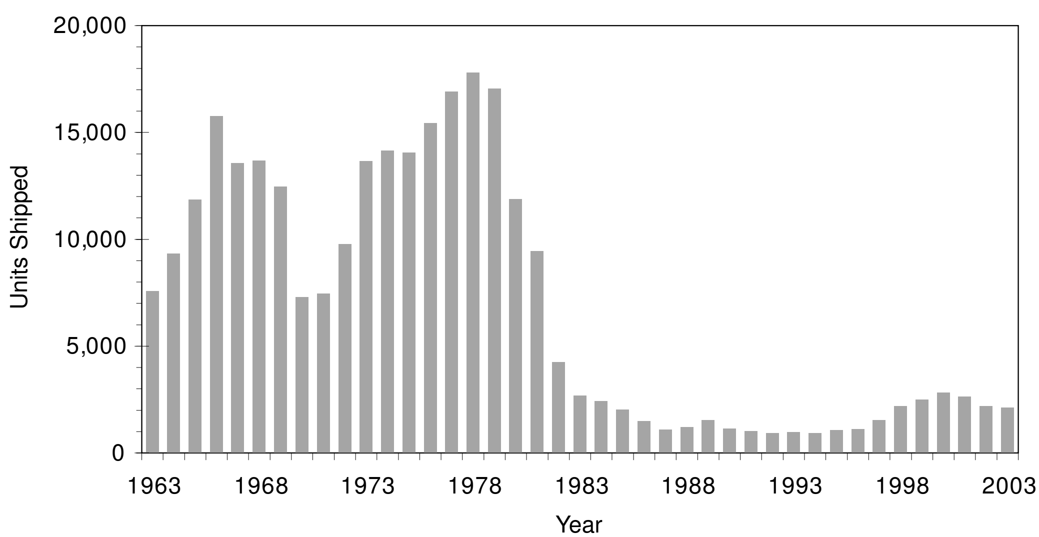


Figure 4: Trend of General Aviation Unit Shipments [Data source: GAMA (2004, p. 5)]

1.1.2 Challenges in PAV Design

A truly personal, on-demand air vehicle has been a subject of interest since the first flight. It has been a dream for aerospace engineers to invent PAVs that can be owned and used by individuals. Thousands of designs have trickled through the public perception, as one can find various PAV concepts scattered in the literature.[3] Some of them have been completely ludicrous; others have not been resolutely practical, as can be seen in Figure 5. Still others have lain hidden waiting for the right time—perhaps not in the immediate future.

Nevertheless, given the amount of years, energy and resources that the aerospace community has invested to obtain a viable PAV, the outcomes have been somewhat disappointing. Although people comfort themselves by saying that dreams eventually come true some

[3] Appendix A.1 offers a brief historical account regarding these efforts.

Figure 5: Various Personal Air Vehicle Concepts

day, none of the alternatives seems viable at this point. What are the reasons that all proposed concepts have failed to achieve their goals?

Obviously, a prime culprit is technology level. One thing learned from history is that technologies have been a driving force of progress: locomotives, cars and airplanes—all were followed by corresponding technology revolutions. The required technologies for a successful PAV system have not reached a readiness level commensurate with various constraints such as performance, cost or environmental compatibility. The era of the PAV system will not be realized until the next technology revolution is ushered into our society, which justifies continued research and development of the prerequisite technology innovations. Development may come in the form of a radical advance by genius engineers and scientists or by looking at the problem from a different perspective. This effort aims to take such an out-of-the-box approach to complement current PAV research and overcome the challenges of the problem.

1.1.2.1 Design and Development Perspective

Regardless of whether an engineering system is a simple commercial product or a sophisticated aerospace vehicle, the early steps of the development process for the system can be posed as in Figure 6. The very first step is, obviously, to establish customer neeeds (Step I). Next, a designer tries to translate the needs and a certain "thing" is formulated to satisfy those needs (Step II). The formulated concept goes through an evaluation process where

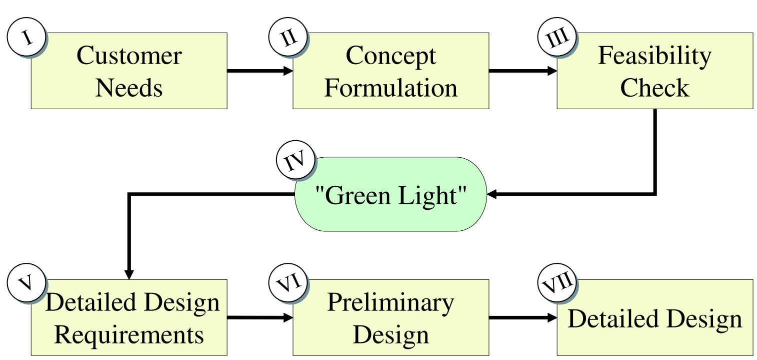
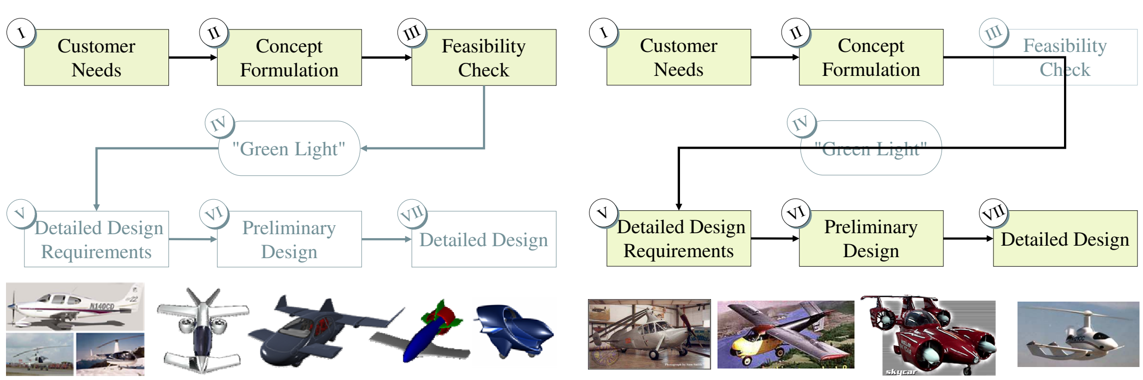


Figure 6: Generic Design and Development Process

feasibility and viability are assessed (Step III). Step IV represents the point at which decision makers decide whether or not to give the "green light". If the decision is "go", detailed design requirements, often called design specifications, are generated and the PAV concept will evolve into a final production design through the work of many engineers from related disciplinary areas (Steps V, VI, and VII).

NASA's progress has reached only as far as Step III as shown in Figure 7(a). The agency has recognized that the PAV design space is huge and ill-defined. This is why NASA's Personal Air Vehicle Exploration (PAVE) program intentionally keeps several concept vehicles as reference baselines and concentrates on identifying synergistic technologies and exploring the design space around each baseline. On the other hand, radical concepts proposed by individual enthusiasts have skipped Steps III and IV and have gone directly to Step V, as illustrated in Figure 7(b). The consequence was not favorable and none of those concepts have found a practical market application. In conclusion, the dream of PAV to date has met a bottleneck at Step III. It is still too early to commit to full-scale PAV development programs. Before a decision maker gives a particular project the "green light" for further development, a means to evaluate heterogeneous concept vehicles must be established. This is not an easy task for many reasons. First, different concepts is vulnerable to subjectivity. For instance, not taking into account other factors, a fast PAV is typically considered superior to slow one. But, could a slow, dual mode PAV be more viable than



(a) NASA's Effort (b) Enthusiasts' Effort

Figure 7: Status of Current PAV Efforts in Industry and in NASA

a fast, single mode PAV? Before attempting to answer such questions, let us look at the problem from a different perspective.

1.1.2.2 System-of-systems Perspective

Few practical PAV systems operate in the current National Transportation System (NTS), but it is still very important to think of a PAV system in the context of its containing systems. The big picture of the NTS is presented in Figure 8 to this end. The NTS on top is divided into a ground transportation system and an air transportation system according to the primary mission space. The air transportation system in turn has multiple, lower-level, constituent systems. Commercial transports and general aviation (including business aircraft) are treated as separate systems for they utilize different vehicles and infrastructure. Similarly, the ground transportation system can be split into several constituent systems as

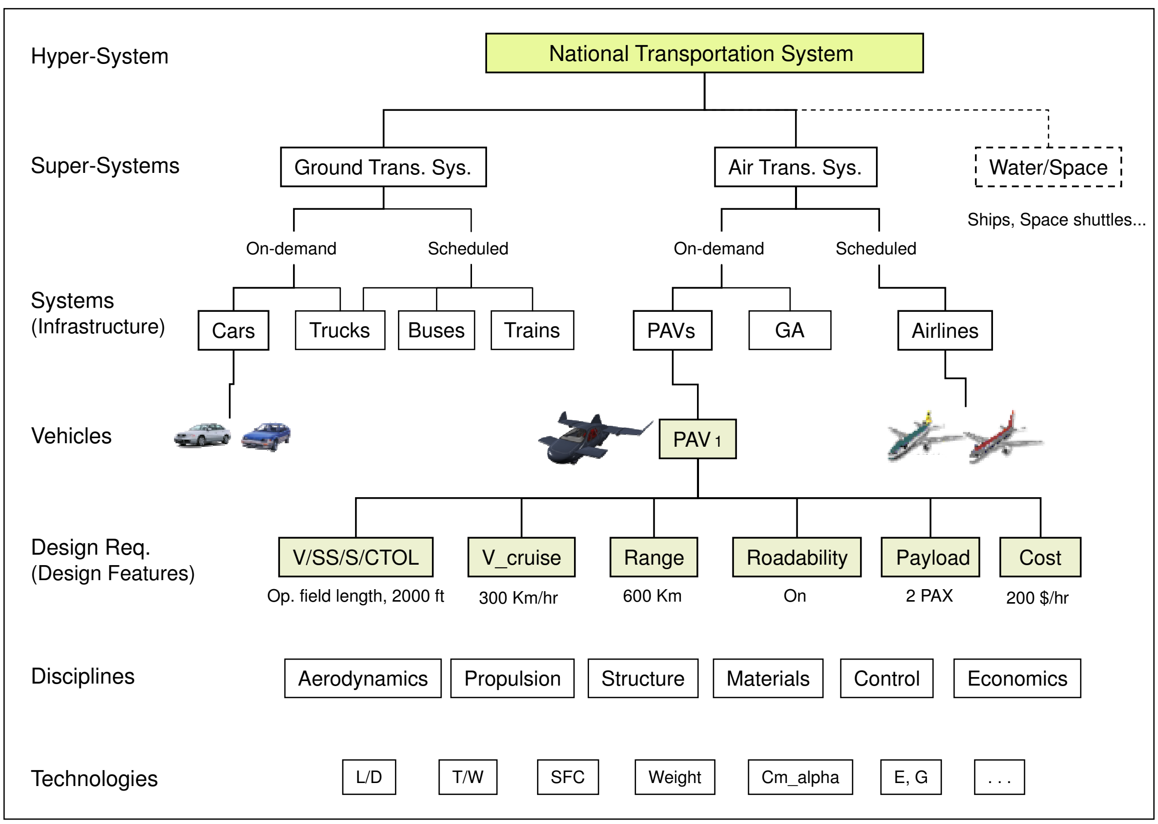


Figure 8: The Hierarchy of the National Transportation System

indicated. In the center of the figure, a hypothetical PAV system is envisioned that has a particular set of design requirements. In fact, any PAV concept can be abstracted through this breakdown. Disciplinary areas should be in place to support specific design requirements through balancing required technologies, as shown at the bottom of Figure 8.

Now imagine how the traveling public would fit within this figure. They interact with the NTS and they are adaptive with respect to any changes in the NTS. If airline ticket prices go down, some travelers who planned to visit a place by car may change their decision. Likewise, when one of the features of the hypothetical PAV concept is altered, a traveler's knowledge about the PAV is updated and the same traveler can then choose a different vehicle. In other words, even a change in PAV design requirements at the very bottom level will propagate all the way up to the top level—the public will interact with a different, new NTS. This mechanism surely endows the NTS with a *system-of-systems*[4] (SoS) character and involves complicated dynamic processes that cannot be completely understood or easily modeled.

1.2 Research Statement

From the design and development phase perspective, it was pointed out that a method is needed to evaluate and compare a wide variety of PAV concepts and to effectively present an analysis result to decision makers. From the system-of-systems perspective, a particular concept vehicle can be decomposed into a specific set of design requirements. Looking at the PAV design problem from the aforementioned perspectives constructs the major axes of the present research.

[4]Recently, this term is increasingly used in reference to the transportation system or network-centric warfare. It seems tautological since the term *system* alone can indicate an integrated whole. Nevertheless, the concept of system-of-systems is still useful, and the confusion rising from the two terms can be mitigated per Hitchins (2003, p. 80) as follows: "The term (system-of-systems) is being applied to the creation of new systems by bringing together existing operational systems under a single umbrella and, presumably, creating or adapting links and interactions between the operational systems, which become subsystems of the higher level umbrella system."

1.2.1 Importance of Design Requirements

In light of the above discussion, the role of PAV design requirements is *to formulate a particular PAV concept vehicle.* Therefore, rather than comparing and/or evaluating different concept vehicles directly, their design requirements (design features or attributes after the design process is completed) can be evaluated instead. In other words, evaluating a concept vehicle revolves itself into a new question regarding the evaluation of its design requirements. At this point, it is clear that the investigation of those requirements should start by placing a PAV system within the evolving NTS. As Charles Darwin indicated, it is not the strongest, fastest nor smartest, but the fittest that survives in an environment.

Another important role of design requirements is *to bridge a "decision-making" domain and an "engineering" domain*, this is best portrayed in Figure 9 where the cores from Figures 6 and 8 are juxtaposed. Once a PAV concept vehicle is selected, then engineers from various disciplinary areas would cooperate to generate a viable product design by improving technologies within specific disciplinary circles. This can be called an "engineering" activity in the traditional sense. However, at this time, the PAV design requirements

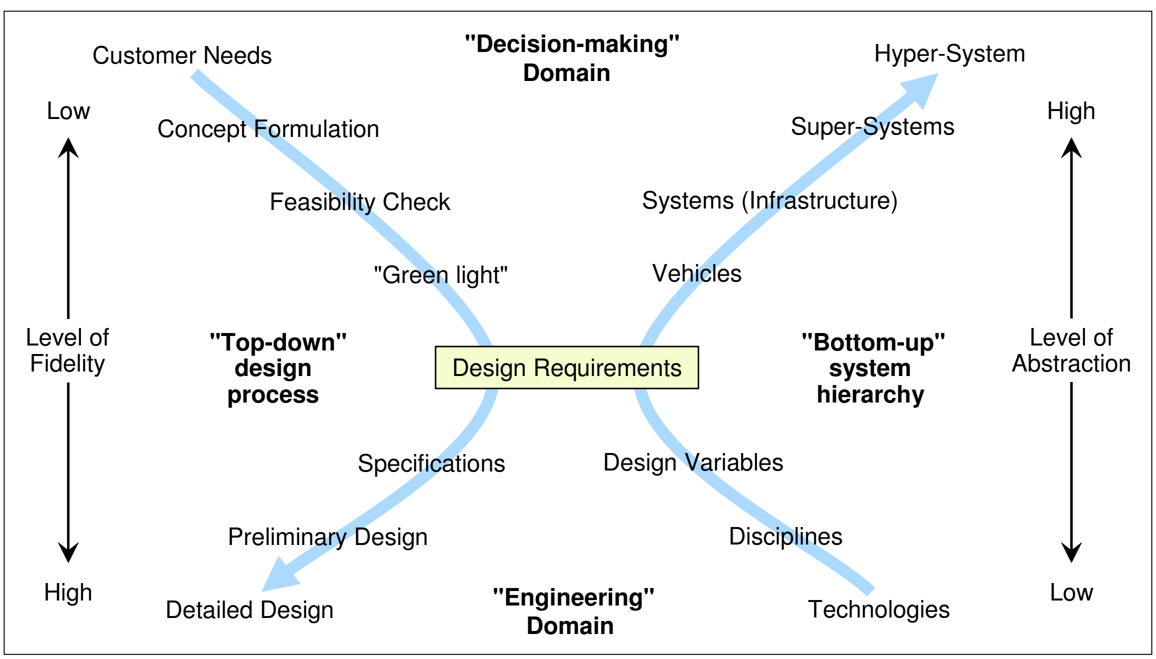


Figure 9: "Decision-making" and "Engineering" Problem Domains

are still at issue. This is one of the reasons why the aerospace community has not been able to freeze the most viable concept vehicle, as evidenced by Figure 5 again. The PAV design problem needs another dimensional wisdom and, unlike conventional aircraft design, must be approached with the same emphasis on "engineering" and "decision-making" aspects alike.

1.2.2 Research Objective Preview

The main body of the present research does not address the design of a specific vehicle nor a particular technological issue. Rather, it intends to look at the larger problem that must precede detailed engineering design phases. The circumstances surrounding the PAV design problem require that "decision-making" issues should be tackled first, generating the design requirements for a viable design. Furthermore, that decision-making process needs active participation from multiple PAV stakeholders: engineers, transportation policy makers, business planners, and customers.

Nevertheless, there is no scientific, systematic and practical way to serve this kind of broad task at this point. Current physics based codes do a fine job of sizing airplanes and helicopters. Existing economic analysis tools can easily determine the break-even point for procurement of commercial transports or fractional ownership of a small business jet. But they are dedicated to solving a problem within a specific boundary. Something else is needed to provide each of the PAV stakeholders with a unified environment to aid the seamless decision-making process.

1.2.3 Thesis Organization

The issues related to PAV and the need for this research have been outlined in this chapter. Chapter II captures previous efforts in the literature relevant to the topic of this thesis. In Chapter III, a theoretical foundation is given to assist in facilitating the course of the thesis. Chapter IV illustrates the main ingredients of the research describing the approaches and methodology proposed. Detailed information on the simulation model generated follows in Chapter V. Chapter VI constructs and explores a set of simulation scenarios as a proof of concept. Finally, a brief recapitulation of the present research and proposed future work is addressed in Chapter VII. The overall structure of this thesis is provided in Figure 10 where the interdependency and flow amongst chapters, sections and appendix materials are pictorially described.

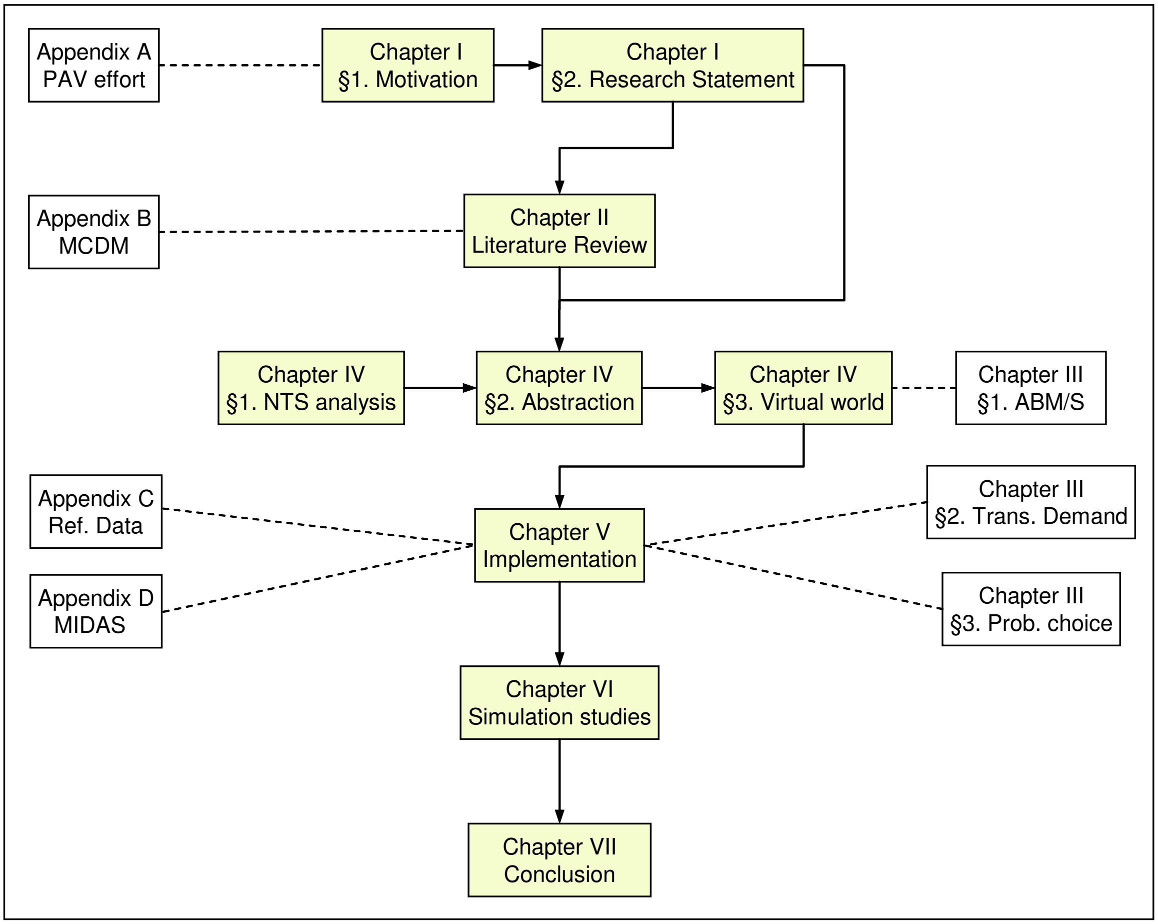


Figure 10: Thesis Structure Overview

CHAPTER II

LITERATURE REVIEW

Contents

This chapter reviews key topics related to the focus of this thesis. Given the vast amount of existing literature, the review is illustrative rather than exhaustive. Generic methodologies to identify and/or elicit a product's design requirements leading to formulating design concepts are reviewed first. The second section examines several attempts addressing evaluation of the concepts found in the aerospace literature. The third section addresses simulation efforts and issues to transportation systems. The last section of this chapter presents the research questions arising from a review of the exiting literature.

2.1 *Tools and Methodologies for Requirement Synthesis*

Given understanding of customers' needs (Step I in Figure 6), a designer attempts to formulate a certain concept vehicle (Step II in Figure 6). Design requirements are elicited and synthesized during these steps. Qualitative approaches that support these tasks are reviewed in this section.

2.1.1 Systems Engineering Tools

The usual approach in determining a product's requirements relies on expert opinion. Typically, a group of specialists who have authoritative expertise participates in a lengthy meeting, finally generating a guideline concerning design requirements.

Systems engineering standards can be thought of as structured descriptions of this type of process, gathering experts experience. These standards provide solid guidelines regarding the development procedure of a product. There exist various standards in the commercial, military and aerospace sectors. A summary of several existing standards is given in Table 1 where the general activities identified in each of the sources surveyed are categorized into five broad concerns. Systems engineering standards are essential in managing a complex and large-scale engineering project. However, it is not surprising that these standards do not direct a specific way to answer the PAV design problem since they are general references applicable across a variety of problems.

In software engineering, the generation of requirements is handled in a formal manner. There is a research thrust called Requirements Engineering (RE) which intends to cover all of the activities involved in discovering, documenting, and maintaining a set of requirements for a computer-based system. (Kotonya & Somerville 1998, p. 8) The key activities of RE include: requirements elicitation, requirements modeling, requirements specification, requirements validation, and requirements management. (Zowghi 1995) The final outcome of these iterative activities is usually referred to as the Software Requirements Specification (SRS). An illustration of the RE process is portrayed in Figure 11.

Table 1: Consensus of Various Systems Engineering Standards [Source: Adamsen (2000, Table 2.1) with modifications]

	What	*How*	*How well*	*Verification*	*Selection*
IEEE 1220-1994	Requirements analysis	Synthesis	Systems analysis	Functional and physical verification	Trade studies and assessments
EIA/IS-632	Requirements analysis, functional analysis / allocation	Synthesis	Systems analysis and control	Verification (defined as a feedback)	
DAU Systems Engineering Fundamentals	Functional analysis	Allocation and synthesis	Evaluation and decision	Analysis, inspection and demonstration	Evaluation and decision
MIL-STD-499A	Mission requirements analysis and functional analysis	Synthesis	Optimization: Effective engineering analysis	Production engineering analysis	Optimization: Trade analysis
NASA Systems Engineering (SP-610S)	Mission, requirements, functional analysis	Synthesis and risk analysis	Systems analysis and modeling	Qualification / acceptance / operational verification	Design optimization and trade study
Systems Engineering at ASDL, Georgia Tech	Requirements and functional analysis	System synthesis through MDO	Systems analysis and control	Robust design simulation	Multi-attribute decision making

The general notion of RE is also applicable to the aerospace field, even though the design philosophy and product development procedures of software engineering are different from those of aerospace engineering. In fact, RE is often treated as a branch of Systems Engineering. Nevertheless, it is difficult to apply RE *per se* to the PAV problem at hand because the primary emphasis of RE is on efficient, systematic and consistent compilation and documentation of a system's requirements.

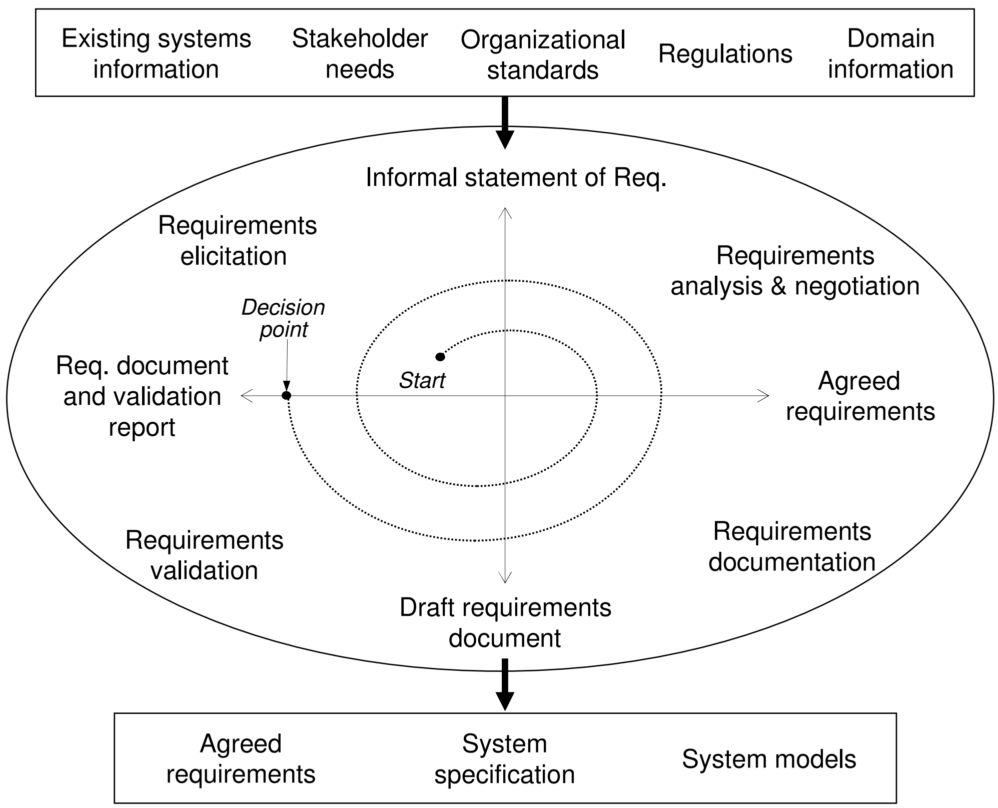


Figure 11: Inputs, Process and Outputs of the Requirement Engineering [Source: Kotonya & Somerville (1998) with modifications]

Quality Function Deployment (QFD) is a powerful tool for supporting product definition and often considered the systematic translation of "the voice of the customer" because it links customer needs to engineering requirements or product/process characteristics. The tool provides a conceptual map for communication across functions and a focus for key product/process design priorities. QFD has extensively supported product development in a range of industries including aerospace, automotive, consumer electronics, clothing, construction and shipbuilding.[1] At the heart of QFD is the House of Quality, shown in Figure 12, which links predetermined customer attributes to specific technical characteristics, built up from interrelated matrices. Most QFD practices use the House of Quality as a stand-alone tool, but it is possible to cascade the matrices to proceed from the customer requirements to the process parameters that need to be controlled. That way, the tool helps structure product planning and design and aims to ensure that customer needs are focused on throughout a project from concept design to manufacture. Although QFD is an useful tool to elicit and rank customer needs in understanding system requirements, it is limited to generating qualitative assessments and has low fidelity in real cases.

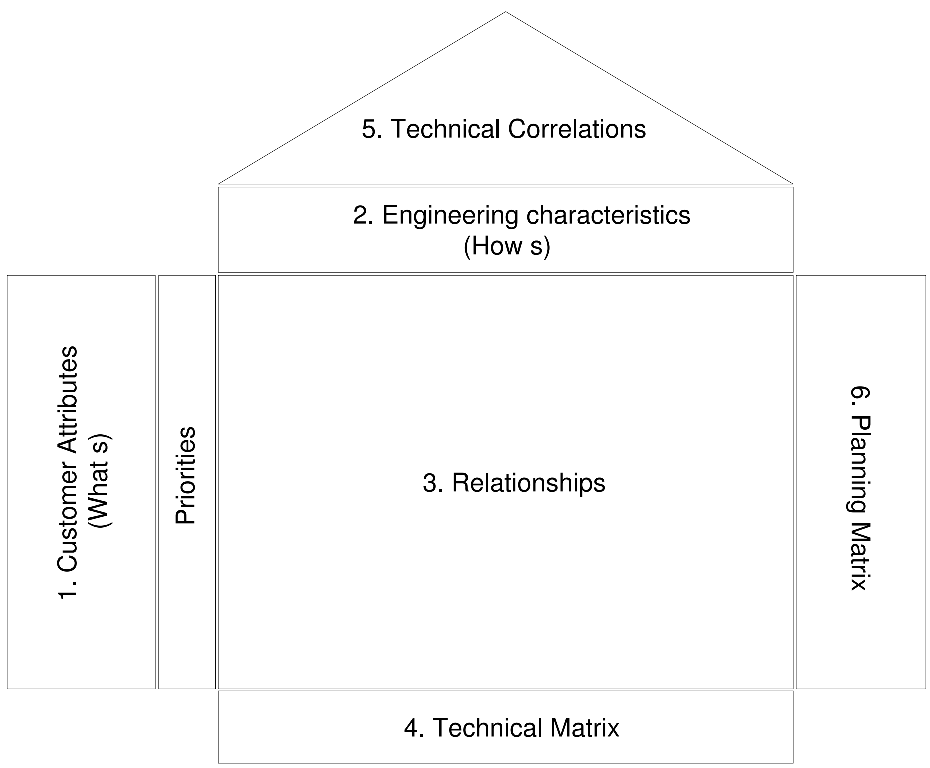


Figure 12: The House of Quality

[1]Originated from Japan in the late 1960s, QFD has become increasingly adopted in the U.S. since the 1980s, and has been credited with the revival of the U.S. automobile industry.

2.1.2 Other Approaches

Genrich Altshuller and his colleagues have created an intriguing method for generating alternatives. He believes that any innovative problem represents a conflict between new requirements and an old system. TRIZ, an acronym for "Theory of Inventive Problem Solving" in Russian, is based on this belief. TRIZ research has proceeded by examining over 2 million patents and classifying them by level of inventiveness to look for principles of innovation. The three primary findings of this research are: 1) Problems and solutions were repeated across industries and sciences, 2) Patterns of technical evolution were repeated across industries and sciences, and 3) Creative innovations used scientific effects outside the field where they were developed. (Altshuller 1984) TRIZ enhances creativity by making individuals think beyond their own experience.

Multi-Attribute Decision Making (MADM) techniques can be creatively used in determining design requirements while obtaining expert opinions. Appendix B.1 provides detailed information on various MADM techniques. Among them, for example, the Analytic Hierarchy Process (AHP) can help decision makers examine a problem with a finite set of alternatives. The AHP evaluates multiple alternatives through a pair-wise comparison process in which experts can gather organized information as well. The limitation with MADM techniques, however, is that they cannot answer the questions necessary to build confidence in the selection of an alternative, especially if a problem structure is complex. (Eagan et al. 2001, p. 27)

2.2 *Analysis-Based Methods for Concept Evaluation*

Design alternatives should be evaluated and/or ranked before a decision can be made. This section reviews a number of methods for this task. Generic methodologies are introduced first followed by utility focused approaches.

2.2.1 Generic Approaches

The most common approach is to define a scalar metric which measures how attractive a certain design is. If the metric can be represented as a function of the design variables of interest, the evaluation problem can be posed as a numerical optimization task with a single objective. For example, Stettner & Schrage (1993) utilized the productivity index (*PI*) as the metric, defined in Equation 1, while investigating tiltrotor aircraft.

$$PI = \frac{\text{Payload} \times \text{Block speed}}{\text{Empty weight} + \text{Fuel weight}} \tag{1}$$

The metric *PI* relates aircraft productivity to total cost because the denominator of Equation 1 can be an indicator for the sum of acquisition and operating cost. In Mavris, Bandte & Brewer (1995), the required average yield per revenue passenger ($/RPM) was selected as an objective for a high speed civil transportation (HSCT) design problem. Since the HSCT must have competitive advantage to current subsonic transportation, economic viability was of utmost importance and $/RPM captures the economic concerns of all interested parties.

When multiple conflicting objectives are in strong presence of a design problem, and a design tradeoff is required, MADM techniques should be employed in the evaluation process to investigate a set of candidate designs. For instance, Mavris & DeLaurentis (1995) looked at a military aircraft concept evaluation problem and they suggested the concept of the Overall Evaluation Criterion (OEC), as shown in Equation 2, to select the most effective design:

$$\begin{aligned} \text{OEC} \;=\;& \alpha\,(\text{Affordability}) + \beta\,(\text{Capability}) \\ &+ \gamma\,(\text{Operational Safety}) + \delta\,(\text{Survivability}) + \varepsilon\,(\text{Readiness}) \end{aligned} \tag{2}$$

where the Greek letters sum to one. This equation is a typical formulation of the Weight Sum Model (WSM) for MADM problems. Taking another example, Mavris & Hayden (1996) tackled a supersonic business jet (SSBJ) design problem combining the OEC and *PI* approaches, as shown in Equation 3:

$$\text{OEC} = \alpha \frac{\text{LCC}_\text{b}}{\text{LCC}} + \beta \frac{PI}{PI_\text{b}} \tag{3}$$

where the Greek letters sum to one as before and the subscript ‘b’ denotes a baseline configuration. *PI* is a productivity index defined in a similar way as in Equation 1 and LCC indicates the life-cycle cost. Altogether, these “scoring” approaches are conceptually simple, but are not free from limitations, especially with the WSM. For example, subjectivity is involved with the weight determination process. In addition, the WSM only finds the “corner” points as the best alternatives, regardless of the weights, if the Pareto front is concave.

Some noticeable advances have been made at requirements synthesis for evolutionary systems. Mavris & DeLaurentis (2000) introduced a novel approach which presents a framework to simultaneously examine the effect of designer’s and/or decision-maker’s inputs, such as design requirements, design and economic variables, and potential technologies, by creative use of a meta-modeling technique with vehicle synthesis tools. In addition, the methodology accommodates probabilistic treatment of the inputs since they are likely uncertain in many cases. This approach is underpinned by the so-called Unified Trade-off Environment (UTE) where the analysis can be performed in a parametric and visual way. Baker & Mavris (2001) expanded the approach and applied it to a future transport rotorcraft design problem.

In summary, these analysis-based methods are found to be very useful and suited for many conceptual design problems. However, they are not directly applicable to the PAV design problem for a number of reasons. First, these methods obviously require physics-based codes to a certain degree, which implies the evaluation process can be performed within a specific vehicle platform or configuration (i.e., rotorcraft or fixed wing aircraft), not across a wide variety of different platforms, let alone revolutionary concept vehicles. This is a serious limitation since the PAV design space is truly open. Even if there exists a 'universal' physics-based code that can simultaneously evaluate a wide variety of PAV concepts, a decision-maker would still be confounded with the *incommensurability* issue. For instance, *PI* is not the most suitable metric to compare a helicopter and a fixed wing airplane—it is inappropriate to compare "apples and oranges" using either a single metric or a MADM technique.

2.2.2 Utility Focused Approach

Since the 1960s, there have been attempts to understand the PAV design problem from a different angle. These attempts look at the problem from the viewpoint of utility (usage) of a vehicle or economic utility to a traveler. For example, Drake, Kenyon & Galloway (1969) focused on characterizing personal mobility solutions in the context of mode choice and the value of time, as shown in Figure 13.

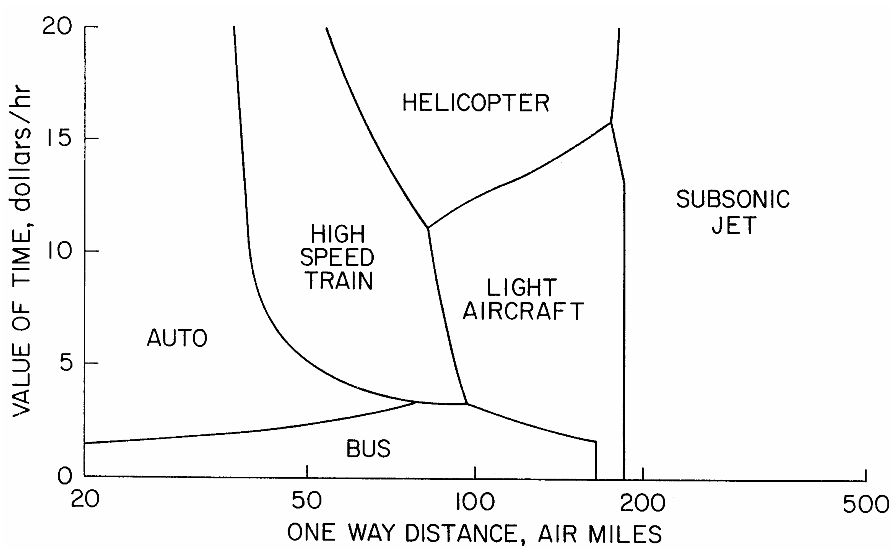


Figure 13: Minimum Cost Transportation Modes [Source: Drake et al. (1969, Fig. 3)]

Downen & Hansman (2003) inherits the same philosophy, essentially recreating similar charts as Figure 13. This study performed a web-based survey of active general aviation (GA) pilots and then developed a mode choice model based on the value of time in an attempt to identify key barriers to the utility of GA. The identified key issues for increasing the utility of GA include: near-all-weather capability of GA aircraft, better accessibility to ground transportation, and modification of the business model for owning and operating GA aircraft.

DeLaurentis, Lim & Kang (2004) emphasized door-to-door trip time in a PAV concept investigation process as well. Through a “cash flow” analysis and a “PAV benefit visualization tool” incorporating the UTE, the benefit from PAV utilization can be quickly visualized compared to existing transportation vehicles. This is based on the premise that travel time saving is converted to monetary “profit” to a PAV user. This line of work is an on-going effort under the framework of NASA’s PAVE project. In Mavris & DeLaurentis (2002), six baseline PAV concepts were selected, ranging from an autogyro to a small jet airplane, and analyzed based on their own economic aspects. The initial findings include that drastic improvement in vehicle speeds and costs will be required before PAV could be considered as an affordable transport means for average travelers.

Compared to the generic methodologies, these approaches have significant advantages for the PAV study since the mobility issue is posted at the center of the problem formulation. In addition, since design requirements are not linked with a particular physics-based code from the outset, simple analyses can be quickly performed to explore various concept vehicles or their effect on the door-to-door trip time. However, these approaches are based on deterministic assumptions (often *ad hoc*) and are designed to tackle the personal mobility issue in a static way. Thus, they are incapable of dealing with the dynamic behavior of the entire transportation system, while maintaining a system-of-systems perspective. For example, it is difficult to simultaneously capture a large amount of diverse traveler characteristics.

2.3 *Simulation of the Transportation System*

The study of transportation systems supports planning, design and operation (or scheduling) of real transportation systems. Driven by exponential growth in computing power (Thomke 1998, Fig. 1), computer simulation of transportation systems has become a basis of increasing the likelihood of achieving a "better" performance, a "faster" acquisition schedule, and a "cheaper" life-cycle cost, as enunciated by the DoD's Simulation Based Acquisition (SBA) Initiative. (Johnson, McJeon & Szanto 1998)

2.3.1 Classification

Over the past decades, a wide variety of transportation models have been developed depending on specific use cases. The vast and cross-linked nature of this subject has now become too diffuse to be covered by any single disciplinary area. Presented in Table 2 is an attempt to classify the various models based on several categories instead of doing an exhaustive survey on this topic, which would be quickly out of the scope of the present research.

Table 2: Classification of Simulation Models in Transportation

Category	Items
Modeling space	Ground / Air
Number of modes	Single-modal / Multi-modal / Inter-modal
Areal resolution	Street / Urban region / Corridor / Nationwide
Time / event	Continuous / Discrete
Modeling fidelity	Microscopic / Mesoscopic / Macroscopic
Building block	Mathematics / System dynamics / Agent-based
Uncertainty	Deterministic / Stochastic

A *mode* refers to a specific travel means for a traveler. It can be a car, transit, general aviation, commercial airline or even walking. Multi-modal problems treat more than one mode, in general, to study management and interaction of multiple modes whereas inter-modal problems handle the shipment of cargo and the movement of people involving more than one mode during a single, seamless journey. (Jones, Cassady & Bowden 1999) Continuous simulation models describe the changes in the state of a system's element over time on a continuous basis in response to continuous inputs. Discrete simulation models have the capability to deal with abrupt changes in states at a point in time. Simulation models can also be classified according to the level of detail with which they represent the system to be studied. Microscopic models consider the characteristics of each individual element and its interactions with other elements of interest. Macroscopic models concern aggregate measures such as flow rate, speed and traffic density. If a model falls in an in-between area, it is called a *mesoscopic* model. Traditional transportation models are built up from mathematical techniques such as partial differential equations, queueing theory or network theory. Meanwhile, System Dynamics and Agent-Based Modeling techniques are gaining popularity. These building blocks, ranging from rigorous mathematical approaches to agent-based techniques, can cover diverse questions and investigate diverse situations.

2.3.2 Large-Scale Models

Reviewed first are large-scale simulation models that are relevant to the air transportation system. Here we will briefly outline several models. Detailed information on each can be found in Odoni et al. (1997), Pritchett et al. (2003) and footnotes.

SIMMOD[2] (The Airport and Airspace Simulation Model) is a network model of the National Airspace System (NAS) by developed by the FAA in the early 1990s. It addresses air traffic control policies and procedures, airport and airspace, and flight schedules. The model tracks the movement of individual aircraft through an airport/airspace system. TAAM[3] (Total Airport and Airspace Modeler) is a popular discrete-event simulation tool for aviation analysis. The TAAM models the entire airport and airspace environment in great detail. It can display realistic three-dimensional color models of the airport, the

[2]http://www.atac.com/prodsvs/simmod.htm
[3]http://www.preston.net/products/TAAM.htm

airspace and individual aircraft. LMINET[4] is a queuing network model that provides analytical solution for estimating delays at airports and enroute sectors of the NAS. It solves differential equations that describe the distribution of delays over a network of airports, given flight schedules, aircraft itineraries and airport capacities. DPAT[5] (Detailed Policy Assessment Tool) is a discrete event simulation that can capture current and future air traffic. DPAT simulates individual flights through a sequence of constrained resources such as airport and airspace capacity. It provides traffic flow predictions including arrival, departure and delay at major U.S. airports. The above models are referred to as *mechanical models* (Pritchett et al. 2003, p. 371) focusing on traffic and trajectory analysis. They are ideal tools to investigate delay and capacity related phenomena, but obviously help little to the PAV design problem.

Recent endeavors in this field have brought a fundamental progress in agent-based modeling techniques.[6] The Jet:Wise model developed at MITRE is such an example (Niedringhaus 2004). Taking airline companies and leisure passengers as agents, the model attempts to explore the evolution of the airline industry within the NAS. In each cycle of simulation, both airline and passenger agents make successive decisions to achieve their respective goals. This cycle is repeated until an "equilibrium"—the state at which the agent behaviors are almost the same as in the immediate prior run—is attained. This iterative mechanism is called an agent-based evolutionary scheme. The Jet:Wise model shows its ability to capture the emergent behavior of the real airlines. For example, the hub-and-spoke system emerged as an airline routing behavior without explicit mechanisms leading to that phenomena. Meanwhile, NASA is enhancing a NAS modeling and simulation capability through the VAMS (Virtual Airspace Modeling and Simulation)[7] project. ACES (Airspace

[4] http://techreports.larc.nasa.gov/ltrs/PDF/1999/cr/NASA-99-cr208988.pdf
[5] http://www.mitrecaasd.org/library/one_pagers/dpat.pdf
[6] In-depth discussion on this subject constitutes §3.1.
[7] http://vams.arc.nasa.gov/

Concept Evaluation System) is part of this effort, which utilizes the High Level Architecture (HLA) and an agent-based modeling paradigm to cover aircraft operations from gate departure to arrival. (Meyn et al. 2004)

Trani et al. (2003) proposes a nationwide, multi-modal, inter-city transport model to investigate the viability of NASA's SATS project, as an extended form of the conventional transportation demand analysis. The overall analysis method revolves around manipulating a county-by-county trip origin/destination matrix (3091$\times$3091). Hence, a key challenge is a proper handling and reconciliation of disparate, huge databases such as Census, American Travel Survey, etc. In contrast to high level of geographic granularity, the model has aggregated demographic and socioeconomic factors by county to strike a balance between a computational efficiency and level of traveler stratification.

On the other side, research in the ground transportation domain also has generated a wide variety of simulation models. Only two representative large-scale models are outlined herein. CORSIM[8] (CORridor SIMulation) is a well-known simulation model. It was developed in the 1960s through the Federal Highway Administration (FHWA). CORSIM is a comprehensive microscopic traffic simulation, applicable to surface streets, freeways, transit operations, and integrated networks with a complete selection of control devices (i.e., traffic signals and ramp metering). TRANSIMS[9] (TRansportation ANalysis SIMulation System), developed at Los Alamos Laboratory, incorporates the most advanced modeling approaches. At the core of TRANSIMS is an agent-based simulation system capable of simulating detailed movements of persons and vehicles through the transportation network. The TRANSIMS models have been tested onto large metropolis including Dallas, Texas and Portland, Oregon.

[8] http://www-mctrans.ce.ufl.edu/featured/TSIS/Version5/corsim.htm
[9] http://transims.tsasa.lanl.gov/

2.4 *Summary*

Under the bold mobility goal, the starting point of the present research is a need to examine various PAV concepts from a system-of-systems perspective, a prerequisite capability in identifying the most viable designs. (§1.1) This need evolved into the principal theme of this dissertation, which is the development of a framework to connect the "decision-making" and "engineering" domains in the early design phases. (§1.2) In light of the theme, the previous section has briefly reviewed relevant approaches in regard to formulation and/or evaluation of design requirements of a new mobility vehicle. The review generates a few observations and corresponding implications as follows.

2.4.1 Synopsis of Literature Review

Usually, design requirements are formulated by intuition, reasoning or some structured methodologies, mostly, in qualitative ways. (§2.1) Although many qualitative tools are powerful and important in properly framing overall design problems, they are not sufficient to form a firm basis for the decision-making process. This leads to the examination of analysis-based methods that can complement and support the process. (§2.2) However, these quantitative methods are limited in their range of usage in the absence of universal synthesis and sizing codes. In addition, traditional metrics used for aerospace vehicles are not sufficient, although necessary, for evaluating a concept PAV. In such circumstances, the use of utility focused approaches is a simple yet useful solution to examine the effects of design requirements while abstractly handling them. They are visually and conceptually elegant and complimentary to the previous methodologies, but still insufficient due to the necessary simplification of the problem structure. Therefore, a need exists for the exploration of modeling and simulation efforts for the transportation system. (§2.3) Each large-scale transportation model surveyed possesses a tremendous capability to track microscopic, even second-by-second, behavior of the computer-generated objects of the investigator's interest. However, the mechanical models and the agent-based models focus on the air transportation system only. Trani's endeavor comes closer to embrace the system-of-systems perspective but the low level of demographic fidelity makes it difficult to study effects such as a wide variety of adaptive behaviors of the travelers. Also, the model is not calibrated yet as it is an on-going effort. Therefore, to date, the capability to tackle the PAV design problems in the context of the entire NTS is elusive. The shortcoming of existing models must be elaborated.

2.4.2 Research Questions and Hypotheses

Given the above discussion, two fundamental research questions are stated herein summarizing the issues raised. These questions spur corresponding hypotheses that embody the major thrust of the present research. In addition, supplementary research questions are presented without accompanying hypotheses to guide and to clarify the development of the research.

Question 1. How can PAV concepts be evaluated to seamlessly integrate the mobility objective with the PAV design process?

This research question is directly related to facilitating the decision-making process in the development of PAV concepts. The following hypothesis is investigated in response to the above research question.

Hypothesis 1. Exploration of PAV design requirements in the context of the NTS provides the necessary linkage to adequately treat a wide variety of issues related to PAV designs and PAV stakeholder concerns.

The preceding summary implied that a tangible model that closely imitates the NTS would be required to provide the foundation on which this hypothesis is tested. The need to build a virtual NTS is established, but is it feasible to integrate heterogeneous objects from the diverse domains that constitute the NTS? and how? The following research questions address these concerns.

Question 2. How can one properly resolve the modeling boundary, the constituent elements within the boundary, and the level of granularity to construct a simulation model?

This is a challenging question. Given the extent and complexity of the NTS, even proper determination of the modeling boundary is a problematic task. Making matters worse, there are a myriad entities in the NTS that exhibit unique behaviors in their own realms and ultimately incur complicated interactions unknown even to themselves. This leads to the subsequent secondary research hypothesis posed as follows.

Hypothesis 2. A proper abstraction framework with a holistic perspective can guide the rapid construction of a useful simulation model and can avoid unmanageable complexity in the model.

If the above hypothesis is efficacious, the end result will be a working computational model of the NTS. A few supplementary research questions are posed to facilitate the verification of the above hypothesis, complementing its broad nature.

1. What are the essential elements in the NTS that should be considered to study the PAV design problem?
   1.1 What is the boundary and scope of a specific modeling target?
   1.2 What modeling approach should the model adopt? Does it require a significant effort for the implementation?
2. Can the simulation model compare a wide variety of PAV concepts? (especially for a revolutionary concept, in the absence of an appropriate physics-based code)
   2.1 What is a quantitative metric to measure the effects?
   2.2 Is the metric compatible with other vehicles with a different platform?
   2.3 For example, can one answer the following question in a quantitative way?
   ▷ How important is roadability in the PAV design requirements?

3. Can the simulation model accommodate multiple stakeholders to aid the group decision-making process?
   3.1 Can it be scalable and non-prescriptive? Can it work as a "living system"? In other words, is it adaptive to changes in assumptions, ground rules and information or data?

Overall, the capability of the computational model to be generated can be examined with the above questions. If they can be answered in an affirmative way, the final outcome will be an approach for an integrated decision-making framework (hence, the title of this dissertation) that serves to solve a "decision-making" problem in the pre-conceptual design phase.

CHAPTER III

THEORETICAL FOUNDATION

Contents

The objective of this chapter is to provide necessary information to familiarize readers with the context of this thesis. The work begins with an exploration of the capabilities and limitations of the agent-based modeling and simulation technique, laying a basis for the present modeling efforts. The next section details theory of transportation demand to establish a mindset in creating the model assumptions, methods and implementations. Probabilistic choice theory is introduced and elucidated in the last section as this theory serves as a fundamental logic in constructing the agent's rudimentary behavior—conditional decision-making under uncertainty.

3.1 Agent-Based Modeling and Simulation

At the heart of the present research is an agent-based modeling and simulation (ABM/S) as a fundamental, enabling technique. This new approach to the study of complex systems has provided researchers with an important theoretical and methodological framework for helping to understand a variety of dynamic, non-linear behaviors. This section begins with the concept of complex systems and agents as understood in this thesis.

3.1.1 Agent

Complex systems often lie beyond our scope of understanding. Complex systems usually have large numbers of defining elements and exhibit non-linear dynamics in their behaviors. Moreover, complex systems also contain interactions between elements that are too complicated to completely analyze and comprehend, making them extremely difficult to deal with analytically or even empirically. Traditional modeling techniques, which are based on the philosophical foundation of reductionism and the top-down approach, are not always an ideal way to treat complex systems.

Agent-based Modeling (ABM), also known as Individual-Based Modeling, takes the other position. It is a bottom up modeling technique that focuses on constructing a virtual world. The idea behind ABM is that the global behavior of a complex system derives from the low-level interactions among its constituent elements. The fundamental building block of models of complex systems is the so-called *agent*, an entity that autonomously fits itself in a certain environment. Various definitions of the term agent can be found in the literature. This is because, in part, the term is from two different lines: software engineering, and modeling and simulation. For example, Wooldridge & Jennings (1995) states that: an agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives. On the other hand, Holland (1995) defines an agent as rule-based input-output element whose rules can adapt to an environment. But common keywords that relate many of the numerous definitions are *adaptive* and *autonomous*. An agent is adaptive in that it can use its experience to continually improve its ability to deal with changing environment. It is an autonomous entity if it operates without external guidance and does not need to follow step-by-step instructions from the modeler who created it. In general, an adaptive autonomous agent can be characterized by the following attributes (Ilachinski 1997, p. 14):

- It is an entity that, by sensing and acting upon its environment, tries to fulfill a set of goals in a complex, dynamic environment.
- It can sense the environment through its sensors and act on the environment through its actuators.
- It has an internal information processing and decision-making capability.
- An agent's goals take from diverse forms: desired local states, desired end goals, selective rewards to be maximized, or internal needs that need to be kept within desired bounds.

The adaptive and autonomous agent operates in the following manner: First, an agent sees the world, and then it makes a decision that entails an action. The world is influenced, however insignificantly, by the action. The same agent now senses a different world, and updates its knowledge, which may then cause a different action or even shift its goal. Accumulated action of the agent produces an *emergent behavior*, which often renders a useful insight to the real world even if a model itself is in a very simple form. This mechanism is portrayed in Figure 14.

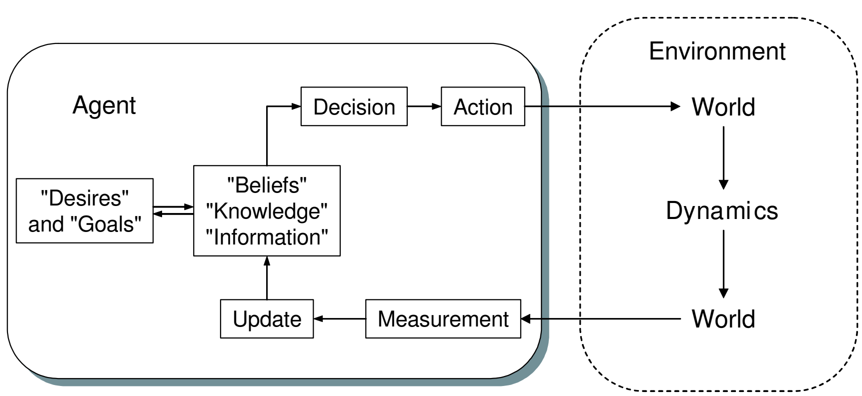


Figure 14: Concept of Agent-Based Modeling

After the rules for guiding an agent's actions and the relationships between the agent and environment are established, a computational model that imitates the real world is simulated. The observation and analysis should follow the simulation invoked by a user: i.e., *let them play and watch it*. For this reason, agent-based modeling and simulation (ABM/S) can be thought of as a scientific reasoning approach that complements deduction and induction. Axelrod (1997) clearly explains this point as follows:

> Like deduction, it starts with a set of explicit assumptions. But unlike deduction, it does not prove theorems. Instead, an agent-based model generates simulated data that can be analyzed inductively. Unlike typical induction, however, the simulated data come from a rigorously specified set of rules rather than direct measurement of the real world. Whereas the purpose of induction is to find patterns in data and that of deduction is to find consequences of assumptions, the purpose of agent-based modeling is to aid intuition.

The major strength of ABM/S comes from the fact that it is a simple, versatile and flexible method that is well suited for studies of complex non-linear systems. Agent-based simulations can reveal both qualitative and quantitative properties of the real system, so ABM/S can be used as a versatile laboratory to perform experiments to test nearly any kind of imaginable hypotheses. ABM/S, however, does have a downside. For a highly realistic model, large amounts of data input and computation may be needed. Another issue for ABM/S is that identifying the "right" rules or behaviors that capture the real dynamics can be a somewhat ad hoc process. (Hood 2002)

3.1.2 Architecting Multi-Agent Systems

It is quite natural that most of agent-based models in practice have multiple number of agents. This is called a multi-agent system (MAS). (Weiss 1999) Since there are more than two agents involved, interaction can occur between agents and environment and between agents themselves. The following discussion gives a few examples regarding how MAS can be constructed to meet a modeler's situation.

The simplest case when "independent" agents interacting with the environment but not one another is portrayed in Figure 15. In the figure, arrows connecting each agent and environment indicate bi-directional interaction, that is, "see and act". This case is just a simple expansion of Figure 14 that has a single agent. For example, people's behavior on purchasing house to own may be the case.

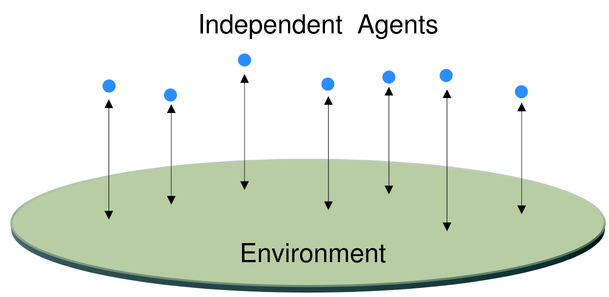


Figure 15: MAS with Independent Agents

However, in general, there are many examples of agents interacting one another. One famous example of this is flocking behavior of birds. Birds, when they fly together, watch obstacles in their flight direction as well as adjacent neighbors to avoid possible collision and to be inside of their group.[1] In this case, agents can communicate with each other and act directly on other agents' states. The *information layer*, shown on top of Figure 16, is introduced to present this mechanism.

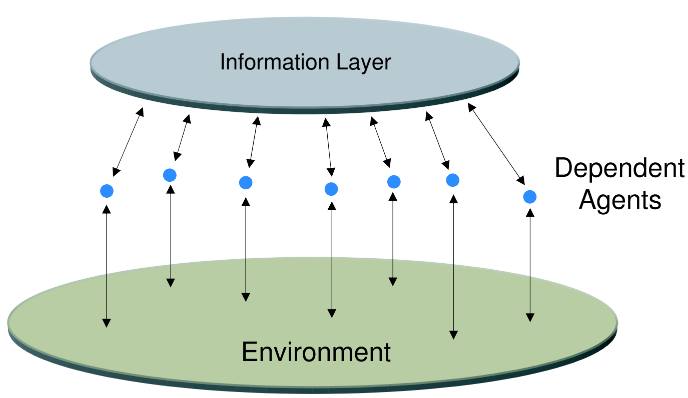


Figure 16: MAS with Information Layer

[1]Reynolds (1987) implemented flocking behavior of birds on computer graphics and named it "boids" (bird + oids), which strikingly resembles the real one and adapted by Hollywood computer animators.

An advanced MAS architecture can be constructed using multiple information layers. Figure 17 shows a multi-agent system in which case agents interact with the environment as usual but part of them collectively behaves as a group. The agents in the figure are divided into two groups: group X and group Y. Agent X_n follows the rules, exchanges information, and affects other members in its group who share the same kind of behavioral patterns and rules through information layer X. Likewise, agents in group Y, which may retrieve a different set of information from the environment and affect the environment in a different way, do so with their own rules without (or with) consulting group X. Many human social behaviors follow this architecture. A representative example would be the stock market where individual and institutional investors form different groups and act differently.

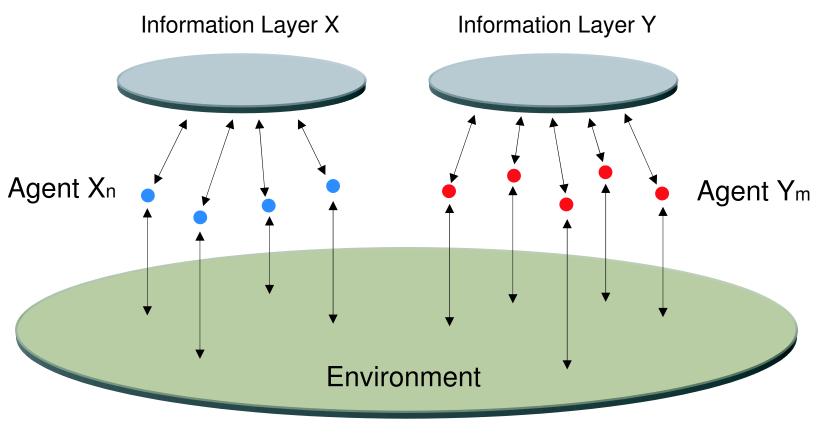


Figure 17: MAS with Multiple Groups

The last architecture of multi-agent configuration is that of organizational architecture, shown in Figure 18. The best example can be found in military organization where the general is the head and the order passes down from top all the way to the bottom level. The lowest level agents, analogous to private soldiers or fighters, directly interact with the enemy and the environment around them. The agents X and Y in the intermediate level gather information and report this, such as attrition, to the super agent. They can control the lowest level agents with their own discretion as long as their goals obey the super agent's instructions.

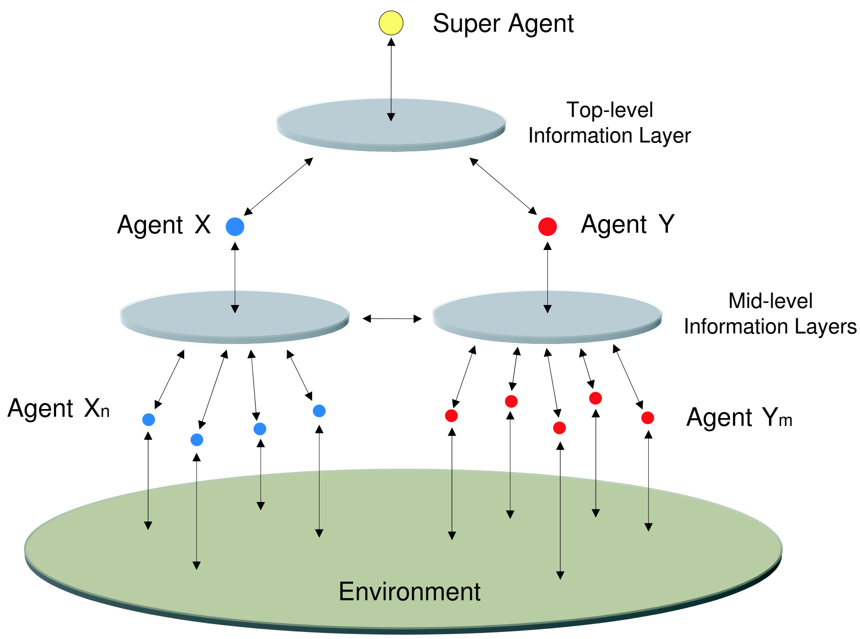


Figure 18: MAS with Hierarchical Organization

The flexibility of ABM has been demonstrated so far through the above examples. Indeed, a modeler can construct virtually any form of agent-based model with the appropriate architecture depending on the nature of the application problem.

3.1.3 Recent Applications of ABM/S

Although present applications of ABM/S have become very numerous, it was not until mid 1980s that the concept of agent based modeling blossomed with the evolution of computer technology. Additionally, as the amount of fundamental research on complex systems grew, the use of agent-based modeling and simulations became more widespread. Below lists a few applications of ABM/S found in the literature.

3.1.3.1 Natural Sciences

Ecology would be one of active application fields within the natural sciences. Fishwick, Sanderson & Wolff (1998) presented modeling method for use in large-scale ecosystem, focusing on the dynamics of species with the Florida Everglades. Hartvigsen & Levin (1997) developed an agent-based model of plant-herbivore interactions to test the interactive effects of explicit space and co-evolution on population and community dynamics.

ABM/S technique is also naturally adapted to the study of animals or plants. Sumpter (2000) presented a mathematical study of determining the relationship between the local interactions of individuals in a population and the global dynamical behavior of that population in the context of honey bee behavior, using a variety of modeling techniques including ABM. Scheffer et al. (1995) tried to model behaviors of fish population with their super-individual approach, an ABM technique which aims to enhance computational efficiency. Deutschman et al. (1997) carried out computational simulations of forest dynamics using an individual-tree model of the forests of the northeastern United States.

Though not an obvious match, ABM/S has even found a fit in chemistry. McMullin & Varela (1997) applied ABM/S to the field of artificial chemistry and presented simulation results which exhibits spontaneous emergence and persistence of autopoietic (*chem.* "it runs by itself") organization.

3.1.3.2 Social sciences

The nature of the social sciences is the study of the human aspects of the world. ABM/S is a promising technique for social scientists in economics, political science, sociology and psychology. Its development has revived old and raised new questions regarding the dynamics of complex human society.

In economics, a new branch of economics—called agent-based computational economics (or ACE)—is fledging with increasing attention recently. Practical applications of ACE can be abundantly found. Basu, Pryor, Quint & Arnold (1996) have build a micro-analytic model, called Aspen, to simulate the U.S. economy. Raberto et al. (2001) carried out an agent-based simulation to look at a behavior of a virtual financial market through a realistic trading mechanism for price formation. An excellent survey paper is given by Tesfatsion (2002) regarding this blooming field.

Other social sciences have actively adopted ABM/S as well. Chakrabarti (2002) is building an ABM to provide a connection between theoretical study of corruption developing micro models of individual acts and empirical study of corruption at the country level. McPhail (1997) discussed collective actions of people in large assembling and dispersal such as crowds, mobs and demonstrations. Helbing, Farkas & Vicsek (2000) also performed a simulation study on aggregate behavior of individuals in a panic situation. Benenson & Omer (2001) are doing an interesting research to demonstrate the dynamics of urban residential distribution of Arab and Jewish people in a few Israeli cities. Merelo, Prieto & Rivas (1997) attempts to forecast the effects of mass media advertising using ABM/S.

3.1.4 Summary

Agent-based modeling is a bottom-up approach to understanding complex systems. It encodes attributes and behaviors at the individual component or microscopic level of the system. The system's macroscopic properties "emerge" as a consequence of these attributes, behaviors and the interactions between them. It is thus a powerful complement to top-down modeling approach, particularly because it allows for the implications of assumptions that are necessarily made in top-down analysis. Whereas agent-based models can be made arbitrarily realistic by capturing the processes, mechanisms or architectures that drive individual components, they can also be made quite abstract in an attempt to understand "the essence of the problem". (Hood 2002)

As pointed out by Bonabeau (2002), ABM/S is particularly useful for modeling of flows, markets, organizations, and diffusion in which system behaviors change due to learning or adaptation. He further states that ABM/S can bring significant benefits when 1) the interactions are complex, nonlinear or discontinuous, 2) an agent is *spatially explicit*, 3) agents are heterogeneous, 4) the topology of the interactions is heterogeneous and complex, and 5) agents exhibit learning and adaptation behavior. However, unlike other approaches based on mathematical formalism, there can be an element of subjectivity in constructing and testing an agent-based model. Also, the simulations of ABM are potentially very intensive from a computational aspect.

3.2 Theory of Transportation Demand

Selected topics related to transportation demand analysis are organized and elucidated in this section, stemming from the literature spread including McCarthy (2001), Kanafani (1983), Labbé et al. (1998) and Noland (2001), among others. The discussion begins with basic foundations of neoclassical economics.

3.2.1 Utility and Indifference Curve

Neoclassical economic theory postulates that human beings are self-interested and highly rational when it comes to making decisions about economic activity. In other words, human beings select the choices that offer them the best possible advantage, given the circumstances they face. Circumstances involve a wide variety of factors: the prices of goods and services, and the constraints on the decisions that may make such as income, regulations and technology limitations. Another premise is that our greed is non-satiable; the more, the better, always. Based on these postulations, neoclassical economists regard that any individual has a capability of comparing alternatives.

Now suppose an individual considers M goods $a_1, a_2, \cdots, a_M$. Then the individual's preference can be represented using a binary operator $\succeq$. If $a_1 \succeq a_2$, the decision-maker either prefers a_1 to a_2, or is indifferent. Let this operator, called the *preference-indifference operator*, have two basic properties: $a_i \succeq a_j$ or $a_j \succeq a_i$ (*Comparability*) and $a_i \succeq a_j \wedge a_j \succeq a_k \Rightarrow a_i \succeq a_k$ (*Transitivity*). Based on these properties, all the alternatives with finite number considered can be sorted by their utility values, so the existence of an alternative preferred to all of them is guaranteed, that is, there exists an $a^\star$ such that $a^\star \succeq a_i$ for all $i = 1, 2, \cdots, M$. The concept of utility function is based on this idea. It is guaranteed that a certain real function $U : A \mapsto \mathbb{R}$, where the choice set A is the list of all the alternatives, exists such that $a_i \succeq a_j \Leftrightarrow U(a_i) \geq U(a_j)$ for all i and j. Hence, the analysis of the individual's preference level for a_i can be studied by evaluating the utility of the corresponding alternative $U(a_i)$ as long as the real function U has the order-preserving property.

In general, each alternative has multiple attributes. When the alternatives represent a family of products, a consumer characterizes the utility of an alternative by considering its price and performance altogether. Further, an alternative can be a bundle of distinct goods $a = (x, y)$, for example, where x and y represent the quantity of foods (X) and clothes (Y), respectively. In this case, the choice set is not necessarily discrete nor bounded. Since human beings are non-satiable, as long as x and y are real positive, any combination of (x, y) belongs to the choice set A, that is, $A = \{(x, y) | \; x \geq 0, \; y \geq 0\}$.

Consider alternatives a_1, a_2, $a_2{}'$, $a_2{}''$ and $a_3 \in A$ in Figure 19(a). It is obvious that a_2 is preferred to a_1 to the consumer since a_2 dominates a_1 in terms of both x and y. The preference between a_2 and $a_2{}'$, however, is vague since two alternatives are compromising each other. Assuming $a_2 \succeq a_2{}'$ and $a_2{}' \succeq a_2$, that is, $U(a_2) = U(a_2{}')$, a consumer feels they are indifferent. One can locate many alternatives that have the same utility value (e.g., $a_2{}''$). Then, a curve can be imagined that connects all the possible alternatives that have the same level of utility in the choice set. This is called an *iso-utility* or *indifference curve*. Different iso-utility curves can be generated and they should be parallel to each other as portrayed in Figure 19(b). From this set of indifference curves, an analyst can decide which alternative is better between alternatives $a_2{}''$ and a_3—whose relationship was ambiguous by looking into Figure 19(a)—for a particular consumer under study.

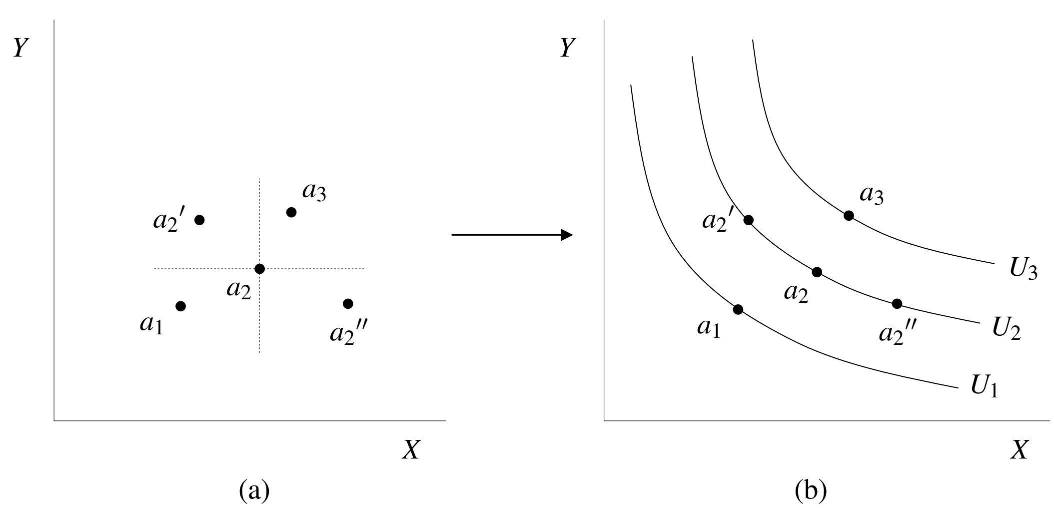


Figure 19: Generation of Indifference Curves

Some mathematical properties can be driven from the definition of the indifference curve. Since the utility is constant along the indifference curve, its total derivative $dU(x,y)$ must vanish:

$$dU|_{U=c} = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy = 0 \tag{4}$$

where c is any constant in $\mathbb{R}$ and this equation gives

$$\left.\frac{dy}{dx}\right|_{U=c} = -\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = -\frac{U_{,x}}{U_{,y}} \tag{5}$$

which indicates any tangent line on indifference curves should have a negative slope since $U_{,x}$ and $U_{,y}$ are always positive. This result can be extended to the general case where n attributes are involved. Then an alternative is a vector $\underline{x} = [x_1, x_2, \cdots, x_n]^T$, and the utility function $U(\underline{x})$ represents a surface in n-dimensional space if the same level of utility is assumed. Likewise, the total derivative vanishes on this *indifference surface*, hence

$$dU(\underline{x})|_{U=c} = \sum_{i=1}^{n} \frac{\partial U}{\partial x_i} dx_i = 0 \tag{6}$$

where $U_{,x_i} \geq 0$ for all i under the non-satiable assumption. Typically, an indifference curve (or surface) is convex to the origin, indicating that the marginal utility from consumption of goods is decreasing. This is called *the law of diminishing marginal utility*. (diminishing returns)

3.2.2 Derivation of Transportation Demand

Suppose a hypothetical situation that transportation and all other goods are given by a commodity bundle (t,x) which contains t units of transportation (T) and x units of all other goods (X). As shown previously, an indifference curve can be generated that gives the consumer the same level of utility. However, this does not necessarily mean the consumer has no limitations in selecting any alternatives on the iso-utility curve. The consumer always has constraints under which she/he seeks to maximize the utility. In microeconomic demand theory, it is common to consider the monetary budget constraints since it is usually the most important.

Now consider a case that a consumer who has the monetary budget B_1. Then the combined expense for purchasing T and X should not exceed the budget B_1. The quantities t and x that a consumer can afford is dictated by

$$P_T \cdot t + P_X \cdot x \leq B_1 \tag{7}$$

where P_T and P_X are the prices of T and X. Any pair (t, x) that meets Equation 7 comprises the *feasible choice set* (or *opportunity set*) for a consumer. A consumer's final alternative can be found by solving a simple optimization problem as follows[2]:

$$\begin{gathered} \text{Maximize } U(t,x), \text{ or minimize } -U(t,x) \\ \text{Subject to: } P_T \cdot t + P_X \cdot x \leq B_1. \end{gathered} \tag{8}$$

A commodity pair is a solution to Equation 8, illustrated in Figure 20(a) where $a_1 = (t_1, x_1)$ is located on the tangent point of the budget line and the utility curve U_1. Suppose the consumer's income increases. This causes a rightward parallel shift in the budget line since the amount of budget will also increase to B_2 $(> B_1)$. The expansion of feasible choice set poses a new constraint and the consumer's choice moves to $a_2 = (t_2, x_2)$ which offers increased utility to the consumer, as shown in Figure 20(b).

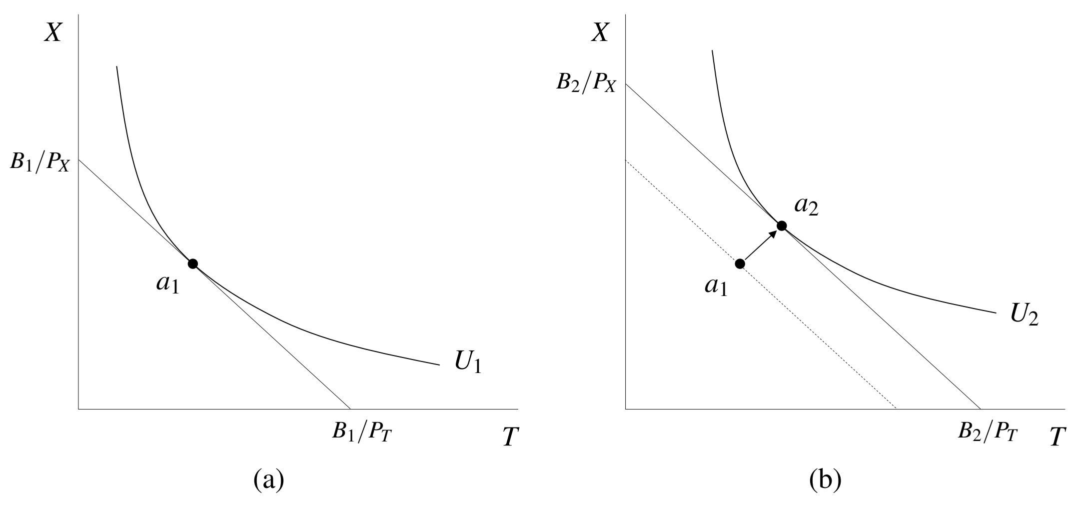


Figure 20: Income Effect on the Consumer's Choice

[2] Alternatively, Equation 8 can also be formulated as: minimize $B(t,x) = P_T \cdot t + P_X \cdot x$; subject to: $U(t,x) \geq U_1$ where U_1 is prescribed amount of utility.

The effect of changes in price of T or X, all else being constant, results in expansion or contraction of a consumer's choice set. Consider the transportation price P_T used to be p_2 but there is an incident that leads to higher transportation prices. The price is p_3 now, so the t-intercept of the budget line shift leftwards, which brings contraction of the feasible choice set. Then, a consumer's selection point also changes with lower amount of transportation with potentially less impact on all other goods as shown in Figure 21(a). If this process is repeated with different prices, a general description of relationship between transportation price P_T and transportation demand t can be obtained, shown in Figure 21(b).

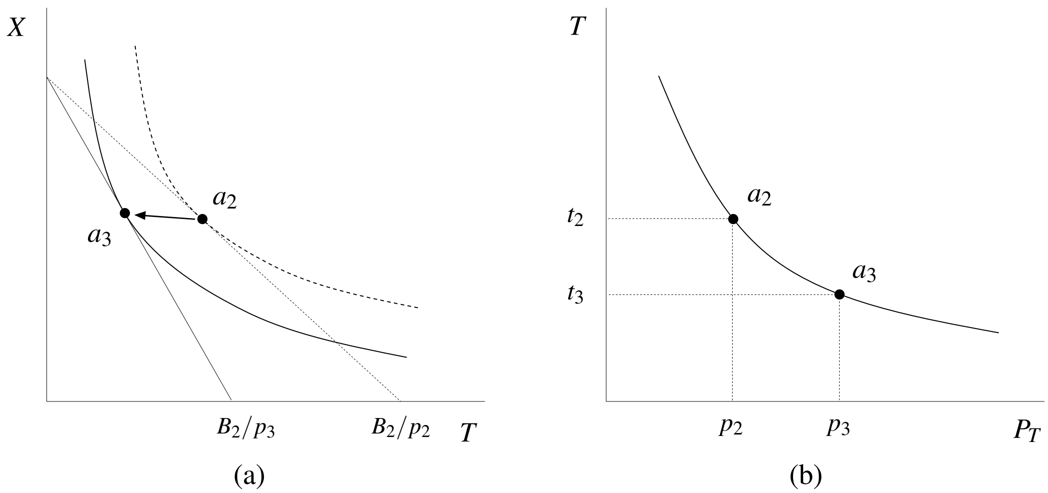


Figure 21: Generation of Transportation Demand

Suppose P_T is replaced with a vector $\overline{p}$. The vector $\overline{p}$ can be thought of as a list of generalized prices which describes 'the amount of resistance' a consumer feels such as the actual price and transportation time of T. These components of $\overline{p}$ have a negative impact on t, but one can think of some components that have a positive impact, for example, the consumer's budget and the price of other competing product X. In economics, the *elasticity* ε is frequently used to measure the consumer's responsiveness of the demand function $t(\overline{p})$ with respect to p, an component of $\overline{p}$. It is defined as the ratio of percent change in p to percent change in t. Mathematically, the elasticity is expressed as

$$\varepsilon_p^t = \frac{\frac{\partial t}{t}}{\frac{\partial p}{p}} = \frac{\partial \log t}{\partial \log p}. \tag{9}$$

A nice feature of the elasticity is facilitating comparisons of the demand to any of the affecting variables, since ε_p^t is dimensionless and independent of the units for t and p. One limitation of the concept of the elasticity is that it cannot accommodate non-cardinal variables. (e.g., high/medium/low, building a new road, technology infusion, etc.) Another weakness is that it is not applicable if more than one variables are simultaneously changed. In order to deal with these situations, the *perturbation* and the *sensitivity* are defined next.

Perturbation ($\mathcal{P}$) The change in system's state ν, from ν_0 to ν_1, denoted by $\mathcal{P}_{0,1}^{\nu}$. If the change involves only a single variable in cardinal number, the perturbation can also refer to the amount of the difference of that variable with respect to its reference (or baseline) value ν_0, i.e., $\mathcal{P}_{0,1}^{\nu} = \frac{\nu_1 - \nu_0}{\nu_0}$

Sensitivity ($\mathcal{S}$) The difference of system's dependent variable η with respect to its reference value η_0 brought by the perturbation $\mathcal{P}_{0,1}^{\nu}$, i.e., $\mathcal{S}_{\nu_0,\nu_1}^{\eta} = \frac{\eta|_{\nu=\nu_1} - \eta|_{\nu=\nu_0}}{\eta|_{\nu=\nu_0}} = \frac{\eta_1 - \eta_0}{\eta_0}$

These definitions are very generic and will be used in later chapters. It is very important to keep track of state ν_0 and state ν_1 when the sensitivity is compared to other sensitivity. The elasticity simply equals to the ratio of the sensitivity to the perturbation as ν_1 approaches ν_0, the reference or baseline state.

$$\varepsilon_{\nu}^{\eta} = \lim_{\nu_1 \to \nu_0} \frac{\mathcal{S}_{\nu_0,\nu_1}^{\eta}}{\mathcal{P}_{0,1}^{\nu}} = \lim_{\nu_1 \to \nu_0} \frac{\frac{\eta_1 - \eta_0}{\eta_0}}{\frac{\nu_1 - \nu_0}{\nu_0}} \tag{10}$$

While this equation is mathematically rigorous in regard of Equation 9, it is not generally required (or possible) to take the limit in practice; instead small perturbation—on the order of a few percent—is adequate enough. The elasticity without taking the limit is frequently used, called the *arc elasticity* e, and it can be calculated through the following equation:

$$e_{\nu}^{\eta} = \frac{\Delta \log \eta}{\Delta \log \nu} = \frac{\log \eta_1 - \log \eta_0}{\log \nu_1 - \log \nu_0} \tag{11}$$

which has an advantage since it gives the same value regardless of perturbation sequence. Still, it is very important to keep track of ν_1 and especially the reference state ν_0. It is meaningless to compare which dependent variable's effect is significant when the baseline state is not the same.

3.2.3 Trip Distribution Models

The transportation activity is essentially to link people and goods that are spatially separated. Therefore, the study of transportation demand should be implemented with distribution of trips in origins and destinations. The first step to study trip distribution is to divide physical space under investigation into mutually exclusive zones or locations. Then, the trip distribution can be conveniently specified in a matrix form. If $t_{ij}(k)$ indicates the number of trips during a given time period k between origin location i and destination location j, an n-by-n matrix $T(k) = [t_{ij}(k)]$ can be constructed where n is the number of locations. This matrix is often called trip table, trip distribution matrix or origin-destination (O-D) matrix. In general, a typical O-D matrix is not symmetric and shown in Table 3.

Table 3: Typical O-D Matrix

$$\begin{array}{cc} & \text{Destination} \\ \text{Origin} & \begin{bmatrix} t_{11}(k) & t_{12}(k) & \cdots & t_{1n}(k) \\ t_{21}(k) & t_{22}(k) & \cdots & t_{2n}(k) \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1}(k) & t_{n2}(k) & \cdots & t_{nn}(k) \end{bmatrix} \end{array}$$

From the O-D matrix, trip production and attraction of each location can be easily obtained. The trip production of i, $P_i(k)$ is the sum of row i representing the total number of trips originating from location i. The sum of column j, $A_j(k)$ is the total number of trips destined to location j and is called the trip attraction of j. The grand total $D(k)$ of the O-D matrix, called total trip demand, represents the total number of trips across all locations. The following summarizes the characteristics of the O-D matrix.

$$\begin{gathered} \sum_j t_{ij}(k) = P_i(k) \\ \sum_i t_{ij}(k) = A_j(k) \\ \sum_i P_i(k) = \sum_j A_j(k) = \sum_{i,j} t_{ij}(k) = D(k) \end{gathered} \tag{12}$$

Creating and estimating a reliable O-D matrix forms a crucial basis for the study of the transportation demand analysis. The simplest approach is to utilize the concept of growth factors.

When a priori knowledge is available, *growth factor models* are most frequently used, though some variations can be found. They basically attempt to predict the future demand by multiplying observed trips by some scaling factor. The simplest model has a constant growth factor for all $t_{ij}(k)$ but in practice each element of the O-D matrix has a corresponding growth factor, that is,

$$t_{ij}(k+1) = g_{ij}(k) \cdot t_{ij}(k) \tag{13}$$

where $(k+1)$ is an index for indicating the next time period after time period k. The concept of the Hadamard product is useful in converting Equation 13 into a matrix equation, and introduced in the followings.

Hadamard Product ($\circ$) Let $M_{m,n}$ denote the set of m-by-n matrices. The Hamadard product of $A = [a_{ij}] \in M_{m,n}$ and $B = [b_{ij}] \in M_{m,n}$ is $A \circ B \equiv [a_{ij} \cdot b_{ij}] \in M_{m,n}$.

Hadamard Exponent Let $A = [a_{ij}] \in M_{m,n}$ then the Hadamard exponent of the matrix A in the order of m is denoted by $A^{[m]}$ and its component is obtained by calculating ${a_{ij}}^m$.

Using the above definitions, Equation 13 is equivalent to the following matrix equation: $T(k+1) = G(k) \circ T(k)$ where $G(k) = [g_{ij}(k)]$. Further, if $G(k)$ is a constant matrix with respect to k, the trip distribution after k periods can be easily computed with the initial trip distribution $T(0)$ as follows:

$$T(k) = G^{[k]} \circ T(0). \tag{14}$$

As seen, growth factor models are a very intuitive and simple approach for estimating the O-D matrix. They are particularly useful for forecasting of near-term demands especially when a clear trend is identified in the past data. However, growth factor models *per se* do not have the capability to account for changes in the system that in turn affects growth factors. Hence, various trip distribution models have been suggested to properly reflect the effects of influencing variables into the O-D matrix. Kanafani (1983) observes that most of trip distribution models have a general form given as follows:

$$t_{ij} = \mu_i \nu_j C_{ij} \tag{15}$$

where μ_i and ν_j are functions of the socioeconomic characteristics of locations i and j, respectively; and C_{ij} is a general function representing the impedance to travel between i and j. He further suggests that trip distribution models can be classified into three major categories: *demand models*, *choice models*, and *spatial interaction models*. The first two groups are rooted in utility theory while the last one is derived from analogies of physical laws.

The most basic approach to modeling trip distribution is to construct a *demand model* for trips between an origin and a destination. This model can be an individual or aggregate function, depending on the information available for calibration. It is derived by using the utility maximization principle discussed previously. Assume a utility function for an individual living location i is represented as follows: $U = U(x_{i1}, x_{i2}, \cdots, x_{in})$ where x_{ij} is the number of trips from the origin i to destination j. Associated with these trips, total travel cost are given by $C = \sum_j c_{ij} x_{ij}$ where c_{ij} is a unit travel cost from i to j. Maximizing U under the constraint C can be solved by constructing a Lagrangian: $L = U - \lambda C$ where λ is a Lagrangian multiplier. Then derivatives of L mush vanish so that U has optima:

$$\frac{\partial U}{\partial x_{ij}} = \lambda c_{ij}, \text{ for all } j \tag{16}$$

which states that constant is the ratio of the marginal utility to the trip cost for all locations. In order to solve the Lagrangian equations, however, an analytic form for U should be given. A common utility function to use is a linear combination of the constant elasticity utility functions (White 1978) which can be expressed as

$$U = \sum_j \alpha_{ij} {x_{ij}}^{\rho_{ij}} \tag{17}$$

where α_{ij} and ρ_{ij} are parameters and $0 < \rho_{ij} < 1$ to satisfy the law of diminishing marginal utility. (i.e., $\frac{\partial^2 U}{\partial {x_{ij}}^2} < 0$) Substitution this into the Lagrange equations makes x_{ij} to be solved for. The result is

$$x_{ij} = \left(\frac{\alpha_{ij}\rho_{ij}}{\lambda c_{ij}}\right)^{\frac{1}{1-\rho_{ij}}} \tag{18}$$

from which a general structure can be observed as follows:

$$x_{ij} = \frac{f(i,j)}{c_{ij}{}^{\gamma_{ij}}} \tag{19}$$

where $\gamma_{ij} = \frac{1}{1-\rho_{ij}} > 1$. The aggregate trip distribution demand t_{ij} is the sum of x_{ij}'s for all individuals in location i and the assumption is that t_{ij} follow the general structure appearing in Equation 19. So the total trip distribution is given as

$$t_{ij} = S_{ij} \cdot c_{ij}{}^{-\gamma} \;\; (\gamma > 1). \tag{20}$$

This result is interpreted as S_{ij} and c_{ij} represent a generalized trip attraction factor and a generalized trip impedance factor between locations i and j, respectively. S_{ij} is often replaced with the product of two functions, each is a function of some socioeconomic characteristics of the corresponding location. Note that the trip impedance c_{ij} can be trip cost, trip time or a combination of both.

While demand models are derived from individual's utility, *choice models* assume that each location possesses some utility that represents the amount of its attractiveness for travelers. Under this assumption, the basic idea is that one can get the probability of location j being chosen as trip destination from location i, $\pi(i,j)$, by utilizing the probabilistic choice theory which is to be discussed in the following section. If the multinomial logit model is adopted (again, to be discussed in the following section), $\pi(i,j)$ is given by

$$\pi(i,j) = \frac{e^{V(i,j)}}{\sum_{k=1}^{n} e^{V(i,k)}} \tag{21}$$

where n is the total number of locations and $V(i,j)$ is called the deterministic utility for trips from i to j which includes a generalized travel cost component and some measures of attractiveness. If the trip production P_i is known, the trip distribution is computed by the following equation: $t_{ij} = \pi(i,j) \cdot P_i$.

Spatial interaction models are derived from analogies with the physical laws. The *gravity model*, derived from Newton's law of Gravitation, is the most common form currently in use and has the longest history. According to Kanafani (1983), the earliest observation that the magnitude of population movements between cites is similar to the gravitational pull between masses was suggested by Ravenstein in 1885. The basic structure is in line with the law as follows:

$$t_{ij} = k_{ij} \frac{\mu_i \nu_j}{{c_{ij}}^{\gamma}} \tag{22}$$

where μ_i and ν_j indicates a function of characteristics of origin i and destination j such as P_i and A_j, respectively and c_{ij} is representing impedance between i and j. Keeping strict gravitational framework would take $k_{ij} = k$ and ${c_{ij}}^{\gamma} = {d_{ij}}^2$ where d_{ij} represents distance between locations i and j.

Another approach based on physical analogy is the *entropy model* introduced by Wilson (1967). In this approach, the entropy of the O-D matrix T is given by

$$S(T) = \frac{\left\{\sum_{ij} t_{ij}\right\}!}{\prod_{ij} \left\{t_{ij}!\right\}} \tag{23}$$

which has the lowest value 1 when all trips are aggregated on a particular cell in the O-D matrix. Just like physical cases, this pseudo entropy is maximized in order to achieve system's equilibrium. Consider that the optimization problem is subject to the following constraints: $\sum_j t_{ij} = P_i$, $\sum_i t_{ij} = A_j$, and $\sum_{ij} t_{ij} c_{ij} = B$. The first two constraints come from Equation 12 and the last constraint mandates that the total transportation cost expended is limited to a budget B. Formulating a lagrangian equation and differentiating with respect to t_{ij} result in the following expression.

$$t_{ij} = k_{ij} P_i A_j e^{-\lambda c_{ij}} \tag{24}$$

where k_{ij} and λ are parameters for model calibration. For detailed derivation, refer to Bierlaire (1996).

Various trip distribution models discussed so far are subject to diverse model assumptions and thus have generated a wide variety of modified forms. Despite the variety, however, all models share the same objective. A good model should be properly responsive to socioeconomic factors and other transportation related conditions, eventually aiming at forecasting the changes in transportation demand.

3.2.4 Natural vs. Induced Growth in Transportation Demand

It would not be exaggeration to say that expansion of transportation volume over time has stressed all the related aspects of the transportation system. The issue is that the growth is not a simple phenomenon. For example, a city administration may endorse a road expansion project to relieve traffic congestion due to population growth. Upon completion, this would obviously help to reduce down the congestion for short term, however, as time goes by, this capacity improvement may bring on extra traffic that makes the expanded road packed again sooner than expected. (See Figure 22)

In a nutshell, the policy maker's dilemma emanates from the fact that the growth is a superimposed effect from *exogenous* and *endogenous* factors (and their coupled power).

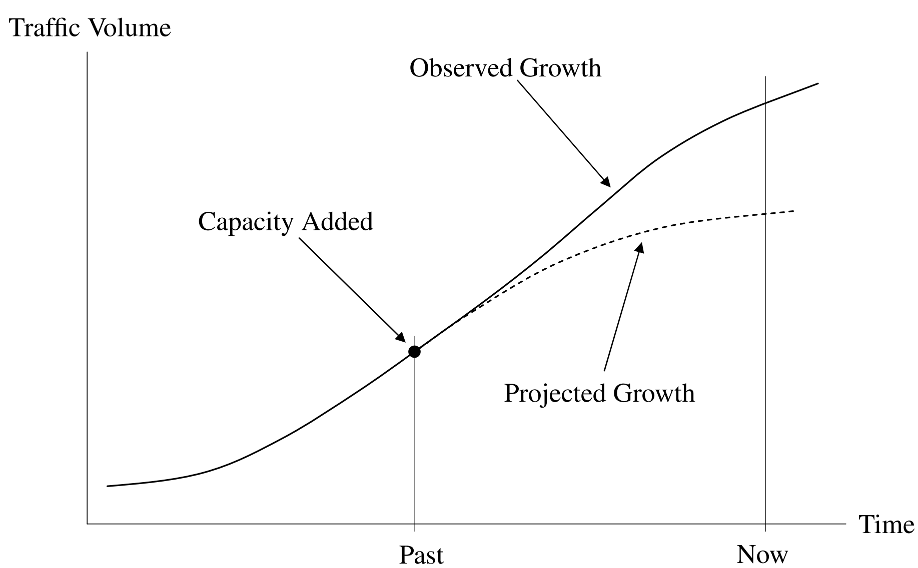


Figure 22: Effect of Improved Capacity on Traffic Volume [Modified from Litman (2001, Figure 1)]

Exogenous factors are derived from population growth, income increase and other socioeconomic changes. In contrast, endogenous factors relate to alterations in transportation supply such as investment in system capacity enhancement or infusion of a new mobility system. In this dissertation, as Lee, Klein & Camus (1999) suggests, the term *induced travel* is broadly defined as the net increase in travel demand derived from endogenous factors. The notion of induced travel was introduced as early as Downs (1962) in order to theoretically justify this phenomenon. (Boarnet & Chalermpong 2000) Since then, theoretical and empirical studies on induced travel have been increasingly recognized in transportation field, mostly in highway/ground transportation sector. The theoretical basis of induced growth in transportation volume can be explained within the supply and demand framework as shown in Figure 23. The demand curve (D_0) at a given time period shows the inverse relationship between the amount of transportation (T) and the generalized cost of transportation (P_T). The supply curve (S_0) has the opposite relationships as increase in traffic volume incurs congestion that causes additional travel time. The two curves D_0 and S_0 intersect at point E_{00}, indicating the equilibrium state of $T = t_{00}$ and $P_T = p_{00}$. In

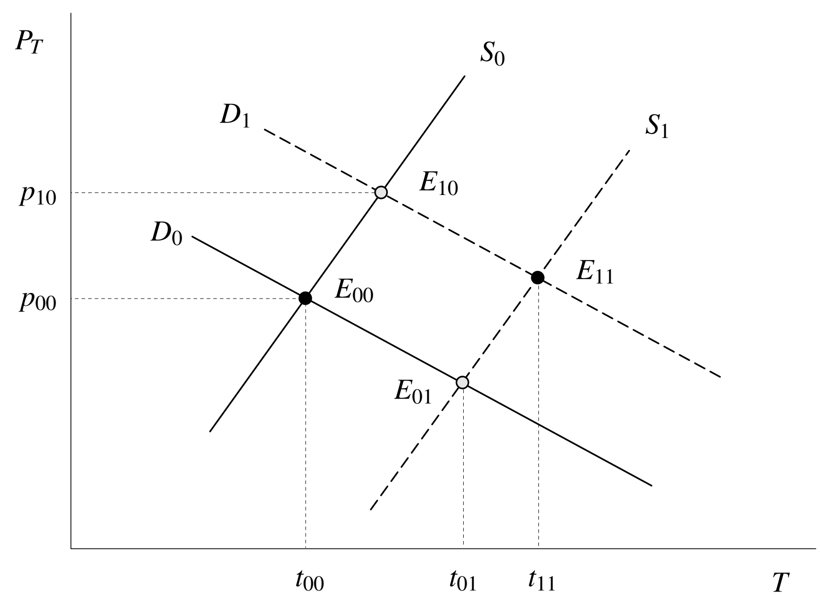


Figure 23: Combined Effects of Induced and Natural Growth [Modified from Noland (2001, Fig. 2)]

general, the transportation demand grows over time naturally. This natural growth in demand makes rightward shift of the demand curve. If nothing is done in response to the growing demand, travelers will pay higher cost (p_{10}) as indicated at point E_{10} because of more sever congestion. This is not desirable; so what if capacity is upgraded? Point E_{01} marks, in theory, a new equilibrium state by increasing the supply only. One can see that additional supply, despite all else being constant, induces additional amount of transportation ($t_{01} - t_{00}$). This situation is, however, purely hypothetical for the transition towards the new equilibrium point cannot occur instantaneously. Instead, the equilibrium state moves gradually while the transportation system accommodates the natural growth as well. An investigator will only observe point E_{11} in the long-run. Both points E_{01} and E_{10} merely lie in the realm of theory, whereas forecasting point E_{10} is more doable when the discernable trend exists. This is why there have existed ontological debates over the induced travel for many years; induced growth is hardly observed in an independent separable way, always confounding with natural growth. But, many influential studies have instantiated the cases of induced travel and argued that the issue of induced travel should be accounted for in the transportation planning phase since its effect would be substantial.[3]

The impact of induced travel is usually measured as increase in vehicle-mile traveled (VMT) since induced travel or induced growth is manifested in the terms of new and longer trips—not only count of trip but also the distance of trip should be considered to indicate induced travel as Noland & Lem (2002) stated. The increased ability for these new and longer trips brings effects on other seemingly unrelated factors as people are adapting themselves to the improved transportation system. For example, a person who usually commutes by bus can consider using an automobile (or even buy one) for commuting purpose and a family may want to move out to suburban area. Similarly, if an efficient high-speed train system can be reached within 15 minutes, a traveler may not fly for long-distance trips. The significance of induced travel lies here. One has to keep in mind that those long-run

[3]See SACTRA (1994), TRB (1995) and Noland & Lem (2002) for details.

effects—on travel mode, vehicle ownership, residential and business location—will recursively configure determinants of transportation demand.

In summation, induced travel is a real phenomenon—in fact, a special case of *rebound effects*[4] cited by many economist—that needs to be taken into account in the planning phase. Although induced travel can result from any transportation system, the evidences and studies are usually documented in the ground transportation literature. It is quite rare in the aerospace field to account for the effect of induced travel and its implications.

3.2.5 Summary and Issues

Starting from microeconomics, a basic foundation for transportation demand analysis has been laid. It should be recognized at this stage the distinction between usual services and goods and transportation. While most services and goods have their own utilities, transportation itself does not possess utility. The role of transportation is to offer opportunity for other economic activities. If there is no economic relation between city A and B, for instance, then no substantial changes will be made in transportation activity regardless of income or price variability. Briefly stated, transportation demand is a *derived demand*. (Kanafani 1983, p. 14) All the discussion in this section hinges on a tacit assumption that socioeconomic changes have direct effect on transportation demand. Having accepted that, this body of theories is solid; nevertheless, it still does not necessarily generate a practical tool to solve the real world problems due to its abstract nature.

Restrictions immediately appear from the outset—little guidance is given concerning the explicit form of the utility function. The function $U(t,x)$ in §3.2.1 cannot readily be characterized. Structuring both abstract arguments t and x is already problematic and even if this task is done finding an analytic form for U is not warranted. For example, it is not explicit how utility depends on the traveler's income. For this reason, most quantitative

[4] Rebound effects refer to increased consumption that results from actions that increase efficiency and reduce consumer costs (Alexander, 1997). For example, suppose a new technology is invented that can save 50% of energy for house heating. Does this technology actually bring 50% of energy saving for a household? The answer is generally negative because people would set a little bit higher temperature than before.

studies regarding transportation demand are mainly found in economics relying on statistical methods, not via the utility function. Regressions on past data focus on discovering the effects of influencing variables, typically hinging on elasticity values. Subsequent inference anchored in the analysis results constitutes the major axis of many relevant studies. Although this methodology has facilitated improved decision-making in the transportation sector, these models of transportation demand involve many sequel challenges.

Clearly, the sheer number of external determinants influences transportation demand. Let alone familiar economic factors such as income and gasoline price, demographic changes including population growth and migration to (or from) urban area, along with other aspects like women's participation in nation's economy, also can fundamentally reconfigure the demand profile. Identifying the key variables and responses is a difficult but important task. Moreover, separating the net effect of a variable of interest is critical by far to reach a meaningful conclusion. Using simple parametric models, however, investigators fail to avoid the confounding issue. In addition, advancements in transportation system—resulted from capacity enhancement or technology infusion—will affect the transportation volume over time. The prediction capability of the existing models is vague to capture this feedback effect caused by internal factors, especially one from technology infusion.

The reviewed theoretical works to date are limited in their accounting for these important issues. This is not because they have a defect. Rather, this is a consequence of the profound fact that transportation demands are only a small, reflected portion of "living things" in our society. It is very difficult to fully capture "living" nature of the entire system, using analytical approaches and statistical approaches by examining handful proxy variables. All theoretical frameworks discussed above are for explaining how the results fit into a larger frame of belief, but they are not complete for forecasting or predicting the future demand, and even the entire system.

3.3 Probabilistic Choice Theory

Evaluation of an alternative is essential for any decision-making process. This theme overarches practically all applied sciences. Engineers and system designers have rested on multi-attribute decision-making techniques augmented with physics-based analysis tools. In social sciences, the choice problem needs a different treatment since it is commonly observed that there exist inconsistent and non-transitive preferences in human behaviors. The development of probabilistic choice theory arose from the need to explain these behavioral inconsistencies. The origin of this theory is in mathematical psychology as appeared in Thurston (1927). Since then, probabilistic choice theory has also been applied in various fields. Its importance is still growing, enunciated by the 2000 Nobel Prize in economics.[5] This section is dedicated to introduce the theory that has a great potential for engineering problems as well.

3.3.1 Random Utility

Envision a choice set having M alternatives and each alternative comprising n attributes. Then an alternative X_m in the choice set can be thought of as a vector in n-dimensional attribute space, i.e., $X_m = [x_{m1}, x_{m2}, \cdots, x_{mn}]^T$. As discussed previously, utility theory postulates that an individual chooses the alternative that offers the highest utility (or the lowest disutility). If the utility for the individual for alternative X_m is given by a utility function $U(X_m)$, the task of maximizing utility under given constraints yields the best alternative. This problem could be tackled from a multi-criteria decision-making (MCDM) perspective, specifically using multi-attribute decision-making (MADM) techniques. Appendix B details the theory of MADM along with its counterpart, multi-criteria optimization (MCO).

The issue is that $U(X_m)$ can be other than a simple, deterministic real function. This can be interpreted as follows: First, the measured value of an attribute may include noise. There is always non-zero likelihood to have a difference between actual value and observed value.

[5] Refer to Manski (2001) for detailed description about Daniel McFadden's contribution.

Second, it is possible for an analyst to omit some attributes into consideration. Perfect understanding of the problem is nearly impossible so this is not an unusual situation. Third, for a certain attribute, it is difficult to measure its value. For example, psychological factors cannot be easily quantified, especially by outside observers. Next, decision-making process is not rational or deterministic in nature, which may seem to violate the assumptions of neoclassical economics. But it is not absolutely contradictory when the term rationality is softened to *bounded rationality*.[6]

In summary, the decision-maker has incomplete or unobservable information and makes a decision with his/her "bounded" rationality, which is a reasonable speculation on human behaviors. Random utility model recognizes this and considers an uncertainty term for an alternative to tackle the issue. Hence, the perceived utility for alternative m is given by

$$U_m = V_m + \varepsilon_m \tag{25}$$

where $V_m = V(X_m)$ represents a deterministic (or systematic) utility term, and $\varepsilon_m = \varepsilon(X_m)$ is a random (or stochastic) utility term that contains uncertain, immeasurable factors. While the uncertainty prevents an analyst from pinpointing which alternative is best, a probabilistic formulation can dictate how likely an alternative is preferred to others. From this perspective, the choice probability of alternative X_m is equal to the probability that the corresponding utility (U_m) is greater than or equal to all the utilities of other alternatives in the choice set, expressed as follows:

$$\pi(m) = Prob\,[U_i \leq U_m, \forall i \neq m] \tag{26}$$

where $\pi(m)$ denotes the probability of alternative X_m being chosen. This idea is perfectly in harmony with utility theory. Equation 26 is further developed as follows:

$$\pi(m) \quad = \quad Prob\,[V_i + \varepsilon_i \leq V_m + \varepsilon_m, \forall i \neq m]$$

[6]Simon (1957, p. 198), the 1978 Nobel laureate in Economics, states the principle of bounded rationality as follows: the capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world-or even for a reasonable approximation to such objective rationality.

$$= \quad Prob\,[\varepsilon_i - \varepsilon_m \leq V_m - V_i, \forall i \neq m]. \tag{27}$$

This formulation offers a foundation in constructing so-called discrete choice models. Depending on the number of alternatives and the assumption regarding ε, the complexity of the choice models will vary.

3.3.2 Binary Choice Models

The basic idea is pursued further by considering the special case where the choice set contains exactly two alternatives. Such situations lead to what are termed binary choice models or dichotomous choice models.

3.3.2.1 Linear Model

Assume the decision-maker considers two alternatives whose utilities are: $U_1 = V_1 + \varepsilon_1$ and $U_2 = V_2 + \varepsilon_2$. The choice probability $\pi(1)$ can be computed according to Equation 27, and $\pi(2)$ is its trivial consequence as shown below.

$$\begin{aligned} \pi(1) &= Prob\,[\varepsilon_2 - \varepsilon_1 \leq V_1 - V_2] \\ \pi(2) &= 1 - \pi(1). \end{aligned} \tag{28}$$

As a simple illustration, let $\varepsilon = \varepsilon_2 - \varepsilon_1$ be uniformly distributed between two fixed values $-L$ and L. The probability distribution function is portrayed in Figure 24. It is obvious that $\pi(1)$ equals to the area of hatched portion in Figure 24 since alternative X_1 is chosen when $\varepsilon \leq V_1 - V_2$. Three cases are generated depending on the position of $V_1 - V_2$ on the x-axis and the corresponding $\pi(1)$ is given by

$$\pi(1) \quad = \quad \begin{cases} 0 & \text{if} \quad V_1 - V_2 < -L \\ \frac{V_1 - V_2 + L}{2L} & \text{for} \quad -L \leq V_1 - V_2 \leq L \\ 1 & \text{if} \quad V_1 - V_2 > L \end{cases} \tag{29}$$

The choice probability $\pi(1)$ is linearly proportional to the amount of $V_1 - V_2$ as the cumulative distribution function (CDF) of ε is linear. The linear model is rarely used in real applications due to the discontinuity at $\varepsilon = \pm L$.

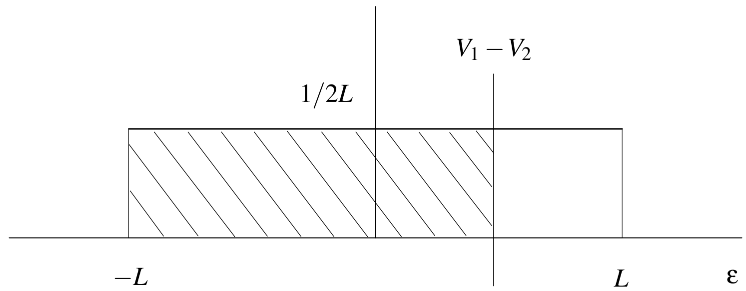


Figure 24: Linear Model

3.3.2.2 *Probit Model*

A fundamental discrete choice model can be generated with the assumption that the random utilities are normally distributed—called the Normal Probability Unit model or Probit model for short. The normal distribution is a very important class of statistical distributions since the "normality" arises naturally in many physical, biological, and social measurement situations. Its probability density function f_N is given by

$$f_N(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{30}$$

where μ and σ are parameters for mean and standard deviation, respectively. Now suppose that ε_1 and ε_2 are normally distributed with zero mean, variances ${\sigma_1}^2$ and ${\sigma_2}^2$, respectively. Let these relations be denoted by $\varepsilon_1 \sim N(0,{\sigma_1}^2)$ and $\varepsilon_2 \sim N(0,{\sigma_2}^2)$. Then, the difference of two random variables also follows the normal distribution, i.e., $\varepsilon = \varepsilon_2 - \varepsilon_1 \sim N(0,\sigma^2)$ where σ^2 is obtained from the following equation:

$$\begin{aligned} \sigma^2 &= Var(\varepsilon_2 - \varepsilon_1) = Var(\varepsilon_1) + Var(\varepsilon_2) - 2 \cdot Cov(\varepsilon_1,\varepsilon_2) \\ &= {\sigma_1}^2 + {\sigma_2}^2 - 2\sigma_{12}. \end{aligned} \tag{31}$$

Note that $Cov(\varepsilon_1,\varepsilon_2) = \sigma_{12}$ denotes the covariance of two random variables ε_1 and ε_2. As in the linear model, the choice probability $\pi(1)$ corresponds to the hatched area in Figure 25 and is given by

$$\begin{aligned} \pi(1) &= \int_{-\infty}^{V_1-V_2} f_N(\varepsilon;0,\sigma)\,d\varepsilon = \int_{-\infty}^{\frac{V_1-V_2}{\sigma}} f_N(x;0,1)\,dx \\ &= \int_{-\infty}^{\frac{V_1-V_2}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2}\,dx = F_N\left(\frac{V_1-V_2}{\sigma}\right) \end{aligned} \tag{32}$$

where $F_N(\cdot)$ denotes the standard cumulative normal distribution function. The Central Limit Theorem states that the sum of a large number of independent random variables follows the normality. Thus, the probit model is rooted in a solid theoretical ground and it has been applied to many practical problems. The downside of the probit model is that $F_N(\cdot)$ is not given as a closed form which causes serious computational problems as a large number of alternatives are considered. This motivated a search for a choice model that is easy-to-use and similar to the probit model.

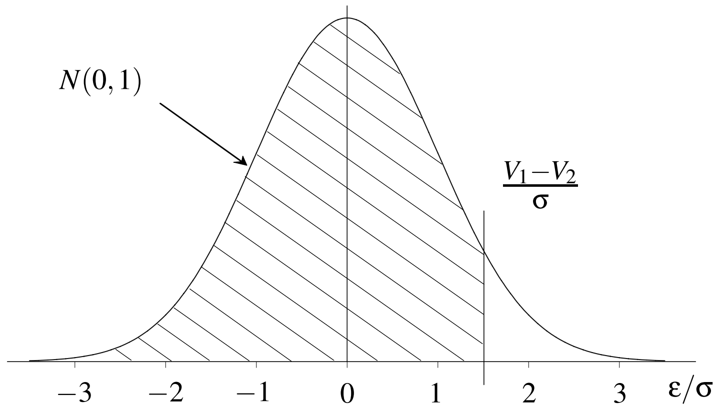


Figure 25: Probit Model

3.3.2.3 Logit Model

The most widely used model in practical applications is the Logistic Probability Unit model, or Logit model. This model is obtained by using the Gumbel distribution whose probability density function f_G and cumulative distribution function F_G are given by

$$\begin{aligned} f_G(\varepsilon;\lambda,\gamma) &= \lambda e^{-\lambda(\varepsilon-\gamma)} e^{-e^{-\lambda(\varepsilon-\gamma)}} \\ F_G(\varepsilon;\lambda,\gamma) &= e^{-e^{-\lambda(\varepsilon-\gamma)}} \end{aligned} \tag{33}$$

where $\lambda > 0$ is a scale parameter and γ is a location parameter. If both ε_1 and ε_2 are independent and identically Gumbel distributed, then $\varepsilon = \varepsilon_2 - \varepsilon_1$ follows the Logistic distribution with the same scale parameter λ and location parameter vanished.

$$f_L(\varepsilon;\lambda) = \frac{\lambda e^{-\lambda\varepsilon}}{(1+e^{-\lambda\varepsilon})^2} \tag{34}$$

The nicest property of f_L is that its cumulative distribution function is the logistic equation which has a simple, closed form as follows:

$$F_L(\varepsilon;\lambda) = \frac{1}{1+e^{-\lambda\varepsilon}} \tag{35}$$

which can be readily confirmed by verifying $\frac{d}{d\varepsilon}F_L(\varepsilon;\lambda) = f_L(\varepsilon;\lambda)$. Therefore, the choice probability $\pi(1)$ can be computed bypassing the integration procedure, as shown below.

$$\pi(1) = \int_{-\infty}^{V_1-V_2} f(\varepsilon;\lambda)\,d\varepsilon = F_L(V_1-V_2;\lambda) = \frac{1}{1+e^{-\lambda(V_1-V_2)}} \tag{36}$$

Let the scale parameter λ set 1 for simplicity. Then the choice probabilities of two alternatives are:

$$\pi(1) = \frac{e^{V_1}}{e^{V_1}+e^{V_2}} \tag{37a}$$

$$\pi(2) = \frac{e^{V_2}}{e^{V_1}+e^{V_2}} \tag{37b}$$

The use of the logistic equation F_L offers not only computational convenience but also another advantage. A plot of F_N and F_L reveals that F_L is a close approximation of F_N as illustrated in Figure 26. The computational advantage together with the close approximation to the probit model have promoted the use of the logit model.

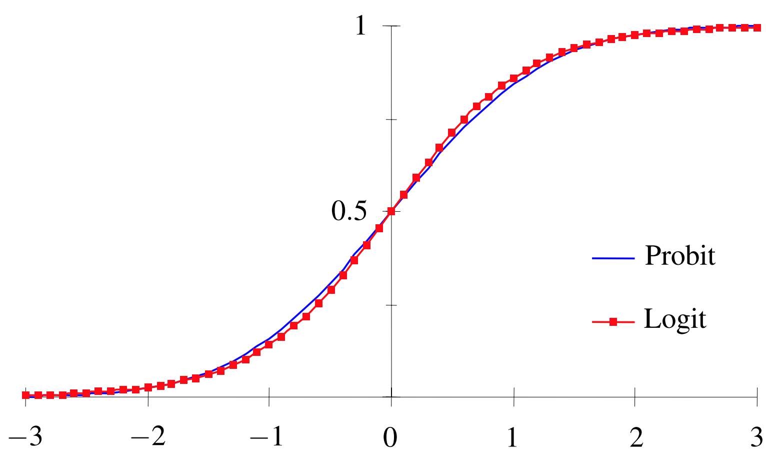


Figure 26: Comparison of Logit and Probit

3.3.3 Multinomial Choice Models

The next discussion treats a more generalized case when more than two alternatives are considered, referred to as a multinomial choice problem.

3.3.3.1 Multinomial Logit Model

The aforementioned logit model can conveniently be extended for multinomial cases. Recall that the decision-maker considers M alternatives and the utility for alternative X_i is represented as $U_i = V_i + \varepsilon_i$. Assuming that the random component ε_i is independent and identically Gumble distributed yields so-called multinomial logit (MNL) model which has the following form:

$$\pi(k) = \frac{e^{V_k}}{\sum_{i=1}^{M} e^{V_i}} \tag{38}$$

which can be inferred as a natural extension of Equation 37. The choice probability $\pi(k)$ is computed with this simple algebraic relation once the deterministic utilities are given. The derivation of Equation 38 is given below.

Proof

Assume that the random component ε_i is independent and identically Gumble distributed. For convenience of proof, set the scale parameter to 1 and the location parameter to 0. This can be done without loss of generality by scaling U_i, i.e., $U_i^* = U_i/\lambda + \gamma$. From Equation 33 the probability density function f_G and the cumulative distribution function F_G are depend on ε_i only and given by

$$f_G(\varepsilon_i) = e^{-\varepsilon_i} \exp(-e^{-\varepsilon_i}) \tag{39a}$$

$$F_G(\varepsilon_i) = \exp(-e^{-\varepsilon_i}) \tag{39b}$$

Now consider the choice probability $\pi(1)$. From Equation 27, $\pi(1)$ is equivalent to

$$\pi(1) \quad = \quad Prob\,[\varepsilon_i \leq V_1 - V_i + \varepsilon_1,\ \forall i \neq 1] \tag{40}$$

Noting that ε_1 is not deterministic entity, Equation 40 is expressed as

$$\pi(1) \quad = \quad \int_{-\infty}^{\infty} \underbrace{Prob\,[\varepsilon_i \leq V_1 - V_i + \varepsilon_1,\ \forall i \neq 1\,|\,\varepsilon_1]}_{(\star)} f_G(\varepsilon_1)\, d\varepsilon_1 \tag{41}$$

where $Prob[A|B]$ denotes the conditional probability. Let $\theta(i)$ denote the event when the expression $[\varepsilon_i \leq V_1 - V_i + \varepsilon_1]$ is satisfied. Since each event $\theta(i)$ is independent, the term $(\bigstar)$ inside the integral in the Equation 41 can be rewritten as

$$\begin{aligned}
Prob[\theta(i),\ \forall i \neq 1\,|\,\varepsilon_1] &= Prob[\theta(2) \cap \theta(3) \cap \cdots \cap \theta(M)\,|\,\varepsilon_1] \\
&= \prod_{i=2}^{M} Prob[\theta(i)\,|\,\varepsilon_1]. \qquad (42)
\end{aligned}$$

Focussing on the probability of the particular event $[\theta(i)\,|\,\varepsilon_1]$ gives the following expression.

$$\begin{aligned}
Prob[\theta(i)\,|\,\varepsilon_1] &= Prob[\varepsilon_i \leq V_1 - V_i + \varepsilon_1\,|\,\varepsilon_1] \\
&= \int_{-\infty}^{V_1 - V_i + \varepsilon_1} f_G(\varepsilon_i)\, d\varepsilon_i = F_G(V_1 - V_i + \varepsilon_1).
\end{aligned}$$

Hence, Equation 42 becomes

$$\begin{aligned}
\prod_{i=2}^{M} Prob[\theta(i)\,|\,\varepsilon_1] &= \prod_{i=2}^{M} F_G(V_1 - V_i + \varepsilon_1) \\
&= \prod_{i=2}^{M} \exp\left\{-e^{-(V_1 - V_i + \varepsilon_1)}\right\} = \exp\left\{\sum_{i=2}^{M} -e^{-(V_1 - V_i + \varepsilon_1)}\right\} \qquad (43)
\end{aligned}$$

using Equation 39b. Substituting Equations 39a and 43 into Equation 41, and collecting terms in the exponent yields

$$\begin{aligned}
\pi(1) &= \int_{-\infty}^{\infty} \exp\left\{\sum_{i=2}^{M} -e^{-(V_1 - V_i + \varepsilon_1)}\right\} e^{-\varepsilon_1} \exp(-e^{-\varepsilon_1})\, d\varepsilon_1 \\
&= \int_{-\infty}^{\infty} \exp\left[-e^{-\varepsilon_1}\left\{1 + \sum_{i=2}^{M} e^{-(V_1 - V_i)}\right\}\right] e^{-\varepsilon_1}\, d\varepsilon_1 \\
&\quad \langle \text{ Let } t = e^{-\varepsilon_1} \text{ and } \Theta = 1 + \textstyle\sum_{i=2}^{M} e^{-(V_1 - V_i)} \ \rangle \\
&= \int_{0}^{\infty} \exp(-t\Theta)\, dt = \frac{-1}{\Theta} \exp(-t\Theta)\Big|_{0}^{\infty} = \frac{1}{\Theta} \\
&= \frac{1}{1 + \sum_{i=2}^{M} e^{-(V_1 - V_i)}} = \frac{e^{V_1}}{\sum_{i=1}^{M} e^{V_i}}
\end{aligned}$$

Therefore, $\pi(k)$ is finally represented as in Equation 38 since the choice probability is not affected by the order of the alternatives. ■

As seen, the MNL model provides the simple, closed-form equation in computing the choice probabilities that are easy to use and intuitively comprehendible—it is not surprising that the MNL model and its variants have been the most widely used in practice. The

MNL model has very important property, frequently referred as independence of irrelevant alternatives (IIA). From Equation 38, the ratio of the choice probabilities of two distinct alternatives is given by

$$\frac{\pi(j)}{\pi(k)} = e^{V_j - V_k} \tag{44}$$

which has nothing to do with on any alternatives other than alternatives j and k. This property holds in presence of new alternatives as well. Let $\pi'(k)$ be the choice probability of alternative X_k when alternatives $X_{M+1}, X_{M+2}, \cdots, X_{M'}$ are added in the choice set. Still, the ratio is the same because

$$\frac{\pi'(j)}{\pi'(k)} = \frac{e^{V_j} / \sum_{i=1}^{M'} e^{V_i}}{e^{V_k} / \sum_{i=1}^{M'} e^{V_i}} = e^{V_j - V_k}. \tag{45}$$

This IIA property comes from the most crucial assumption of logit model: the random component ε_i is independent and identically Gumble distributed. This assumption may not be valid in some situations, which implies the MNL model should be used with care. First, it is possible that ε_1 and ε_2 are correlated each other. The most famous case is the "red bus / blue bus" paradox.[7] Next, for instance, ε_1 can have larger scale parameter than ε_2 does. In other words, the MNL model cannot treat variations of random utilities.

The presence of the IIA property *per se* does not necessarily bound a choice model to be either good or bad since some choice problems show that the property is a realistic representation. The important point is that the modeler should consider alternative methods if the IIA property of given problems is violated or suspicious.

3.3.3.2 Nested Logit Model

One solution is to consider a multinomial choice context as one with hierarchic structure. As an example, Figure 27 describes choice structure of transportation modes for a commuter where the circled modes constitute the elements of the final choice set. The modeler

[7] Assume that a car and a red bus each with its own choice probability are in a commuter's choice set. All else unchanged, let a blue bus be included in the choice set, which exactly has the same service attributes as the red bus except for its exterior color. This would not change anything but a choice model having the IIA property will predict that the choice probability of taking the car decreases.

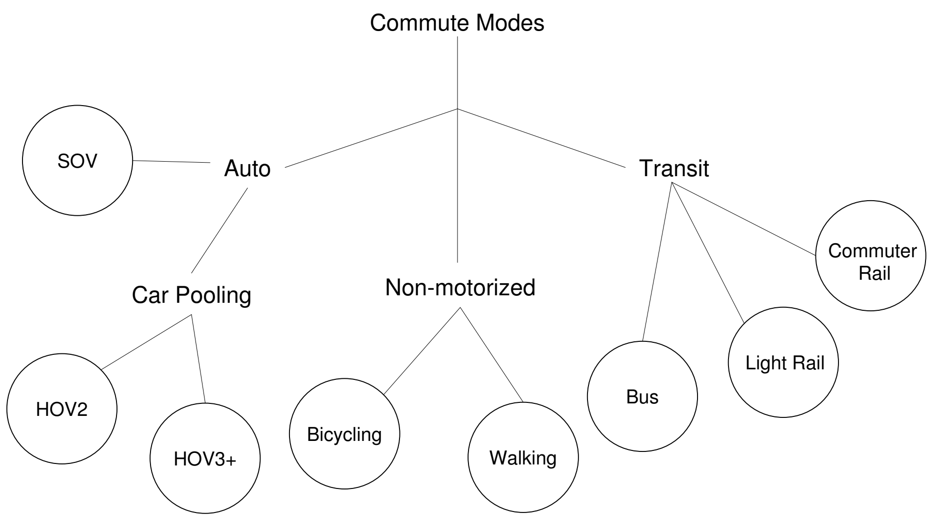


Figure 27: Nested Logit Model Sample [Source: Beimborn & Kennedy p. 24]

can group alternatives at different levels as long as the grouped alternatives share common attributes. The resultant model is called the nested logit model which is an easy and feasible way to resolve the IIA issue of the MNL model. Properly constructed, the IIA property holds within each nest and does not hold in general for alternatives in different nests. For example, suppose the choice problem has two levels and alternative X_m belongs to nest N_k among $N_1, N_2, \cdots, N_K$. Then the choice probability of alternative X_m can be obtained by the following decomposition:

$$\pi(m) = \pi(N_k) \cdot \pi(m|N_k) \tag{46}$$

where $\pi(N_k)$ denote the choice probability of nest N_k and $\pi(m|N_k)$ is the conditional probability of alternative X_m being chosen among the alternatives inside of nest N_k. The simplest MNL model takes the following expressions for $\pi(m|N_k)$ and $\pi(N_k)$.

$$\pi(N_k) = \frac{\exp(\tau_k I_k)}{\sum_m \exp(\tau_m I_m)} \tag{47a}$$

$$\pi(m|N_k) = \frac{\exp(\frac{1}{\tau_k} V_m)}{\exp(I_k)} \tag{47b}$$

where the τ_k are called dissimilarity parameters with the inclusive values I_k defined as

$$I_k = \ln \sum_{l \in N_k} \exp\left(\frac{1}{\tau_k} V_l\right). \tag{48}$$

3.3.3.3 Multinomial Probit Model

Multinomial probit is an extension of binary probit models to more than two alternatives. If Equation 25 is rewritten in a vector form, then the utility vector $\underline{U} = [U_1, U_2, \cdots, U_M]^T$ can be given by

$$\underline{U} = \underline{V} + \underline{\varepsilon} \tag{49}$$

where $\underline{V} = [V_1, V_2, \cdots, V_M]^T$ and $\underline{\varepsilon} = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_M]^T$. The model assumes that each random utility follows the normal probability distribution: $\varepsilon_i \sim N(0, {\sigma_i}^2)$. Further these random components can be correlated each other so the random utility vector $\underline{\varepsilon}$ is defined as a multivariate normal (MVN) distribution:

$$\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix} \sim \mathrm{MVN} \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} {\sigma_1}^2 & \sigma_{12} & \cdots & \sigma_{1M} \\ \sigma_{21} & {\sigma_2}^2 & \cdots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \cdots & {\sigma_M}^2 \end{bmatrix} \right) \tag{50}$$

which can be denoted by $\underline{\varepsilon} \sim \mathrm{MVN}(\mathbf{0}, \Sigma)$ for short where the symmetric matrix Σ is called the dispersion matrix, the variance-covariance matrix or simply the covariance matrix. Using this notations, the total utility vector can be rewritten as $\underline{U} \sim \mathrm{MVN}(\underline{V}, \Sigma)$. If the covariance matrix Σ is known, the joint probability density function is given as follows:

$$f(\underline{\varepsilon}) = \frac{1}{(2\pi)^{M/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} \underline{\varepsilon}^T \Sigma^{-1} \underline{\varepsilon} \right). \tag{51}$$

Per Equation 40, the choice probability of alternative X_1 can be obtained by computing the following equation:

$$\pi(1) = \int_{-\infty}^{\infty} \int_{-\infty}^{V_1 - V_2 - \varepsilon_1} \cdots \int_{-\infty}^{V_1 - V_M - \varepsilon_1} f(\underline{\varepsilon})\, d\varepsilon_M \cdots d\varepsilon_2\, d\varepsilon_1 \tag{52}$$

where M integrals should be carried out. Computationally this equation is very intensive so many multinomial probit models in practice have less than four or five alternatives. Avoiding direct integration, one naïve approach is to rely on the Monte Carlo simulation technique, which is straightforward but obviously requires large computational resources.

Another technique is to use an approximate but analytic equations, which attempts to transform a trinomial choice problem into a binomial case. Suppose that there exist three alternatives and define U_* such that $U_* = \max(U_2, U_3)$. Clark (1961) found the approximate relation that $U_* \sim N(V_*, {\sigma_*}^2)$ if $U_2 \sim N(V_2, {\sigma_2}^2)$ and $U_3 \sim N(V_3, {\sigma_3}^2)$ where V_* and ${\sigma_*}^2$ are given by the following equations:

$$V_* = \left[F_N\left(\frac{V_2 - V_3}{\sigma}\right), F_N\left(\frac{V_3 - V_2}{\sigma}\right), f_N\left(\frac{V_2 - V_3}{\sigma}\right) \right] \cdot \begin{bmatrix} V_2 \\ V_3 \\ \sigma \end{bmatrix} \tag{53a}$$

$${\sigma_*}^2 = \left[F_N\left(\frac{V_2 - V_3}{\sigma}\right), F_N\left(\frac{V_3 - V_2}{\sigma}\right), f_N\left(\frac{V_2 - V_3}{\sigma}\right) \right] \cdot \begin{bmatrix} V_2^2 + {\sigma_2}^2 \\ V_3^2 + {\sigma_3}^2 \\ \sigma(V_2 + V_3) \end{bmatrix} - V_*^2 \tag{53b}$$

where $\sigma^2 = {\sigma_2}^2 + {\sigma_3}^2 - 2\sigma_{23}$. F_N and f_N indicate the standard normal cumulative function and the density function, respectively. Hence, $\pi(1)$ is given by

$$\pi(1) = Prob\,[U_i < U_1, \text{ for } i = 2, 3] = Prob\,[U_* < U_1].$$

A choice problem that has more number of alternatives can also be treated by recursive uses of this trinomial case. For example, if five alternatives are considered, one needs to compute $U_* = \max(U_2, U_3, U_4, U_5) = \max[\max\{\max(U_2, U_3), U_4\}, U_5]$. However, this recursive algorithm eventually triggers large amount of computations as the number of alternatives increases. The use of multinomial probit models would be expanded with increasing computing powers and improving computational methods.

CHAPTER IV

APPROACH

Contents

The discussion in Chapter I generated the research questions that can be essentially boiled down to how the National Transportation System is modeled. For this purpose, a brief analysis of the modeling target is presented in Section 4.1 with an emphasis on passenger transportation. Sections 4.2 and 4.3 bear upon the central ingredients of the present research. The entity-centric abstraction framework is proposed, which serves as the top-level blueprint in tackling the modeling task. Following is the tactical blueprint for modeling and simulation scheme as the idea of the virtual world is crystalized, built from the abstraction framework.

4.1 Analysis of the National Transportation System

A basic point of departure for the development of modeling and simulation is to understand the world being modeled. Exploration of the National Transportation System (NTS) begins with a concise look into two defining facets: the physical extent and the traveling public. Next, the recent status of the transportation system is illuminated in the context of the nation's economy.

4.1.1 The Physical Extent

The United States has the world's most extensive transportation system which offers personal and cargo mobility for the nation's nearly 300 million residents and 10 million business establishments. Currently, over 230 million motor vehicles, transit vehicles, and railroad cars were available for use on the over 4 million miles of highways, and railroads that connect all parts of the nation. The capacity of the air and transit systems in the U.S. is also phenomenal. The system includes about 213,000 aircraft and over 19,000 public and private airports. In 1999, this air transportation network supported about 4.6 trillion passenger-mile of travel and 3.8 trillion ton-mile of commercial cargo. Besides these primary modes, rail and transit play an important part in cargo and intra-city mobility. The remaining elements are water and space transportation system. Detailed analysis of these secondary modes are not presented herein since their passenger throughput are almost negligible especially for long-distance travel.

4.1.1.1 Ground Transportation

Motor vehicles include any type of self-propelled land vehicles and are the most frequent means to move payload in the nation. According to BTS (1999*b*, p. 14), they carry over 90 percent of all passenger trips and over half the freight tonnage. Defining elements of this system include public roads and streets, automobiles, vans, sport utility vehicles, trucks, motorcycles, taxis and buses operated by households, transportation companies, governments and other businesses. The relevant infrastructures other than roads are garages, truck terminals, and other facilities for motor vehicles. Public roads in the nation were composed of interstate highways, other National Highway System (NHS) roads and non-NHS roads (total 4 million miles). The roads are used by 132 million cars, 75 million light trucks (vans and sport utility vehicles), 7.8 million commercial trucks, 729,000 buses, and 4.2 million motorcycles. The average distance traveled by each car and light truck annually is about 12,000 miles, or added together, about 0.42 light years that is about one-tenth of the distance to Proxima Centauri, the nearest star beyond our solar system. Service providers that carry passengers and freights using these vehicles include 4,000 private motorcoach companies and 511,000 interstate freight motor carriers and numerous taxi or rental car companies. The railroad system is a bit slow but still efficient transportation means especially for large cargo. The transit system carries daily, short-haul traffic on a scheduled basis. It uses various types of vehicle such as buses, heavy train and light train. The table below shows the summary of the extent for each ground transportation system.

Table 4: Ground Transportation System [Source: BTS (2002*a*)]

Mode	Components
Motor vehicle	Public roads
	46,747 miles of interstate highway
	114,790 miles of other National Highway System roads
	3,820,134 miles of other roads
Train	Miles of railroad operated (2001)
	97,631 miles by Class I freight railroads
	17,439 miles by regional freight railroads
	27,563 miles by local freight railroads
	23,000 miles by Amtrak (passenger)
Transit	Directional route-miles (Stations)
	Bus: 160,506
	Commuter rail: 5,209 (986)
	Heavy rail: 1,572 (1,019)
	Light rail: 897 (613) *Note: $\bigoplus = 24680mi$

4.1.1.2 Air Transportation

Major elements of the air transportation system include National Airspace System (NAS), airports, aircraft and air carriers. According to FAA (2002, p. 7), the nation is covered with 5,314 public-use airports and 13,992 private-use airports as of January 2001. These airports are classified into several categories by FAA's National Plan of Integrated Airport Systems (NPIAS), which identifies airports that are significant to national air transportation and to which FAA allocates funding. The breakdown is shown in Table 5.

Table 5: NPIAS Categories of Airports

Category	Subcategory	Count	Sum
National Airport System	Commercial Service*	546	3,364
	Reliever*	260	
	General Aviation*	2,558	
Low Activity Landing Areas	General Aviation*	1,950	15,942
	Closed to Public	13,992	
Total Airports in the U.S.			19,306

* Open to public, total 5,314 airports

Commercial service airports are further broken down into large, medium, small hubs and non-hubs depending on the size of total enplanements[1]. Combined, they handled nearly 640 million enplanements in the year of 2000. (BTS 2001) Noticeable differences among the categories of commercial service airports (as of 2001) are described in Figure 28. Reliever airports are high capacity general aviation airports in major metropolitan areas. The 2,558 general aviation airports in the NPIAS tend to be distributed on a one-per-county basis in rural areas. The airports that are not included in the NPIAS have little activity in general. They have an average of 1 based aircraft, compared to 32 based aircraft at the average NPIAS general aviation airport.

Air carrier aircraft carry passengers or cargo for hire under 14 CFR (Code of Federal Regulations) 121 (large aircraft–more than 30 seats) and 14 CFR 135 (small aircraft–30

[1]FAA definition: Passenger boardings at airports that receive scheduled passenger service.

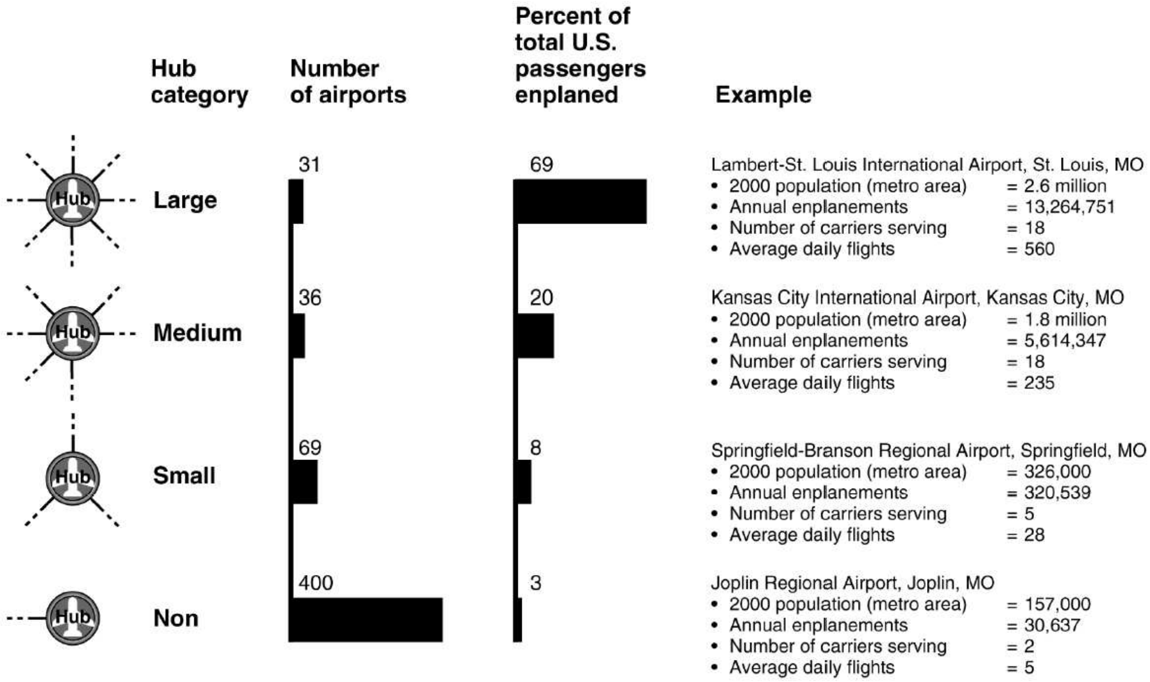


Figure 28: Commercial Service Airports [Source: Hecker (2003, p. 4)]

seats or less). These commercial aircraft are operated by 13 major carriers, 30 national carriers and 38 regional carriers as of 1999. The other aircraft which do not belong to air carriers fall into the general aviation category. There are about two hundred thousand active general aviation aircraft which flew 3.9 billion miles in 1997. The majority of general aviation aircraft is occupied by 149,422 fixed wing airplanes with one piston engine. The number of each type of aircraft is given in Figure 29.

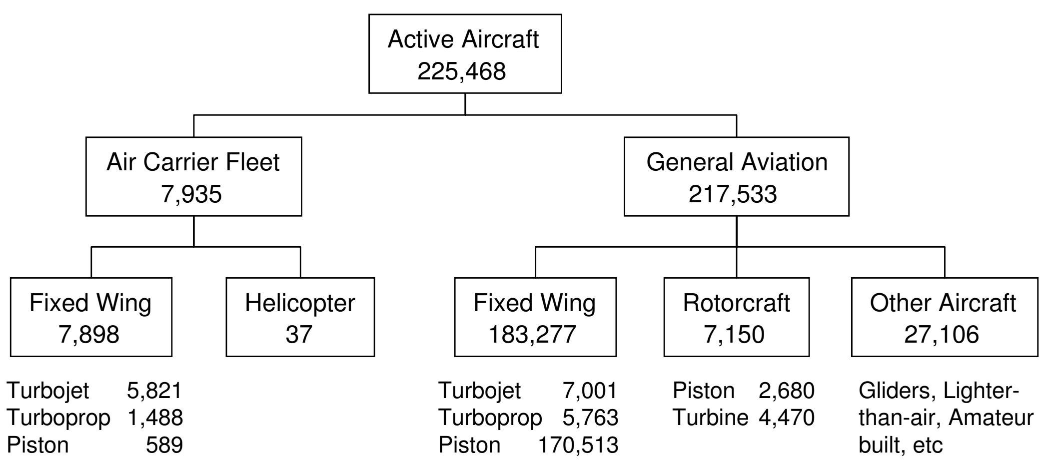


Figure 29: Types of Active Aircraft in Service

4.1.2 The Traveling Public

The nation's travelers are indeed chief players of the NTS. The government is always interested in their characteristics to assist in generating a better transportation investment and policy. In doing so, the Bureau of Transportation Statistics (BTS), part of the Department of Transportation (DOT), has carried out nationwide traveler surveys on a regular basis, called American Traveler Survey (ATS). In the 1995 ATS, approximately 80,000 randomly selected households were interviewed. The survey collected information about all trips of 100 miles or more, one way, taken by household members in the year of 1995. The following exhibits are mostly based on the survey results.

4.1.2.1 Population

Summarized are a few demographic data including the number of population, the number of households, and the mean income from the year of 1995 to 2000 in Table 6. It can be observed that Americans are getting richer and the number of households increases a bit faster than that of population as the average number of household size decreases.

Table 6: Recent Trends of Demographic Statistics

Year	Number of Population (in thousand)	Number of households (in thousand)	Mean Income (chained in 2001 U.S dollars)
1995	262,804	99,627	51,835
1996	265,228	101,018	52,934
1997	267,784	102,528	54,653
1998	270,248	103,874	56,240
1999	272,691	104,705	58,254
2000	281,422	108,209	58,639

4.1.2.2 Travel Origin and Destination

According to BTS (1999*a*), American households took nearly 685 million long-distance trips (100 miles or more, one way) in 1995. About 656 million of those trips were to destinations in the U.S. These domestic trips are broken down in the form of the origin destination matrix as shown in Figure 30.

The first column and row of the table are for trip origin and destination breakdown, respectively. Label Top-50 indicates the largest 50 metropolitan areas aggregated of the nation; Top-160 denotes the next 51st through 160th metropolitan areas. N-metro means the remaining non-metropolitan regions. It is observed that 74.2% (= 100% – 15.8%) of all domestic trips involve at least one metropolitan area either in trip origin or destination. This figure is likely to increase with the continual trend of the urbanization in the nation.

O \ D	Top-50	Top-160	N-metro
Top-50	16.7%	8.2%	15.6%
Top-160	9.6%	5.2%	9.1%
N-metro	12.9%	6.9%	15.8%

Figure 30: Trip O-D Matrix

4.1.2.3 Travel Motivation

Of all household trips in 1995, 29.3% were for business and 70.7% were for personal purposes. The travels for personal purposes can be further broken down into 29.8% for visiting friends and relatives, 27.0% for leisure activities, and 13.9% for personal business.

4.1.2.4 Travel Distance

It was found that difference in travel purposes (personal or business) significantly affects many attributes of trip. For example, personal travel tends to be shorter than business travel in terms of one-way trip distance, as illustrated in Figure 31 showing a natural trend that the longer the distance the less travel for both cases. Also, an analysis with respect to income brackets—which is not presented here—revealed that average trip distance is positively correlated with household income.

4.1.2.5 Travel Party Size

The size of a travel party is also affected by travel purposes. Over 60% of business travel is done by a single person while nearly three quarters of personal travels have at least two persons in a unit trip party. Detailed information is given in Figure 32. The amount of the personal travel portion will change with the current trend of decreasing average number of household members.

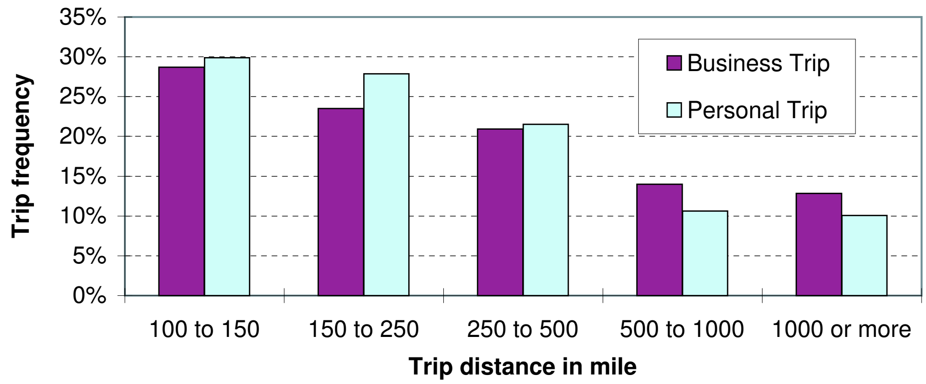


Figure 31: One-way Trip Distance Distribution by Travel Purpose

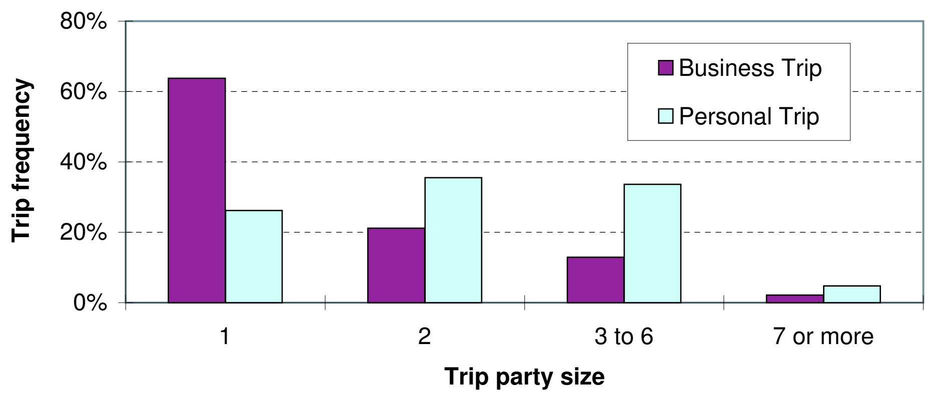


Figure 32: Travel Party Size Distribution by Trip Motivation

4.1.2.6 Modal Choice

Another characteristics of special interest concerns how Americans travel. The 1995 ATS data shows that 77.0% of all long-distance household trips were done by cars, vans and light trucks. 19.7% were done by airplanes of which 96.4% by commercial transport and 3.6% by corporate or personal aircraft. The remaining 3.3% were picked up by the other secondary modes such as buses, trains and ships.

Travel purpose and household income is a big factor on how people travel. The apparent reason is, of course, "who pays for it? and how much?". Table 7 correlates mode choice with travel purposes. The other secondary transportation modes are excluded in this analysis. Travel distance is also an important factor on a traveler's modal choice as well. Based on 1995 ATS data, the bar chart in Figure 33 is prepared to show primary transportation modes with one-way trip distance brackets. Again, the secondary modes are excluded here. It can be observed that the extent of the air travelers becomes larger as travel distance increases, which can intuitively be deduced.

Table 7: Transportation Mode Choice

Mode	Motor Vehicle	Commercial Air
Personal travel	60.3%	10.0%
Business travel	19.8%	9.8%

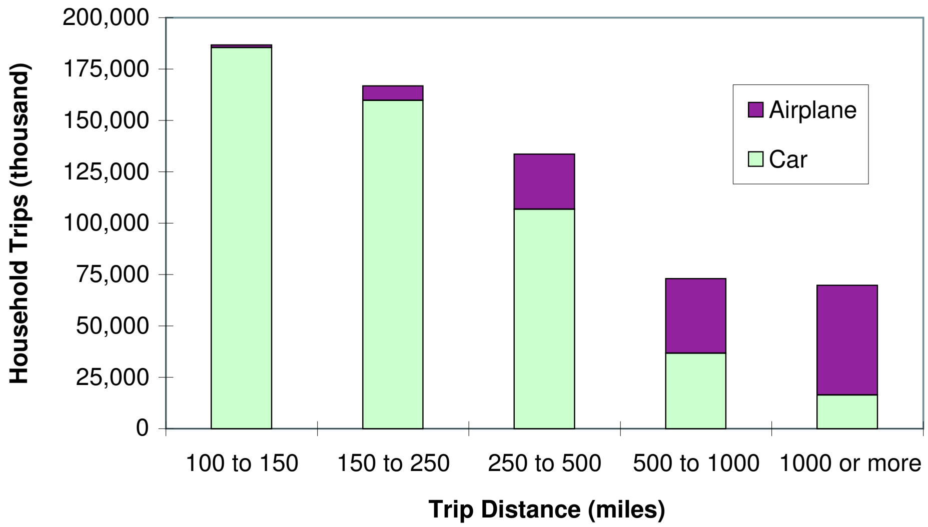


Figure 33: Travel Distance and Modal Choice

4.1.3 Economic Trends in Transportation

While the above examination provided a snapshot of the key elements of the transportation resources, it is worthwhile to analyze transportation related trends from an economic perspective with a view towards a comprehensive understanding of the entire system with the national scale. The first topic begins with all transportation expenditures, as portrayed in Figure 34. The expenditures in passenger transportation and freight transportation are indicated separately. Both expenditures were of a similar magnitude back in the 1960s but the gap between them was noticeably expanding around the airline deregulation in 1978. Since then, passenger transportation had been flourishing in comparison to freight transportation with a nearly constant ratio of three to five.

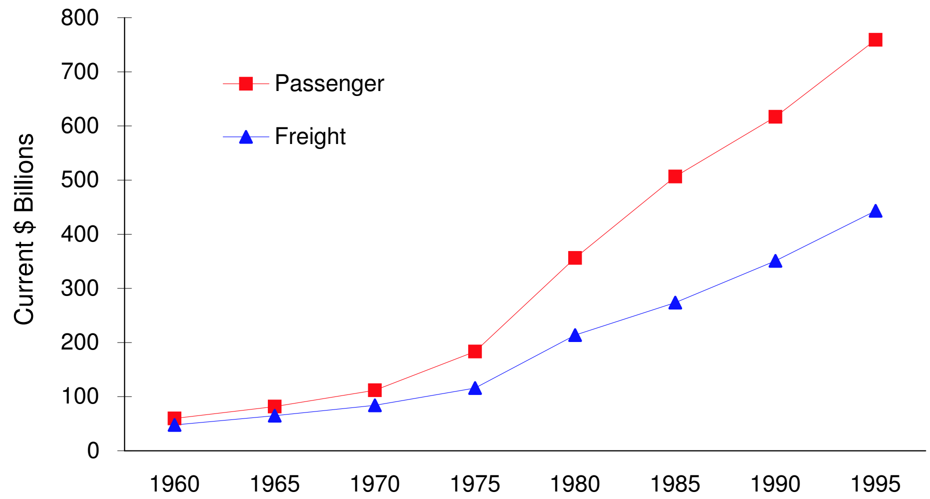


Figure 34: Passenger and Freight Transportation Expenditures at 5-year Intervals

It is clear that these expansions are credited to growing population in the nation. Additionally, the nation's rising economic affluence has further boosted transportation expenditures. In order to substantiate this assertion, it is necessary to figure out how individual consumers have spent their income within the transportation sector. Table 8 shows the detailed distribution concerning how a reference household allots its annual expenditure to major expenditure categories from 1990 to 1999. The annual expenditures in the second column are represented in terms of current dollars.

Table 8: Expenditure Distribution of a Reference Consumer Unit

Year	Expenditures	Food and Apparel	Housing	Transportation	Other Expenses
1990	28379	20.84%	30.67%	18.04%	30.45%
1991	29614	20.28%	31.24%	17.39%	31.08%
1992	29846	20.05%	31.75%	17.52%	30.68%
1993	30692	19.79%	31.40%	17.77%	31.04%
1994	31733	19.08%	31.85%	19.05%	30.03%
1995	32262	19.25%	32.42%	18.64%	29.70%
1996	33798	19.08%	31.80%	18.88%	30.24%
1997	34820	18.75%	32.37%	18.54%	30.33%
1998	35536	18.25%	32.96%	18.62%	30.18%
1999	36996	18.31%	32.59%	18.95%	30.15%

During the 10-year period, the portion of daily necessities (essential goods or basic staples) has been persistently decreased while housing related expenditures are picking up. Transportation expenditures also tend to increase except for the years of 1991, 1995 and 1997. It does not belong to the author's discretion to say whether or not it is coincident but years of 1991 and 1995 recorded the lowest growth rates of real Gross Domestic Product (GDP) over 10 years (-0.2% and +2.5%, respectively). For more detailed analysis, presented is the history of percent of annual expenditure against real GDP of the nation in Figure 35. The horizontal axis represents billions of 2000 chained dollars while the vertical axis indicates the percent of expenditures from the above table. The numbers around the marks in the plots indicate the last digits of the 1990s. While the first two categories from the left have relatively high regression coefficients, transportation expenditures grow with an erratic trend for it is likely susceptible to other economic and societal factors. This observation agrees with the hypothesis that people allocate money so as to improve the quality of life with increased economic wealth.

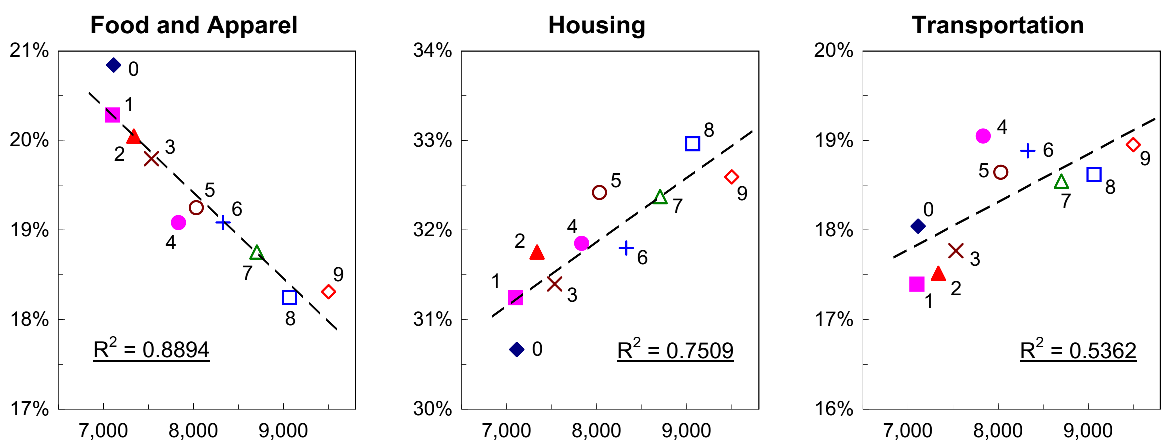


Figure 35: Expenditure by Major Expense Category

Given this observation, the next question would be about where the money is going in the end. To investigate this point, the 10-year trend of each transportation mode's utilization is shown in Figure 36 where each mode's vehicle-mile traveled (VMT) was normalized to the year of 1990. As seen, transit is recently ramping back up but train travel still fails to sustain itself. While both air and car transportation are growing with a steady tendency, the growth of air travel is notably faster than that of car travel.

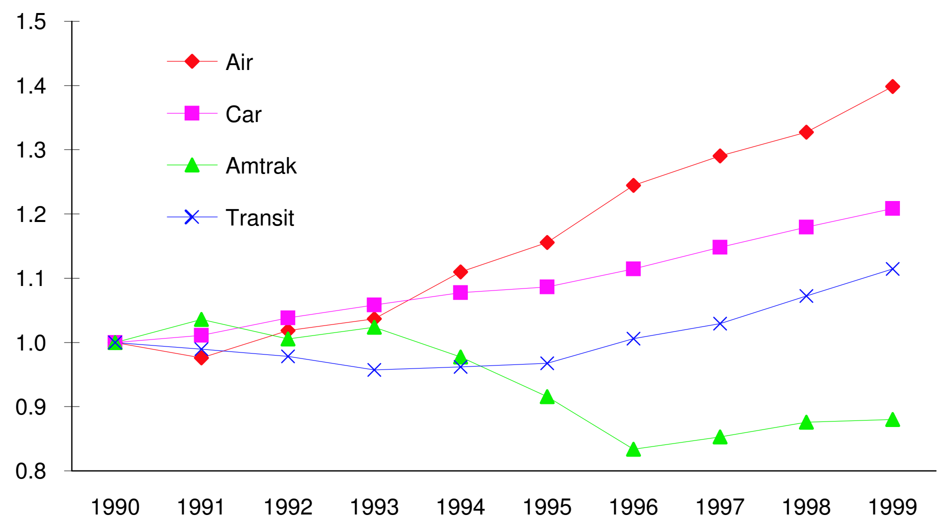


Figure 36: Normalized VMT Trend of Transportation Modes

In summation, Figures 35 and 36 testify that economic affluence makes the public spend more money on transportation: purchasing new vehicles, traveling more and longer, and using a high-speed mode more often. Further, it can be deduced that, under the assumption that this trend continues, efficient public intra-city transportation (transit) is becoming an important issue and a combined need for on-demand (car) and high-speed (air) transportation is soon ahead. Nevertheless, the nation's transportation infrastructure has changed very little in recent years. Road lane-miles, for instance, have grown by about 4% between 1980 and 2000, while cars and light trucks have increased by 40%. In air transportation, the number of aircraft operated by air carriers has increased by more than 35% since 1990, while the number of certificated airports (those serving scheduled operations seating more than 30 passengers) has shrunk. (BTS 2002*b*, p. 35) As the heavy use of the nation's infrastructure offers a challenge in the future, improved management of the system is one method being used to keep traffic flowing. An additional potential solution is to infuse a new transportation supply, like a PAV system.

4.2 *Entity-Centric Abstraction*

The objective of this section is to propose the entity-centric abstraction framework serving as a strategic blueprint to tackle the aggressive challenge of modeling the NTS. Further, it aims for creating a frame of reference to spur the inter-disciplinary and trans-domain research involved in contemplating a future *transportation architecture*—a holistic, hypothetical representation of the National Transportation System (NTS) reconfigurable by cascading changes in economic, societal and technological development.

4.2.1 Motivation and Conceptual Foundation

The NTS is indeed a complex system both in the colloquial and technical sense of the word, which governs the overarching difficulty in modeling of the NTS itself. Complexity in the NTS stems primarily from three properties: the heterogeneity of constituent systems, the distributed nature of these systems (Carson & Doyle 2002), and the presence of "deep uncertainty" (Lempert, Popper & Bankes 2003) in exploring its future state. In light of these general properties, the major consequences of each source of the complexity can be examined.

4.2.1.1 The NTS as a Complex System-of-systems

The NTS is composed of many heterogeneous elements breeding a wide variety of issues both within traditional disciplines (engineering: aerospace, civil, mechanical, etc) and across domains (engineering vehicles, business enterprises, governmental policy/regulation, etc). This source of complexity presents challenges to understanding (different languages), modeling (different design variables and time scales), and assessment (different stakeholders). Treatment of the second source of complexity, the distributed nature of the systems, is also a significant challenge. As opposed to a lumped system, the important system interactions among the elements are poorly understood. Further, in the presence of the third source—deep uncertainty— complications arise more often which leads an investigator to confronting chaotic system behaviors. Deep uncertainty inherent in the distributed systems results in, at best, imprecise models. Finally, uncertainty resident in the environment gives rise to un-modeled feedback dynamics associated with the ultimate 'control' (a policy, a vehicle, etc.) chosen. Combined with the irreversibility of many decisions, the result is a partially controlled process with path and time dependency. The NTS as a complex system, then, may best be conceived as a 'living system': a sum of diverse things which possesses a collective sentience, evolves over time and is organized at multiple levels to achieve a range of (possibly) conflicting objectives, but never quite behaving just as planned.

The traditional approach to modeling a large, complicated system takes the position of reductionism, the philosophical dogma that has dominated the development of the modern sciences since Descartes. (Grosholz 2001) This approach is anchored in a seemingly impeccable postulation that a system can be described in terms of its components and accordingly, integration of many small-scale, hierarchically decomposed models leads to understanding of the whole system. Although the felicitous achievements over hundreds of years testify to its success, the reductionism strategy is not complete for the study of complex systems. The reductionism strategy creates "box-inside-a-box" mentality and becomes simply impractical when a system is composed of unmanageable number of heterogeneous elements. This is, however, only a superficial reason. The fundamental shortcoming of reductionism comes from the fact that, as commonly noted, a whole is more than the sum of its parts.

4.2.1.2 Quest for the Holism through Abstraction

An alternative to the reductionism is the holism, the view that one must first understand the whole. Such a perspective is particularly critical for study of the NTS, where the various interrelated facets can be understood uniquely as an integrated system-of-systems (SoS). The essence of the problem—the hard-to-grasp insight—likely appears only at this elevated perspective. However, while a holistic approach may facilitate one's intuitive understanding of problem structure, it does not necessarily make communication between the involved facets any easier. An effective lexicon is needed to provide the bridge the gap between understanding and communication. The lexicon bridge is critical since professionals from the various domains have one thing in common: they are typically trained to solve problems using methods and ideas prevalent to their own domain. This legacy is the source of the often-used term ‘stovepipe’ in reference to the narrow scope thinking in a particular area of specialty knowledge. It is also clearly an artifact of the reductionism mindset. The real NTS, however, can only be fully understood via ‘across’ stovepipes, spanning various columns of knowledge, so the holistic frame of reference is required for such a trans-domain endeavor.

Therefore, before browsing a variety array of modeling tactics, and resorting to the laborious modeling task with the “box-inside-a-box” practice[2], the initiative is taken to come down to an elementary level and contemplate on how to properly approach the problem. Thus, the modeling task at hand for simulating the NTS—deemed as a “grand challenge” (Wieland, Wanke et al. 2002)—can be put aside for now since what is really needed is a new way of thinking. This leads to engaging the power of abstraction for it requires a rigorous mental activity that enables to achieve the holistic perspective while also breaking lexicon barriers.

The whole idea of abstraction is, in essence, the notion of both classifying things (creating sets) and representing organization (forming networks) using articulate lexicon for the purpose of being able to conceive and examine at the holistic level. For study of a SoS like the NTS, the lexicon must ensure that 1) all parties understand the description, and 2) all relevant portions of the problem are covered. Proper abstraction aims for generic, universal, uniform semantics. Levels of abstraction should be employed to adjust the vantage point of the holistic perspective. In the end, the ultimate goal of the abstraction framework should serve to allow practitioners and theorists of this field to navigate, communicate, model and design collaboratively as well as produce a useful product to the decision makers.

[2]For example, Garcia (2003) described the difficulty and limitation in modeling the entire National Airspace System (NAS) by integrating available computer analysis codes.

4.2.1.3 Three Approaches for Organization

The organization of things is just as important as the things to be organized. The first step in an abstraction process is representing organization under this belief, which can be carried out from different viewpoints. Depending on the implementation approach, the abstraction process may generate variations in the mental model, the extent of traceability, and the scope of understanding about the target system. Further, a distinct computational model will be created. Hitchins (2003) enumerates some of the many viewpoints including entropy models, poached egg model, recursion model, open system model, cybernetic model, process model, transport model and queuing model. On the other hand, Rouse (2003) offers four views: hierarchical mappings, state equations, non-linear mechanisms, and autonomous agents. The next discussion attempts to simplify these leading thinkers' viewpoints into the following three approaches to organization.

Hierarchy-centric approach

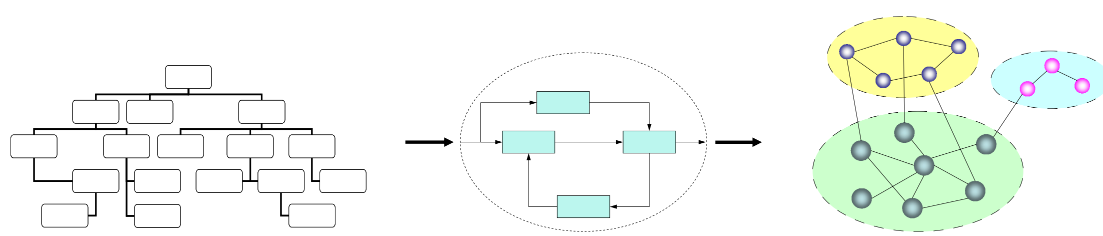
The most common and simplistic way of organizing is the hierarchy-centric approach. A system of interest is divided into its sub-systems and a sub-system is divided into its components and so forth. Conversely, the system can be an element of a higher-level system. The higher-level system can also have other system instances running parallel to the system of interest. This approach has several strengths. It is intuitive and shows the structure of the system with clarity. However, this clarity comes at the price of flexibility since this approach is fundamentally founded on the reductionism, inheriting the limitations discussed before. Also, the creation of a lexicon could be problematic when a large number of strata are involved (e.g., repetitive extensions as in the term system-of-systems).

Flow-centric approach

The flow-centric approach has a somewhat different perspective, one in which the quantification of relationships is paramount. This approach emphasizes flows within the system rather than the components. Inside the boundary or control volume, the elements of the system are organized to reflect generating, dissipating and processing of the flows as shown in Figure 37(b). The flow medium can have a variety of formats. For example, a model from a physics-based field can take from energy, current, or other time-space variables. Recently, the use of System Dynamics to applied economics problems has been gaining in popularity. (Sterman 2000) In this case, the flows can be money, information, or materials. Strengths include ability to capture dynamic behavior at high levels of abstraction, capturing so-called primary feedback phenomena. This capability is obtained at the expense of insight at the component level due to aggregation.

Network-centric approach

Barabási (2002) contributes towards the third approach having recently gained momentum. The network-centric approach focuses on building nodes and links. For example, nodes are places (origins or destinations) while links are the characterization of what flows between places and how. This approach can flexibly define the elements in the system as well as their relationship. Unlike the hierarchy-centric approach, the network-centric approach does not necessarily call for the monotonous nesting structure. Instead, it only needs topological information between nodes. Also, it becomes quite natural to introduce the concept of the interface and layer and to embrace the object-oriented philosophy. The whole ingredients construct a body of network as shown in Figure 37(c).



(a) Hierarchy-Centric View (b) Flow-Centric View (c) Network-Centric View

Figure 37: Three Approaches for Organization

The three approaches described are not exhaustive of the possibilities nor are exclusive of one another. The entity-centric abstraction, thus, attempts to accommodate all whenever appropriate as abstraction proceeds, although the basic mindset is invoked from the network-centric approach in framing overall organization.

4.2.1.4 The Concept of Entity

As discussed earlier, the holistic perspective is essential to understand the whole without prescribed boundaries. But what exactly is being abstracted? Here, embracement of the holistic perspective begins by adopting the assumption that (colloquially) “everything is on the table”. Two subsequent questions may follow from this assumption: What is *everything*? What is the size of the *table*? These two questions are circuitous; in other words, the size of the table constrains the extent and sort of things in the transportation environment, and vice versa. If the questions were posed to a supreme transportation architect—a hypothetical individual who wishes to shape the transportation architecture under her/his design—she/he would surely realize that not only physical factors, such as vehicles and infrastructure, but also organizational elements, such as public interest groups and industrial firms, should reside inside of their problem boundary. Otherwise, a wide variety of elements in the NTS cannot be examined together, which loses the vantage point of the holistic perspective.

Under the entity-centric abstraction framework, all of these factors “on the table” find themselves a home, unified through the concept of *entity*.[3] Entity is analogous to *object* in the computer science domain. Rumbaugh et al. (1991) defines the term object as a concept or thing with crisp boundaries and meaning for the problem at hand. In objected-oriented programming like Java or C++, the internal view of any object uncovers *states* (or variables) and *behaviors* (or methods) as the defining elements. Similarly, an entity is composed of *attributes* and *functions*, which correspond to states and behaviors, respectively. Moreover, the entity can have *sentience* and *interfaces*. The role of these four key rudiments of the entity is to symbolize its being (attribute), doing (function), thinking (sentience), and linking to ‘externalities’ (interface). Anchored in this conceptual foothold, the entity-centric abstraction is instantiated with particular entity characterizations.

[3] In modeling and simulation field, the term entity generally refers to a structural components of a discrete-event simulation. Entity is understood as something that has attributes and that causes changes in the state of the simulation. (Ingalls 2002) This is slightly different but not contradictory to the present definition.

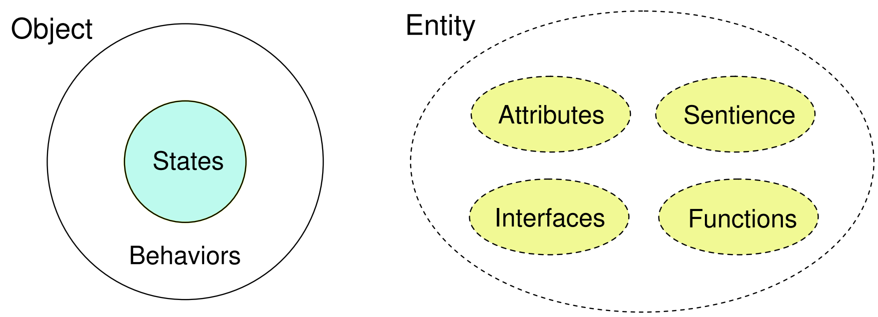


Figure 38: Comparison of Object and Entity

For example, a car is modeled as an entity that has attributes, functions and interface, without sentience. Attributes of a car contain certain characteristics that are unique to (or that defines) the car: make, model, vehicle identification number, gas mileage, etc. However, speed and pos ition at a particular time belong to the interface since the values of those variables result from interaction with other entities: road conditions, other cars, the driver, etc. Taking more examples, an organization or individual is an entity with a complete set of the four rudiments. Likewise, the concept of entity can encapsulate various kinds of events and conditions in the transportation environment.

In this context, a modeler can flexibly determine the boundary of an entity, elaborating the rudiments to describe the target system as needs arise. Therefore, an entity can be thought of as an extended form of object, though not necessarily having the 'crisp' boundaries for the purpose of obtaining inherent flexibility. Anchored in this conceptual foothold, the entity-centric abstraction captures any instance among *everything*, and upon completion of identifying things of interest, the transportation architects simply include the corresponding entity or entity groups. If they want to add or subtract some entity, the size of the *table* expands or shrinks accordingly.

4.2.2 Modeling Entities

Given the foundation, the abstraction process begins by identifying and hypothesizing key entities in the NTS. Two pairs of entity descriptors emerge: *explicit-implicit* and *endogenous-exogenous*. Based on these descriptors, four entity categories are generated: *resources*, *stakeholders*, *drivers*, and *disruptors*. All these entities are inter-webbed by networks that define the linkages amongst themselves and are described in further detail next.

4.2.2.1 Resources and Network

The first group of entities is embodied as transportation resources. Vehicles and infrastructure are examples of resources that consumers physically experience—thus having *explicit* nature—when traveling or sending shipments. Traditionally, resources within a general category have been treated in their own realm. Further improvement in mobility will nevertheless demand an integration of these distinct dimensions. Exploring a new mobility resource in this larger context can reveal its competitive advantage relative to existing resources and uncover the extent to which it is in harmony with a future transportation architecture.

Consequently, a view that encompasses all resources in the NTS together is useful. The decomposition of the NTS follows the hierarchy-centric approach. In doing so, usual practices were adopting prefixes like sub-, super- and hyper- as in Figure 8 of Chapter I as well as using the circuitous phrase: i.e., system-of-systems. To avoid the confusion with ambiguous derivations, the lexicon employs the use of Greek letters to delineate from strata of the hierarchy as described in Table 9.[4]

Table 9: Hierarchy Descriptors

Levels	Descriptions	Examples
α	The base level of entities.	Cell
β	Collections of α-level systems, organized in a network.	Organ
γ	Collections of β-level systems organized in a network.	Human
δ	Collections of γ-level systems organized in a network.	Society

The basic building blocks are designated α-level for which further decomposition will not take place. This notation facilitates the nesting structure of the resources in the NTS concisely unfolded. As shown in Figure 39, if vehicle is designated as α-level, the collection of all resources has the hierarchy descriptor δ. One distinction from Figure 8 is

[4]The idea of adopting Greek letters is credited to Robert Calloway in NASA Ames research center.

that a hypothetical new mobility resource is positioned deliberately not attached to any existing β-level system. The power of abstraction enables this hypothetical generalization, which is important since future innovation does not solely hinge on the air transportation sector. Further, the advantage of adopting the hierarchy descriptors is that a wide spectrum of decisions can now be unambiguously labeled facilitating trans-domain communications. This is significant because from design point of view, each level entails its own technologies, economics, and operational rules. For example, engine selection would be an α-level decision-making activity whereas the deregulation is a good example of a γ-level policy.

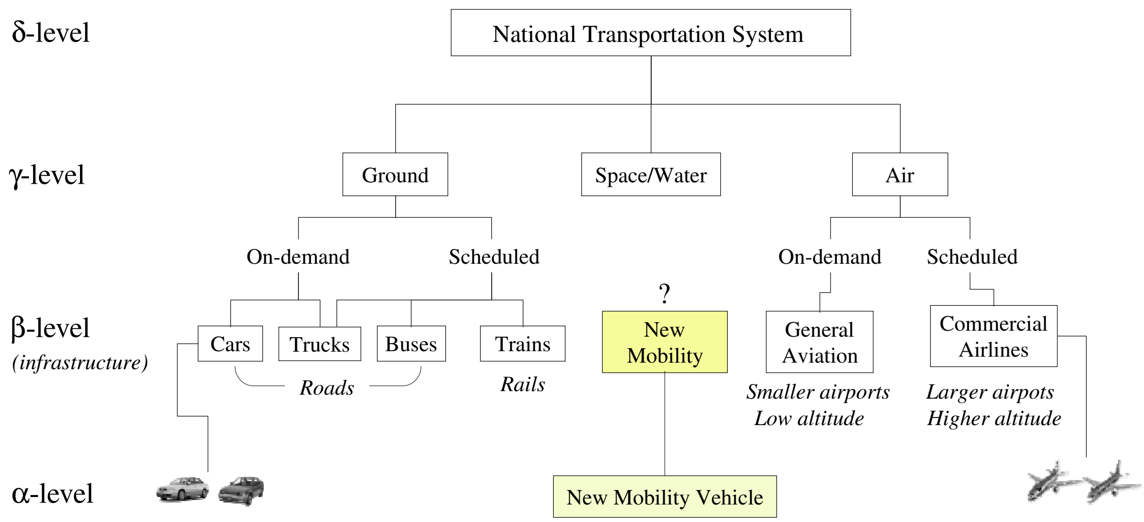


Figure 39: NTS Resource Hierarchy

Inside of transportation resources is a connected-ness in the sense that a perturbation at any lower level (e.g. vehicle's attributes) will result in an impact on many stakeholders and thus permeate into the entire system. This is so partly because all resources are bonded together via a topographical network that defines the physical connection between resources in which material (people or products) can flow. Additionally, trains, buses, automobiles and airplanes (and their respective infrastructure components) are connected in an economic sense, facilitating the intermodal and multimodal nature of transportation. Thus, proper abstraction should embrace the concept of the network that connects resources. Within a network perspective, then, the flexibility and degree of interoperability between resources becomes extremely important. Different types of infrastructure will offer varying degrees of flexibility. For example, a major hub airport may be viewed as a highly inflexible piece of infrastructure because it is difficult for such an airport to adapt to new vehicle types and operational schemes. Thus, the degree to which infrastructure resources are reconfigurable is an important design consideration. The combined consideration of resources and their network is vital to achieve significant improvements in future transportation architectures. These explicit entities, however, are not sufficient to completely formulate the problem. There are subtler, yet still important issues.

4.2.2.2 Stakeholders and Network

The National Research Council pointed out that NASA's Small Aircraft Transportation System (SATS) concept, with massive numbers of small aircraft operations, could entail adverse societal consequences including safety concerns and inefficient energy consumption per unit distance traveled per capita. (TRB 2002) This case points to the need for consideration of 'other-than-physical' factors—certain entities are present that desire to exert forces on the architecture for their own interests.

These entities are called stakeholders, and in most circumstances their behaviors and decisions are not manifested in an explicit manner to the consumers. The relevant stakeholders are identified in Table 10, where a broad abstraction has resulted in a collection of stakeholders in both private and public sectors, ranging from the actual consumers of transportation services to those involved in technology R&D. Each stakeholder has objectives representing their interests that dictate the manner in which they influence the transportation architecture. Indirect stakeholders influence the NTS by their outputs or goals being accepted or filtered by other direct stakeholders.

An intangible network can be imagined that defines the connection between stakeholders. This connected-ness comes in two forms. Firstly, one particular stakeholder may interact with another directly. Secondly, if a stakeholder influences a particular resource, after permeating through the resource network, the state of the transportation architecture

Table 10: Transportation Stakeholders

Category	Stakeholders	Descriptions	Objectives
Public	Consumers	Individual travelers or shippers (for commercial goods) that are the end user for the transportation system.	Min: travel time, expense; Max: comfort, safety, mobility reach
	Society	Represents the aggregate interests of citizens to the national level.	Min: noise, emission; Max: quality of life.
Industry	Service Providers	Owners of resources who sell transportation services to consumers.	Max: profit, market share, consumers' satisfaction.
	Manufacturers	Design, produce and sell transportation resources to service providers and/or customers.	Max: profit, market share, service providers' satisfaction.
	Insurance Companies	Provide protections against damage or loss of transportation resources by imposing insurance fee.	Max: profit, market share, customers' satisfaction.
Government	Regulatory Agencies	Impose rules on the system that restrict stakeholder activity and resource characteristics.	Max: safety, security.
	Infrastructure Providers	Plan, approve and execute employment and enhancement of infrastructure resources.	Max: capacity, Min: budget, delay.
Indirect Stakeholders	Media	Report information, forecast and plan from/to the public.	Varied, but vague.
	Research Agencies	Develop and provide transportation related technologies.	Provide firm foundation for transportation development.

will be modified. A consequence of the new state is a perturbation back to the originator and/or other stakeholders.

This stakeholder network can be hypothesized as a complicated web linking distinct organizations as nodes. Each link between the stakeholders possesses its own characteristics that depict the nature of an interaction: medium, polarity, strength and so forth. For example, the research agencies-to-manufacturer link may be expressed by monetary funding for research programs with developed product designs provided in return, which can be investigated by a System Dynamics model with the flow-centric mindset as in Kang et al. (2003).

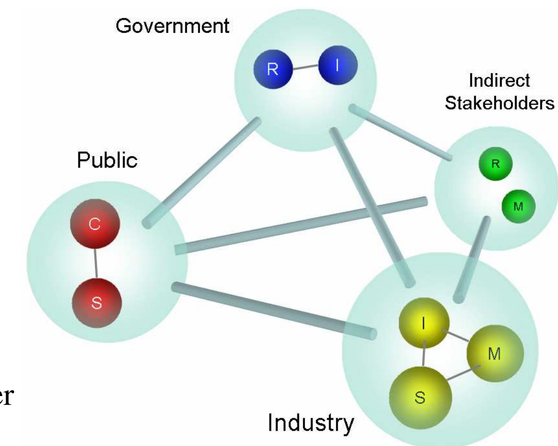


Figure 40: Stakeholder Network

Another important point is that, as indicated in Table 10, different stakeholders have different objectives. Individual travelers want to spend less time and money (or maximize mobility credit) with acceptable safety and reliability. For society, the aggregate noise, emissions, energy, cost and security are paramount. For service providers, the ability to make a profit while satisfying both consumers and society is the challenge. This dictates that the objectives of different stakeholders may be in conflict with each other. Thus, as shown in Figure 41, a Pareto front for consumers and societal perspectives emerges, for example, since a system that maximizes mobility for individuals is likely to increase total transportation-related energy expended.

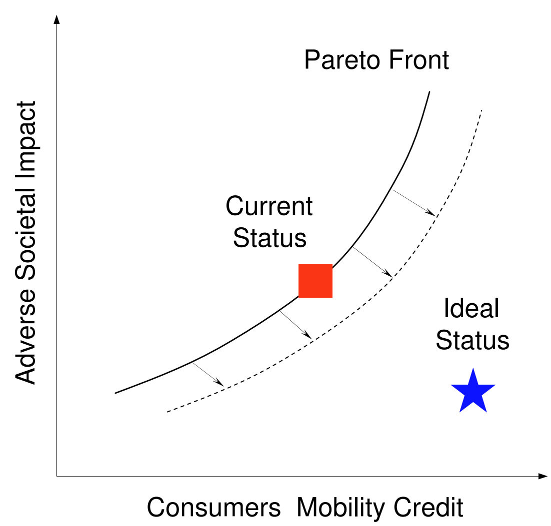


Figure 41: Conflicting Criteria—Notional Pareto Fronts

Traditionally, the scope of a particular resource design problem included only a subset of the stakeholders (e.g., Regulator–Consumer–Manufacturer–Researcher link). However, in an evolving SoS, the concern of all the stakeholders and the sensitivities between them must be tracked. While there has been no shortage of innovative air vehicle concepts proposed in the past, very few come to fruition partly due to the disregard of the broader group of stakeholders. For concrete improvements to be made, each stakeholder must realize value. While certain myopic policy decisions may merely perturb the current state along the existing Parieto front (solid line), the goal of the transportation architects should be to shift the entire Pareto front in the direction of the ideal state. Future innovations in transportation are unlikely to lie solely in radical resource designs[5], but also in understanding the complicated interactions stemming from the implicit entities and their networks.

4.2.2.3 Exogenous Entities

The description of the two entity groups has emphasized their generic, ontological characteristics. Actual transportation activities occur when those two groups have meaningful ties with the transportation environment—an arena where all transportation-related events occur. In that environment, however, there exist even other entities within its boundary.

If stakeholders and resources are considered endogenous building blocks in that they are under partial or full control of the transportation architects, within the transportation environment can be juxtaposed many *exogenous* entities of different types, traditionally considered given assumptions, circumstances and constraints about the transportation environment (e.g., population, weather). For the transportation architects, and from a design perspective, there is no control variable within exogenous entities since they have unidirectional influence (e.g., weather). Also, these entities typically have wide-reaching effects and take imperceptible feedbacks from the transportation architecture, if any (e.g., population). The exogenous entities can be categorized into two groups according to how they

[5]Interesting example: the failure story of 3-wheel car (tricar) back in the 1920s—found at [http:morgan3w.de/literature/magazine/lcc191122.htm], a lesson for PAV designers

affect the transportation environment.

Driver entities are largely concerned with economic, societal and psychological circumstances that influence the stakeholder network. In market-driven world, a great measure of transportation phenomena is governed by many economic factors. These include household income and gasoline price, demographic-related issues including the population distribution profile and the trend of population concerning growth and migration to (or from) urban areas. Further, much more than quantifiable factors go into the transportation stakeholders. A large portion of transportation activities are motivated by cultural and psychological reasons. Some trips are made as a result of lifestyle choice and are influenced by specific cultural events: summer vacation, Thanksgiving, etc. Psychological factors are also important. The surge in air travel after Lindbergh's successful transatlantic crossing is a prime example. With perturbation in any of the driver entities, each stakeholder seeks to adapt to changed circumstances, which brings reconfiguration of the transportation architecture.

On the other hand, *disruptor* entities affect the resource network and/or a portion of the driver entity group. They reduce the efficiency of the resource network, disable particular nodes and links of the network, or even bring the entire system down. Weather influences the resource network on a real-time basis: visibility problems, icing, and thunderstorms are primary issues that degrade punctuality and safety. Natural disasters also have their place in the transportation environment. These natural events affect the local environment, and the influence may cascade into the remainder of the national system. In contrast, there exist artificial disruptors under two categories. The first group influences the resource network directly (e.g., traffic accident, mishap operation). The second category of events affects psychological concerns, an element of the driver group. The drop in air travel after the 9/11 attacks is a primary example. Table 11 summarizes the two exogenous entity groups discussed.

In summation, in an analogy of the electrical circuit, drivers are akin to voltage sources which generate voltage and disrupters are akin to impedances which change the magnitude

Table 11: Exogenous Entities

Category	Drivers	Disruptors
Effect	Determining overall demand profile for transportation activities.	Causing delay and/or cancelation of transportation activities.
Examples	- Economic factors: GDP, household income, fuel price. - Societal factors: demographic characteristics, urbanization trend. - Psychological factors: culture, perception of safe/secure system.	- Artificial disruptors: accident, terrorism, pollution. - Natural disruptors: weather related events that affect operational condition of resources.

and phase of the voltage. These two groups together determine circumstances and constraints for all transportation activities. While difficult to describe and often too transient to predict, drivers and disruptors are significant parts of the NTS. The entity-centric abstraction framework embraces these "externalities" just like other "internalities" in an attempt to describe the whole.

4.2.3 Synthesis into Network of Networks

The previous discussion was devoted to abstracting the rudimentary elements of a transportation architecture of which the parsimonious statement using these rudiments proceeds as follows: *stakeholders employ particular resources organized in networks to achieve an objective under the various exogenous entities*. The subsequent task in completing the entity-centric abstraction framework is to properly establish the connection between the four entity classes to enable synthesis of the final form.

4.2.3.1 Transportation Architecture Field

In a nutshell, a transportation architecture results through the union of a particular resource network, stakeholder network and set of exogenous entities. The description between these entity groups and the time-variant transportation environment can be concisely portrayed in a pseudo three-dimensional space, called the *Transportation Architecture Field* (TAF), as illustrated in Figure 42.

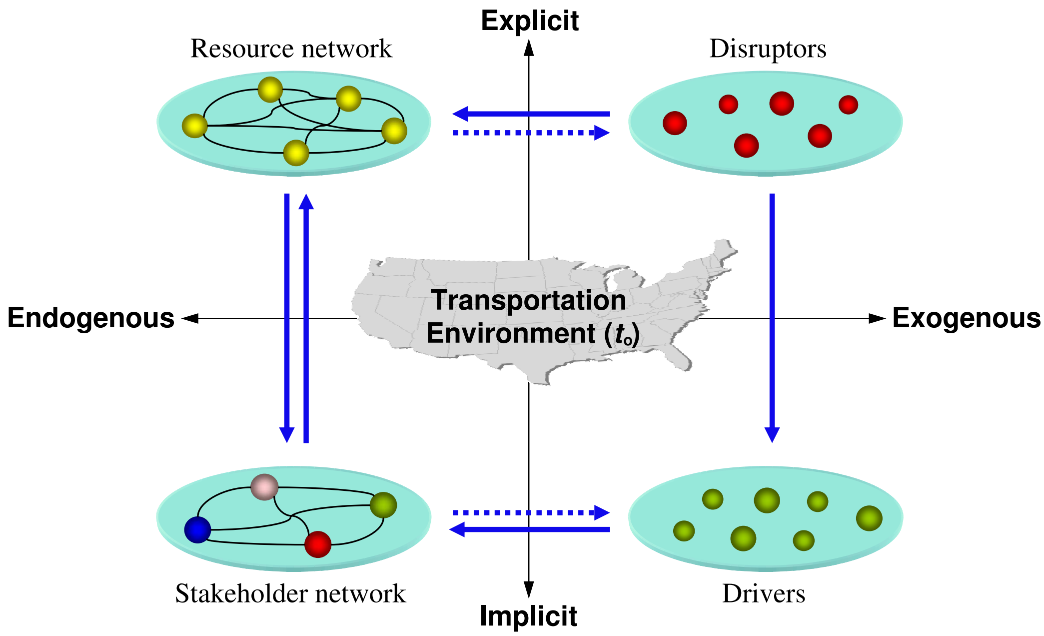


Figure 42: Transportation Architecture Field (TAF)

This depiction summarizes the entity-centric abstraction for synthesizing a transportation architecture. The figure represents a slice of the TAF with respect to given time $t = t_o$ where time axis (not shown) is out of the plane of the figure. Two axes parallel to the plane generate four quadrants where the corresponding entity groups are situated based on their entity descriptors, i.e., explicit-implicit and endogenous-exogenous. Unlike in the reductionism mindset, the role of the descriptors is not to facilitate break-down of the entities into smaller pieces. Instead, it only intends to organize them by articulating their generic, endowed natures. The purpose of the arrows is to show the connections between the quadrants, or entity groups, where primary influences that have unambiguous relationships are indicated by the solid arrows while the dotted arrows characterize indistinct or secondary influences. For example, good economy has a direct influence in the resource network through the stakeholder network, but not exactly vice versa. Also, a vulnerable resource network will likely scale up the probability of disrupting incidents.

The TAF is constructed through organizing (i.e., "networking") the networks and in particular, the networks for resources and stakeholders give the TAF a system-of-systems character. The transportation stakeholder network embodies the decisions concerning the status of the NTS, while the resource network determines how the NTS is actually configured when accessed by consumers. The dual network effects are co-mingled and *evolve* over time with the *evolving* TAF. Then, the type, structure and attribute of the networks can be treated as the architecture design parameters to the extent that such freedom is consistent with reality. Overall, the centrality of constructing the TAF stems again from the recognition that the organization of things can be just as important as the nature of things to be organized.

4.2.3.2 Modeling Hypothesis

While there has been many recognitions pointing out the importance of the holistic perspective[6], not many become reality in tackling a system-of-systems problem. This is partly because it is difficult within the holistic perspective to formulate the necessary computational model in order to generate a useful, quantitative output to make an informed decision for the decision-makers. Richard Bellman insightfully put it like this: the right problem is always so much harder than a good solution (Bellman & Landauer 2000). So, the next focus for the research should be on how the entity-centric abstraction framework realizes its value. In response to bridging the gap between the proposed abstraction framework and a working computational model, the following modeling hypothesis is proposed:

> *A modeling methodology treating the four major classes of transportation architecture entities can be created to synthesize alternative conceptual solutions and facilitate evaluation of the alternatives against multiple criteria.*

For testing the above hypothesis, the following four criteria are essential and they are:

Efficacy: The methodology must lead directly to required products and support efficient decision-making.

[6]For example, Keating et al. (2003) attempts to expand the scope of the traditional Systems Engineering while developing the concept of System of Systems Engineering (SoSE).

Flexibility: The methodology must be amenable to change in response to new customer requirements, new modeling constructs or new dynamics that emerge.

Comprehensibility: The methodology must be understandable, usable and interpretable by non-experts.

Traceability: The methodology must make transparent the rationale and paths taken towards decisions reached.

The efficacy of the methodology can be evaluated by how well it represents the characteristics of the SoS. For example, it must capture the time variant nature of the problem, including the simulation of latent effects due to the distributed nature, feedback mechanisms and the consequences of uncertainty. The resultant model must also embrace sufficient flexibility to support the emergence of revolutionary resource entity designs, the ability to impose or remove constraints easily and the capturing of all types of architecture design variables (vehicles, travelers, infrastructure, etc). Overall, the decision-support method must be able to adaptively employ the balanced level of abstraction that gives meaningful results without becoming overburdened by confounding detail—i.e., it must be comprehensible. Finally, an often overlooked trait, but one that is generally found to be very important, is decision traceability. The ability to present rationale and trace the history of decisions reached can increase the legitimacy to external parties. For the near term, the modeling hypothesis will remain unproven until significant research can be conducted. It does, however, point to the specific requirements that can guide the search for confirmation. To this end, initial research investigations should begin through the exploration of subsets of the governing problem.

4.2.4 Recapitulation

The proper comprehension, and then design, of a transportation architecture represents a tremendous challenge, one that surely requires the wisdom and innovation out of many different research communities. To meet this challenge, academia should provide frames of reference, thought processes and problem formulations. It is from this motivation that the present section has been conceived.

The primary premise was the necessity of the holistic perspective. Under the expected high degree of complexity for this problem, the entity-centric abstraction framework was proposed to study potential transportation architectures without prescribed boundaries. The four instances of entities abstracted are the network of resources, the network of stakeholders, the drivers and the disruptors. The concept of the transportation architecture field (TAF) was set forth to properly connect them. It appears that the abstraction described is universal, covering any conceivable particular architecture. The modeling hypothesis are directed towards the ultimate purpose of an ability to compute a wide variety of value metrics to delineate between alternative architectures. But more broadly, it is the hope of the author that the ideas will also spur interest and facilitate research collaboration between normally disconnected disciplines: aerospace engineering, civil / transportation engineering, business, public policy and so forth.

The ultimate goal of the abstraction framework is to guide the transportation architects in enabling the characterization of the time dependent nature of problem, the existence of feedback with the system, and possibly emergence of visionary designs. Further, the model must be expanded to cover the entire spectrum of stakeholder behavior, network structure and resource characteristics. This expansion will be based upon an object-oriented philosophy. The modular objects (*everything*) on the problem domain (*table*) can then be linked together easily, in order to examine the properties of the larger architecture. Complex behavior can then be analyzed within individual modules or from a SoS perspective.

4.3 Construction of the Virtual World

The modeling hypothesis posed in §4.2.3.2 has provided motivation to integrate all entities that constitute the transportation architecture. It ultimately leads to formulating a working computational model—the missing "link" of the holistic perspective. This computational model is called a *virtual world*[7] in reference to integrated digital programs with its own physical and/or non-physical rules and circumstances. It is obvious that the validity of the modeling hypothesis hinges on how much this artificial world can be constructed in a complete and realistic way. Such a formidable task should take benefits from a large amount of substantial research in various disciplines and domains.

Therefore, the main theme of this section revolves around laying out a tactical blueprint to facilitate a collaborative construction of a virtual world with necessary modeling fidelity and granularity. Guided by the entity-centric abstraction framework, the idea of representing the modeling entities is described first, and the methodology on integration of these entities is addressed next. The section concludes by examining the issue of verification, validation and accreditation (VV&A) for the present approach, which is critical for any modeling efforts.

4.3.1 Entity Representations and Integration

The abstraction process produced the concept of entity—a set of attributes, functions, interfaces and sentience with a flexible boundary. In representing the entities of the transportation architecture, these four basic rudiments will be elaborated as needed. In doing so, the object-oriented thinking will play an important role since entity can be thought of as an extended form of object as discussed earlier. One benefit from this exercise is that the natural connection from modeling to implementation is guaranteed, which is truly important when it comes down to programming.

[7] In computer science, this term is usually related to *virtual reality* and *artificial life*, concerned with design of three-dimensional graphical spaces and simulation of living organisms on computers, respectively. (Heudin 1998)

4.3.1.1 Resource Network

The resource network is a complicated web of vehicles and infrastructure, providing the means to transport people or packages from origin to destination, essentially enabling door-to-door trips. The function of the resource network is supported by the operations of vehicles, portals, and enroute space that connects spatially separated points. These most essential entities of the resource network are visualized together in Figure 43. Three transportation modes are portrayed envisioning a unit travel mission profile where the term mode refers to a particular choice vehicle combined with the corresponding infrastructure. Two existing modes (airline and car) are most important in terms of traffic volumes as the emphasis is on long-distance passenger trips. A "new mobility mode" is infused into this unit network as the focal point to explore future transportation architectures. The capability of entity representations should aim towards expressing not only well-defined existing modes but also a hypothetical mode as shown here.

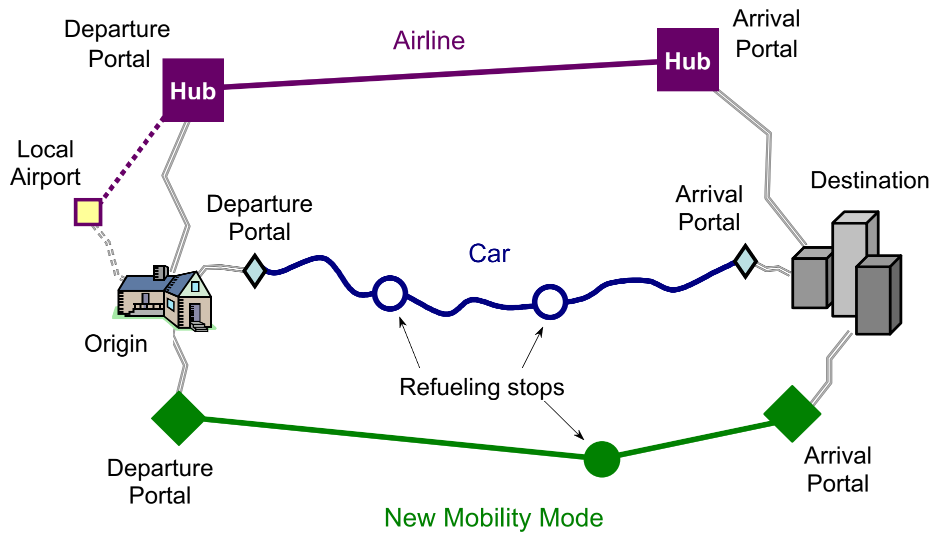


Figure 43: Simplified Resource Network

Vehicle is a primary entity of the resource network to a traveler. Portals and enroute space, often called infrastructure together, provide the settings in which a vehicle operates. A ground vehicle is on-demand, cost-effective and suitable for daily, short-haul trips while a business jet aircraft offers the most time-efficient method to travel from coast to coast.

Despite their distinct characteristics, each vehicle can be regarded as an object that encapsulates its own functions and attributes including technological / economic characteristics. A brief synopsis of vehicle's attributes is listed in Table 12. This is a template for representation of any type of vehicle. Other 'soft' factors—including vehicle comfort, perceived prestige and safety, emissions, 'coolness' factor, security concerns, or practically anything else—can be qualitatively modeled and added with a certain ordinal scale. Finally, the interfaces should prepare connections to account for the influence on vehicle's attributes and functions from various factors in the transportation environment (e.g., fuel price, labor cost, traffic).

Table 12: Vehicle Attributes

Technical Performance	Economic Characteristics	Infrastructure Compatibility
Cruise speed	Acquisition cost	Types of portal
Maximum range	Direct operation cost	Types of enroute space
License requirement	Insurance / maintenance cost	Need of secondary mode
Payload capacity	Price / fee schedule	
Weather resistance		

Portals refer to the transition points between modes of transportation (except for car mode). They can be airports, bus and train stations, highway on-ramps, or whatever portal types are required by new forms of travel. A portal can be characterized by the type of vehicle it accommodates, location, maximum throughput per given time period, construction time/cost and required resources for operation. The operational scheme of portals varies: e.g., airports operate under the centralized control system and on a scheduled basis whereas highway ramps accept on-demand traffic without a control tower. All these features constitute a portal entity, defined by its own attributes, interfaces, and functions. Among the various attributes, the most important ones include time-related characteristics regarding transportation activities such as processing time for boarding a travel method, waiting times and portal delay. These characteristics combine to take up the majority of the non-moving portions of travel. Some representative attributes related to time are broken down in a generic way in Table 13. The combined time at the destination portals are less than those at the origin portals since the wait-ahead portion becomes negligible.

Table 13: Portal Time Attributes

Time-breakdown	Descriptions
Mode connection	Required time to transfer from/to secondary mode
Wait-ahead	Required time for most scheduled services
Wait-in-line	Required time for processing ticketing, baggage claims & security check
Portal delay	Undesirable waiting time due to capacity limit, weather, etc.

The *enroute space* of the infrastructure is made up of air routes, highways, rail roads, etc. Also, part of enroute space are support points *en route* (for rest and refueling) that have their own effects on *block speed*—the ratio of trip distance to combined travel time. The enroute space can be conceptualized through an entity representation as well. For example, a path-length parameter can be introduced to account for non-linear trajectory between points due to topographic, operational circumstances as travelers make ways through. The time-related interfaces can also be constructed to allow the inclusion of an array of delays and slowdowns possible in the course of traversing any physical portion of the NTS. Refueling time, climb profiles, intra- and inter-city traffic and other transient factors are several examples. Each enroute space has a particular degree of construction cost required, autonomy level, disruptor susceptibility and so forth. A portion of their characteristics is listed in Table 14.

Table 14: Enroute Space Entity

Attributes	Interfaces
Types of portals and vehicles	Refueling/rest points
Path-length parameter	Enroute delay effect (inter- and intra-city)
Construction cost	Influence from weather effect
Operation cost & rule	The amount of vehicles that accommodates

The portals and enroute space share common characteristics: they are stationary, expensive to build and many stakeholders in multiple levels have to draw consensus to construct them. Also, they have their own secondary properties. For example, a non-towered, rural airport is more susceptible to adverse weather than a hub airport at a large metropolitan area. Similarly, unexpected catastrophic events may have different effects at different locations and for different types of portal and enroute space.

In summation, it has been discussed that each element in the resource network can be described by the entity representation. Such generic abstract approach offers the capability of synthesizing the previously mentioned “new mobility mode”—any conceivable transportation system regardless of its being advanced fixed wing aircraft, high-efficient magnetic leviation train, or roadable rotorcraft. In formulating a new mode, one should keep in mind that a harmony of the three basic entities is much more important than their capabilities alone to achieve overall transportation goals.

4.3.1.2 Stakeholder Network

The entities in the resource network have a set of properties or attributes as they exist in the tangible form. In contrast, the organizational entities introduced earlier—the stakeholders—need a different treatment since representation of their sentience as well as their interconnections are the key challenge.

Various approaches can be exploited for the stakeholder network. Firstly, it is possible to adopt a System Dynamics approach as mentioned in §4.2.2.2. The internal working of a stakeholder can be captured by using a causal loop diagram and stock-and-flow analysis. Then, linking this internal web to other stakeholder webs completes the overall representation of the network structure. This approach is not, however, an amenable one since it requires a priori specifications of the entire network as a monolithic structure. The use of agent-based modeling (ABM) is a well suited approach for manifesting the complicated behaviors of a collection of sentient entities. The stakeholders in the NTS are “agents” by any sense, and can be modeled as such through the analysis of goals and behaviors—i.e., manifesting sentience and functions.

Taking an example for illustration, the consumer stakeholder’s (traveler’s) ultimate goal is to complete trips comfortably and safely with less travel time and money spent. A series of actions should occur to fulfill this goal. The most explicit one is moving the travelers themselves in a vehicle with their own route choices on the journey.[8] However, it should be recognized that the travelers must go through some sort of mental activity to actually "get there". They have to select the most appropriate mode from the many available transport options; or perhaps even before that, they need to plan (or cancel) the upcoming trips based on monetary constraints or other changing situations. These underlying mental activities are the embodiment of sentience and are established by implementing a set of logics with decision-making algorithms.[9]

Likewise, different types of stakeholders can be treated in the same fashion. The mental model of each stakeholder entity can be analyzed and modeled independently. Then, a collection of these mental models can be organized by constructing a multi-agent architecture (MAS) as illustrated in Figure 44. The role of the information layer is to provide a route through which the stakeholders can communicate and interact. In essence, the information layer encompasses the interfaces of stakeholder entities of interest. Then, in theory, the final form of the whole organization—i.e., the stakeholder network—emerges as simulation progresses. Benefitting from ABM's inherent flexibility in constructing the model, the modeling boundary of the stakeholders and their network can be easily extended with a varying degree of the fidelity, on a need basis for the problem scope of a modeler's interest.

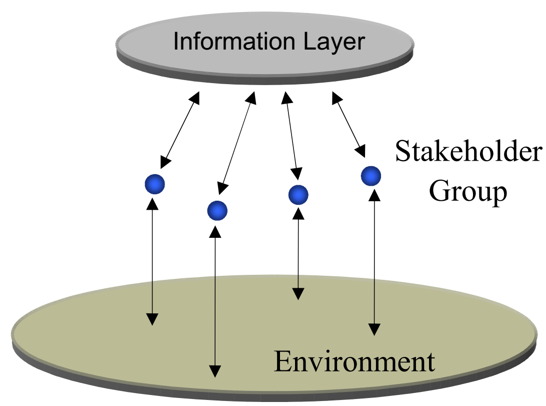


Figure 44: Representation of the Stakeholder Network with ABM

[8] To understand this behavioral pattern in detail, an investigator needs to perform discrete event simulations attached to a physical environment model.

[9] If this is the case, an abstract representation of a physical environment may be sufficient since information is the real currency of interest. This motivates to formulate the concept of *locale* to be introduced in §4.3.1.4.

4.3.1.3 Exogenous Entities

The previous discussion highlighted that entity attributes are the most salient rudiment for the resource network while the key issue for the stakeholder network concerns representing sentience. The focus of representing exogenous entities is on building interfaces in the sense that, as implied in Figure 42, each exogenous entity has its ontological meaning when it has the right connection to other entities in the transportation architecture. For example, the economic situation governs household income, which becomes a basic attribute of an individual consumer. Hence, the modeler not only identifies key exogenous entities that govern a great measure of transportation phenomena but also presents them without (or with minimal) loss of the intrinsic contents and connections.

Drivers

As explained in §4.2.2.3, the driver entities are the underlying sources of the stakeholders' behaviors from economic, societal and psychological motivations. The economic drivers consist of various factors in a wide variety of forms. For example, in order to aggregately measure the nation's economic condition, both Gross Domestic Product (GDP) and Consumer Price Index (CPI) may be used as the simplest top-level scalar metrics. Gasoline price is a simple scalar as well, but time- and location-specific variation should be considered to investigate microscopic behaviors of the stakeholders. In contrast, there exist some factors with disaggregate nature. Household income varies across an individual household, so this entity must be treated in a table or distribution format as shown in Table 15. Besides purely economic concerns, demographic factors exert a significant force driving the overall demand profile of transportation activities: the locations in which people live, the number of members of an individual's household, level of education and age/sex/worker composition of population—all factors can be encapsulated in a multidimensional table. The model granularity heavily depends on the level of aggregation of these driver entities.

On the other side of the driver group, cultural and psychological entities offer a challenge at a different dimension. A working model rarely exists showing quantitative effect

Table 15: Historical Household Income in Percent Distribution

Year	Income brackets in thousand 2001 adjusted U.S. dollars [0,5)	[5,10)	[10,15)	[15,25)	[25,35)	[35,50)	[50,75)	[75,100)	100+
1991	3.1	7.8	7.5	14.3	13.6	16.7	18.7	9.4	9.0
1992	3.5	7.8	7.9	14.7	12.9	16.5	18.6	9.1	9.0
1993	3.6	7.7	7.9	14.4	13.3	16.4	17.9	9.2	9.7
1994	3.3	7.5	7.8	14.8	12.7	16.4	17.9	9.5	10.0
1995	3.1	6.9	7.8	14.5	12.9	16.3	18.4	9.8	10.3
1996	3.0	7.0	7.8	14.2	13.0	15.6	18.8	9.9	10.8
1997	3.1	6.7	7.3	13.9	12.5	16.0	18.5	10.2	11.8
1998	3.0	6.3	7.0	13.4	12.6	15.4	18.7	10.7	12.7
1999	2.7	5.8	7.0	13.3	12.3	15.5	18.6	10.8	14.0
2000	2.9	5.8	6.9	13.0	12.5	15.4	18.6	10.8	14.1

resulting from them, but the basic methodology of representing these entities is not complicated. The strength of their attributes can be mapped by ordinal scales, and the modeler needs to define the relationship between the scale and each stakeholder's interface under the proper assumption. For example, perceived security can be represented on a one-to-ten scale, and this scale can be used in shifting a consumer's desire to travel. While this work might be arbitrary, it is significant that modeling of "less-quantifiable" factors can be accomplished by perturbing the strength of the relationships, which is a relatively straightforward calibration process. Survey studies and subsequent analysis would be easily accommodated through this avenue if such results become available.

Disruptors

The primary effect of the disruptors is related to the efficiency of the resource network. These undesirable entities can be considered as an instance of discrete events, although their cascading consequences are likely to resonate over time. By and large, these events boil down to a few elements: location, strength, duration, and locality (coverage). All these elements are associated with uncontrollable nature for which probabilistic treatment can account. The Monte Carlo Simulation can be employed to invoke disruptor entities in the transportation environment, supported by calibration from empirical data. In actual simulations, each disrupting event will have varying degree of effect and influence a particular portion of the resource network. For example, automotive travel is generally resistant to inclement conditions, while air travel is sensitive to short-term changes in weather, especially for one using general aviation aircraft. This whole mechanism can be completed with incorporation of thresholds within the interface of the resource entities, shown in Figure 45.

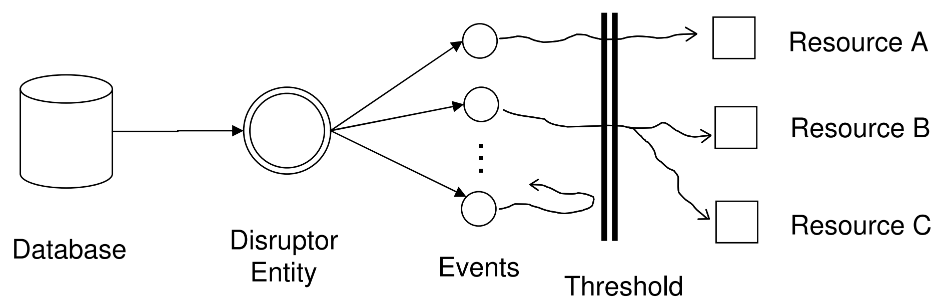


Figure 45: Disruptor Mechanism

The influence of the disruptor entities is two-pronged as one recalls Figure 42. Some events directly affect the psychological driver, which is especially evident when the extent of harmful effect is far-reaching. Representing this mechanism follows in the same fashion as in the above figure. If the psychological driver is modeled with the scales that directly affect the stakeholders' behaviors, the inclusion of thresholds in the driver interface can properly model the relationship between the disruptor and the driver, eventually towards alternating the stakeholders' behaviors.

In summation, the exogenous entities encompass a wide variety of elements in the transportation environment affecting the resource network and portion of the drivers. Hence, treatment of this heterogeneity is not straightforward but possible to the extent the modeler desires, supported by the entity concept. The proper use of databases and extraction of relevant data are the key tasks in representing exogenous entities, especially for economic and societal drivers. The nature of these externalities—uncontrollability and ambiguity themselves and over time—can be tackled through stochastic treatment. Further, 'soft' factors within drivers and disruptors can be modeled with incorporation of interfaces that should be placed so that a relevant entity can access entity of its interest. For example, again, while GDP has a valuable meaning to the policy makers, the amount of disposable money is much more important to individuals. Finally, striking a balance between the amount of detail and the scope is very important in adjusting to fidelity level of entity representations.

4.3.1.4 Integration with Locales

Formulating a virtual version of the NTS, where all entities reside, is the final goal in creating a computational model. As noted previously, an arena is needed where all entities—regardless of being in the explicit/implicit or endogenous/exogenous category—are synthesized altogether in order to have concrete, physical, real meanings.

The concept of locale is introduced as a top-level building block of final integration. The locale is an abstract representation of a unit geographic environment, encapsulating its own transportation resources and stakeholders, economic and societal circumstances, and disruptors as portrayed in Figure 46. It can represent a state, a county or an area with the same zip code depending on desired level of granularity. Overall model fidelity depends on how accurately locales are modeled, so sufficient detail is desired but balanced against the need for operational (computational) feasibility.

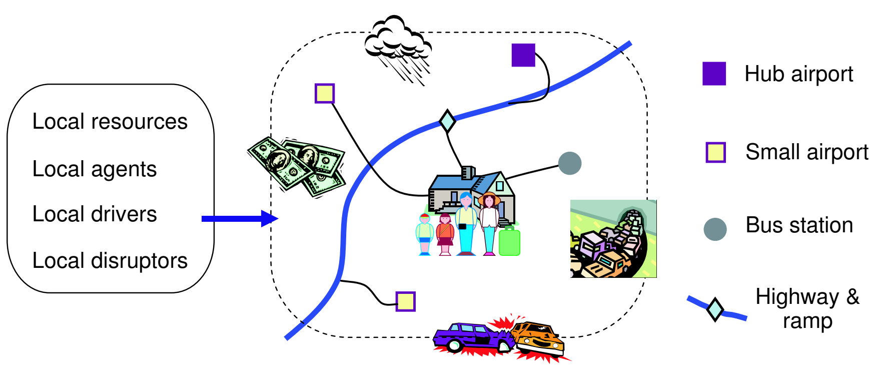


Figure 46: A Unit Locale

The construction of a particular locale is very similar to the process of creating an object from class in object-oriented programming. First, four entity groups should be defined as global components. In doing so, these global components should be tied with the real world for a particular time period of interest. As discussed, they can be any objects: simple scalars, matrices, probability distributions or a real and/or logic functions depending on their nature. A unit locale is a composite instance built from the global components, and thus inherits most of their original properties.

Now, the transportation environment is a set of N discrete locales with appropriate topological information with many heterogeneous objects that the stakeholders can interact with directly: vehicles, portals, events of delay, and so forth. Care should be taken since, for each locale, some properties should be tailored to reflect specific conditions for its respective entities. The overall resource and stakeholder networks are synthesized upon completing the creation of locales. As simulations progress, the collective behavior over the entire system can be fed back into the global components. This information then affects and changes the global components themselves, which in turn updates the locales where new sets of local agents are populated. Just like the real NTS, the whole model will possess its own collective sentience. This completes the conceptual mechanism of the virtual NTS, or the simulation "universe" as portrayed in Figure 47.

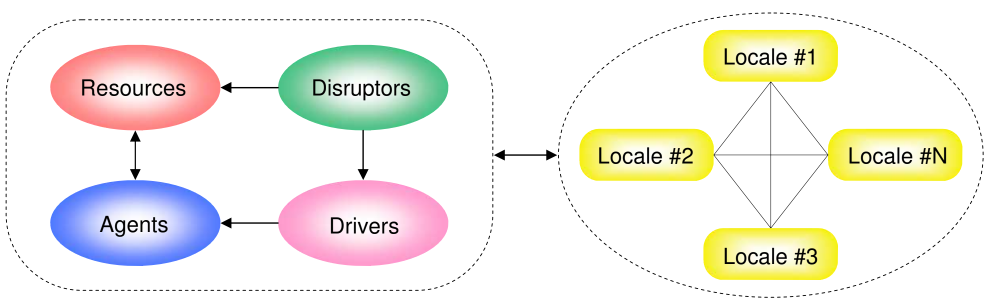


Figure 47: Simulation "Universe"

This interlocking locales and the four entity groups are the final form of the virtual world. The organization of this simulation "universe" offers the simplicity while allowing a remarkable amount of information processed. A wide variety of interactions and elements within the transportation environment at a host of different levels can be treated with enhanced flexibility and traceability. This feature will offer the manageable complexity of implementing a better simulation granularity and fidelity, or even a "new universe" in response to a need for examining a totally different situation.

4.3.2 The Legitimacy of the Virtual World

The structure of the adopted analysis, the integrity of the developed logic and the achieved empirical relevance establishes the legitimacy and originality of any theoretical or methodological study. It is, however, an unavoidable fact that any conceivable virtual world is merely 'a model' of the true transportation architecture despite the best combination of strategy and effort. This proposition generates an immediate issue pertaining to the confidence boundaries and limitations of the computational model. This section is devoted to contemplate this issue before implementing a working virtual world.

4.3.2.1 Verification, Validation and Accreditation: Definitions and Paradigms

It is obvious that the modeler, the developer and the user together are concerned with the degree to which the virtual world and its results have credibility. They may ask: "Are we building the model right? Are we building the right model?" This concern is universal for any modeling and simulation activity and it eventually revolves around the terms *verification*, *validation* and *accreditation* (VV&A), although mistakenly used interchangeably on some occasions. The present research accepts the definitions of these terms as stated in DoDI 5000.61[10] and they are:

Verification: The process of determining that a model implementation accurately represents the developer's conceptual description and specifications.

Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Accreditation: The official certification that a model, simulation, or federation of models and simulations is acceptable for use for a specific purpose.

The specific meanings of VV&A become even clearer considering the simple modeling and simulation (M&S) paradigm, illustrated in Figure 48. Starting from *reality*, an investigator

[10]Department of Defense Instruction 5000.61. The instruction adopted the definitions suggested by Williams & Sikora (1992).

builds a *base model* that is the investigator's image or mental model of the real world (*analysis & modeling*). It should be, at least conceivably, capable of accounting for the complete behavior of the real world. The base model can be represented by mathematical equations, logical expressions or a composite of both. The *computer model* is built from the base model through an iterative process (*programming & debugging*). After the computer model is up, the investigator executes the code and tunes up certain parameters within the computer model to match reality (*simulation & calibration*).

Within this paradigm, verification refers to the activity of keeping the base model and the computer model consistent as the computer code is being built. Verification is a required premise for validation that has two different implications. The *operational* validation is an attempt to make the computer model and reality isomorphic, after which the *conceptual* validation is automatically guaranteed due to the transitivity law. Accreditation is the final, authoritative confirmation after verification, validation and other necessary processes are completed with satisfaction.

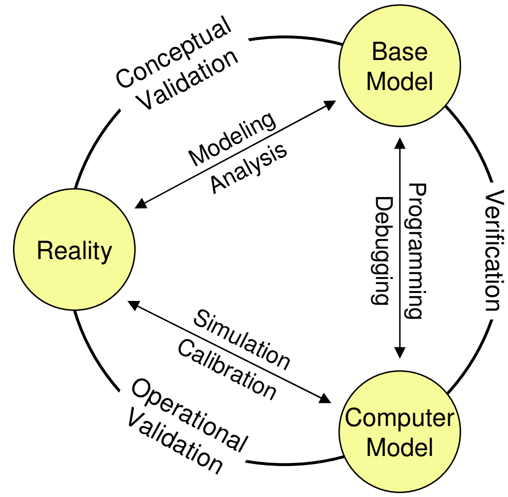


Figure 48: Simple Modeling and Simulation Paradigm [Based on AIAA (1998) and Sargent (2000, Figure 2)]

Usually, the cost of VV&A is quite significant, especially when extremely high levels of model confidence is desired. (Sargent 2000) Hence, the level of confidence should be compromised against the overall cost. The user decides the level of VV&A activity necessary to ensure that the computer code yields *acceptable* results.

We may add that it is not sufficient to completely clear up the VV&A issue since a critical question concerns what reality is. One must be very prudent in characterizing reality because human beings possess insufficient capabilities for grasping its entirety. In other words, one can only see a subset or partial image of the real world. In addition, in gathering raw data, measurement errors can occur and outliers can exist, so a certain processing is required to filter the raw data into a refined form. Therefore, the base model is constructed not from the real world but a proxy world, a collection of refined data obtained from experiments, observations and other available methods. Hartley (1997) and Sargent (2001) recognize this point which is reflected in Figure 49.

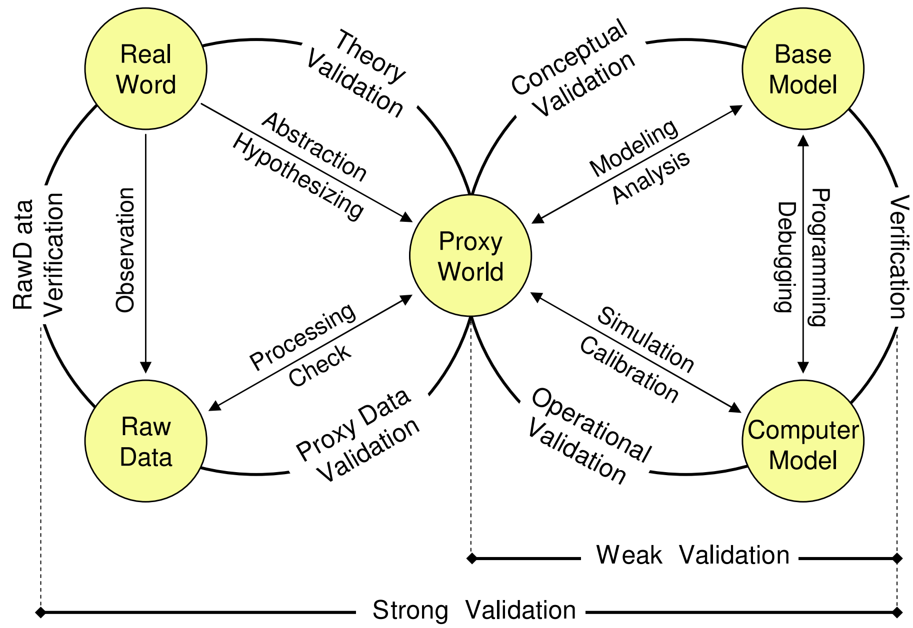


Figure 49: Expanded Modeling and Simulation Paradigm [Source: Hartley (1997, Figure 1) and Sargent (2001, Figure 2) with re-interpretations]

Under this expanded paradigm, the majority of the modeling and simulation activity is based on "weak" validation where the computerized model becomes logically equivalent to the proxy world with compromised cost and confidence. If the primary interest of the investigator is on a physical system, an observation is typically done by prescribed experiments which are supposed to produce consistent results, all else being equal. For example, in the study of natural sciences, the proxy world is equivalent to the real world under some assumptions and conditions. If this is the case, a further extension can be made. That is, "weak" validation spans to "strong" validation meaning that the base model now is equivalent to the real world. This generalization yields a complete structure of knowledge that academia often accepts as a law or a theory.

4.3.2.2 Validation of Agent-Based Models

Given the principles of VV&A, it is evident that the scope of the present research is not assuming accreditation or "strong" validation. Verification and "weak" validation, however, must be in place. Perceiving verification as a doable task, although significant effort is required throughout the life cycle of the model, validation of the virtual world is by far a difficult task. Unfortunately, the difficulty starts from the fundamental premise of agent-based models: Is it possible to validate agent-based models? How do we validate them? This concern has not yet been fully concluded among researchers, despite the widespread acceptance of the agent-based approach. In brief, two schools of thought exist regarding this issue.

Lane (1993) asserts that "[...] it is possible that the causal mechanism hinted at in the CAS (complex adaptive system) is swamped by the additional 'turbulence' in the real world, and some entirely different sets of interactions of direct effects drive the formation of the feature of interest." This viewpoint, although logically sound, generates skeptics towards ABM/S, especially among those trained with and used to rigorous mathematical formalisms. To the skeptics of ABM, the approach is too mystic for their scientific taste and thereby the simulation result would be "fragile", meaning that "it is often not easy to find out whether model result is a mere artifact of specific parameter configurations or the really meaningful result." (Cederman 1997)

On the contrary, other researchers argue that a new perspective is desired as the complexity science grows in accordance with "the death of positivism". (Henrickson & McKelvey 2002) These proponents of ABM/S insist that is not appropriate to find out whether the postulated process operates in the real world because the agent-based approach produces 'pattern predictions' or 'robust processes'[11] rather than point-like predictions of single events. (Srbljinovc & Skunca 2003) By the same token, they believe that validation of

[11] Goldstone (1991, p. 57) uses this term in reference to a sequence of events that has unfolded in similar (but neither identical nor fully predictable) fashion in a variety of different historical context.

agent-based models is less critical yet keeping a model simple and attaining a certain degree of veridicality is much more important since the question "can you grow it?", instead of "can you explain it?", is better suited for the study of a complex system. (Epstein & Axtell 1996)

Such controversies surrounding the validation issue of ABM do not necessarily justify that the burden of validation should be relieved from the modeler's responsibility, and does not mean that the present research should formulate a generic theory for validating any ABM. Under such circumstances, the best policy is to come up with practical ground rules to lay down the reasonableness of the virtual world.

4.3.2.3 Principles for Validating the Virtual World

A set of guiding principles for proper construction and sound interpretation of the virtual world is established. These principles—although the term may sound pretentious—have the form of the imperative and each principle has the following noteworthy implications detailed next.

1) Keep it simple.

For any modeling effort, this principle is undeniably an essential tenet, but is even more important for grounding a forthcoming validation task of the virtual world. As Carley (1996) puts in her paper, a simpler model obviously demands less levels of validation standards especially when the purpose of the model is to show a proof of concept or to illustrate the relative impact of basic explanatory mechanisms. Also, the simplicity increases the transparency of the model, which in turn secures a better potential for a third party examination and thereby generalization later. Hence, modelers fight against temptations to infuse "modeling artifact"[12] or "stipulative patches"[13], often formulated on an ad hoc basis, just to obtain a 'good' or a 'pleasing' model.

[12] A process, a routine, or a set of parameters implemented in a computational model derived from a common sense for use of model calibration.

[13] This term is coined by Lustick (2000) to indicate portions of programming lines based on some assumption when translating a concept into a computational code.

2) Seek "weak" validation.

This principle is unavoidable for the present research to accept the following assumption—the real world's various aspects are accurately (at least fairly well) represented by the data used in constructing and/or calibrating the model. Unless there are clear and compelling reasons, the modeler should set the various parameters, distributions and procedures such that they match "as is" data. This stringent initialization eventually pays off later, even if a slight manipulation leads to more agreeable simulation results, for accumulated variations in the model parameters likely form fewer opportunities for extended use of the model. Accepting this principle, "weak" validation can be completed by an individual (or a small group of individuals), and will be served as a basis for "strong" validation that requires trans-domain collaborations.

3) Avoid use of mechanism and data that lead to the observed data.

As discussed in §4.3.2.2, if the model has a sufficient number of parameters, some combinations of the parameters can supposedly match the observed data. Thus, it is difficult to conclude if the result is from the real dynamics captured by the model or from a specific parameter configuration, i.e., the model is doing nothing but a non-linear regression.[14] One logical way to evade this concern starts with a prudent construction of the virtual world. The modeler needs to keep the model away from influencing mechanisms that enforce the results to (coercively or unconsciously) match the observed data. To meet this goal, two conditions should be met: First, the model is described, preferably, by microscopic level data or theories only. Second, the model is calibrated against field data that were not used in constructing the model. Satisfaction of these criteria secures the essence of agent-based models—the virtual world will reveal not a preprogrammed result but an emergent behavior, generated from low level interactions.

[14] It is worth mentioning the following counter-argument. Carley (1996) states that "[...] For models where the process is represented [...] by rules, interactive processes, or a combination of procedures and heuristics, [...] there is no guarantee that a sufficiently large set of procedure and heuristics [...] can be altered so that they will generate the observed data."

4) Utilize sensitivity analysis.

This last principle is directly concerned with the extent to which the simulation results has credibility. The use of sensitivity analysis of simulation scenarios and the theory of ordinal optimization underpin the basic rationale of the principle. The ordinal optimization theory, by Ho, Sreenivas & Vakili (1992), is developed to offer an efficient way to tackle mathematically hard optimization problems. The theory is based on the concept of "goal softening", combined with a theorem that comparison of relative orders of performance metrics converges much faster than the performance metrics themselves.[15] The most important corollary of this theory is that a crude "surrogate model" can estimate the relative rank order of various alternatives with a sufficiently high level of confidence.

Now imagine a finite set of simulation scenarios that the investigator formulates to study the effects on a certain scalar metric. The results of the simulation scenarios are plotted as illustrated in Figure 50 where the error bars indicate that the "actual" values may not be equal to those observed from the virtual world. The size of the error bars reversely depends on the level of model fidelity. Regardless of the fidelity level, however, if the investigator's objective is to get the order alignment of the scenarios (i.e., goal softening), the ordinal

[15]On the order of e^{-n} vs. $n^{-\frac{1}{2}}$ where parameter n roughly corresponds to the amount computational time.

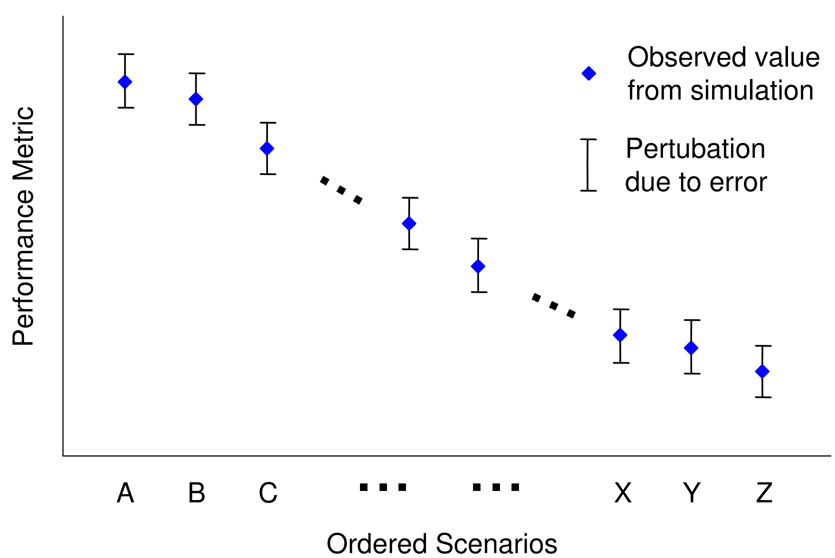


Figure 50: Ordered Performance Chart

optimization theory says that the outcome from this sensitivity analysis has a highly degree of validity. In other words, even the skeptics would feel comfortable to say that it is very likely that Scenario A outperforms Scenario C, not to mention comparing Scenarios A and Z. (See Figure 50.)

CHAPTER V

IMPLEMENTATION

<u>Contents</u>

The central theme of this chapter is the process of making the virtual world operational and, more precisely, translating the conceptual virtual world into the computational model. This translation begins with a brief review of applicable databases on which the scope of the modeling exercise depends. Then, details of the model components are described and their implementations are illustrated, culminating in a presentation of the unified simulation model, ***Mi***.

5.1 Data Sources and Model Scope

As discussed in §4.3.2.1, the basis for a proxy world is formed by reviewing and interpreting available field data. This task may take a small portion of the entire modeling effort, but cannot be overemphasized since the proper proxy world leads to the sound base model and thereby the computational model.

5.1.1 Database Review

The foundational data for model construction can be obtained from the U.S. Census Bureau. The census databases have accumulated a wide variety of features regarding people, businesses and geography in extensive breadth and depth since as early as 1790. The latest effort, Census 2000, provides information about the 115.9 million housing units and 281.4 million people across the nation. Census 2000 is the most applicable database for constructing the virtual world since it not only counted the population, but also sampled the socioeconomic status of the population, capturing a dependable snapshot of the nation's socioeconomic state.

Meanwhile, other government agencies have tracked the trends and characteristics of the NTS. Within the U.S. Department of Transportation (DOT) alone, various databases have been generated through the Bureau of Transportation Statistics (BTS), Federal Aviation Administration (FAA), Federal Highway Administration (FHWA) and so forth. The most useful database from the DOT agencies for this research, as hinted in §4.1.2, is the 1995 American Travel Survey (ATS). The subject of the survey is long distance round trips (100 miles or more, one way) during the year 1995 by all modes of transportation. Although the database is based on a relatively small number of samples[1], it contains wide-ranging information to sufficiently comprehend the big picture of the NTS and the behavior of the traveling public—the consumer stakeholder—on a national scale. Another authoritative source of national data can be found in the Nationwide Personal Transportation

[1] Approximately 65,000 respondents and 260,000 vehicle trips

Survey (NPTS) that predominantly treats the amount and nature of daily, short-range personal travel. This biased scope provides, however, little insight towards the aviation-related issue. Besides these surveys, much of the detailed information about the NTS can be obtained from TranStats[2], providing "one stop shopping" for various transportation data. It federates various multi-/inter-modal transportation databases that can be used to construct the model—including significant effects within the NTS, such as airline delays caused by weather and other factors.

While the above databases constitute primary sources for the present modeling exercise, different sources will be referenced as needs arise, and accordingly cited in place. One difficulty in using disparate sources in the conception of the proxy world is that not all of the data agree on certain characteristics. Household income distribution does not match, for instance, between the 1995 ATS data and the census databases. While this makes sense from the perspective of which households make up the traveling public, it certainly adds a certain amount of ambiguity, so care must be taken with the interpretation and selection of the field data. For this particular situation, the census databases will take priority since a large potion of the descriptive statistics in them comes from the complete enumeration of the nation. This policy remains as a general guideline for perception of the proxy world, which also resonates with the third principle established in §4.3.2.3.

5.1.2 Model Scope

The overall extent of the modeling boundary—the size of "the table"—should be in accordance with the databases to be referenced, and should be sufficient enough to satisfy the need to study the PAV design problem in the context of the NTS. This concern is directly related to the first supplementary research question posed in §2.4.2. In an attempt to answer the question, Table 16 provides a summary of the target modeling scope with respect to several categories.

[2]Extensive web-based database, http://www.transtats.bts.gov/

Table 16: The Modeling Scope of the Virtual World

Category	Description
Unit Period of Time	1 Year
Geographic Boundary	The Continental United States
Transportation Activity	Passenger travel (round trip)
Travel Distance	Long distance (100 miles to 2,750 miles)
Class of Modes	Car, Airline, General Aviation
Type of Stakeholders	Consumer, Service Provider

At a glance, the time and physical boundary of the modeling exercise is quite large and may not be feasible for a microscopic mechanical model. The model will simulate one year as a unit period of time since most statistics on the databases are aggregated measures with one year interval. The continental United States (CONUS) is the modeling boundary in terms of physical space, which excludes transportation activities that involve international areas, the states of Alaska and Hawaii, and so forth as either origin or destination.

Recalling that the main theme of the present research revolves around the mobility goal with a potential PAV system, passenger trips in the CONUS are set as the target transportation activities. Consequently, the lower bound of great circle trip distance is 100 miles due to the 1995 ATS data, and the upper bound is 2,750 miles.[3] The lower bound seems appropriate enough since, for less than 100 miles, a car is an undoubtedly best travel method. Single-destination round trips are examined because one-way trips are sparse and it is hard to extract information on multiple-destination trips out of the databases.

Three classes of transportation modes are included in the model. Personal automobiles, commercial air transport, and general aviation will be considered along with their corresponding portals and enroute spaces. The other secondary modes (trains, buses, etc.) are disqualified as a meaningful transportation mode, as examined in §4.1.2.6 and Figure 36, for long-distance passenger transportation. This is true for general aviation as well, but

[3] Approximate great circle distance from Miami, FL to Seattle, WA. (airport MIA ↔ airport SEA: 2,722 miles)

general aviation cannot be omitted because it is widely considered as a leading indicator for an advanced, on-demand, point-to-point transportation system, and is the main subject of the present research.

There are two types of stakeholders taken into account: consumers and service providers. The importance of the consumers cannot be overemphasized: They are the source that generates all transportation activities and transportation related issues. A unit consumer can be generated from households and enterprises. Other types of stakeholders, such as manufacturers, government, and research agencies, are not included in the model since there does not exist specific data to analyze their behavior in a unified form. It is understandable because their behaviors are too complex, often appearing erratic.

The virtual world is couched in the form of an agent-based model and is focused on passenger transportation, which boils down to key behaviors of the interacting agents (and observable quantities): trip generation (amount) and modal split (market shares). Therefore, the following sections are devoted to description of the agents and transportation environment. First off, we construct and illustrate the environment in which the agents can reside and live.

5.2 *Transportation Environment*

As mentioned in §4.3.1.4, an arena where all transportation related events occur is required to embody the transportation environment. The key building block is locale, an abstract geographic unit in which all entity groups are synthesized together.

5.2.1 Establishment of Four Locales

The simplest description of the structure of the virtual world would be a union of a finite number of locales, where the overall model granularity is directly determined by the number and the size of locales. Given the extent of the NTS, however, it is too demanding, especially for an individual, to describe all the airports, highways, streets and geographic conditions in the nation even though they have well-defined physical characteristics. A simplification process, therefore, is required to attain a feasible model, complying with the proposed scope.

Analysis of Census 2000 data served that purpose. From the database, 465 urbanized areas and 3,170 urban clusters were identified. An urbanized area (UZA) and an urban cluster (UC) are geographic categories defined by the Census Bureau. Loosely speaking, the lower limit of the population size of an urbanized area is at least 50,000; that of an urban cluster is at least 2,500. The third category is the rural area that does not belong to any urban areas, thus typically containing fewer than 2,500 people. Each category breaks down further by population count, adding up to 15 sub-categories (geography codes [1] through [15]) as shown in Figure 51. Sequentially, the value of average household income of each sub-category is plotted in Figure 52 to show the respective economic status.

As mentioned above, a certain degree of aggregation is required to abstractly represent the physical space of the nation in order to have a feasible model. To this end, three population cut-offs were set: 2,500,000, 500,000 and 50,000. Figure 53 shows a log-log plot of the number of population against population rank of 3,564 urban areas in the CONUS.[4] The

[4]Excluding HI, PR, VI, Guam, etc. Including 450 urbanized areas and 3,114 urban clusters only since detailed breakdown for the rural areas was not available.

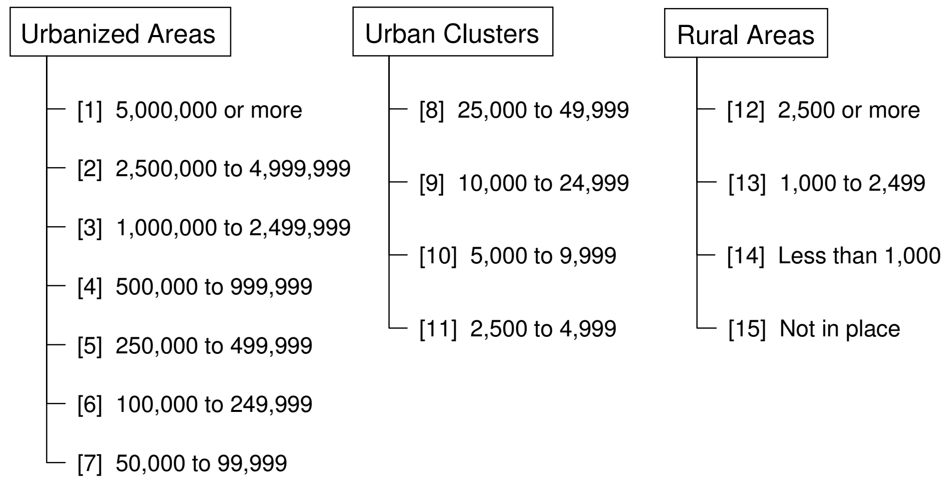


Figure 51: Geography Classification

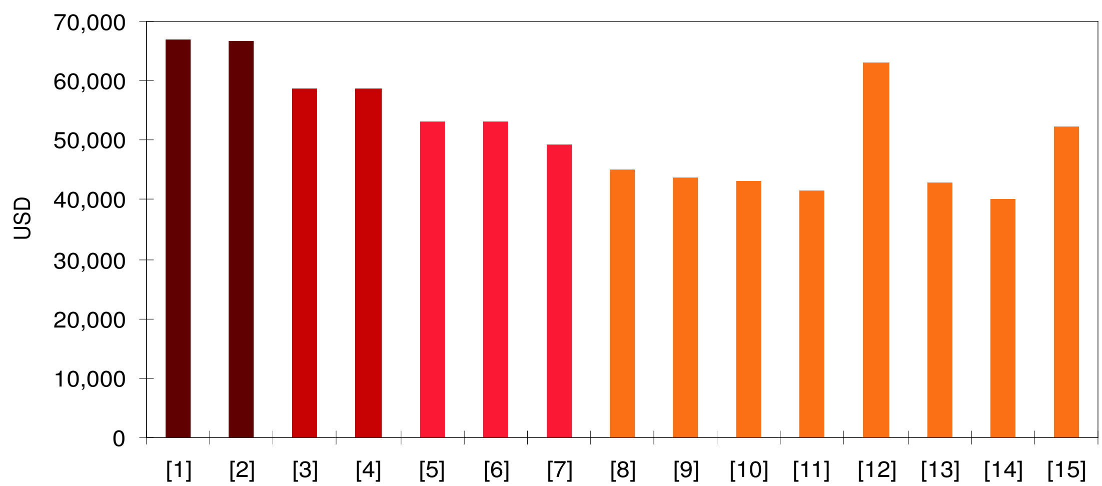


Figure 52: Average Household Income vs. Geography Code

rank-size distribution of the urban agglomerations asymptotically follows a power law.[5] The cut-offs that separate each locale are indicated by the dotted lines.

The rationale behind this classification is grounded with observation of Figure 52. With only two exceptions (codes [12] and [15]), the cut-offs categorize distinct areas fairly well such that economic characteristics in the same group are evenly distributed. Therefore, all geographic areas can be ascribed one of the following locales: 15 Large-Metro (LAR), 55 Medium-Metro (MED), 380 Small-Metro (SML) and 3114+1 Non-Metro (NOM) in a

[5]Often called Zipf's law, considered as one of the most striking features of complex self-organized systems. (Bak 1996)

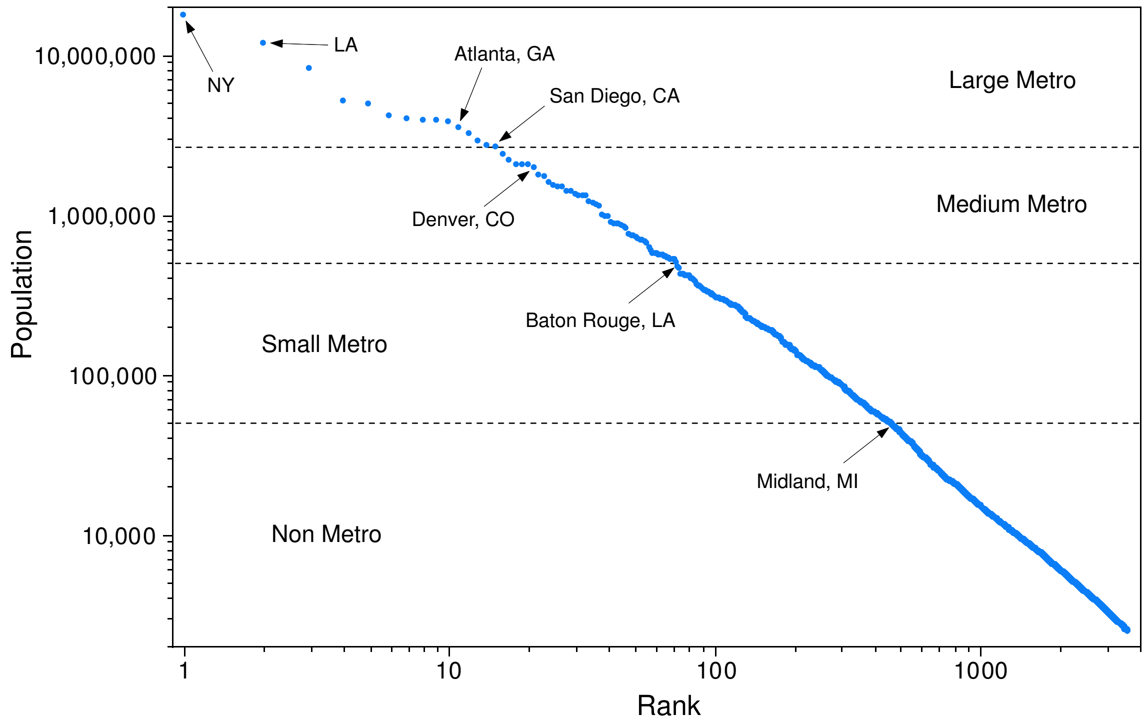


Figure 53: Log-Log Plot of Population against Rank for the Urban Areas

descending order of population size. Detailed information on each locale category is given in Appendix C. Three cities that belong to locales LAR, MED and SML are sampled and scaled in Figure 54. Table 17 summarizes the aggregated demographic and economic characteristics of the four locales.

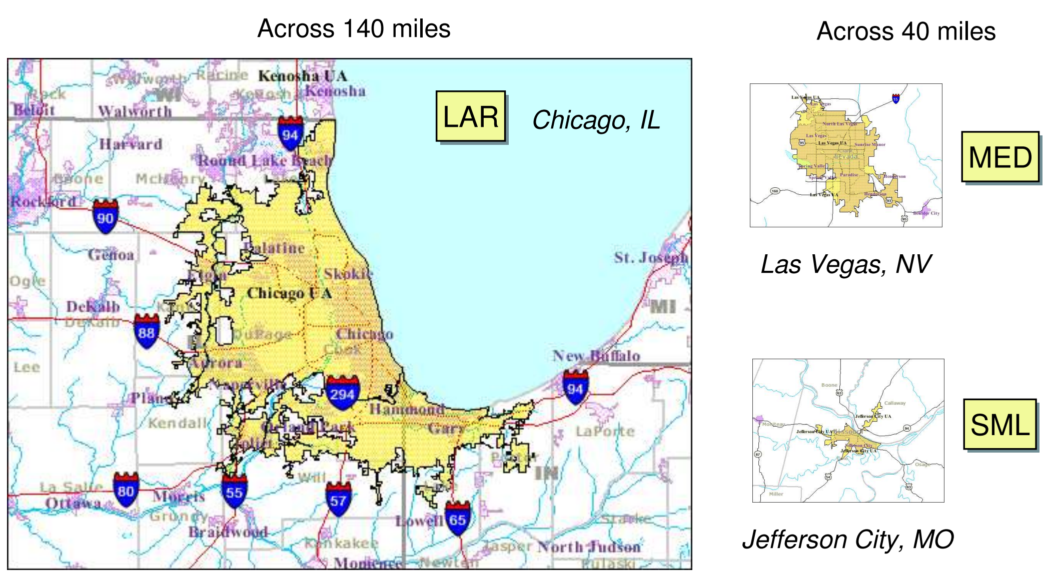


Figure 54: Example of Locales LAR, MED and SML

Table 17: Summary Statistics of Four Locales

Codes	# Population [a]	# Household [b]	[a]/[b]	Average Income
LAR	82,592,615 (29.3%)	30,098,938 (28.5%)	2.74	$66,906
MED	57,429,442 (20.4%)	22,122,360 (21.0%)	2.60	$58,703
SML	52,301,767 (18.6%)	20,011,976 (19.0%)	2.61	$52,150
NOM	89,098,082 (31.7%)	33,305,848 (31.6%)	2.68	$48,701
Total	281,421,906 (100%)	105,539,122 (100%)	2.67	$56,643

Besides close correlation between demographic and economic characteristics, there are other factors that can be shared within the same locale such as accessibility of transportation resources and other transient factors including traffic delays and congestions. For example, a Large-Metro has large-hub airports conveniently located within its territory while a Non-Metro does not have even small-hub airports close to its vicinity. Figure 55 pictorially outlines this concept where a description of each locale is portrayed by the basic elements of the transportation resources, along with a few example cities.

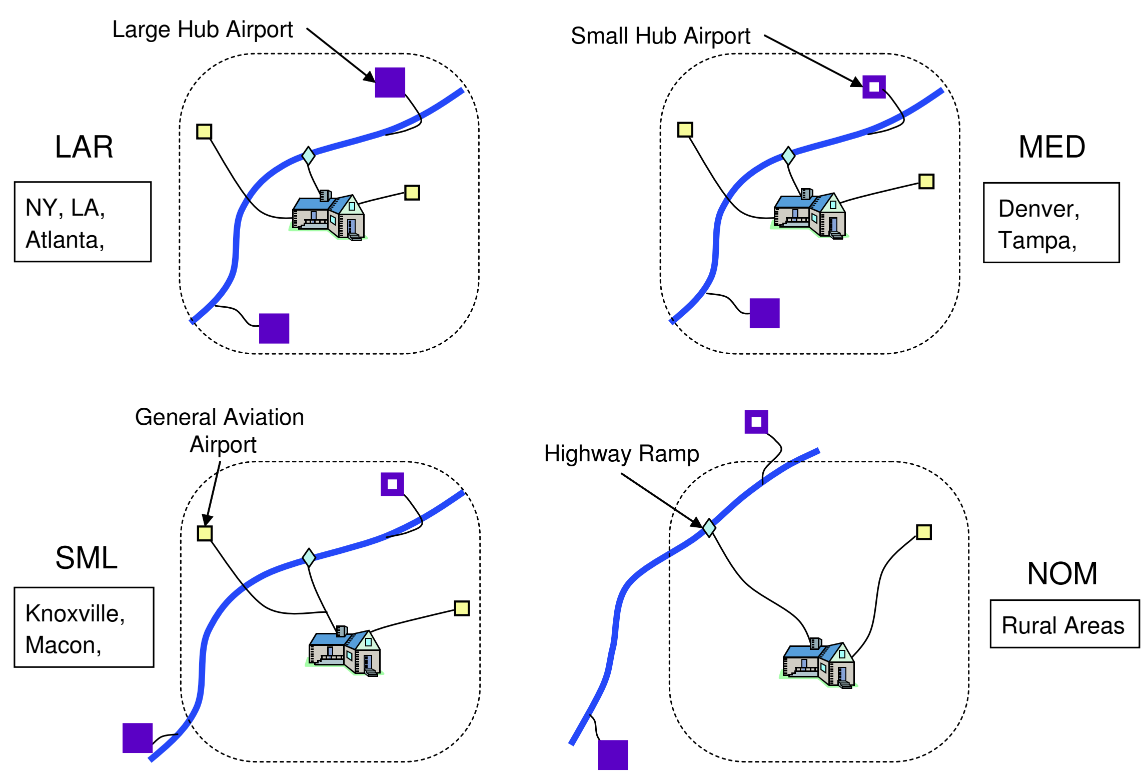


Figure 55: Pictorial Description of Four Locales

5.2.2 Macroscopic Modeling of Transportation Activities

As all the geographical areas in the CONUS are aggregately represented in a finite number of discrete locales, so are the flow and amount of the transportation activities. Another piece of information needed to mimic the physical environment is distance among these locales. These three quantities belong to the driver entity group that governs macroscopic characteristics of the model. Simple theoretical models are constructed and described next using transportation demand theory outlined in Chapter III.

5.2.2.1 O-D Matrix Model

As explained in §3.2.3, information on the flow and amount of the transportation activities is given by an origin/destination (O-D) matrix. Since only four locales are on the modeling scope, a 4-by-4 O-D matrix is generated. The only difference from a conventional O-D matrix is that the diagonal term has a meaning, thus has a non-zero value. Equation 15 is rewritten in a matrix format with the Hadamard product notation, substituting population of locale i and j (p_i and p_j) for μ_i and ν_j:

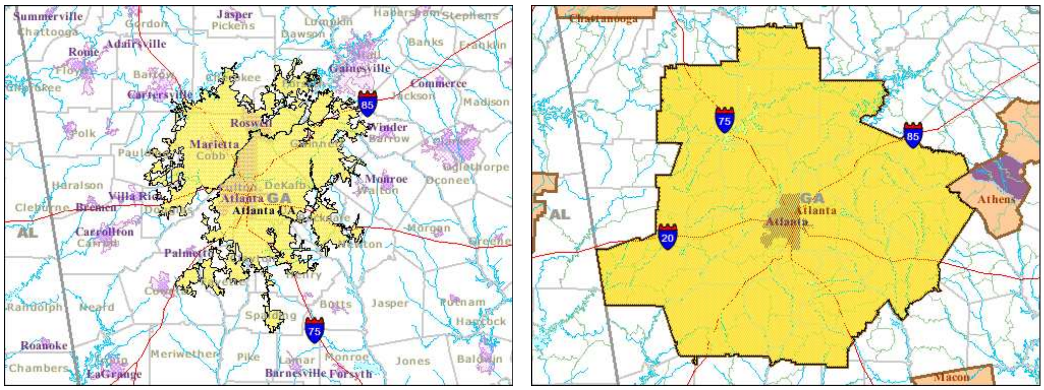
$$[t_{ij}] = [p_i \cdot p_j \cdot C_{ij}] = [p_i \cdot p_j] \circ [C_{ij}] \tag{54}$$

where C_{ij} denotes a generalized impedance and t_{ij} indicates the amount of trips over a period for $i, j =$ LAR, MED, SML and NOM. Given t_{ij} and p_i, the equation contains 16 unknown C_{ij}'s. After getting p_i from Table 17, and normalizing the matrix $[p_i \cdot p_j]$ such that the sum of all elements adds up to 100(%), the result is the following symmetric matrix.

$$[p_i \cdot p_j] = \begin{bmatrix} 8.61 & 5.99 & 5.45 & 9.29 \\ 5.99 & 4.16 & 3.79 & 6.46 \\ 5.45 & 3.79 & 3.45 & 5.88 \\ 9.29 & 6.46 & 5.88 & 10.02 \end{bmatrix} \tag{55}$$

The 1995 ATS data was examined to find out the degree to which the matrix is close to reality. Unfortunately, the data was collected based on Metropolitan Statistical Area (MSA) as a geographical unit, including 162 MSA's and 1 aggregated non-metropolitan area, which does not make an one-to-one match to the Census Bureau's definitions adopted herein (UZA, UC and Rural Area). The following figure is a sampler for this point.

Vicinity Maps around Atlanta, Georgia



(a) Urbanized Area (UZA) (b) Metropolitan Statistical Area (MSA)

Figure 56: Comparison of Geographical Units

Hence, the O-D matrix model contains a certain degree of inevitable discrepancy, even though a large amount of work was put to re-create the 163-by-163 matrix to follow Census definition as much as possible. The final result is shown below.

$$[t_{ij}] = \begin{bmatrix} 8.82 & 7.86 & 3.88 & 11.15 \\ 6.08 & 4.25 & 2.52 & 9.07 \\ 2.62 & 2.55 & 1.13 & 4.48 \\ 6.92 & 8.42 & 4.40 & 15.83 \end{bmatrix} \tag{56}$$

Under the assumption that the above values reflect reality fairly well, all C_{ij}'s can be computed by substituting Equations 55 and 56 into 54, producing the result in Table 18.

Table 18: C_{ij} Values for the Four Locales

O\D	LAR	MED	SML	NOM	Average
LAR	1.025	1.314	0.712	1.200	1.062
MED	1.015	1.022	0.665	1.404	1.027
SML	0.481	0.672	0.328	0.762	0.561
NOM	0.745	1.303	0.748	1.579	1.093
Average	0.816	1.077	0.613	1.236	N/A

The row average value can be considered an indicator of how active transportation consumers in the corresponding locale are. In contrast, the column average shows the attractiveness of the locale as a trip destination. These average values in general are close to one with a notable exception for locale SML. Also, the column average for NOM is relatively high and that of LAR is relatively low. This may be due to a majority of leisure trips being bound for locale NOM and a lack of desire on the part of the consumers in locale SML for long-distance travel. The C_{ij} values computed are taken as constant in the model. That is, if population changes occur, a new O-D matrix should be generated with variations in the matrix $[p_i \cdot p_j]$ only.

5.2.2.2 Trip Distance Profile

Specification of distance between origin and destination completes the macroscopic model of transportation activity. One thing for sure is that the number of short-distance trips far exceeds that of long-distance trips. The gravity model captures this effect and Equation 22 is rewritten here:

$$t_{ij} = k_{ij} \frac{p_i p_j}{{d_{ij}}^{\gamma}}. \tag{57}$$

Let the subscripts i and j be dropped out with an assumption that the size and the population density is constant throughout all geographical units. Also, suppose that all locales are distributed evenly in a corridor and the attraction factor k_{ij} is the same. When the locale at the end of the corridor is considered only, we have

$$t = G \frac{P^2}{D^{\gamma}} \tag{58}$$

where the constants G and P^2 replace k_{ij} and $p_i p_j$ terms and D represents one way great circle distance between locales. The equation has only one unknown G while simplifying the trip amount to have dependency only on D. Suppose that the total amount of trip is T and the distance bracket of interest is $[L, U]$, then the constant G can be obtained from the following constraint:

$$\int_L^U G \frac{P^2}{D^{\gamma}} dD = T \tag{59}$$

which can be used to solve for $t(D)$, and the result is

$$t(D) = T\frac{D^{-\gamma}}{\int_L^U D^{-\gamma} dD}. \tag{60}$$

Further, if T is normalized to 1, the above equation becomes a frequency or probability distribution function.[6] Figure 57 illustrates the probability distributions and the cumulative distributions with $L = 100$, $U = 2,750$ and three γ values: 1.2, 1.6 and 2.0. The value of γ is related to the strength of physical tie between an origin locale and its closer vicinity. As γ approaches zero, which is only possible for an imaginary situation, the driving force of transportation activity no longer depends on physical distance.

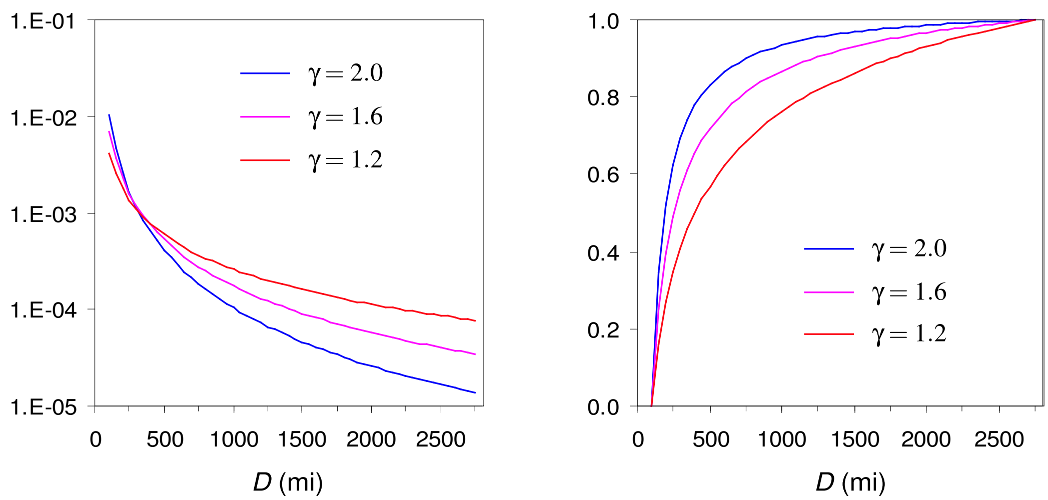


Figure 57: Trip Distance Profiles by Difference γ Values

This distance distribution model is very straightforward, yet deserves some attention. First, in an attempt to complement the above hypothetical assumptions, a scaling function of D can be introduced to more closely imitate the real trip distance profile.[7] Second, the gravity model, in a conventional sense, concerns the captured or observed transportation demand. So, if the modeler wants to consider the effect of unconstrained desires, a smaller value of γ should be adopted. Clearly, it is one of model parameters for calibration use.

[6] In statistics, this is also known as the Pareto distribution.

[7] Although to be addressed in the future, a rigorous approach formulated from an analogy of a multi-body problem can generate a more accurate model.

5.2.2.3 Abstract Treatment of Mission Time-Space

Under the aforementioned macroscopic structure, the next step is to connect spatially separated locales with a microscopic view. This connection is accomplished by utilization of the transportation resources through multiple mission segments that Figure 58 portrays for a door-to-door journey in a conceptual mission time-space.

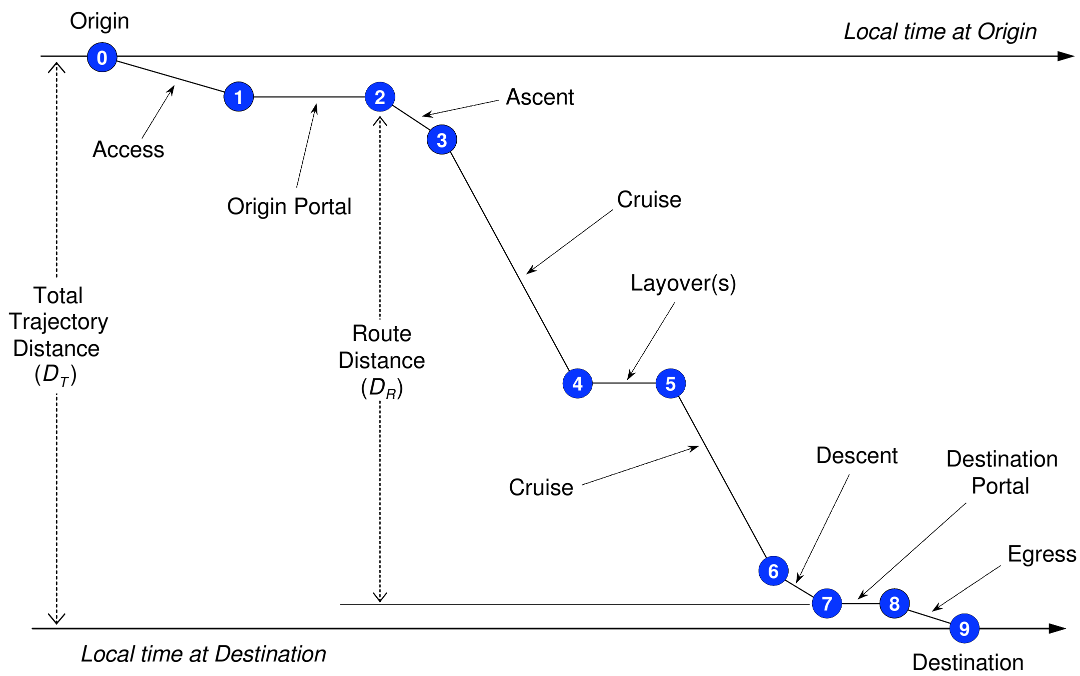


Figure 58: Mission Segments on O-D Time-Space

The choice mode illustrated pauses once to refuel (5^{th} segment, from nodes 4 to 5) during the course and spends a certain amount of wait time at the origin portal and destination portal (2^{nd} and 8^{th} segments). Unless the mode has a dual-mode capability, a certain form of ground transportation, called access / egress mode, should be accompanied to get passengers to and from the portals (1^{st} and 9^{th} segments). This conceptual decomposition helps in organizing time dependency between each mission segment and the resource entities of any mode, as illustrated in Table 19.

Time spent on segment i (t_i) is not only a function of mode chosen. For instance, even with the same origin and departure portals, route distance D_R—measured along with the projected trajectory—can vary since multiple itineraries can exist. Portal attributes

Table 19: Resource Entity Dependency on Mission Segments

Segment	Notation	Vehicle	Portal	Enroute
Access / Egress	t_1 / t_9	✓	✓	
Origin / Destination Portal	t_2 / t_8		✓	
Ascent / Descent	t_3 / t_7	✓		✓
Cruise	t_4 / t_6	✓		✓
Layover	t_5	✓		✓
Delay	τ		✓	✓

which depend on mode as well as locale determine t_1 and t_2. Specifically, probabilistic treatment of distance to origin portal and ground access speed is employed for t_1 calculation to account for diverse situations. As an example, cumulative trip percentage against airport access time by automobile is portrayed in Figure 59 where the dotted line shows the field data gathered around metropolitan area of Boston, MA. The solid line, an approximate match to the data, is a result from a Monte Carlo simulation taking triangular distributions for distance and ground speed.

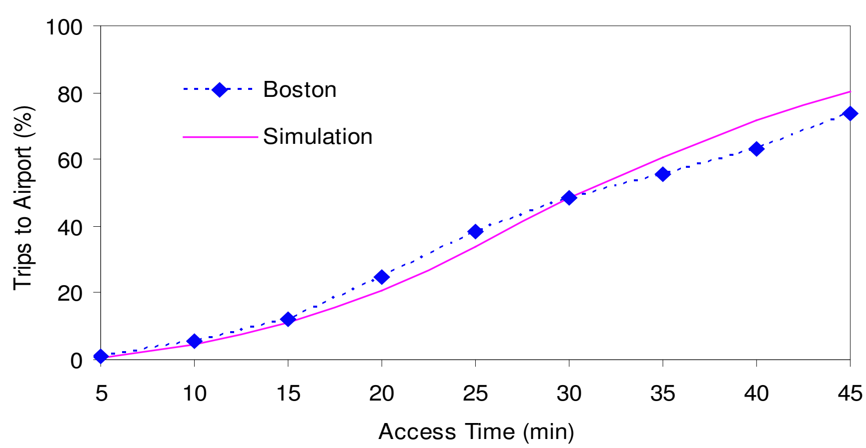


Figure 59: Distribution of Access Time to Airport

Similarly, waiting time at origin portal t_2 can be calculated as a sum of probabilistic time attributes of a portal entity, following the breakdown scheme as previously suggested in Table 13. As opposed to the "pre-flight" and "post-flight" phases, characteristics of vehicle and enroute space entity directly affect time during in-vehicle segments (t_3 through

t_7). For a high-speed mode, a non-negligible amount of time is required to achieve cruise altitude and/or speed, which is illustrated in Figure 60 where a constant acceleration is assumed during the transition periods at both ends.

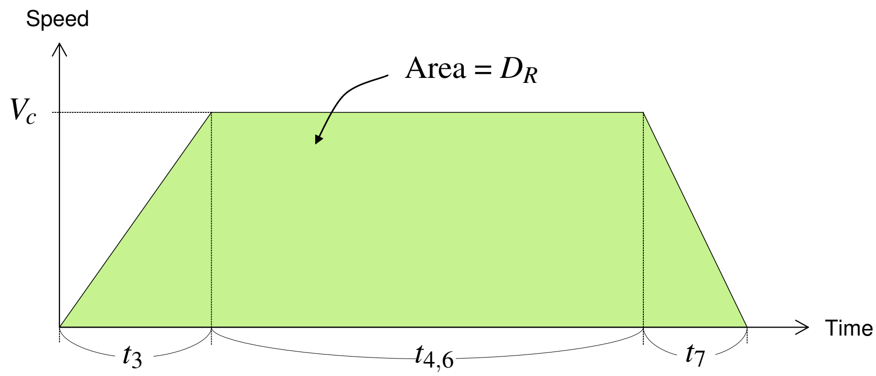


Figure 60: Flight Portion

The lengths of t_3 and t_7 is governed by vehicle performance and, if these quantities are given and denoted by $t_{A/D}$ $(= t_3 + t_7)$, in-vehicle time excluding t_5 can be computed by the following equation.

$$t_{3,4,6,7} = \frac{D_R}{V_c} + \frac{t_{A/D}}{2} \tag{61}$$

where D_R and V_c indicate route distance of flight and cruise speed, respectively. Time spent during layovers, non-moving part of in-vehicle portion, is also related to attributes of vehicle and enroute space. For example, refueling range (R_{fuel}), time for refueling (T_{fuel}), distance interval (R_{rest}) and rest time (T_{rest}), and T_L for either enforced or required layover can determine t_5 as follows:

$$t_5 = \left[\frac{D_R}{R_{fuel}}\right] T_{fuel} + \left[\frac{D_R}{R_{rest}}\right] T_{rest} + T_L \tag{62}$$

where the brackets imply that only the integer quotients are taken. Determination of the remaining segments (t_8 and t_9) mirrors origin portal access and wait (t_1 and t_2). Then it is trivial to compute the total time required to complete a trip mission, one-way:

$$t(m) = \sum_{i=1}^{9} t_i(m) + \tau(m) \tag{63}$$

where the letter m indicates an index of the mode considered and each $t_i(m)$ inside of the summation term can be obtained by the aforementioned algebraic equations. The last term $\tau(m)$ denotes delay from endogenous and exogenous causes combined. Quantification of delay, however, is not a straightforward matter at all and will be addressed further in the later section.

5.2.3 Transportation Resources

As outlined in §5.1.2, four modes for this study consist of personal automobiles (CAR), commercial air transport (ALN), general aviation with piston engine (GAP) and general aviation with jet engine (GAJ). Each mode is presented as a combination of relevant transportation resource entities. The following discussions are devoted to concisely describe each mode such that, regardless of involving air or ground transportation modes, the summation term in Equation 63 can be addressed.

5.2.3.1 Mode CAR

Mode CAR is the most convenient, affordable travel method commensurate with a wide range of needs of the general public despite its (relatively) low speed. The strengths of CAR would be much eclipsed without superb accessibility along with the nation's full fledged road system. Major highways and interstate freeways are considered the enroute space of CAR. Consequently, the portal is their ramps / exits that connect local roads to the enroute space. Table 20 summarizes the portal accessibility condition and wait time stratified by locale.[8]

Table 20: Portal Attributes of Mode CAR

	LAR	MED	SML	NOM
Portal distance (mi)	$\square(0.5,5)$	$\square(0.5,5)$	$\square(0.5,10)$	$\square(0.5,40)$
Ground speed (mph)	$\triangle(20,25,30)$	$\triangle(25,30,35)$	$\triangle(30,35,40)$	$\triangle(40,45,50)$
Portal wait (hr)	$\triangle(0,0.1,0.2)$	$\triangle(0,0.05,0.1)$	$\triangle(0,0.03,0.05)$	$\triangle(0,0.01,0.02)$

[8]Throughout this dissertation, the notation $\square(a,b)$ means the uniform distribution and $\triangle(a,c,b)$ indicates the triangle distribution where a, b and c denote lower bound, upper bound and most likely value, respectively.

The portal wait distributions are set such that they are notionally proportional to the degree of congestion of the corresponding locale. Other attributes in vehicle and enroute space entity are required to evaluate Equation 63. For example, for calculation of Equation 61, a circuitry factor (S) is needed to reflect the actual route distance D_R. Table 21 shows key variables and their representative values.

Table 21: Vehicle and Enroute Space Attributes of Mode CAR

V_c (mph)	R_{rest} (mi)	T_{rest} (hr)	R_{fuel} (mi)	T_{fuel} (hr)	$t_{A/D}$ (hr)	S
65	150	0.2	300	0.1	0.05	1.22

Notice that mode CAR has very small $t_{A/D}$ value and R_{rest} is assigned such that a driver stops at about every 2.5 hours driving in the cruise segment to prevent intolerable fatigue. In real simulation, all these attributes are subject to a small variability. For instance, cruise speed can vary depending on an individual driver, so the value of V_c is obtained from the distribution $\triangle(60, 65, 70)$.

5.2.3.2 Mode ALN

Mode ALN is the most prevalent form of scheduled public transportation suitable for long-haul trips that entails check-in and check-out at the portals for mode changes. Hence, as opposed to mode CAR, t_1 and t_2 (thus t_9 and t_8) are not negligible. Driven by economic forces, the infrastructure of mode ALN has evolved into the hub-and-spoke system, and accordingly the model should properly capture this salient feature.

The portals of mode ALN were grouped into four categories as explained in Figure 28 showing that 90% of total enplanements are captured by Large- and Medium-hub airports. If these two airport categories are collectively designated as AP:HUB, then the rest of them can be called AP:SML for simplification. Table 27 differentiates two types of “big and busy” and “small and empty” airports with respect to time attributes.

These airports are not uniformly scattered in the CONUS. Most residents in locale LAR and MED have convenient access to AP:HUB while locale NOM does not encompass any

Table 22: Airport Processing Time in Hours

	Elements	AP:HUB	AP:SML
Check-in	Mode change	0.15	0.1
	Wait-ahead	0.5	0.4
	Wait-in-line	0.3	0.25
Check-out	Wait-in-line	0.25	0.2
	Mode change	0.45	0.3

within its community boundary. Locales MED, however, possesses some diversity. For example, a few cities in locale MED, like Denver, CO and St. Louis, MO, have Large-hub airports (DEN and STL) whereas Albany, NY is equipped with a Small-hub airport (ALB). Adopting the simplified classification scheme, the accessibility model for four locales is illustrated in Table 23 showing the correlation with locale categories.

Table 23: ALN Portal Accessibility

	Within Locale	Beyond Locale	%
LAR	AP:HUB, $\triangle(1, 15, 40)$	N/A	100
MED	AP:HUB, $\triangle(1, 8, 20)$	N/A	62
	AP:SML, $\triangle(1, 8, 20)$	AP:HUB, $\square(40, 80)$	38
SML	AP:SML, $\triangle(1, 6, 15)$	AP:HUB, $\square(40, 80)$	59
	N/A	AP:HUB, $\square(50, 100)$ AP:SML, $\square(20, 40)$	41
NOM	N/A	AP:HUB, $\square(60, 120)$ AP:SML, $\square(30, 60)$	100

The numbers in the table indicate percent of population of a row locale obtained through an cross-examination of the Census data and the FAA's airport list. The distributions adopted come from an empirical study similar to that shown in Figure 59. Although not shown in the table, access speed distributions are set slightly higher than that in mode CAR because these airports are directly connected to highways. According to FAA (2002, Figure 4), 67% of the population of the U.S. reside within 20 miles of commercial service airports. The present accessibility model turns out it closely matches this fact (68%).

Given this accessibility model, access and egress time can readily be computed. But it should be pointed out that this task leads to a route choice problem regarding the pair of origin and destination airports. A maximum of four route (or airport pair) choices is possible, and it can be assumed that a route choice which involves only AP:HUB directly flies to the destination to imitate the hub-and-spoke configuration. However, in order to reflect diverse cases in reality, a number of connection stops is given with discrete distributions as described in Table 24. Now, time for the intermediate stops can be computed by: $T_L = N_C \cdot \triangle(0.5, 0.75, 1.0)$. The distributions in the equation result from the current airliners' operation policy, thus belonging to their decision variables.

Table 24: Number of Connection Stops (N_C)

Airport Pairs	0	1	2
AP:HUB–AP:HUB	80%	20%	0%
AP:HUB–AP:SML	20%	70%	10%
AP:SML–AP:SML	0%	20%	80%

In Equation 61, a vehicle's performance characteristics can solely determine time for the flight portion, which is not exactly true for mode ALN. Since airlines operate their vehicles on a scheduled basis, it is important to look up the airline schedules to figure out time information for their flights. Table 25 lists published flight times and distances of various market pairs from an Internet travel agency. A simple regression analysis of column 2 and column 5 data reveals that there exists an almost perfect linear relationship ($R^2 = 0.9989$) between flight distance and flight time. From the regression equation, V_c and $t_{A/D}$ can be obtained. The result equation is:

$$t_{3,4,6,7} = \frac{D_R}{505.3} + \frac{1.12}{2} \tag{64}$$

meaning that $V_c = 505.3$ (mph), and $t_{A/D} = 1.12$ (hr). Typical state-of-the-art commercial airliners can fly at over 650 mph when they reach the top of troposphere, but it turns out that ascent / descent / loiter segments influence actual flight time significantly.

Table 25: Flight Time Data for Selected Market Pairs

Market pairs	Distance (miles)	Flight Time (hours) West-bound	East-bound	Average
ATL–MCN	79	0.72	0.62	0.67
DFW–HOU	233	1.02	0.97	0.99
ATL–WAS	541	1.92	1.50	1.71
ATL–ORD	599	1.83	1.78	1.81
ATL–HOU	691	2.07	1.70	1.88
SFO–DEN	954	2.58	2.31	2.44
ATL–DEN	1,196	3.20	2.69	2.95
DEN–LGA	1,626	4.30	3.40	3.85
ATL–LAX	1,943	4.73	3.93	4.33
JFK–LAX	2,459	5.95	5.00	5.48
JFK–SFO	2,572	6.19	5.17	5.68

Note: All data are from non-stop flights and there exist differences between east bound and west bound flights due to the westerlies.

5.2.3.3 Modes GAP and GAJ

As seen in Figure 29, general aviation (GA) spans various platforms. But over 95% of "transportation-worthy" vehicles are fixed wing aircraft, of which 93% are dominated by piston engine aircraft. This type of airplanes—usually slow, small and for personal-use—are designated as GAP. The other types of fixed wing aircraft—designated as GAJ—have turbine engines enabling faster speed and larger payload capacity. The primary use of these vehicles is for business purposes, capturing a high-end market.

Advantages in using GA over mode ALN come from that the current hub-and-spokes system does not pose hard constraints on GA operation. As a result, mode GA can easily accommodate on-demand and point-to-point trips, which means correlation between locale and portal will become much simpler than that of mode ALN. According to FAA (2002, Figure 4), 85% of the population resides within 20 miles of 1061 GA airports with over 25 aircraft. This implies that even 50% of residents in locale NOM have GA airports nearby. Thus, the accessibility model, as shown in Table 26, is implemented such that residents in locales LAR, MED, SML have convenient access to these under-utilized GA airports.

Table 26: Portal Accessibility for Mode GA

Locale	Within Locale	Beyond Locale	%
LAR	$\triangle(1,4,10)$	N/A	100
MED	$\triangle(1,4,10)$	N/A	100
SML	$\triangle(0.5,3,5)$	N/A	100
NOM	$\triangle(0.5,3,5)$	N/A	51
	N/A	$\square(20,40)$	49

Another advantage of GA travel is that airport processing time is shorter than that of mode ALN, offering further convenience to travelers who have struggled with long waiting lines. The following table shows the assumption adopted in the present model. Separation of AP:GAP and AP:GAJ is needed to reflect streamlined process of check-in/check-out for travelers that can afford premium services.

Table 27: Airport Processing Time in Hours

	Elements	AP:GAP	AP:GAJ
Check-in	Mode change	0.10	0.05
	Wait-ahead	0.20	0.10
	Wait-in-line	0.20	0.10
Check-out	Wait-in-line	0.10	0.05
	Mode change	0.30	0.05

The number of aircraft to be used should be kept at minimum to simplify the model since it would be extremely time consuming to model all GA aircraft and to browse their performance data. Hence, each mode is represented by one type of vehicle, which can be of course extended later. Turnbull (1999) reports that the typical GAP is a four-place, single engine piston with a cable-operated flight control system. Under that guideline, the five most prolific models are selected and shown in Table 28. Cessna 182 was chosen as the representative GAP vehicle since its max takeoff weight ($TOGW$), wing span (b), top speed (V) and range (R) were found to be very closed to the weighted average values.

Table 28: Representative GA Aircraft Models

	GA M/S*	$TOGW$ (lb)	b (ft)	V (kt)	R (nm)
Cessna 172	12.3%	2300	35.0	125	575
Piper PA28	11.18%	2550	35.0	133	443
Cessna 182	7.21%	2550	36.0	140	582
Beechcraft 35	3.39%	2725	32.8	165	673
Mooney M20	3.38%	2575	35.0	171	890
Weighted Average	N/A	2486	35.0	138	574

*M/S: Market Share

Through a similar process, Cessna 550 Citation II, the top seller light corporate jet, was chosen as the representative vehicle of mode GAJ. The following table summarizes the vehicle attributes of baselines GAP and GAJ that are needed to compute flight time.

Table 29: Vehicle and Enroute Space Attributes of Modes GAP and GAJ

	V_c (mph)	R_{rest} (mi)	T_{rest} (hr)	R_{fuel} (mi)	T_{fuel} (hr)	$t_{A/D}$ (hr)	S
GAP	160	400	0.2	660	0.2	0.4	1.05
GAJ	441	2000	0.0	2000	0.4	0.5	1.01

5.2.4 Exogenous Entities

The two exogenous entities—drivers and disruptors—and their models should have a close relevance to the database available. The following section outlines several mathematical attempts to describe these entities.

5.2.4.1 Driver Entities

The Census database spans numerous characteristics and classification variables. Since the Bureau continually records, updates and verifies the database in each subsequent interview and survey, it is the most important resource to employ in the model construction. The present model focuses on the most influential demographic and economic attributes of people and business entities that govern the transportation activities.

Household Attributes

Table 30 lists the income distribution of the U.S. households in 1999. Since the whole transportation environment is grouped into four locales, each column separates its respective distribution by income brackets.

Table 30: 1999 U.S. Household Income Distribution by Locale

Income Group ($)	LAR	MED	SML	NOM
Less than 10,000	8.82%	8.63%	9.94%	10.55%
[10,000, 15,000)	5.22%	5.78%	6.80%	7.35%
[15,000, 20,000)	5.12%	5.91%	6.79%	7.18%
[20,000, 25,000)	5.49%	6.36%	7.11%	7.37%
[25,000, 30,000)	5.54%	6.34%	6.85%	7.09%
[30,000, 35,000)	5.62%	6.36%	6.65%	6.88%
[35,000, 40,000)	5.33%	5.87%	6.11%	6.35%
[40,000, 45,000)	5.17%	5.62%	5.80%	6.03%
[45,000, 50,000)	4.57%	4.97%	5.06%	5.28%
[50,000, 60,000)	8.64%	9.14%	9.08%	9.31%
[60,000, 75,000)	10.78%	10.83%	10.17%	9.99%
[75,000, 100,000)	11.94%	10.92%	9.59%	8.62%
[100,000, 125,000)	6.98%	5.64%	4.55%	3.71%
[125,000, 150,000)	3.68%	2.73%	2.07%	1.59%
[150,000, 200,000)	3.41%	2.37%	1.68%	1.30%
200K or more	3.71%	2.52%	1.77%	1.43%

The raw data are converted into cumulative formats in percentile and then the value of the income can be obtained through solving a closed form equation. The base function adopted is the Richards' generalized logistic equation (Richards 1959) and the curve fitting is significantly improved by adding extra parameters into the conventional Richards' model. The modified equation is:

$$u = \beta_1 \{1 - \beta_2 \exp(-\beta_3 I)\}^{\frac{1}{1-\beta_4}} + \beta_5 \tag{65}$$

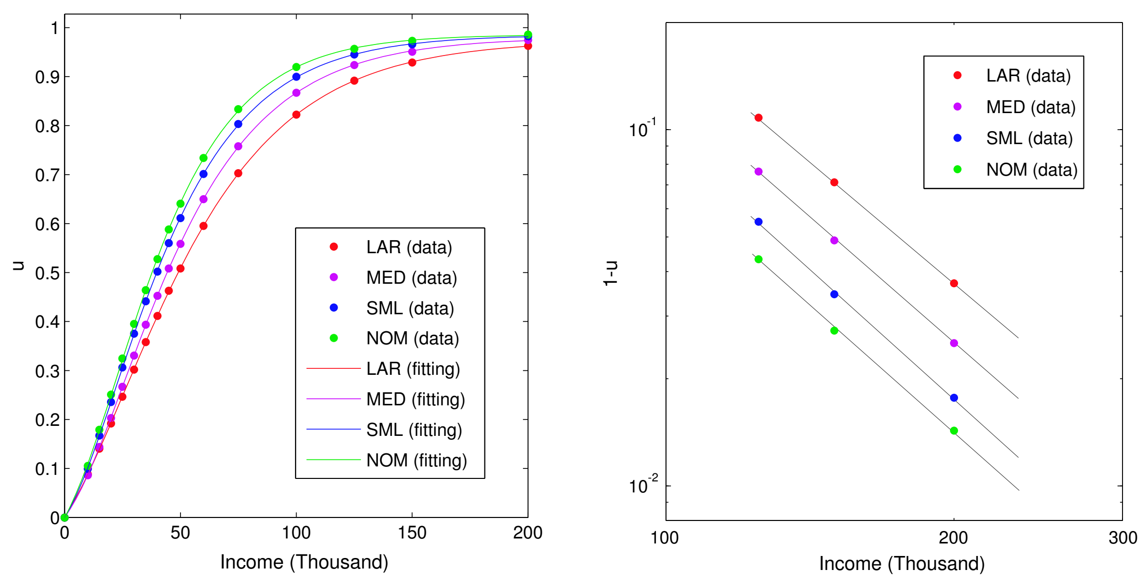
where u indicates the percentile, I stands for income and the added regression parameters are β_1 and β_5. The satisfactory non-linear regression results with Equation 65 are shown in Figure 61(a).

This near-perfect fitting, nevertheless, does not guarantee extrapolating beyond \$200K —the most likely income bracket for potential PAV user group or early adopter. Since no further information on the highest income bracket was available, a mathematical model to describe the highest income group is needed instead, which is potentially critical to the simulation outcome. Close examination on the higher income regions produced Figure 61(b) where the linear trends are clearly observed on a log-log scale, which strongly suggests that the following power function is the most suitable form for extrapolation use:

$$u = 1 - \alpha I^{-\beta} \tag{66}$$

where α and β are positive regression parameters. In economics, this model is accepted as the Pareto Law of income that provides a good fit to the distribution of high incomes. (Reed 2003)

Having Equations 65 and 66, the inverse transformation method (Law & Kelton 2000, pp. 440–448) is employed to compute the value of I. Adopting extra parameters in Equation 65 brings the small price. The lower bound of random variable u (u_0) should be solved in such a way that no negative income values are allowed. The following algorithm summarizes the final mathematical model spanning the entire income distribution.



(a) Richards Model

(b) Pareto Model

Figure 61: Household Income Model

1. Obtain the percentile value of u from a uniform distribution $\square(u_0, 1)$.
2. Compute I from the following equation:

$$I = \begin{cases} -\frac{1}{\beta_3} \log\left(\frac{1-U^{1-\beta_4}}{\beta_2}\right), \text{ where } U = \frac{u-\beta_5}{\beta_1} & \text{if} \quad u_0 \leq u < u_1 \\ \left(\frac{1-u}{\alpha}\right)^{-\frac{1}{\beta}} & \text{if} \quad u_1 \leq u < 1.0 \end{cases} \tag{67}$$

The cut-off value of the high income region is set to \$125,000 and its corresponding percentile u_1 need to be calculated for each locale. All parameter sets needed to describe each locale's income distribution are given in Table 31.

Table 31: Parameter Settings for Income Generation

	LAR	MED	SML	NOM
u_0	3.5637E-04	-3.1400E-05	-4.3360E-04	-9.5070E-04
u_1	9.6203E-01	9.7365E-01	9.8150E-01	9.8401E-01
β_1	1.0717E+00	1.0255E+00	1.0378E+00	1.0802E+00
β_2	3.6031E-01	5.7714E-01	5.6572E-01	1.6235E-01
β_3	2.5498E-02	2.8991E-02	3.1593E-02	3.5666E-02
β_4	8.1363E-01	7.2397E-01	7.2122E-01	9.2766E-01
β_5	-9.7142E-02	-4.5397E-02	-5.2525E-02	-9.4282E-02
α	6.3571E+03	6.5041E+03	6.2496E+03	3.5220E+03
β	2.2749E+00	2.3527E+00	2.4126E+00	2.3446E+00

Another key attribute considered in populating an individual household is the number of household members or household size. Table 32 lists the relationships between household income and household size. It may seem, at first glance, that mathematically modeling the listed data would be an infeasible task. Nevertheless, a relatively simple process can be adopted as follows. After determining the income value, one can create a discrete distribution function by looking up the table given income range row. If this process is repeated, a collection of households that makes up the real population in the US can be readily obtained.

Table 32: Household Size Distribution in Thousand

Income Bracket	Household Size							Mean
	1	2	3	4	5	6	7+	HH size
Under $2,500	757	514	217	127	65	14	23	2.06
$2,500 to $4,999	609	312	167	132	53	13	7	2.07
$5,000 to $7,499	2141	530	299	159	81	39	15	1.69
$7,500 to $9,999	2143	640	318	156	73	36	16	1.7
$10,000 to $12,499	2326	943	353	244	121	45	18	1.83
$12,500 to $14,999	1775	1060	333	214	130	62	37	1.98
$15,000 to $17,499	1677	1207	515	328	162	73	45	2.15
$17,500 to $19,999	1286	1240	456	250	148	62	34	2.18
$20,000 to $22,499	1497	1238	527	338	179	62	34	2.21
$22,500 to $24,999	1069	1312	440	291	149	65	39	2.3
$25,000 to $27,499	1243	1402	516	315	186	99	45	2.32
$27,500 to $29,999	881	1137	414	358	199	52	41	2.46
$30,000 to $32,499	1111	1242	483	451	231	69	46	2.46
$32,500 to $34,999	731	974	440	316	167	62	58	2.57
$35,000 to $37,499	967	1188	541	427	223	80	59	2.55
$37,500 to $39,999	491	959	419	366	190	79	28	2.71
$40,000 to $42,499	706	1111	527	478	208	81	45	2.68
$42,500 to $44,999	427	964	396	370	169	57	26	2.7
$45,000 to $47,499	502	898	497	473	226	73	47	2.84
$47,500 to $49,999	357	795	476	376	162	52	24	2.8
$50,000 to $52,499	474	1000	507	453	232	62	35	2.8
$52,500 to $54,999	301	755	368	315	176	70	41	2.9
$55,000 to $57,499	331	801	442	390	171	42	53	2.89
$57,500 to $59,999	190	678	385	362	142	38	39	3
$60,000 to $62,499	323	845	496	458	189	63	35	2.93
$62,500 to $64,999	206	577	370	348	114	58	29	2.99
$65,000 to $67,499	179	657	396	384	186	50	31	3.06
$67,500 to $69,999	150	610	298	270	94	48	25	2.9
$70,000 to $72,499	155	609	343	348	159	41	25	3.03
$72,500 to $74,999	91	401	275	308	117	34	22	3.19
$75,000 to $77,499	157	550	351	363	164	67	30	3.14
$77,500 to $79,999	100	411	276	299	71	48	14	3.07
$80,000 to $82,499	111	505	328	315	107	32	15	3.01
$82,500 to $84,999	76	394	253	245	120	15	10	3.08
$85,000 to $87,499	72	370	282	255	104	24	14	3.14
$87,500 to $89,999	45	339	164	214	94	26	20	3.25
$90,000 to $92,499	95	299	188	254	99	30	16	3.2
$92,500 to $94,999	21	273	154	184	86	25	13	3.33
$95,000 to $97,499	35	248	195	239	102	25	21	3.41
$97,500 to $99,999	37	242	148	176	61	23	15	3.27
$100,000 and over	879	4437	2597	2962	1271	444	242	3.22
Total	26724	34666	17152	15309	6981	2445	1428	2.62

Enterprise Attributes

The remainder of the economic drivers are concerned with enterprises. The Census database keeps track of information on these firms that are classified into five categories from Class I to Class V according to size of sales or receipts. The cut-off amounts are 1 million, 10 million, 100 million and 1 billion in an ascending order. As such, Class I firms encompass small businesses earning less than 1 million dollars annually and the largest companies in the nation belong to Class V recording more than 1 billion. Table 33 sums up the size, total revenue, and employment of the nation's firms as of 1992.

Table 33: Five Classes of Enterprise

	Class I	Class II	Class III	Class IV	Class V
# Firms	3,807,253	704,535	90,201	7,743	1,097
Yearly Sales ($)	9.74e+11	1.92e+12	2.26e+12	2.05e+12	4.41e+12
Employment	1.67e+07	1.96e+07	1.51e+07	1.39e+07	2.17e+07
Payroll ($)	2.58e+11	4.43e+11	3.88e+11	3.68e+11	6.68e+11
Average # Employees	4.4	28	167	1,791	19,788
Average Salary ($)	15,512	22,643	25,744	26,531	30,771
Average Sales ($)	2.56e+05	2.73e+06	2.50e+07	2.65e+08	4.02e+09

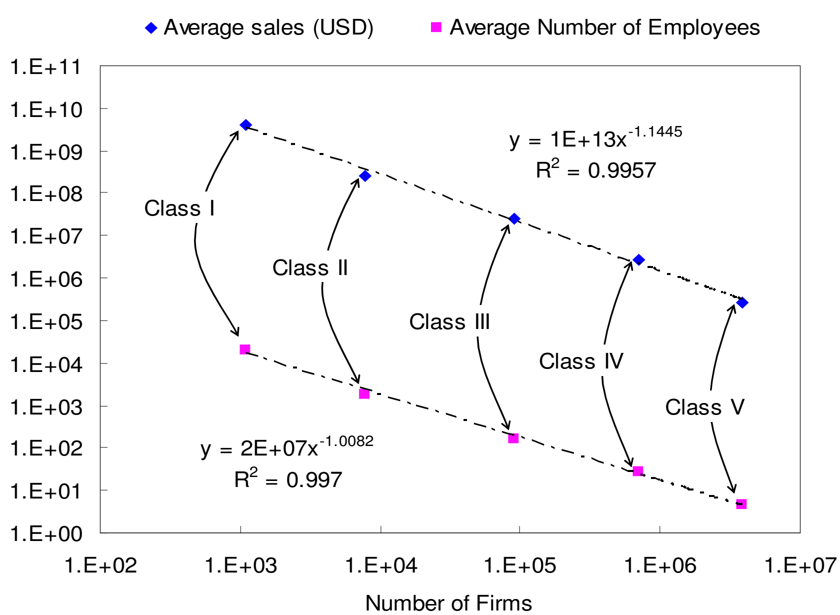


Figure 62: Zipf's law

Interesting findings can be highlighted when the aggregated characteristics of each class are plotted on a log-log scale. It is noticeable that average number of employees and amount of sales against the number of firms obey a power law as shown in Figure 62.

Miscellaneous Drivers

Besides the most influencing driver entities described above, many socioeconomic indicators are included in the driver category. Most of these indicators are time variant and can be taken from many government agencies other than the Census Bureau, such as the Bureau of Labor Statistics, the Energy Information Administration and the Bureau of Economic Analysis. Table 34 lists selected indicators that have a direct relevance to the model.

Table 34: Historic Trends of Selected Economic Indicators

Year	AvGas	Jet Fuel	Motor Gas	GDP in billions		CPI
	(Nominal USD per Million BTU)			(Nominal USD)	(2000 USD)	(1995 = 100)
1990	9.32	5.68	9.12	5,803.1	7,112.5	85.8
1991	8.71	4.83	8.93	5,995.9	7,100.5	89.4
1992	8.54	4.52	8.96	6,337.7	7,336.6	92.1
1993	8.24	4.29	8.83	6,657.4	7,532.7	94.8
1994	7.96	3.95	8.96	7,072.2	7,835.5	97.2
1995	8.36	4.00	9.22	7,397.7	8,031.7	100.0
1996	9.29	4.82	9.85	7,816.9	8,328.9	103.0
1997	9.39	4.53	9.81	8,304.3	8,703.5	105.3
1998	8.11	3.35	8.45	8,747.0	9,066.9	107.0
1999	8.81	4.01	9.31	9,268.4	9,470.3	109.3
2000	10.87	6.60	12.01	9,817.0	9,817.0	113.0

In addition, factors that directly affect the amount of trips also make up the driver group, e.g., previously explained C_{ij} matrix (Table 18), total population and its distribution (Table 17). Yet, other factors exist that relate to cultural or lifestyle patterns. The number of nights at the destination and the ratio of personal to business trips are pulled out from the 1995 ATS database and are included in the model.

5.2.4.2 Disruptor Entities

A wide variety of disrupting events directly affect the resource network. As for modeling the disruptor entities, balancing of simplicity and comprehensiveness had to prevail. The present model only includes weather since it is the most visible aspect of the natural disruptors and has a major impact on the efficiency and safety of vehicle operations in the NTS. Taking air transportation, for instance, as much as two thirds of the airline delays are attributable to adverse weather. (Lindsey 1998)

Instead of comprehensively gathering weather data on a national scale, the estimated number of delayed days per year at twenty of the busiest U.S. airports was used instead as shown in Figure 63. The number of days on which delay occurred is categorized according to three types of weather conditions: thunderstorms, heavy fog (visibility less than 0.25 miles), and reduced visibility (visibility greater than 0.25 miles but less than 7 miles).

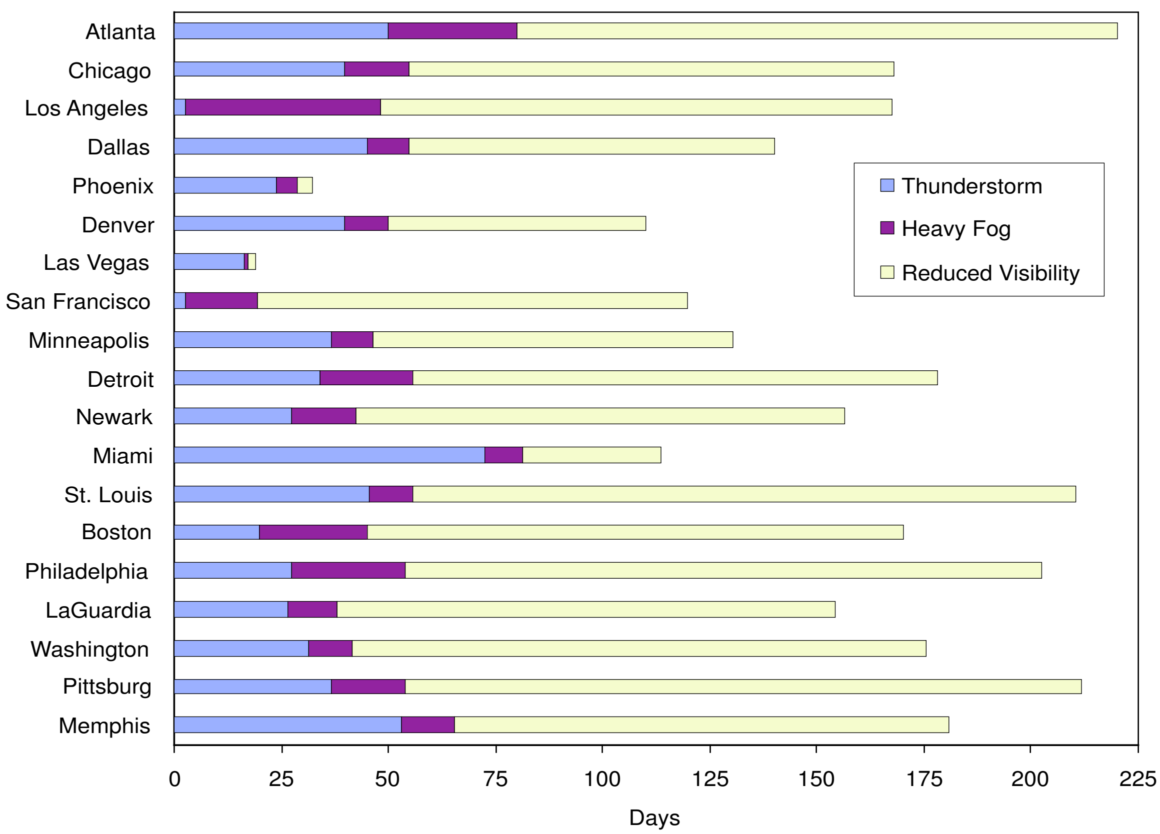


Figure 63: Weather Distribution at Major Airports [Source: Weber et al. (1991)]

The airports are listed by annual enplanements in the descending order. Table 35 lists the aggregated percentage of days of clear weather as well as the three adverse conditions. Each airport's data was proportionally weighted with its enplanement.

Table 35: Percent of Weather Category

Clear	Reduced Visibility	Heavy Fog	Thunderstorm
58.5%	27.6%	4.8%	9.1%

A weather entity is modeled upon these findings with two attributes: scale and frequency. It is readily admitted that other details are missing. The weather scale is logarithmic and indicates the severity on a 0 to 9 scale—any transportation activities cannot be conducted on the largest scale 9, which would be a rare event. It is assumed that the frequency of the weather scale follows the geometric distribution, a discrete counterpart of the exponential distribution and the specific values are listed at the top of Table 36. The frequency-to-scale behavior of the model bears an analogy to those of natural disasters.[9]

Table 36: Weather Model and Influence Matrix

Scale	0	1	2	3	4	5	6	7	8	9
Frequency	0.355	0.231	0.150	0.097	0.063	0.041	0.027	0.017	0.011	0.007
Ground	0	0	0	0	0	0	20	40	80	100
HiAlt	0	0	0	0	0	0	20	60	100	100
LoAlt (VFR)	0	0	20	60	100	100	100	100	100	100
LoAlt (IFR)	0	0	0	0	20	60	100	100	100	100

Having the weather model, the next step is to define its influence on the resource entities, or more precisely, the enroute spaces. Each enroute space has its own influence array as an interface to weather events, as shown in the bottom portion of Table 36. The influence value indicates the percentage of inoperable condition. For example, when the weather scale is 3, 60% of the pilots who have VFR ratings would wait until the weather improves whereas no restrictions are posed on operating mode GAP with IFR ratings.

[9]Earthquakes are a prime example. (See http://neic.usgs.gov/neis/eqlists/eqstats.html.)

5.3 Transportation Stakeholders as Agents

The essence of ABM lies within the design of the agents that play a key role in the computer-generated microcosm. Of all the sentient entities in the transportation architecture, the consumer and service provider stakeholders are chosen as the agents in the present research. Detailed descriptions of these active players are presented next.

5.3.1 Transportation Consumers

It is an understatement to say that the transportation consumers are the most important sentient entity among the stakeholder types identified in Table 10. Other types of stakeholders, to say the least, are passively responding to the consumers' individual or aggregate behaviors. Similar to how generic consumers are not visible until they participate in the market, the transportation consumers become recognizable as they purchase tickets or make travel arrangements. Then, a traveler or a group of travelers that make up a trip party can be uniquely considered as a single consumer agent. Although a unified model to describe every detail of the traveler would be hardly obtainable, the modeling can start from well-accepted theory that they exhibit behavioral distinctions depending on primary trip purpose.

5.3.1.1 Two Reference Groups as Traveler Generator

The distinction between non-business travelers and business travelers is desired as they identifies themselves with different reference groups. The model accordingly incorporates two reference groups: households and enterprises that generate non-business travelers and business travelers respectively. This partition is also consistent with the economic convention, which takes households and enterprises as basic units of most economic activities.

Household and Personal Travelers

The aforementioned driver entities formulated from the Census data are directly linked to populate a unit household instance of which the residential locale, household income, household size are the defining attributes. Besides these usual characteristics, transportation related attributes are also tied to each household instance.

Based on the resident locale, the portal accessibility for all modes is obtained first. The household agents have their own conditions regarding access speed and distance to the nearest highway ramp, AP:HUB, AP:SML and GA airports. Some attributes are established in conjunction with demographic and economic status, such as vehicle ownership and self-pilot capability for all transportation modes excluding mode ALN. Mode CAR is the most available means if transportation to the general public. Hence, a simple assumption was implemented in the model: each household has a car for personal use and a driver's license. This assumption may seem rather lenient, yet except for a small percentage of low income families, it is largely true. Related to mode GAP, the FAA estimated that 519 thousand individuals had pilot licenses (excluding student certificates) and 300 thousand pilots had instrument ratings in 1995. At the same year, adult population with income was about 183 million. Hence, the model assumes that 0.285% of the income earners have self-pilot capability and 60% of the pilots can fly on IFR condition. One economic condition posed is that the households having GAP pilots must earn more than $35,000 a year, which make up about 50% of all households.

Table 37: Number of Pilots: 1980 to 2000 [Source: Census Bureau (2003, p. 687)]

Year	Private	Commercial	Air transport	Student
1980	357,479	183,442	69,569	199,833
1985	311,086	151,632	82,740	146,652
1990	299,111	149,666	107,732	128,663
1995	261,399	133,980	123,877	101,279
2000	251,561	121,858	141,596	93,064

A stronger correlation may exist between economic affluence and GAP ownership. The FAA estimated that there were 183,000 active GA fleets and 100,000 of these airplanes belonged to individuals. In order to be eligible as a GAP owner, 1) the household should have a pilot and 2) the household income is over \$70,000. As a result, the GAP ownership model dictates that nearly 40% of those who meet the two conditions possess GAP. Likewise, a much stronger economic constraint is posed for mode GAJ, as it is undoubtedly clear that only the wealthiest people can own such an expensive vehicle. The model assumes that 5% of the households that earn more than \$2,200,000 possess GAJ. Ownership percentage of these GA vehicles can go up due to fractional ownership.

The rest of the attributes are related to people's habit with respect to transportation activities. The first attribute considered is weight of time (w_t) that, when multiplied by hourly income of the household, speaks to the household's perception on the value of time. The consensus regarding the amount from transportation economists is about 0.7 when a trip is motivated by non-business purpose. (Kruesi 1997) Therefore, the triangular distribution $w_t = \triangle(0.6, 0.7, 0.9)$ is used to reflect people's diverse appreciation on the value of time.

As presented in Table 8, the American consumers spend about 18.5% of their income on transportation over a year. This figure includes all transportation related expenses such as car purchasing, financing cost and commuting. With the absence of more detailed information, it is assumed that an average household spends 4% of its income on long distance travels. The model uses the distribution $\triangle(0.0, 0.04, 0.08)$ to obtain the transportation budget percent for an household since the amount varies across each individual household.

Enterprise and Business Travelers

The enterprise model is also linked with the Census database from which business travelers are generated.[10] Determination of the resident locale and portal accessibility condition follow the similar process as in the household model with some necessary adjustments. For example, the portal distances are reduced because most firms usually are located nearby convenient transportation infrastructure. However, noticeable distinctions exist regarding

[10]The other sources include the federal, state and local governments as well as military. But these public organizations are ignored in the model due to lack of data set, which may not cause a big difference since the civil side consumers obviously make up the majority of all passenger transportation activities.

income and size of the trip party. For income estimation, the partitioned regression approach is the same but the personal income data is used instead of household income data to reflect employee's position in a firm. When a trip party is composed of two or more persons, each member's income is aggregated. Nonetheless, unlike in the household case, the Census database does not contain information on size of business travelers. The ATS data shown in Figure 32 is analyzed further for this purpose, and the result is given in Table 38 where the size of more than seven is ignored due to its small portion. This is the first time that the ATS data is directly referenced in building the model.

Table 38: Size Distribution of Business Travelers

Trip Party Size	1	2	3	4	5	6
Percent	62.65	22.20	7.29	4.23	2.16	1.08

The GAP pilot model should also be adjusted accordingly. In order to have self-pilot capability, the average income should be over \$17,400, a median value of all income earners. Then, the probability of having at least one pilot is equal to the converse probability of not having any pilot in the group: i.e., $p(n) = 1-(1-p_1)^n$ where n is the size of the trip party and p_1 takes $2 \times 0.285\%$. Also, it is assumed that 60% of the agents having self-pilot capability can fly on IFR condition. For the monetary budget, the business travelers use the distribution $\triangle(0.0, 0.05, 0.10)$.

The other key attributes have a strong dependency on the type of enterprise class, which is obtained through a discrete distribution taking the row data of employment in Table 33. If the traveler, for example, is a member of a family-owned small business, a decision-making process and outcome would be quite similar to that of personal travelers, whereas the highest officers of the largest company operating a private business jet as the company's asset put much higher value on time. Following this logic, Table 39 is prepared to describe the assumptions adopted for different enterprise classes. The second and third rows of the table indicate the probabilities of being mode GAP and mode GAJ owners, respectively.

The probabilities of having higher weight of time by enterprise class $w_t = \triangle(1.0, 1.1, 1.2)$, as opposed to the previous $\triangle(0.6, 0.7, 0.9)$, are described in the fourth row. Having been intuitively derived, the probability values in the table serve as the model parameters for the developer at this stage. Even so, they suffice for representing collective behaviors of the business travelers by enterprise class. Substantial research from other disciplinary areas can remove any subjectivity issue.

Table 39: Probability of Vehicle Ownership

	Class I	Class II	Class III	Class IV	Class V
GAP owner	0.0%	0.0%	0.0%	0.2%	2.0%
GAJ owner	0.0%	0.0%	0.0%	1.0%	10.0%
Higher Weight of Time	0.0%	30.0%	50.0%	70.0%	100.0%

5.3.1.2 *Behavioral Rules*

A consumer agent requires information and then exhibits a behavior based upon its own decision making rule. The basic behavioral pattern is the same regardless of the reference group of the consumer. It is attempted herein to lay out a generic structure underlying the consumer's sentience. As a result, a multi-step process is postulated and a pictorial description of the flow given in Figure 64.

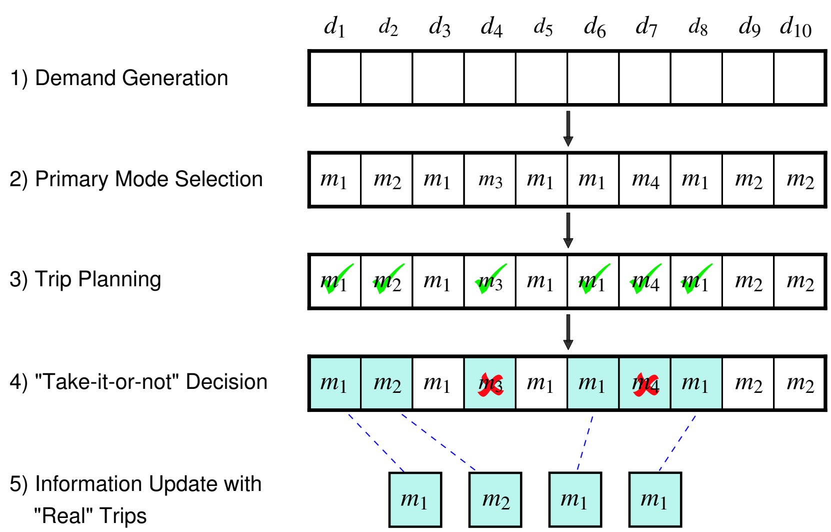


Figure 64: Behavioral Process of Consumer Agent

1) Demand Generation

Every transportation activity simply starts with a desire to go somewhere. At the top of Figure 64, a particular consumer expresses ten demands, d_1 through d_{10}. Each cell hosts a set of detailed trip information. The primary data includes the purpose of the trip, destination, size of the trip party and travel distance. Prospective travel date, number of nights away at destination, type of lodging planned on the way, etc. make up the secondary information. The size of the demand list is set to 40 per year for the personal travelers and 60 per year for the business travelers.

2) Primary Mode Selection

Base on the created wish-list, the next behavior of the consumer is to contact service providers and gather detailed information on time and cost. Then the consumer goes through a mode selection process. The mode selection model is underpinned by the probabilistic choice theory (§3.3), to be discussed in detail in §5.3.3. The outcome of the selection process includes means of transportation, intermediate stops, trip time and trip cost. Figure 64 shows that this particular consumer finishes the process for each cell and chooses from modes m_1 through m_4 as indicated inside of the cells. But the wish-list is still nothing more than an aspiration since the consumer cannot afford every wish turning into a real trip.

3) Trip Planning

The consumer begins to plan or eliminate some trips in the list at this step. In general, the monetary budget puts a constraint on consumer behaviors as discussed in §3.2.2. For the transportation consumer, the concept of budget should expand to account for other concerns such as time and psychological reasons. These constraints can be represented together in a multi-dimensional space, called the mobility budget space. Consider a two dimensional case as illustrated in Figure 65 where the consumer's time budget (B_t) and the cost budget (B_c) bound a feasible space. The coordinates of the numbered dots represent cumulative time and cost. For example, T_2 sums up the expected times for the first and the second trip taken from the list, not necessarily in the left-to-right direction. The consumer repeats this until the aggregated expenses do not exceed the budgets. Figure 64 shows that this particular consumer eliminates d_3, d_5, d_9 and d_{10}. The consumer is willing to spend his time and money on the rest, termed the captured demands.

4) "Take-it-or-not" Decision

In a perfect world, the captured demands turn into real trips. This is not always true for a sentient—often unpredictable—creature in the real world fraught with disrupting events. For example, the consumer can rescind a planned trip due to inclement weather. The other example is the terrorist attacks on September 11, 2001, which brought a complete shutdown of the National Airspace System for days. The algorithm of this step can be implemented with the mobility budget space. Suppose that the space below has a third axis out of the plane, called the ψ-axis, and the consumer has a certain amount of psychological budget B_ψ on a 1 to 10 scale. Notice that the ψ-axis does not take cumulative value like the other two axes. Initially, ψ values of all dots are 0. But the imagined traveler cancels d_4 and d_7 (Figure 64) at the presence of isolated events that triggered Points 3 and 5 (Figure 65) elevated over B_ψ. This algorithm is easily established by a set of logic equations describing those triggering events.

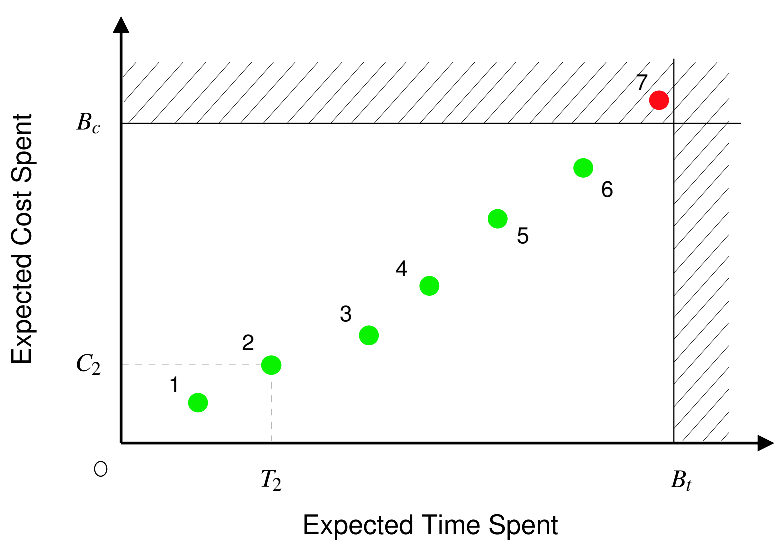


Figure 65: Mobility Budget Space

5) Update information

The initial wish-list has been subject to double screenings so far. When millions of other consumers go through the same procedure, the virtual world is filled with a vast amount of "real" transportation activities. While the above four steps focus on the consumer's individual behavioral patterns, this step aims to describe how the resultant transportation activities affect the consumer stakeholder's collective sentience. The selected influence concerns "real" delay (τ_r) and "perceived" delay (τ_p). This topic requires in-depth investigation and is addressed next.

5.3.1.3 Probabilistic Delay Model

In today's transportation system, delay is increasingly making up a significant portion of the total time spent on a trip. The amount of delay is determined by the system's demand and capacity. Quantifying this delay time can be tackled with a combination of queuing theory and simulation if the problem structure is of a manageable scale. However, when the total system is networked at a massive scale, factors beyond the local level become relevant. For example, a flight from ORD to STL can be affected by a seemingly irrelevant event such as a malfunction in ATL. Furthermore, the demand and capacity are not static but time variant and unpredictable. The system throughput at an airport experiences short-term changes due to weather conditions, and highway drivers are often puzzled by why there are so many people leaving so early. The delay is, therefore, an *a posteriori* quantity—the exact amount can be obtained only after it happens. Given these circumstances, probabilistic treatment based on analysis of the field data is a feasible approach.

As for inter-city travel, the relevant delay data can be found only for mode ALN since on-time performance of every scheduled flight has been tracked by the FAA. Taking the calendar year of 1995, on-time performance data of all 5,327,435 flights are examined. The analysis result is shown in Figure 66 where both extreme ends are truncated because there seems to exist some anomalies in the raw data stemming from the database being exhaustively huge (+1GB).

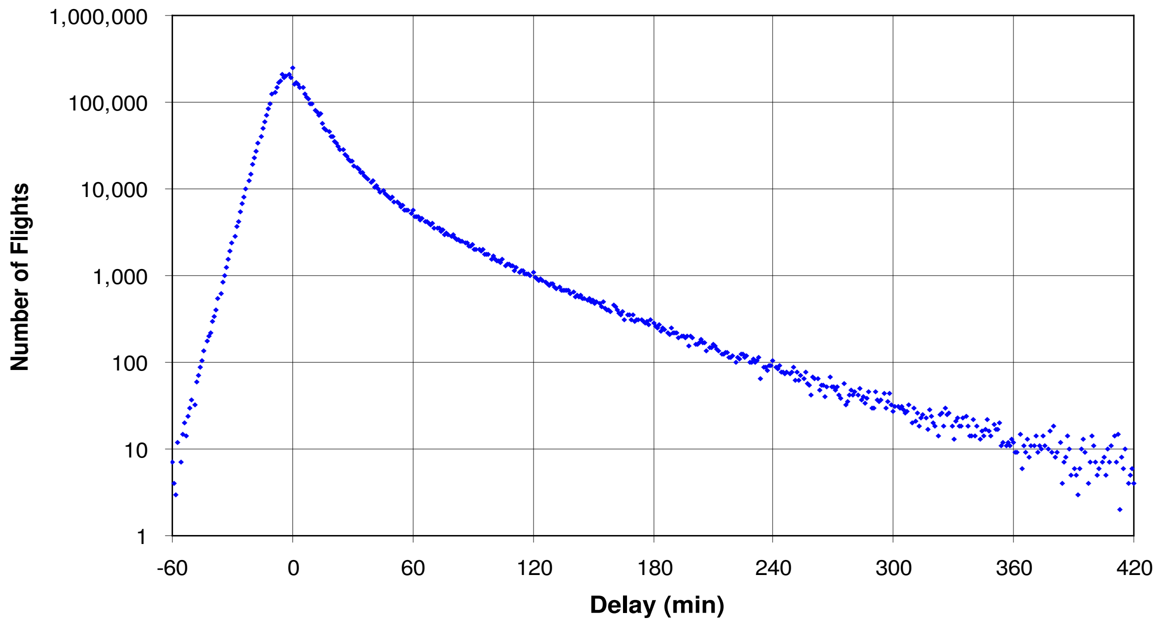


Figure 66: Airline Delay in 1995

The FAA counts a flight as on-time if it arrives less than 15 minutes later than the time scheduled in the carriers' computerized reservations systems (CRS). Table 40 summarizes the delay data by time brackets with inclusion of cancelled/diverted flight category.[11]

Table 40: On-time Performance Summary

Category	Time bracket (hr)	Number of flights	Frequency	Cumulative Freq.
On-time	Less than 0	2,319,450	43.5%	43.5%
	0.0 to 0.25	1,866,338	35.0%	78.6%
Delayed	0.25 to 1.0	843,345	15.8%	94.4%
	1.0 to 2.0	148,129	2.8%	97.2%
	2.0 to 3.0	32,775	0.6%	97.8%
	3.0 or more	15,001	0.3%	98.1%
C/D*	—	102,397	1.9%	100.0%

* Cancelled/Diverted

As it can be seen, 78.6% of all flights were operated on-time in 1995. The "delayed" category makes up 19.5%, of which 18.9% suffered from more than 1 hour holdups. To

[11]Cancelled flight—a flight listed in a carrier's CRS during the seven calendar days prior to scheduled departure but not operated. Diverted flight—a flight that is required to land at a destination other than the original scheduled destination for reasons beyond the control of the pilot/company.

obtain a quantitative model, a non-linear regression was performed, focusing on the delayed category. Equation 65 was adopted again as the fitting function. The fitting parameters are shown in Table 41, which concludes the construction of the probabilistic model of the real delay for mode ALN.

Table 41: Delay Model Parameters

u_0	u_1	β_1	β_2	β_3	β_4	β_5
14054.863	1.19358	0.0232	-27071.821	-14053.865	0.037	0.9979

While the real delay (τ_r) constitutes an important aspect of the resource network, another interpretation of delay is necessary as the sentient entities in the TAF are part of the modeling equation. The transportation consumers gather information on delay through direct/indirect experiences. Based on that information, the consumers formulate and update a perception on each mode's delay. The "perceived" delay (τ_p) is not measurable because it concerns a relative and fuzzy degree of punctuality in people's mind whereas its counter part, the real delay, is hard to predict yet possible to measure. Thus, the perceived delay model is constructed with two attributes: on-time reliability and delay distribution. The on-time reliability refers to the punctuality of a mode, given in percentage. A right-angled triangle is used to represent the shape of the delay distribution, which is reasonable given how ordinary people do not know the exact form of the distribution. The perceived delay models for all modes are shown in Table 42. Inspired by the real delay model of mode ALN, the triangular distribution has the same lower bound of 15 minutes.

Table 42: Parameters for Perceived Delay Model

Mode	On-time Reliability	Delay Distribution
CAR	80.0%	$\triangle(0.25, 0.25, 1.0)$
ALN	70.0%	$\triangle(0.25, 0.25, 3.0)$
GAP	95.0%	$\triangle(0.25, 0.25, 1.0)$
GAJ	99.0%	$\triangle(0.25, 0.25, 1.0)$

The consumers may initially frame perceptions drawn from personal experiences, but consolidate and generalize them through communications with others and reports from mass media—one of the indirect transportation stakeholders. Then, as far as the perceived delay is concerned, the same model is applicable to all consumers as if they have a collective sentience.

Lastly, the basic idea behind "Update Information" is that the consumers can perform a balancing process. That is, if the consumers are dissatisfied with too much delay from mode m, they would self-adjust their on-time reliability or delay distribution. Suppose an investigator performed a simulation with multiple consumers and the results were available. Then, the following algorithm describes this behavior.

1. Construct a set $\Theta(m)$ that contains the trips done by mode m.
2. Compute the difference between the perceived delay and the real delay:

$$\delta(m) = \sum_{i \in \Theta(m)} \left\{ \tau_p(m) - \tau_r(m) \right\}_i \tag{68}$$

3. Adjust the amount of on-time reliability of mode m, $R(m)$:
   - Increase $R(m)$, if $\delta(m) > \varepsilon$.
   - Decrease $R(m)$, if $\delta(m) < -\varepsilon$.
4. Repeat the simulation until $|\delta(m)| \leq \varepsilon$.

The convergence parameter ε has a small positive value and needs to be determined prior to the simulation. Ideally, this algorithm should be processed for all modes. Since the mode choices are not independent of each other, it would require many iterations accompanied by simultaneously changing each $R(m)$ to obtain the final converged status. This procedural complication, however, does not make a big concern with the assumption that $\tau_p = \tau_r$ for modes CAR, GAP and GAJ. It is inevitable due to the absence of the real delay data excluding mode ALN at this point.

5.3.2 Transportation Service Providers

The transportation service providers interact with the consumers while coordinating their own resources. The basic function of the providers in the model is to look at a trip and to offer price and time information to the consumers. Implementation of this task requires a set of logics to discern trip and consumer attributes to which a service provider suitably responds. Three business models of the providers are identified: SELF, RENT and FARE. Proper combination of the model and the resource constructs a distinctive service provider. All providers share the generic structure explained in Equation 63 for the time calculation, but the cost calculation procedure requires tailoring depending on the provider type.

5.3.2.1 Business Model SELF

Vehicles can be owned by an individual or by an enterprise on either a full or fractional ownership basis. If the consumer has a vehicle, the consumer itself takes the role of a service provider of the mode. Hence, the business model SELF is best suited for multi-destination and/or on-demand trips since no premium cost is posed for the flexibility.

For cost calculation, the concept of perceived cost must be understood first. Usually, economic analysis of a certain vehicle focuses on life cycle cost, which includes direct/indirect operating cost and ownership cost. This approach is reasonable for accessing the viability of vehicle acquisition and for modal share problem of intra-urban commuting, but is not applicable for assessing long distance trips. People have already bought cars for everyday use and owner of a GAP vehicle does not care about acquisition cost anymore. Hence, the business model SELF takes the assumption that perceived cost is equal to out-of-pocket cost, excluding insurance, maintenance, and acquisition-related costs in comparison with costs of the other modes.

The business model SELF can employ vehicles CAR, GAP or GAJ. As such, three providers are defined and Table 43 describes how to calculate the perceived cost for each of the three providers. The total perceived cost is composed of vehicle operating cost and other miscellaneous cost such as landing / parking fees incurred to GAP:SELF and GAJ:SELF. Lodging cost is applicable to CAR:SELF and GAP:SELF, when driving or flight time exceeds a daily limit of 8 hours.

Table 43: Perceived Cost Composition for SELFs

	Vehicle Operating Cost (one-way)	Other Cost
CAR:SELF	GasolinePrice × Distance × 1/MPG	Lodging cost
GAP:SELF	AvgasPrice × GPH × FlightTime	Lodging cost, landing/parking fees
GAJ:SELF	$360/hr* × FlightTime	landing/parking fees

MPG: mile per gallon, GPH: gallon per hour

* This amount is taken from Hoffer et al. (1998).

The provider CAR:SELF adaptively changes vehicle operating cost calculation when the consumer is a business traveler. Most companies pay reimbursements for the use of personally owned cars while conducting business. Reimbursement for mileage is computed at 1995 IRS Standard Mileage Rate of 30 cents per mile. It is likely that this policy is not applicable to family-owned, small businesses. Table 44 shows percentage of accepting this policy by enterprise class.

Table 44: Acceptance Rate of Mileage Reimbursement Policy

Class I	Class II	Class III	Class IV	Class V
25%	75%	95%	99%	99%

5.3.2.2 Business Model RENT

The business model RENT receives payment from the consumers for the use of leased transportation resources. The providers CAR:RENT, GAP:RENT, GAJ:RENT belong to this model but CAR:RENT is omitted for simplification purposes. For mode GAP, further classification is useful. Depending on self-pilot abilities, a traveler can choose from the providers GAP:RENT and GAP:HIRE. While GAP:RENT provides only a vehicle for lease, GAP:HIRE offers pilot services as well. In line of this classification, the provider

GAJ:HIRE is the only type associated with mode GAJ. Common features of these providers are that they require certain minimum hours for rental and daily charges even if the consumer does not use their resources. The following table summarizes specific settings modeled regarding the business policy and rates for each of the providers in the RENT category. Small fees such as parking and landing fees are accounted for in addition to the basic rates.

Table 45: Rates Policy of the Business Model RENT

	Hourly Rates	Min. Rental	Overnight Charge
GAP:RENT	$90 wet hobbs	3 hours	Equivalent to 3-hr rental
GAP:HIRE	adding $80 for a pilot	3 hours	Equivalent to 3-hr rental
GAJ:HIRE	$1200	2 hours	$400/day

5.3.2.3 Business Model FARE

The business model FARE charges the price for a ride on a scheduled public transportation service. Only commercial airlines, called ALN:FARE, fall under this type of business model in the present research. The problem is that the price scheme of airline fares is notably complex to a degree that an ordinary consumer fails to see the mechanism behind it. This is because each airline has its own proprietary Revenue Management System (RMS) to control the availability and/or pricing of the seats in different booking classes with the goal of maximizing revenue. (McGill & Van Ryzin 1999) Any attempt to design an RMS is quickly going beyond the scope of the present research, so an alternative way needs to be taken to model the provider ALN:FARE.

First, the baseline fare was constructed as a function of flight distance D based on the market data from an internet-based travel agency. Regression on the data points results in the following second-order polynomial equation:

$$P = \sum_{i=0}^{2} \alpha_i D^i \tag{69}$$

where P indicates the round trip ticket price per a customer. The values of the parameters α_0, α_1 and α_2 are 8.60×10^1, 1.77×10^{-1}, and -2.46×10^{-5} in that order. Second, a set of assumptions has been made based on rationalization. For leisure trips, the consumer waits for a deal or discounted price. For business trips, time schedule and convenience is more important. Also, smaller companies are more price sensitive than their bigger conterparts. These behavioral distinctions can be alternatively modeled if the provider ALN:FARE has the capability of discerning the customer type. As such, Table 46 is created to imitate seemingly erratic price scheme.

Table 46: FARE Schedule by Enterprise Class

Consumer Type	% Consumers	Price Multiplier
Household	20	$\triangle(0.6, 0.8, 1.0)$
	60	$\triangle(0.8, 1.0, 1.2)$
	20	$\triangle(1.0, 1.2, 1.4)$
Class I Enterprise	40	$\triangle(1.0, 1.2, 1.4)$
	60	$\triangle(0.8, 1.0, 1.2)$
Class II Enterprise	20	$\triangle(1.0, 1.4, 1.8)$
	80	$\triangle(0.8, 1.0, 1.2)$
Class III Enterprise	20	$\triangle(1.0, 1.6, 2.2)$
	80	$\triangle(0.8, 1.0, 1.2)$
Class IV Enterprise	20	$\triangle(1.0, 1.8, 2.6)$
	80	$\triangle(0.8, 1.0, 1.2)$
Class V Enterprise	20	$\triangle(1.0, 1.8, 2.6)$
	80	$\triangle(0.8, 1.0, 1.2)$

The provider ALN:FARE adjusts the final price, taking the baseline price P and the price multiplier from the table depending on the consumer type and the enumerated probability. Lastly, as pointed out in §5.2.3.1, the consumer can choose from a multiple route choices depending on departure and destination locale. Hence, unlike the other providers, as many as four ALN providers co-exist during the simulation, competing each other with different price and time offers to the same consumer.

5.3.3 Mode and Route Selection

Now that all service providers are established, the consumer can ask each service provider for time and cost information. Then, the consumer adds time and cost of his portion such as portal access cost and time at the origin, and ground transportation cost at the destination when necessary. This concludes the computation of expected time and cost for each mode, often referred to as disutility of the mode, which is an opposite concept of utility.

5.3.3.1 Formulation of Mode Utility and Choice Structure

The disutility of a travel mode m comprises the amount of cost $C_i(m)$, time $T_i(m)$ as well as nuisance $N_i(m)$—level of impedance due to the consumer's concerns over safety, comfort, etc. It is a common practice to use the additive assumption, i.e., the utility for the consumer i of mode m can be expressed as

$$U_i(m) = -\alpha\{c_iC_i(m) + t_iT_i(m) + n_iN_i(m)\} \tag{70}$$

where negative α is taken since large cost, time or nuisance is not desirable to the consumer. The three positive weights c_i, t_i and n_i reflect the agent's perception on the importance of money, time and nuisance, respectively. Notice that t_i is a product of weight of time (w_t) and hourly income of the consumer. The problem is that the modeler has incomplete or unobservable information on $N_i(m)$ that cannot readily be quantified, especially by outside observers.

Previously introduced probabilistic choice theory can resolve this issue by taking the immeasurable $N_i(m)$ terms as the random utility (ε) in Equation 25. (See §3.3) Then, the deterministic (or systematic) utility is now given by

$$V_i(m) = -\alpha\{C_i(m) + t_iT_i(m)\}. \tag{71}$$

Simplifying the assumption is that α is a constant and $c_i = 1$ for all i. Hence, α serves as the only parameter used for model calibration. Having defined the systematic utility, the next step is to choose a proper discrete choice model in the context of the problem.

The Multinomial Logit Model is widely accepted because of its simplicity. However, the transportation practice usually adopts the Nested Logit Model (NLM) when the possibility of having correlated modes cannot be eliminated completely. As such, a three level nest structure is created for analysis of modal split as shown in Figure 67.

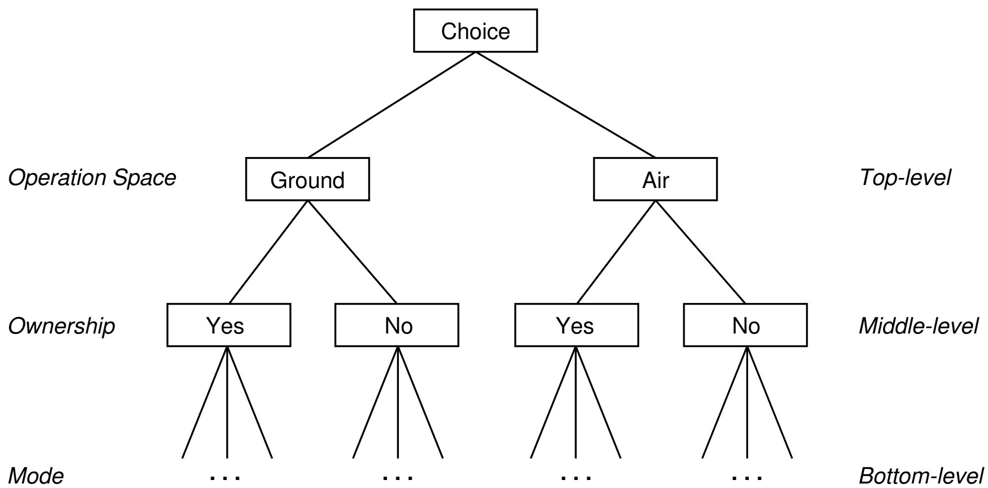


Figure 67: Nested Mode Structure

The top level describes the choice between ground and air transportations. Within each top level nest, the middle level creates the choice of sub-nest according to ownership. The bottom level contains the choice mode within each sub-nest. As a result, the modes within the same sub-nest share the random utility or nuisance term.

5.3.3.2 Development of Tournament Logit Model

Although this structure can accommodate any mode, a few technical complications exist with the problem at hand. First, as indicated in Equations 47 and 48, the dissimilarity parameter for each nest as well as for each sub-nest should be estimated. Second, the choice modes can float across sub-nests, meaning that depending on ownership and self-pilot capability a mode can belong to a different sub-nest. Moreover, it is possible for PAV problems that an alternative can be a member of different nests. For example, a dual mode PAV would belong to both ground and air nests.[12] Lastly, ALN route choices should be

[12]This overlapping situation can be tackled with other choice models such as the Paired Combinatorial Logit model (Koppelman & Wen 2000), however, at the price of extra complexity.

plugged in under the mode ALN, making mode ALN itself the sub-sub-nest.

Under such circumstances, it is necessary to have a choice model that is both flexible and expandable. The Tournament Logit Model (TLM) is formulated in an attempt to address this challenge. As the name implies, the algorithm makes an analogy to a tournament. Suppose alternative A_o is chosen over other alternatives A_i within a nest. A_o obviously advances to the next tournament or the upper-level nest just like the NLM, but the difference is that it takes the maximum utility among A_i. Mathematical representation of the TLM is quite straightforward. Envision a case with two levels. The probability of choosing mode m given that a mode in nest N_k is chosen equals to

$$\pi(m|N_k) = \frac{e^{V_m}}{\sum_{j \in N_k} e^{V_j}}. \tag{72}$$

And the marginal probability of choosing a mode in nest N_k is expressed as

$$\pi(N_k) = \frac{e^{V_k^*}}{\sum_{n=1}^{K} e^{V_n^*}} \tag{73}$$

where $V_k^* = \max_{j \in N_k} V_j$. Then, the product of the marginal and the conditional probability concludes the probability of mode m being chosen.

$$\pi(m) = \frac{e^{V_m}}{\sum_{j \in N_k} e^{V_j}} \cdot \frac{e^{V_k^*}}{\sum_{n=1}^{K} e^{V_n^*}} \tag{74}$$

As shown, no extra parameters are needed for characterizing the overall nest structure and the nests at different levels. This simple algorithm reduces the problem complexity to a manageable degree in spite of the aforementioned situation. Also, it resonates with the consumer's mindset. When a seemingly inferior mode is chosen, it has a certain reason—which cannot be captured in the deterministic utility—to the consumer, who would hang on to the reason all the way to the top-level.

5.4 *Development of Simulation Framework*

This section outlines the tangible simulation framework developed for this dissertation effort. Detailed information on the main simulation code is described first, followed by a brief introduction to its distributed computing environment and GUI environment. Verification of the code and the environment is presented last.

5.4.1 Simulation Code: *Mi*

The simulation code ***Mi***[13] was originally developed in Microsoft® Excel, which made manipulation of the initial idea quite easy despite the slow simulation speed. This incipient version focuses on building a mode choice model with the weighted sum method. (Lewe et al. 2002) ***Mi*** has seen numerous updates, and the latest version employs the Java™ programming language for several reasons. The principle of objected-oriented programming (OOP) is harmonious with that of the entity-centric abstraction. Also, implementation of the ABM in an OOP language is much more natural than through traditional languages such as FORTRAN or C. Additionally, the object-oriented nature of ***Mi*** allows for easier expansion as need arises. The platform independence is another attractive feature of the Java language, considering the practical goal of this research is to provide a simulation model that can be shared with the transportation stakeholders.

5.4.1.1 **Mi** *Class Diagram*

Programming of ***Mi*** means the conversion of each entity in the TAF into a Java class. In OOP terminology, a class is loosely defined as a template from which corresponding object instances are populated. All classes of ***Mi*** are constructed to meet the specifications explained in the previous sections. With the Unified Modeling Language (UML) notation, the relationships amongst classes can be depicted in a class diagram. The class diagram of ***Mi***, shown in Figure 68, can be considered a class of the virtual TAF as a whole.

[13] The name of ***Mi*** is originated from a fusion of the East and the West languages. The Roman numeral **M** means 1000, **Mi** stands for mile, ***i*** is for eye. In the East, the term 'thousand-mile-eye' means a clairvoyant.

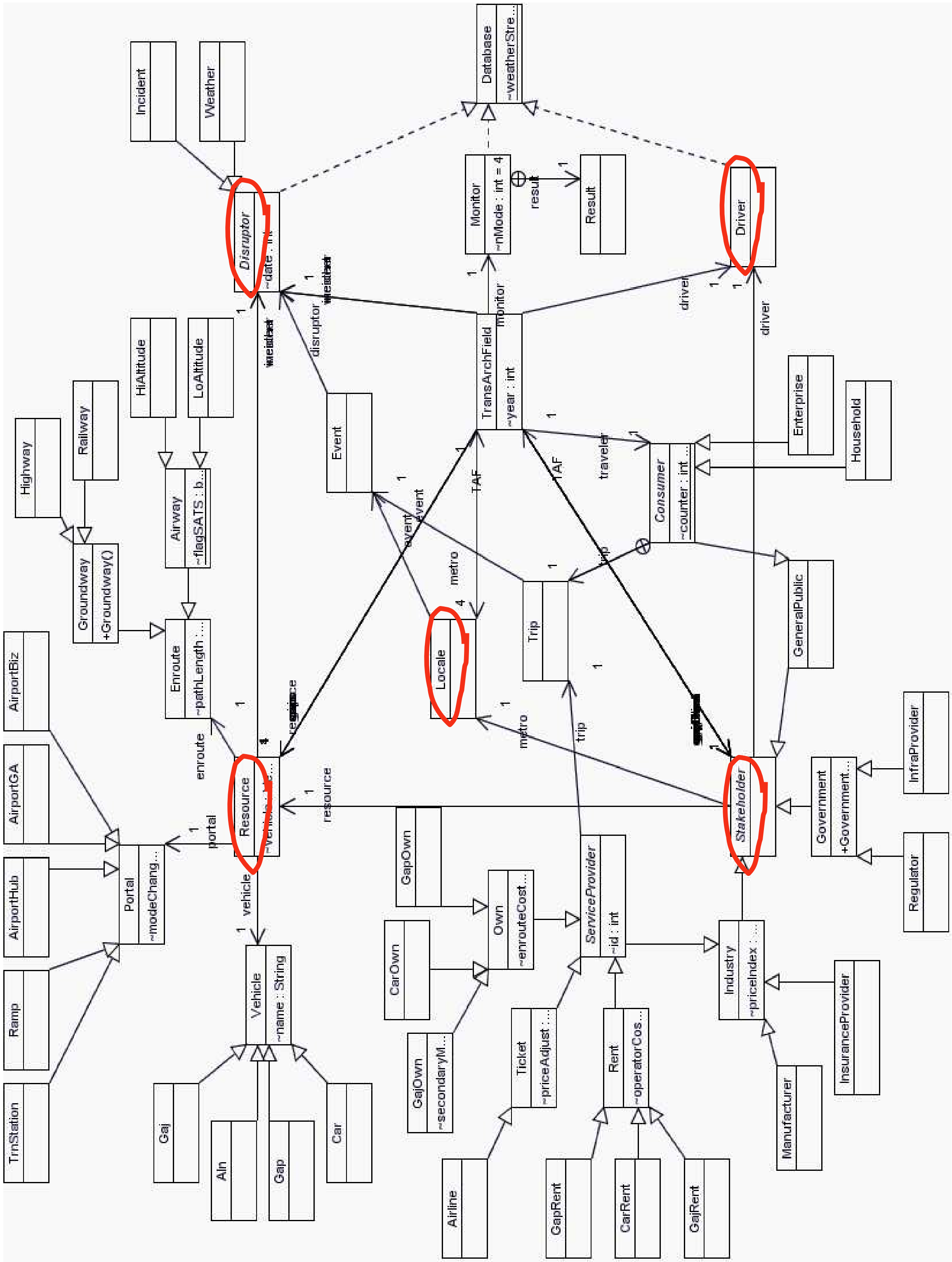


Figure 68: ***Mi*** Class Diagram

Each rectangle indicates a Java class, with solid lines indicating the interdependency between the classes. Once a class is established, it is quite easy to generate a child class that inherits the attributes and behaviors of the parent class, indicated by empty arrows. The classes wrapped by the red curves are pivoted in positions to point out a correspondence to the four quadrants and the transportation environment shown in Figure 42. The class diagram generically shows how the TAF can be interpreted and implemented, but is subject to change in accordance with the evolving ***Mi***.

5.4.1.2 ***Mi*** *Flowcharts*

The top-level sequence of the program flow of ***Mi*** is shown in 69(a). After the definition of a simulation scenario, ***Mi*** generates an "instance" of the virtual TAF, and triggers that instance to run itself. Figure 69(b) is an enlarged depiction of the "Generate & Run TAF" steps as the virtual world's most simplistic formulation is the environment and the agents. The agents' environment is defined when the drivers, the disruptors, the infrastructure, and the locales are formulated in association with the databases necessary as indicated in Figure 69(c). The consumer and the service provider stakeholders "live" in the environment, and they interact with each other as illustrated in Figure 69(d).

A series of such procedures constitute a unit cycle of ***Mi*** simulation. On a typical desktop computer, this task takes about ten minutes of computation time for a simulation with one million consumers. After one cycle is completed, the consumers learn and update their information collectively. The same sequence is iterated until the "equilibrium" state—when $\delta(m)$ in Equation 68 approaches zero—is reached, after which the simulation of a scenario is concluded. Such a simulation approach is known as the agent-based evolutionary scheme. (Niedringhaus 2004) Essentially, the scheme allows for the agents to live in the same environment over and over until they "get it right."

If the arrived-at-result is not satisfactory (e.g., calibration), the user has to modify the scenario and/or model parameters. The whole sequence is repeated, implying that ***Mi*** simulation calls for a nested loop iteration—the autonomous agents invokes the inner-loop simulation, and the user performs the outer-loop simulation at his/her discretion. As a result, the whole execution of ***Mi*** simulation can become quickly computationally expensive. Initially, it was envisioned to assign one Java thread for each consumer. But the computational overhead of managing threads on top of already prohibitively demanding tasks makes this plan infeasible.

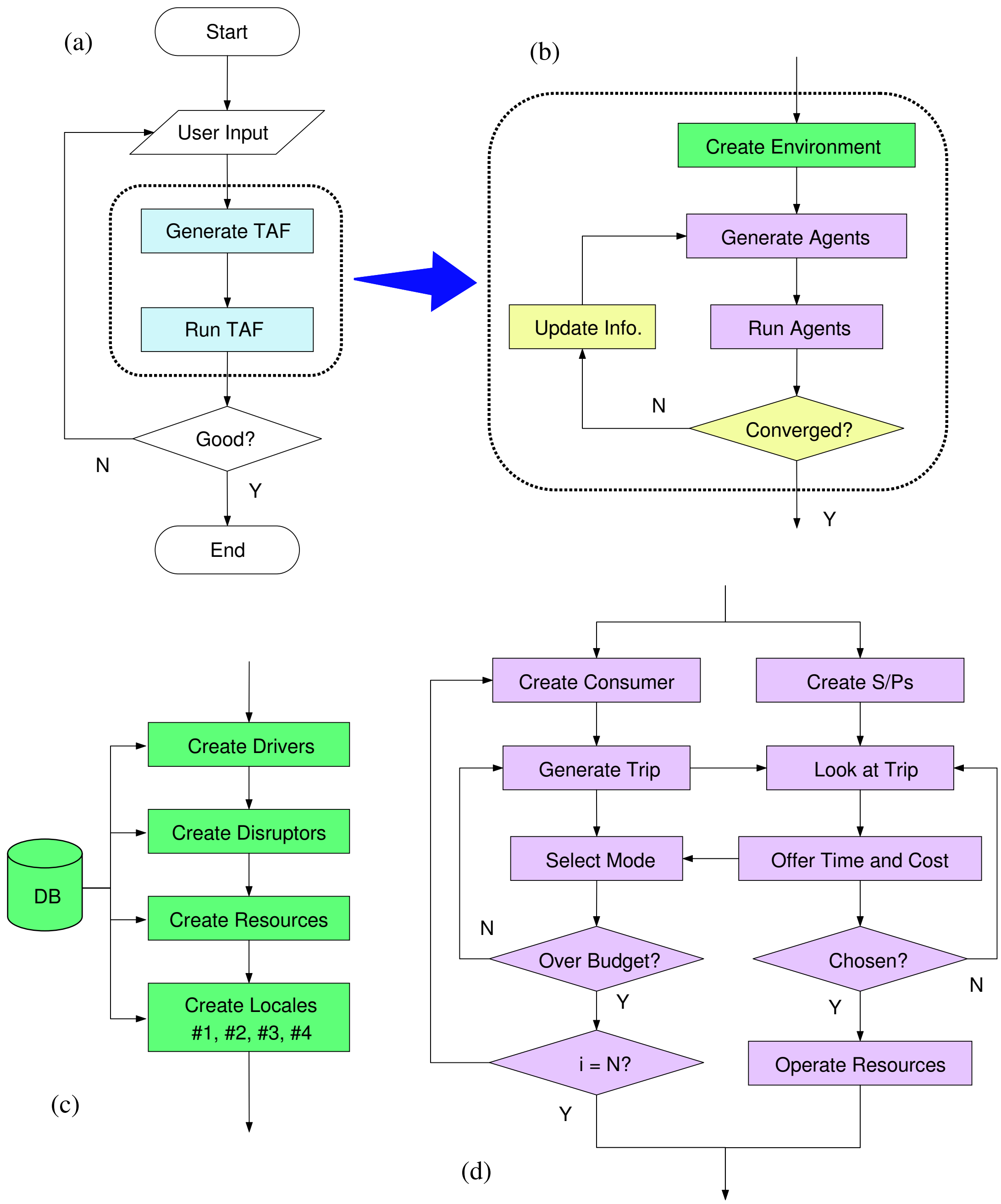


Figure 69: ***Mi*** Flow Charts

5.4.1.3 Distributed Computing Environment

As pointed out in §5.4.1.2, substantial amount of computation time is required for a unit simulation scenario, which would eclipse the usefulness of the simulation and thus would become the major bottleneck of the decision making process. To alleviate the computational burden, the Multi-platform Integrated Development Aid System (MIDAS) was developed. The objective of MIDAS is to conveniently interweave available computing resources across a Local Area Network (LAN), regardless of the operation systems.

The foundational building block of MIDAS is JavaSpaces™ technology, a high-level Application Programming Interface (API) of the Java network technology. Using JavaSpaces technology, one can flexibly create and manage a "space"—a logically shared memory where data can be stored, accessed, and updated in real time that naturally facilitates a master-and-worker type distributed computing environment, as shown in Figure 70. The space acts as a medium that connects the master and the workers, and where all storage and interchange of task and result entries take place. Another benefit of creating such a space is that the resulting environment becomes inherently self-load balancing. More in-depth technical details of MIDAS are given in Appendix D.

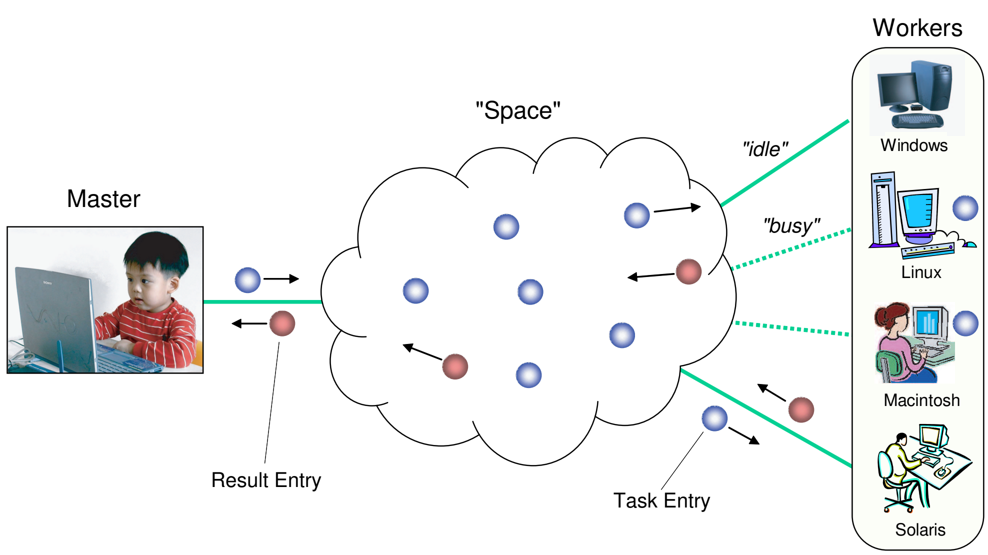


Figure 70: MIDAS Concept of Operations

5.4.1.4 Graphical User Interface

A Graphical User Interface (GUI) provides easy access to the simulation program for inexperienced external users as well as developers. A good GUI should facilitate parameter tuning for calibration, feeding scenarios to the simulation engine and interpreting the results of the simulation. To maximize the usability, a web-based GUI implementation and distribution is of paramount importance. As such, a branched effort to develop a GUI version of ***Mi*** was initiated. Development of the GUI, however, put an additional burden on the programmer. The implementation of the basic functionality was not difficult but, to say the least, time consuming. In addition, up-to-date revision of the GUI needs much attention along with the ever evolving source code of ***Mi***. Therefore, an older version ***Mi*** is currently mounted to the GUI environment.

Screen shots of ***Mi*** GUI are captured in Figure 71. This Java applet is operational and can be accessed via the PAVE web site at http://www.asdl.gatech.edu/teams/pave at the time of this writing. A user can investigate the effect of a new mobility vehicle by changing its design attributes at the input fields located in Figure 71(a). The five bar charts on Figure 71(b) are for calibrating the model with black hairlines indicating calibration target points. The three bar charts on the left indicate the usage distributions of the three modes by distance bracket. The two-colored bar chart on the bottom-left corner shows the mode split between modes ALN and CAR. The last bar chart on the top-right corner shows the market share by modes ALN and CAR and by trip purposes. Also, the simulation result is visualized in a "market space" plot, showing the distribution of the agents' mode choices over annual income and travel distance, as shown in Figure 71(c). Each mark on the space represents a single trip and a three color scheme of blue/cyan (CAR), red/magenta (ALN) or yellow (GAP) is employed to indicate a consumer's choice mode. Thus, a decision-maker is able to observe the potential GAP market region along with other modes in a highly visual and dynamic manner.

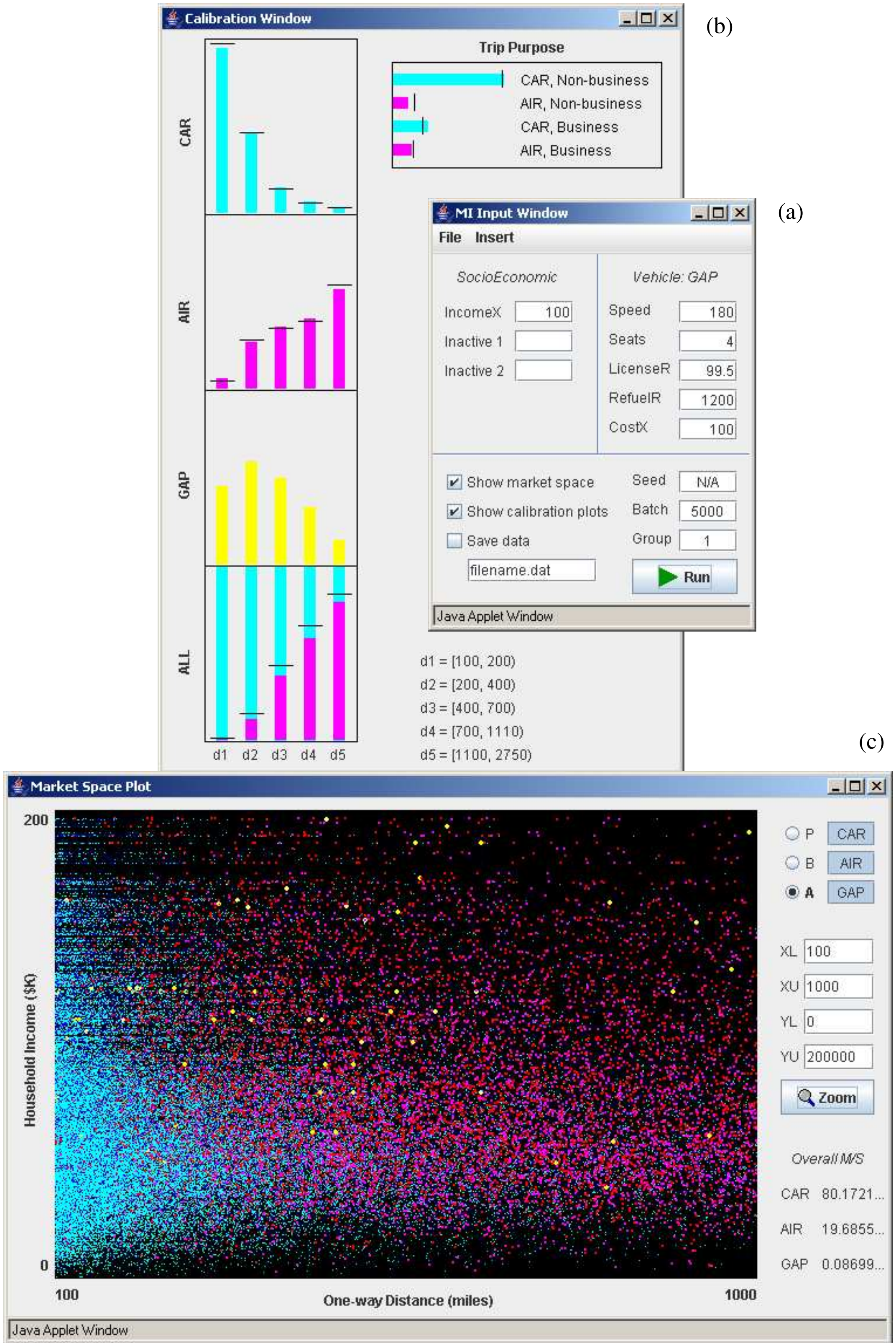


Figure 71: ***Mi*** GUI Screen Shots

5.4.2 Verification

As explained in §4.3.2.1, model verification is a mandatory premise before calibration and simulation, insuring the code ***Mi*** actually performs what it is intended to do. The primary focus in this effort was on the agents' behavioral pattern since it is the fundamental mechanism in the agent-based model.

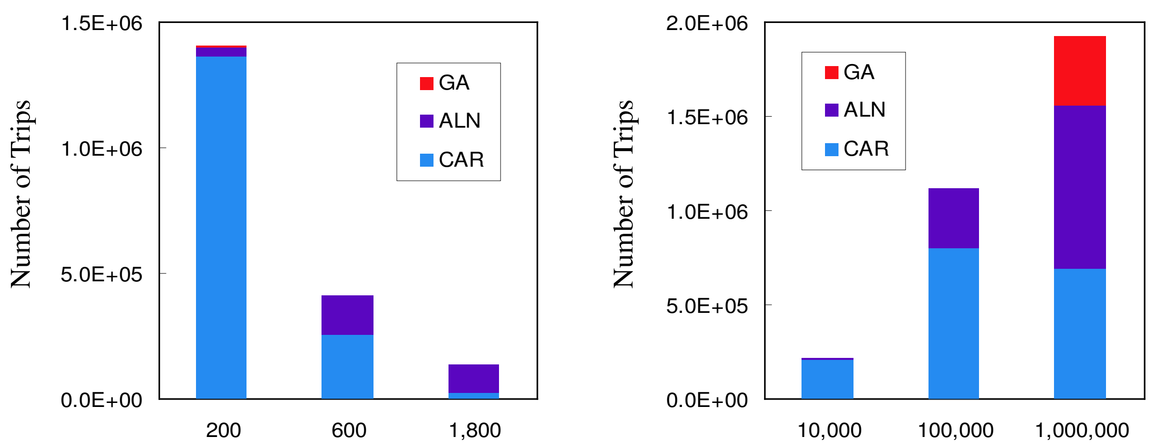
First, accurate agent generation was monitored based upon the specifications outlined previously. Of particular interest is the partitioned income model postulated in §5.2.4.1 since the overall income model significantly affects the simulation outcome. Moreover, any simulation result regarding modes GAP and GAJ will be very sensitive to how the high income model is configured. One million households were generated and their income data were extracted by resident locale. It was found that the overall distributions and percentile statistics were very close to the real data. Even so, this is not sufficient to confirm the high income model adequately reflects reality. Despite occupying small portion of all households, the high income group can change the average income to a significant degree. Table 47 underscores that the postulated high income model is satisfactory where the fourth column indicates the percent difference in average income to the real data in Table 17.

Table 47: Household Income Comparison

Locale	# Households	Average Income	Difference
LAR	284,896	66,563	-0.51%
MED	209,467	58,738	0.06%
SML	190,008	52,131	-0.04%
NOM	315,629	48,353	-0.71%
Total	1,000,000	56,434	-0.37%

Subsequent effort focused on verifying whether ***Mi*** or not fulfilled the qualitative requirements identified in regards to the generic behavioral rules of the consumers. To succinctly monitor the responses of the agents, six cases were run and the results are shown in Figure 72. The first three cases were for monitoring the trip distribution by trip distance:

200, 600, and 1,800 miles, one way. Observation of Figure 72(a) reveals two trends. First, the amount of trips is decreasing as trip distance is getting longer and second, the modal split trend seems to be reasonable. Notice that modes GAP and GAJ are not discernable enough, so they are indicated by code GA with red color. The next three cases were run fixing the income at three levels: \$10,000, \$100,000 and \$1,000,000. As shown in 72(b), households with higher income are more likely to travel more frequently than a lower income household. The members of the lowest income group cannot travel frequently and, if they do, they take shorter trips or by the most economic mode—mode CAR. The wealthiest group has much more degrees of freedom in choosing a suitable mode. To summerize, all verification results were satisfactory; the basic agent decision-making algorithm and implemented steps were prompt and responded quite well without need for user interventions.



(a) by Distance (miles) (b) by Household Income (USD)

Figure 72: Mode Choice Model Verifications

Lastly, the performance of MIDAS was also verified, which showed an expected linear scalability with respect to the number of workers connected to the space. For detailed documentation of this process, interested readers are referred to Appendix D.

CHAPTER VI

SIMULATION STUDIES

Contents

The calibration process is the next step after verification of the simulation code and the distributed computing environment. The first section of this chapter details how this task is accomplished. In the following sections, the capabilities of the integrated simulation framework are tested and showcased through a number of empirical studies, accompanied by discussion of the simulation results and their implications.

6.1 Baseline Scenario and Calibration of the Model

A first step for any simulation based study is to develop a solid baseline to which a formulated simulation scenario can be compared against in a coherent way. This section attempts to articulate the baseline model and its calibration result, followed by the prospective simulation scenarios outlined for this dissertation effort.

6.1.1 Preparation of the Calibration Datums

The baseline would describe a reasonable approximation of the system's present state. Such a description can be best found in the 1995 ATS database. Hence, the calibration datums should be prepared in such a way that the virtual world can be tuned to it.

6.1.1.1 The Rugged Trip Matrix

Suppose that a unit simulation is completed and the real trips for each consumer (the last step in Figure 64) is stacked together in a list as conceptually represented in Figure 73. This is called the rugged trip matrix where the squares contain trip-specific information such as destination locale, distance, purpose, number of a trip party, mode choice, weather condition, delay, etc. When these entries are combined with the attributes of the consumer C_i such as origin, income, and mobility budget, the investigator can retrieve comprehensive information of all transportation activities in the virtual NTS.

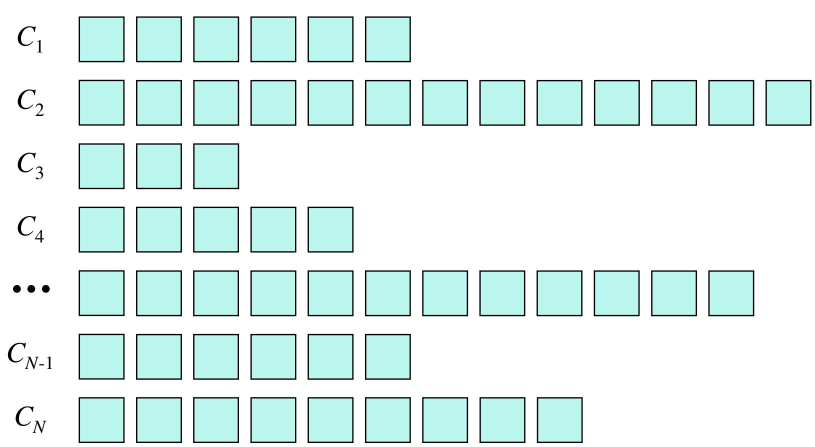


Figure 73: Rugged Trip Matrix

The best way to treat this massive rugged trip matrix would be to employ a database engine since a typical simulation generates one million consumers. Instead, ***Mi*** is equipped with a built-in data monitor module that counts the number of trips by choice provider and by a 50-mile distance bracket. The outcome of this procedure is recorded in the template form shown in Figure 74.

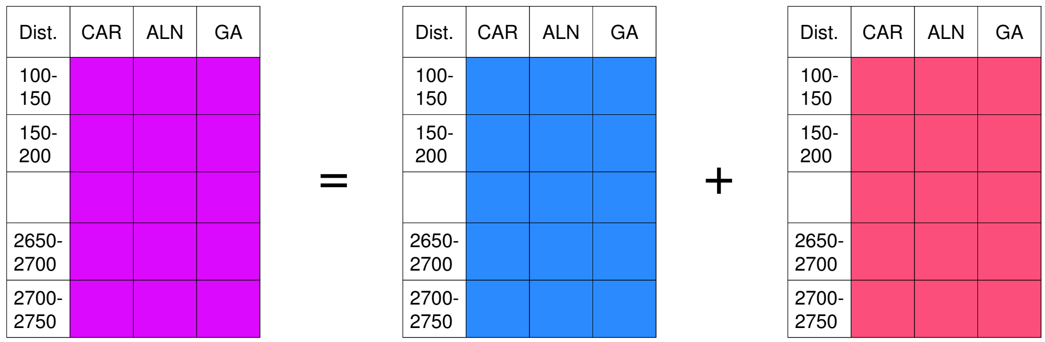
Distance Bracket	CAR	ALN	GAP			GAJ	
			SELF	RENT	HIRE	SELF	HIRE
100 to 150							
150 to 200							
200 to 250							
2,600 to 2,650							
2,650 to 2,700							
2,700 to 2,750							

Figure 74: Postprocessing Format

As mentioned previously, the scope of 1995 ATS database is in accordance with that of this research, from which much information on the NTS can be retrieved with sufficient levels of detail. The only missing attribute, however, is the detailed distinction regarding general aviation. The database aggregates all general aviation modes as a "corporate/personal airplane" mode. Hence, the finespun calibration on modes GAP and GAJ are regrettably omitted from the outset. If this deficiency is overcome in future surveys, the virtual world will have much improved credibility for general aviation.

6.1.1.2 The Calibration Datums

The construction of the datums began by screening the database. It was a necessary step to exclude irrelevant data such as international destined trips, bus / train / ship trips and trips greater than 2,750 miles to abide by the scope of the study. The screening result should be processed according to the format shown in Figure 74. Since aggregation of columns GAP:SELF through GAJ:HIRE into a single column GA was unavoidable, the screening outcome would be recorded in a "mode-distance matrix" with 53 rows of distance brackets and 3 columns of modes. The last process is the decomposition of the mode-distance matrix by trip motivation as two types of the consumers are included in the model, as shown in Figure 75.



(a) Total Trips (b) Non-Business Trips (c) Business Trips

Figure 75: Decomposition of the Mode-Distance Matrix

Let $[ALL]$, $[PSN]$ and $[BIZ]$ stand for the mode-distance matrices for all trips, non-business trips and business trips, respectively. The goal of calibration is each of mode-distance matrix from the ATS database and calibration simulation is proportionally equivalent, i.e.,

$$[PSN]_{ATS} = [k] \circ [PSN]_{CAL} \tag{75a}$$

$$[BIZ]_{ATS} = [k] \circ [BIZ]_{CAL} \tag{75b}$$

where the symbol $\circ$ indicates the Hadamard product and the 53-by-3 matrix $[k]$ is filled up with the same constant k. When Equation 75 suffices, the last condition $[ALL]_{ATS} = [k] \circ [ALL]_{CAL}$ becomes automatically trivial.

Figures 76(a), (b) and (c) show the final results by taking each column of the matrix $[ALL]_{ATS}$ and normalizing to the column sum, where the y-axis indicates the relative frequency and the x-axis represents the distance bracket. Subsequently, Figures 77 and 78 describe the mode-distance matrices $[PSN]_{ATS}$, and $[BIZ]_{ATS}$ in the same format. As shown, the difference is not dramatic for mode CAR whereas it is more noticeable for modes ALN and GA. These charts are visually useful yet not appropriate in the present

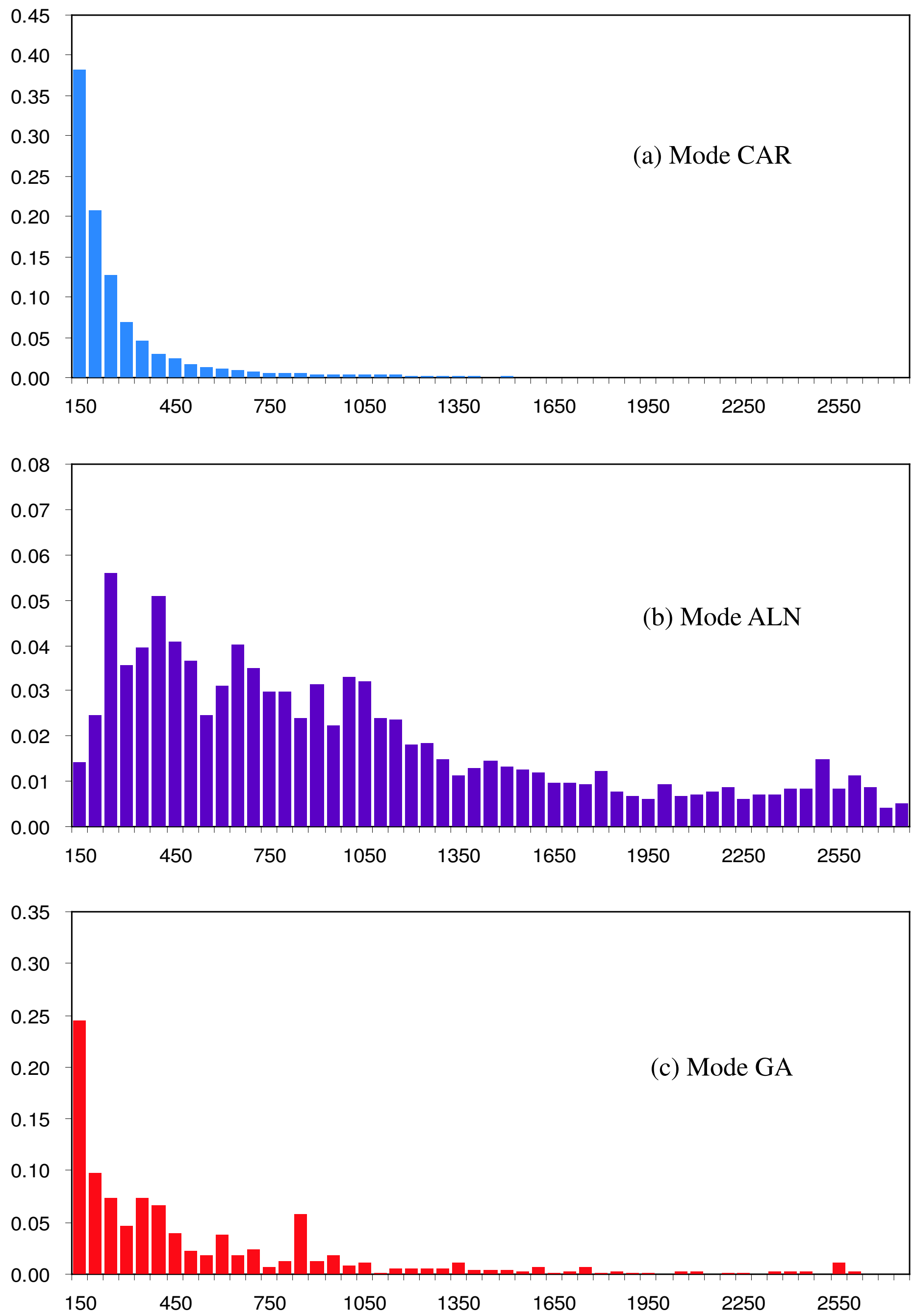


Figure 76: Trip Distance Distribution by Mode from $[ALL]_{ATS}$

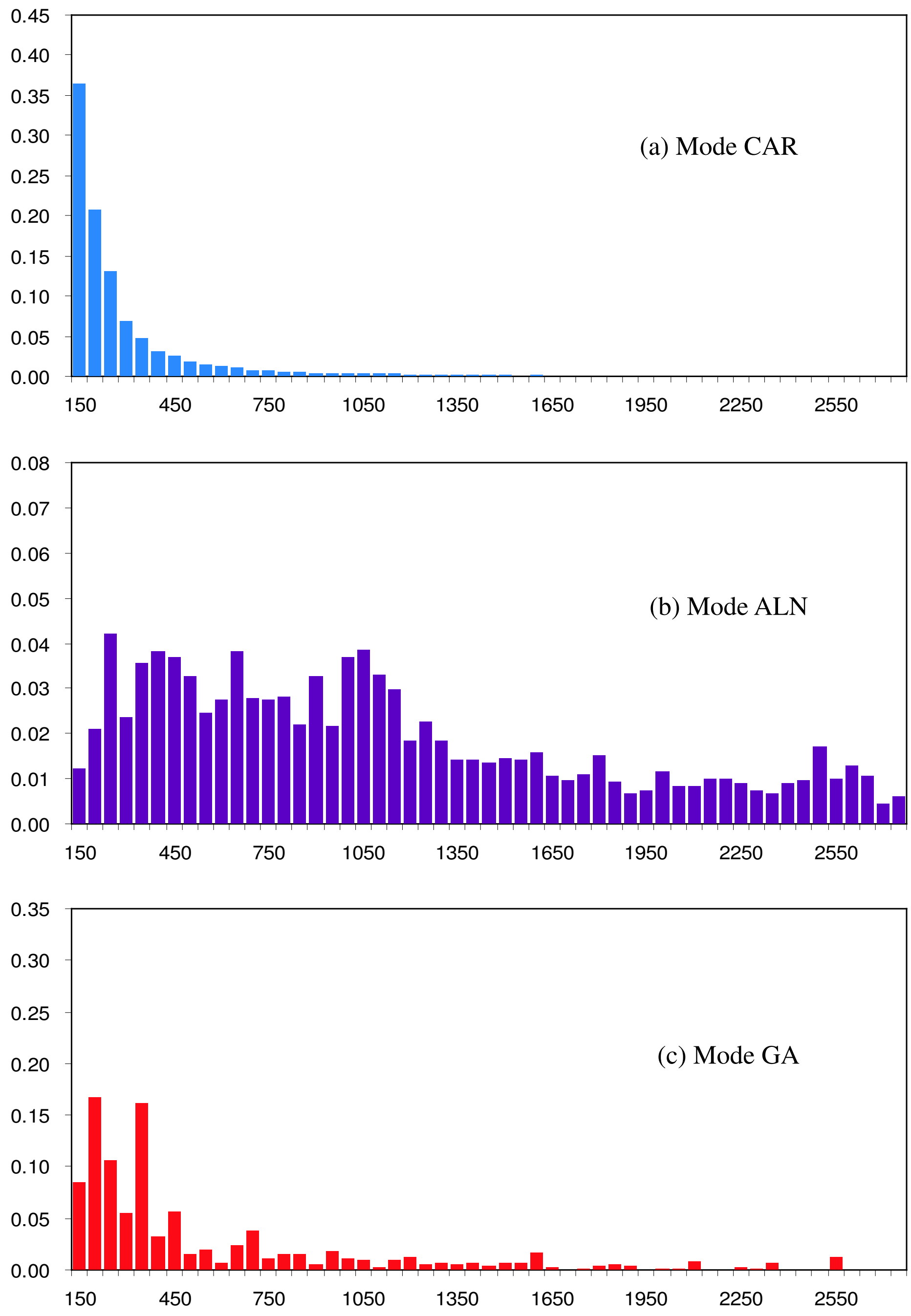


Figure 77: Trip Distance Distribution by Mode from $[PSN]_{ATS}$

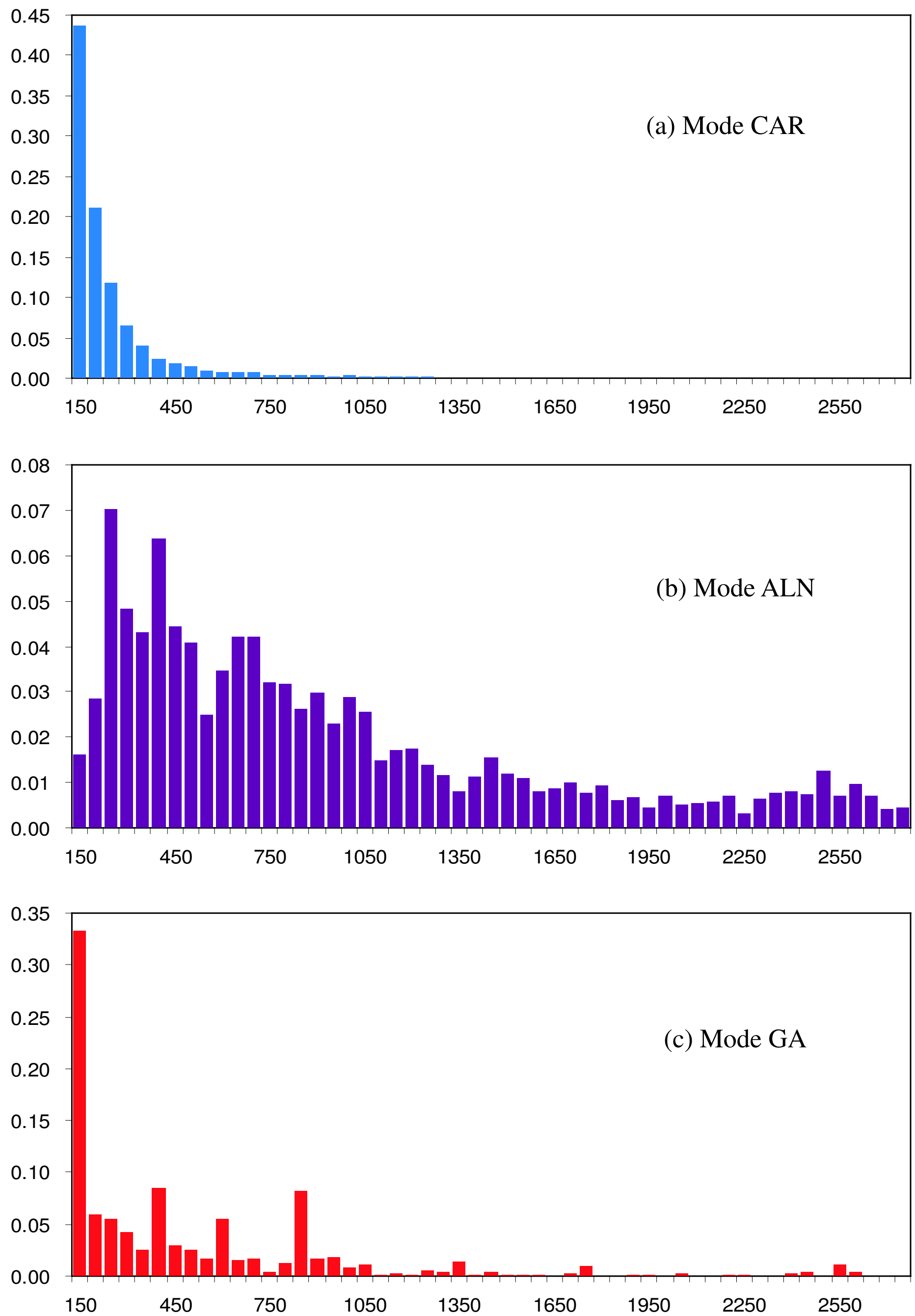


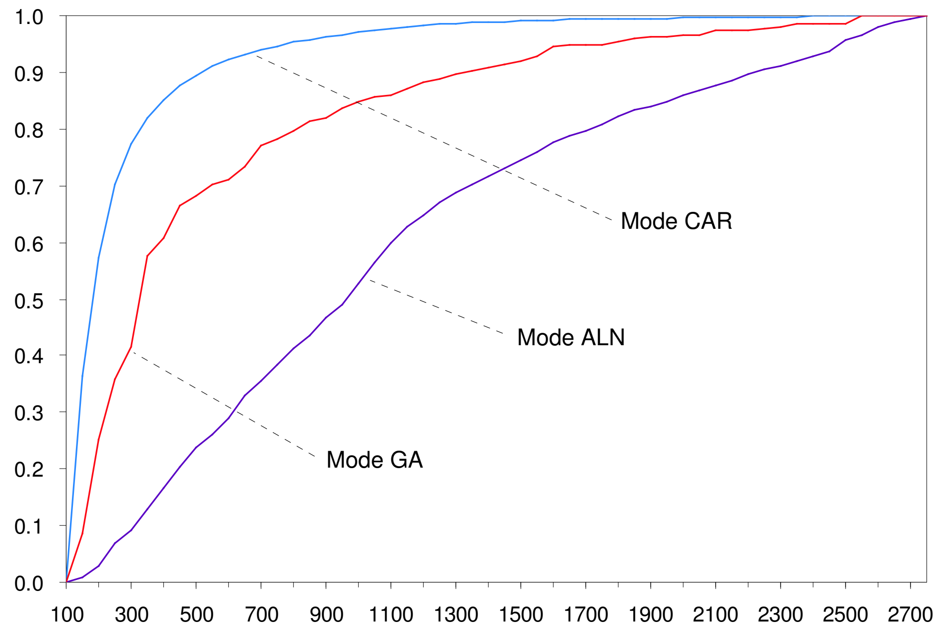
Figure 78: Trip Distance Distribution by Mode from $[BIZ]_{ATS}$

conditions for the calibration datums with the above stringent calibration objective as the bars of both mode ALN and GA charts are extremely irregular to follow for any simulation model. Instead, the cumulative distribution format is adopted to present the datums. The cumulative frequency against trip distance of each mode is shown in Figure 79, where mode GA features the rough curves. This implies that the sample size of the ATS survey was not sufficient for the GA trip.

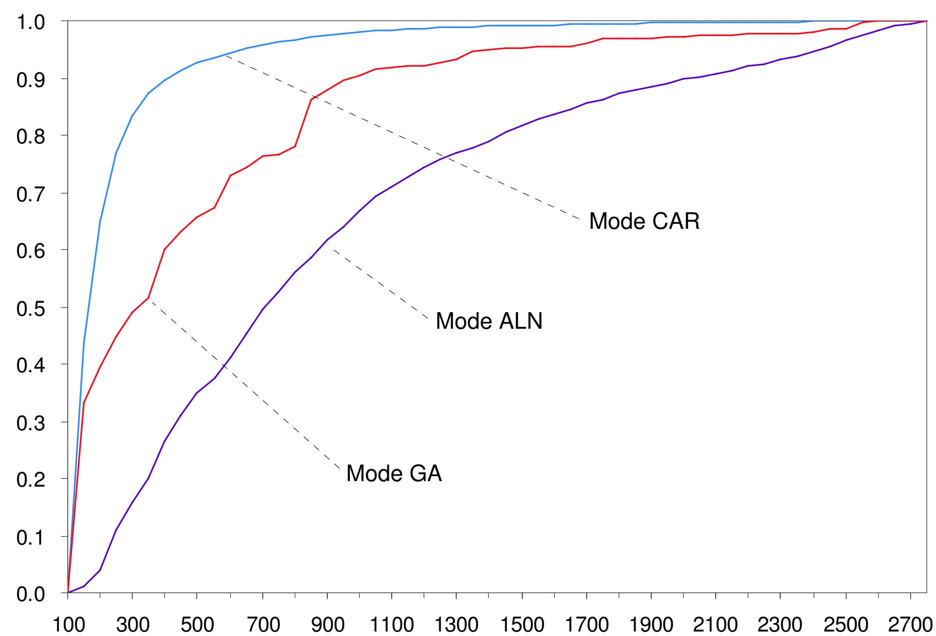
So far, each mode has been independently examined by looking at the column data alone. In order to satisfy Equation 75, interaction between the modes should also be analyzed by examining the row data. Figure 80 is prepared where the three-colored area shows the changes in the relative market share of the modes according to the corresponding distance bracket. The distinct behavior of the household consumers and the enterprise consumers becomes much clearer with these charts. For the use of calibration datums, only the market share of mode CAR (blue) is taken, as that of mode GA is negligible. In an actual calibration practice, Figures 79(a) and 80(a) speak to the condition of Equation 75a. Likewise, Equation 75b is examined through Figures 79(b) and 80(b). Lastly, the constant k should be equal, so all trip data are counted and grouped by mode and by purpose as shown in Table 48, where the values in parenthesis shows the percent to the total amount of trips. This table, along with Figures 79 and 80, constitutes the calibration datums to be used later.

Table 48: The Amount of Trips by Mode and by Purpose

	Personal Trips	Business Trips	All Purposes
Mode CAR	295,701,927 (57.38%)	91,506,292 (17.76%)	387,208,219 (75.14%)
Mode ALN	62,835,578 (12.19%)	61,331,079 (11.90%)	124,166,657 (24.10%)
Mode GA	1,392,778 (0.27%)	2,550,262 (0.49%)	3,943,040 (0.77%)
All Modes	360,354,497 (69.85%)	155,823,761 (30.15%)	516,178,258 (100.0%)

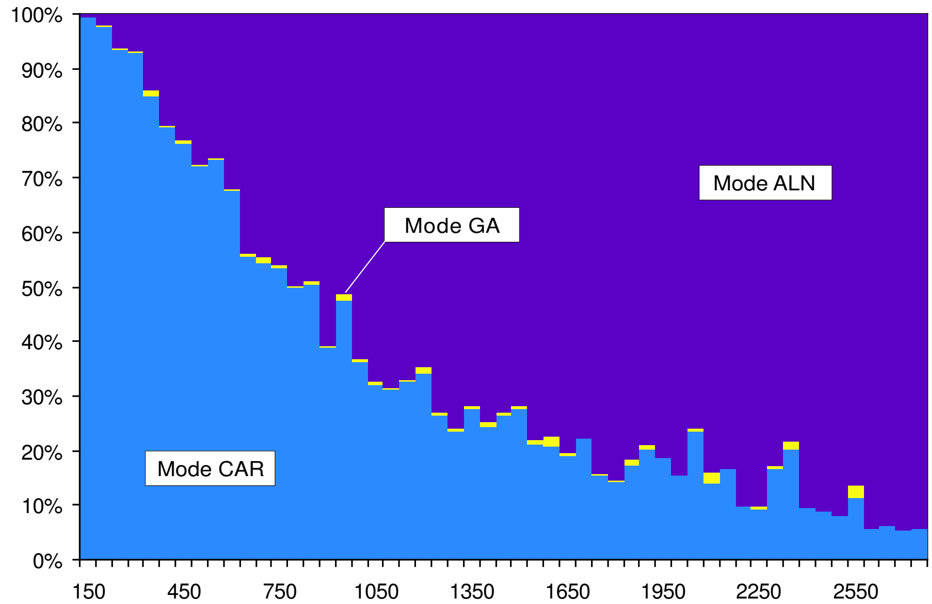


(a) Non-Business Purpose

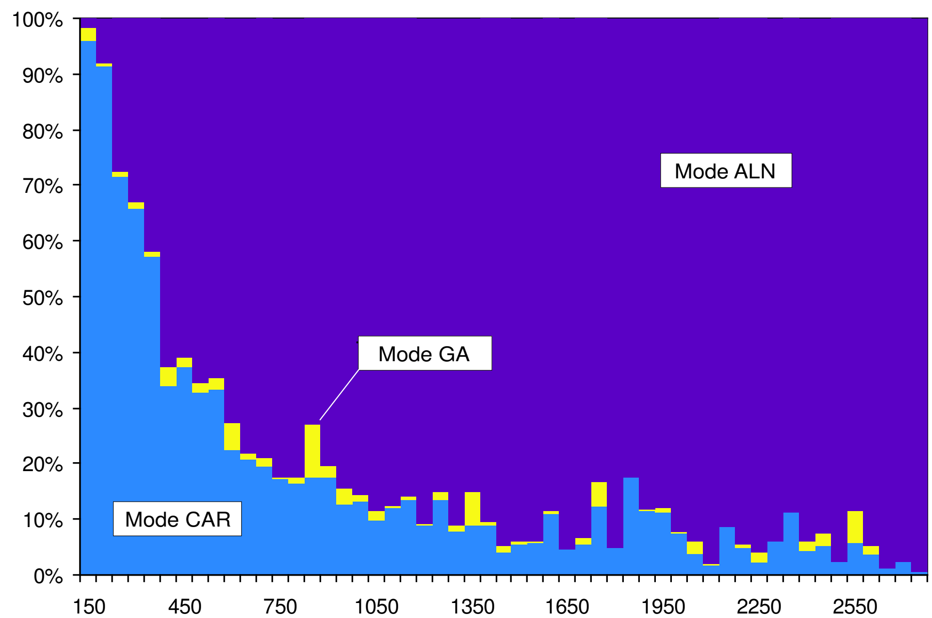


(b) Business Purpose

Figure 79: Cumulative Distribution by Mode



(a) Non-Business Purpose



(b) Business Purpose

Figure 80: Market Share Profile by 50-mile Distance Bracket

6.1.2 The Baseline Model

The un-calibrated model ***Mi*** requires parameter adjustments as well as supplementary modules. As pointed out in §4.3.2.2, the construction of an agent-based model revolves around the question "can you grow it?", instead of "can you explain it?". Hence, understanding and learning the target system is in the course of the calibration of ***Mi***. In that light, this subsection documents necessary "touches" for completing the baseline model, followed by the final calibration results.

6.1.2.1 Model Tuning

Numerous yet necessary adjustments are related to economic parameters. All monetary values in the baseline model are scaled to 1995 nominal dollar. Commodity prices were referenced from various historical data when available, otherwise, calculated by considering present data and inflation rates. For example, the average prices of motor gasoline and aviation gasoline are set to $1.20 and $2.20 per gallon. The living expenses and labor costs are largely correlated to the size of the metropolitan area, so these price data take some variability with an average consumer price index of the corresponding locale. The income model already incorporates this distinction with respect to locale, but needs additional scaling with GDP deflator besides the inflation adjustment as the income model was formulated based on Census 2000 data that describes income in 1999.

It is assumed that the macroscopic behavior of the transportation activity follows the gravity model postulated in §5.2.2.2. To obtain an estimate of γ in Equation 60, Figure 81 is prepared where the data points indicate the relative frequency of all transportation activities from the mode-distance matrices $[PSN]_{ATS}$ and $[BIZ]_{ATS}$ by distance bracket. As shown, the fitting curves verify that the postulated gravity model is close to reality. The baseline model takes slightly reduced values of 1.6 and 1.4 for non-business and business trips respectively, to account for the latent demands that cannot be captured due to the mobility budget constraint.

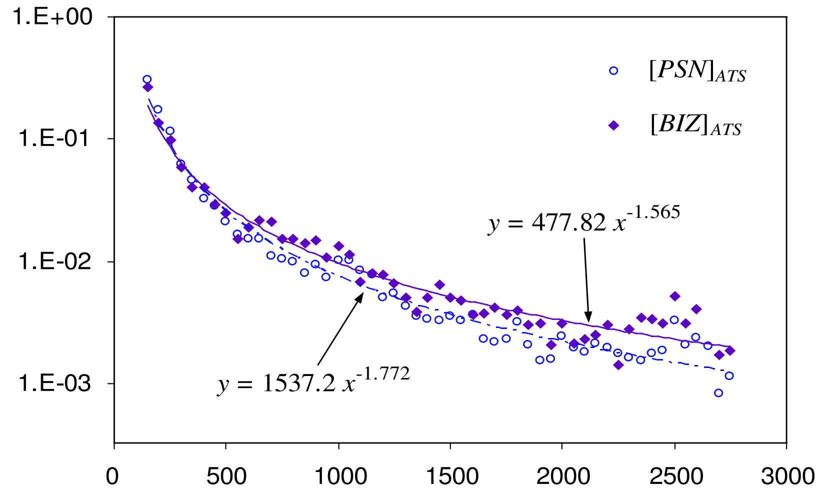


Figure 81: Validity of The Gravity Model

The mode choice model established in §5.3.3 is very influential to the simulation outcome. For instance, if the parameter α in Equation 71 is extremely low, consumers become indifferent to all choices. The baseline model takes the values of 0.10. Other alterations in the consumer model involve psychological factors. For most non-pilot travelers, choice of small aircraft as the primary mode would be psychologically challenging. To reflect this barrier for non-pilot consumers, the TLM is repeated once when the first choice is mode GAP. On the contrary, a GAP owner has more desires for long distance trips, inducing the vehicle purchase in the first place, and vice versa. This propensity is implemented with boosting the transportation budget and the trip distance. The percentage of the transportation budget is increased to $\triangle(0.04, 0.08, 0.12)$, and trip distance is still assigned from Equation 60 but with $\gamma = 1.1$.

An interim model with the above modifications produced Figure 82. The model outcome is represented by the dots whereas the datum from the matrix $[ALL]_{ATS}$ is represented by the solid lines. ***Mi*** is capable of reproducing the empirical data to high proximity, without any coercive parameters and mechanisms. Similar trends were observed with the matrices $[PSN]$ and $[BIZ]$ as well. Given these satisfactory results that clearly testify to the validity of the model, the final adjustment was made such that all the established model components were fixed.

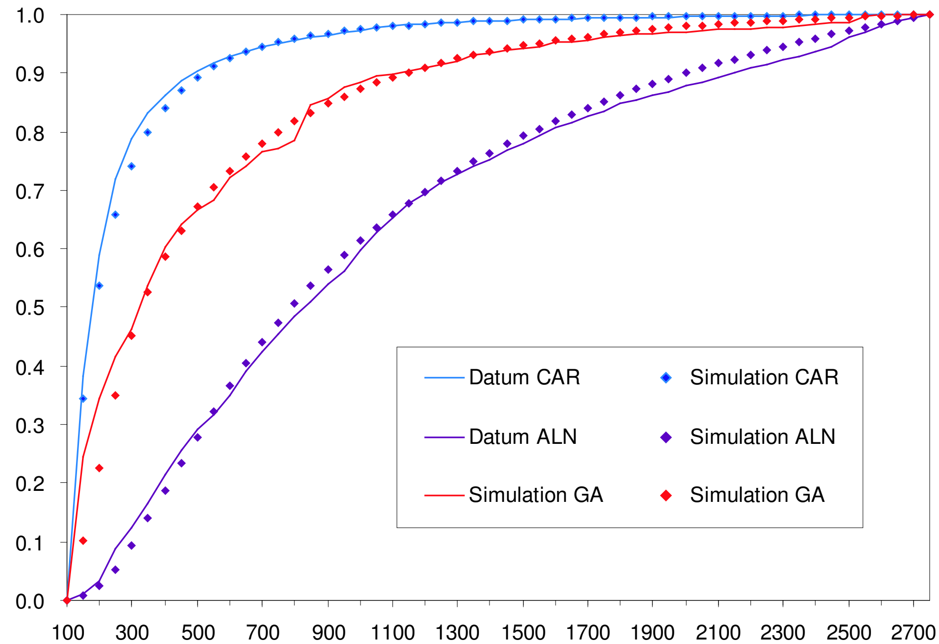


Figure 82: Cumulative Frequency by Mode

Figure 83 compares the real distribution and the model outcome in terms of the ratios of $[ALL]_{ATS}$ to $[ALL]_{CAL}$ by distance. Around 2,500 miles region, these ratios are unusually high. This phenomenon is due to the fact that a sizable portion of the U.S. population resides in the East and the West coast. If these ratios are used as weights, the total trip distribution would be much closer to the reality. Hence, the ratios of $[PSN]_{ATS}$ to $[PSN]_{CAL}$

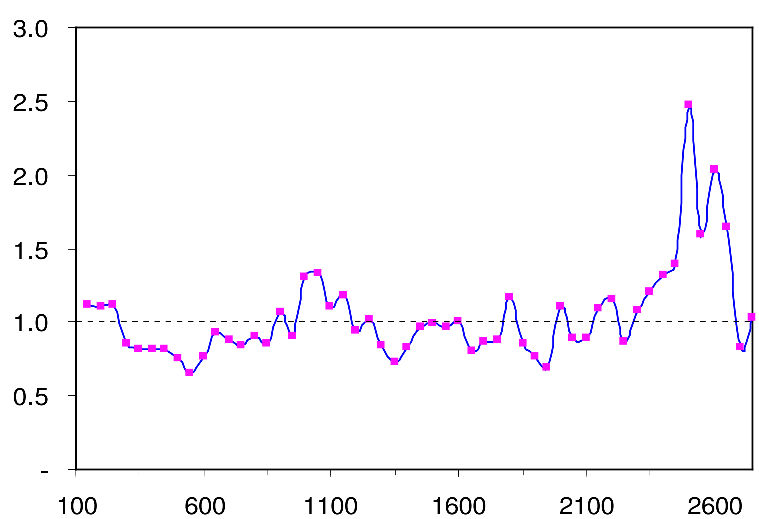


Figure 83: Correction Factors of The Gravity Model

and $[BIZ]_{ATS}$ to $[BIZ]_{CAL}$ are applied to the mode distance matrices. Figure 84 shows the results after this process, where each subplot presents the relative frequency of each mode by distance bracket.

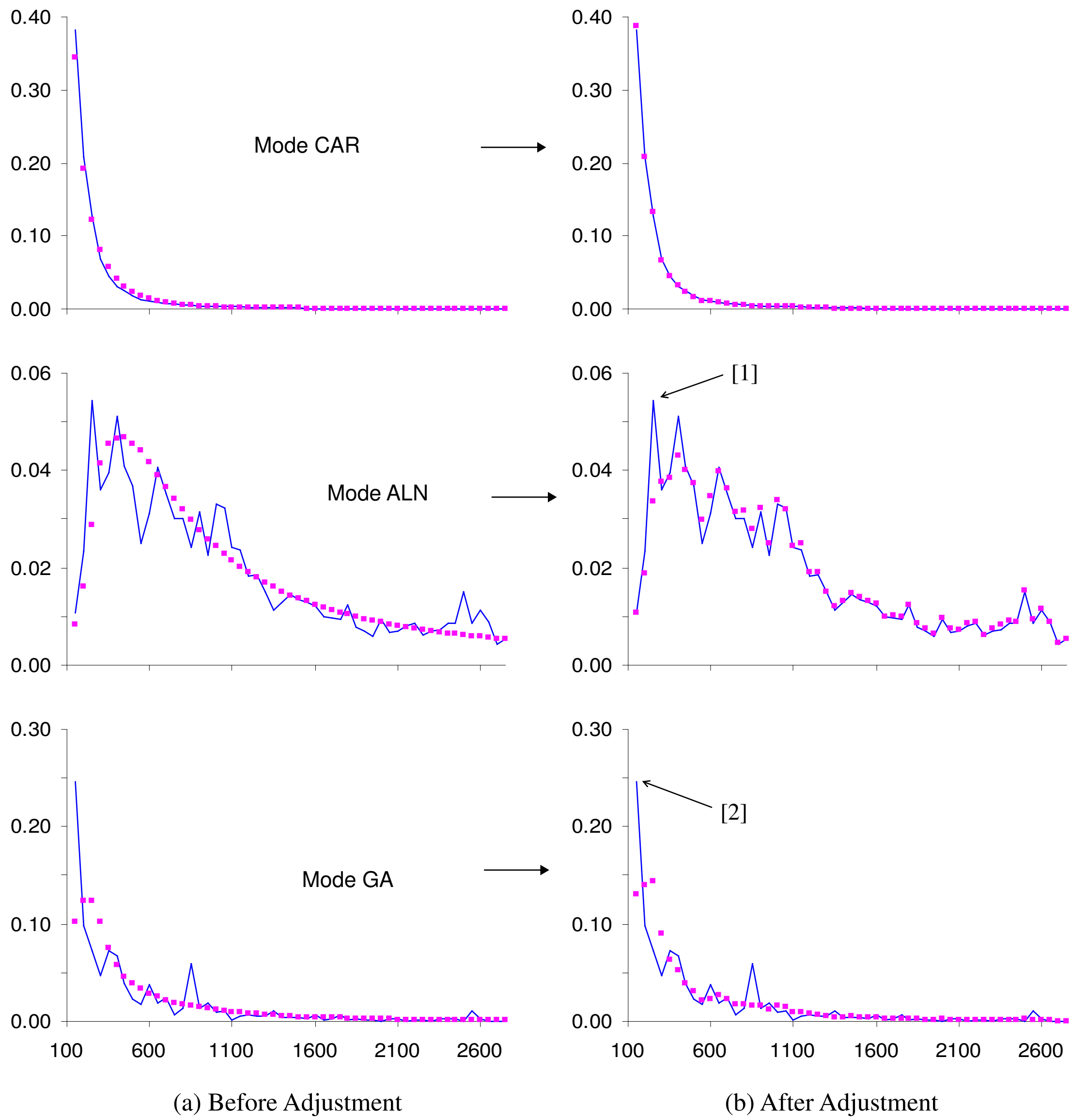


Figure 84: The Effect of the Gravity Model Adjustment

These figures confirm that the modified model (the purple dots) brings an even closer imitation to the real trip distribution profile (the blue lines). It is often taken for granted that some mismatches are inevitable in any simulation study, but the salient mismatches indicated by the markers [1] and [2] in Figure 84 were analyzed, leading to interesting findings. In the range of 200 to 250 miles, surprisingly, there exist three city pairs in top 10 airline markets. (Refer to Table 69 in Appendix C for details.) A close examination on GA trips less than 150 miles through the ATS database revealed that about 80% of them are solely originated from the state of Maryland, consisting of business trips only.

6.1.2.2 The Result

After adopting the aforementioned corrections and adjustments, the code ***Mi*** was executed on MIDAS that enables the code to run a clustered 128 computers (with 256 CPUs) simultaneously. In addition, the iterative process for the feedback algorithm detailed in §5.3.1.3 was automated using a commercial wrapper.[1] Figure 85 captures a screen shot of the iteration history. The time spent to reach this equilibrium state was about 10 minutes with 2 million consumers.[2] The baseline scenario converged when on-time reliability of mode ALN equals to 73.48%. The public in the virtual world perceive that about a quarter of mode ALN operations are not punctual, which is a quite reasonable estimate although it is difficult to quantify.

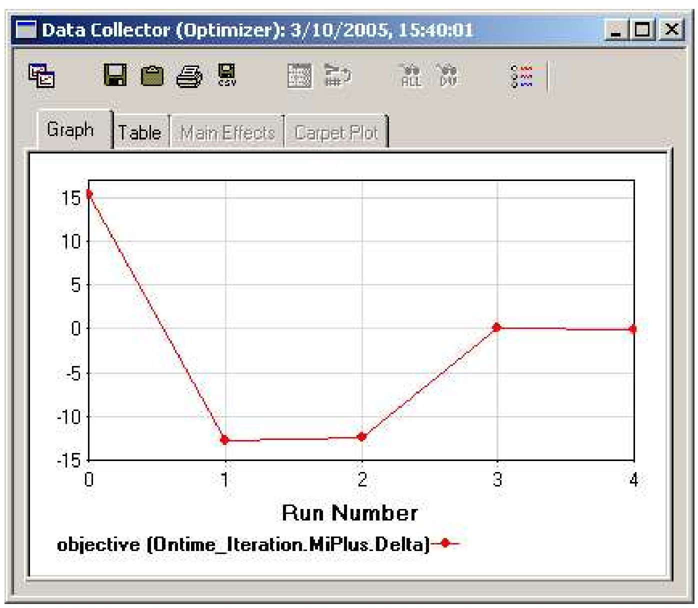


Figure 85: Convergence History of the Baseline Scenario

[1]ModelCenter™ for process integration and design optimization. Refer to http://www.phoenix-int.com/.
[2]Without the formidable computing power by MIDAS, it would take two days on a typical desktop.

Table 49 summarizes the final modal split as a result of the baseline by mode and by trip motivation, where each number in the parenthesis indicates the difference to the calibration datum described in Table 48. Improved matches observed in Figures 86, 87 and 88 ascertain that the calibration objective set forth in Equation 75 is accomplished. The small disparities between the datum and simulation are credited to such hard-to-catch factors with the present macroscopic model indicated by the markers [1] and [2] in Figure 84. Note that Figures 86(b), 87(b) and 88(b) deliver the same information as in Figure 80 except for modes ALN and GA being aggregated.

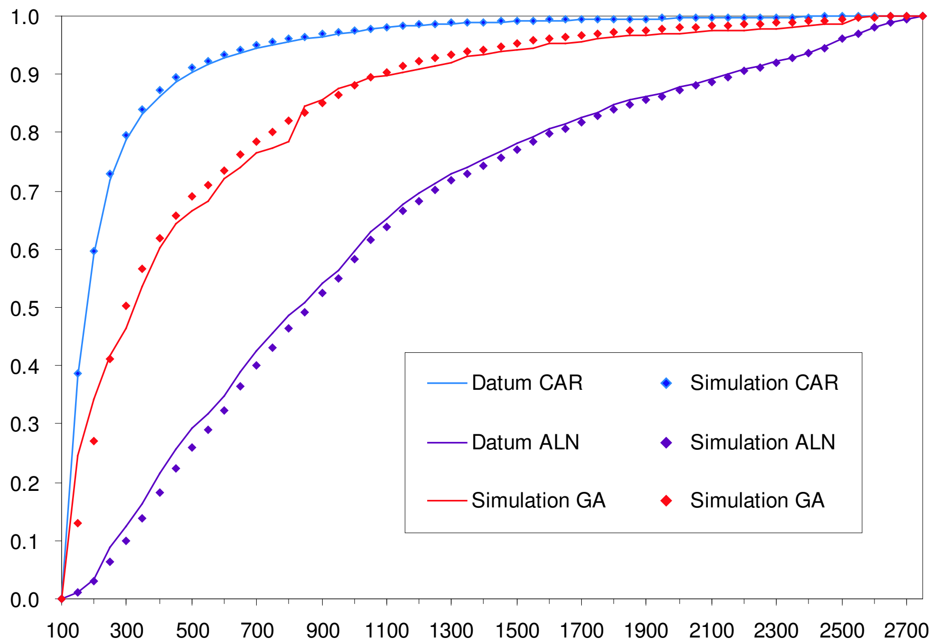
Table 49: The Summary

	Personal Trips	Business Trips	All Purposes
Mode CAR	57.41%	17.74%	75.14%
	(0.03%)	(-0.02%)	–
Mode ALN	12.18%	11.93%	24.11%
	(-0.01%)	(0.03%)	(0.02%)
Mode GA	0.26%	0.49%	0.74%
	(-0.01%)	(-0.01%)	(-0.02%)
All Modes	69.85%	30.15%	100.0%
	–	–	–

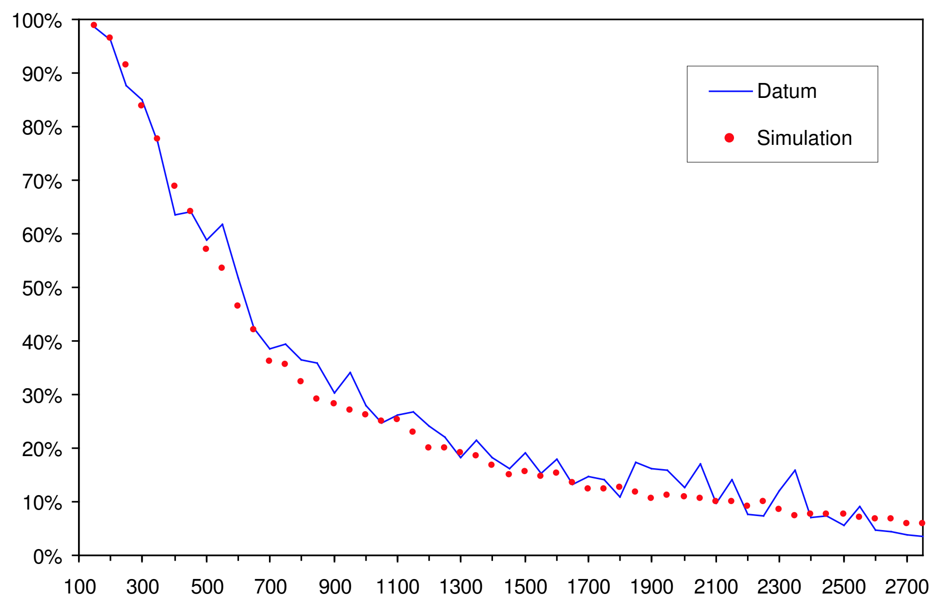
Finally, the route choice result of mode ALN is shown in Table 50. Although the airports models are simplified and aggregated, the close representation of the enplanement distribution stemming from the hub-and-spoke system implies another convincing aspect that the baseline and the model ***Mi*** have achieved a sufficient level of validity.

Table 50: Route Choice Result

<table>
<tr><th>Airport Category</th><th colspan="2">% Enplanement</th><th>Airport Category</th></tr>
<tr><td>Large-Hub</td><td>69</td><td rowspan="2">87.2</td><td rowspan="2">AP:HUB</td></tr>
<tr><td>Medium-Hub</td><td>20</td></tr>
<tr><td>Small-Hub</td><td>8</td><td rowspan="2">12.8</td><td rowspan="2">AP:SML</td></tr>
<tr><td>Non-Hub</td><td>3</td></tr>
<tr><td colspan="2">FAA Data</td><td colspan="2">Simulation Result</td></tr>
</table>

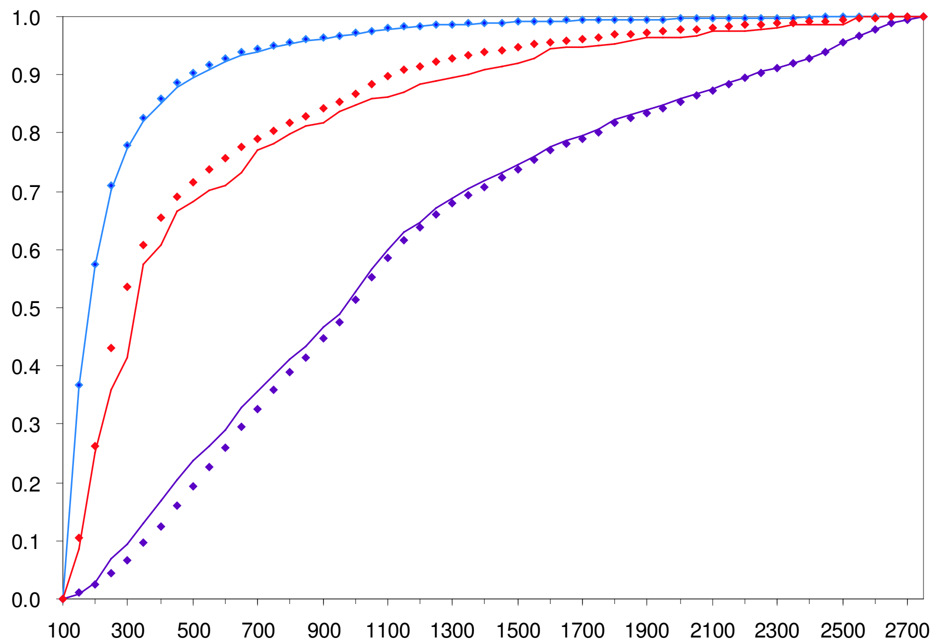


(a) Cumulative Distribution by Mode

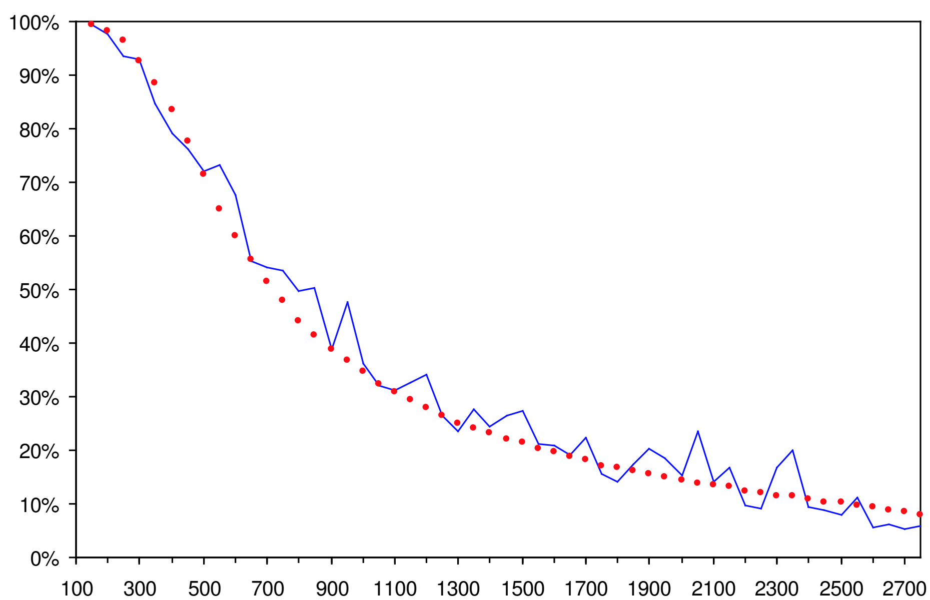


(b) Market Share Profile by 50-mile Distance Bracket

Figure 86: The Calibration Result from $[ALL]_{CAL}$

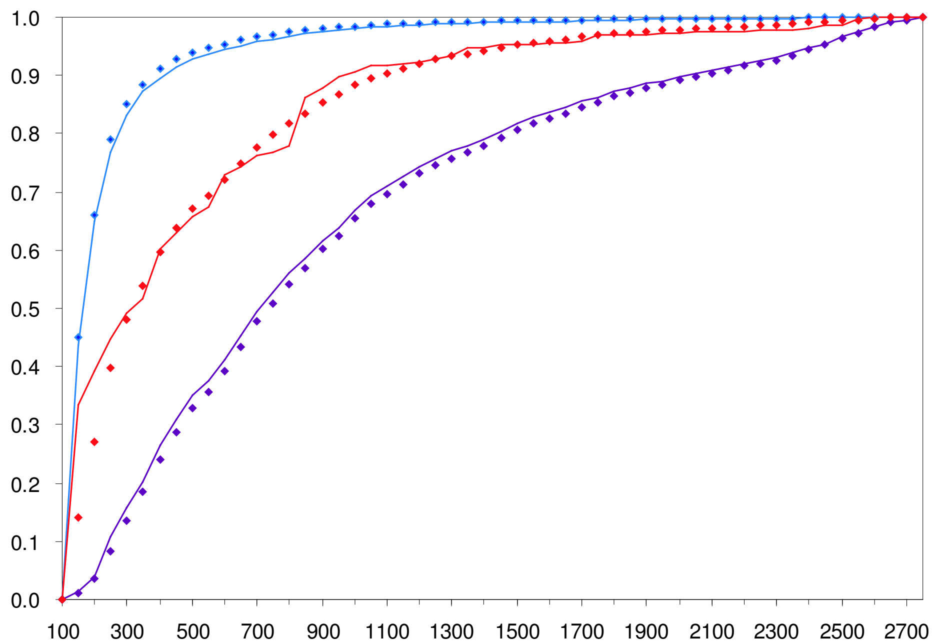


(a) Cumulative Distribution by Mode

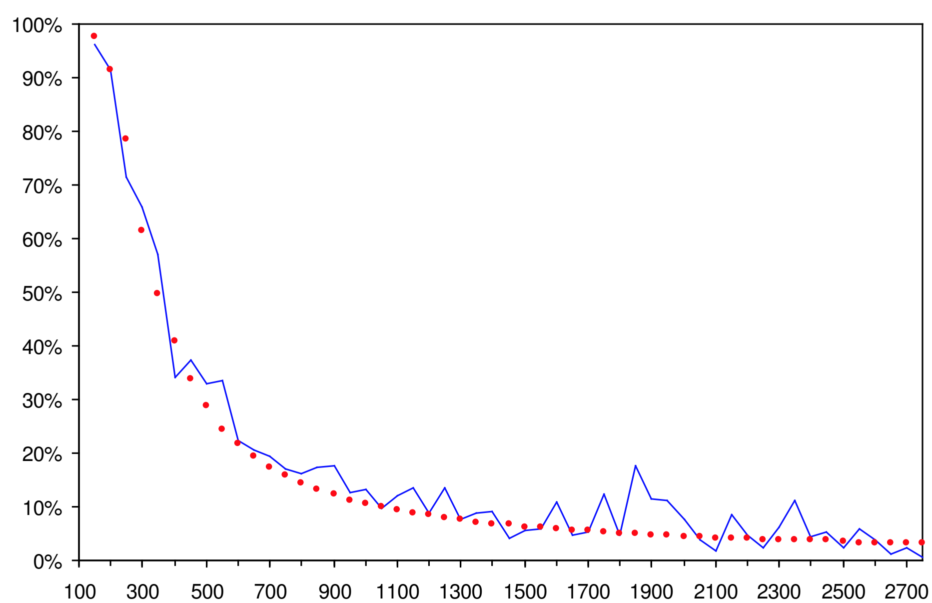


(b) Market Share Profile by 50-mile Distance Bracket

Figure 87: The Calibration Result from $[PSN]_{CAL}$



(a) Cumulative Distribution by Mode



(b) Market Share Profile by 50-mile Distance Bracket

Figure 88: The Calibration Result from $[BIZ]_{CAL}$

6.1.3 Scenario Preview and Simulation Strategy

The calibrated model has established a reasonable and close approximation of the existing NTS, laying a foundation for the further study driven by simulation scenarios. Several scenarios of interest are formulated and introduced below, after being grouped into four broad themes.

A prudent investigator not only utilizes capabilities but also recognizes limitations of the model. ***Mi*** can claim to achieve the "weak" validity, meaning that the model incorporates many model parameters at the investigator's discretion while interpreting the proxy world. This inherited model ambiguity needs to be examined by perturbing the assumptions; seeking to understand a range of the model response rather than being satisfied with a point response. The first group of the scenarios investigate this concern with respect to the model parameters and the socioeconomic conditions. The remaining themes ties directly to the original objective of the modeling effort as a hypothetical personal air vehicle is infused in the virtual world. The second group of the scenarios treats vehicle design requirements, which has a strong relevance to the traditional realm of aerospace engineers. Meanwhile, ***Mi*** is capable of accommodating more sophisticated scenarios. The third group aims to formulate and investigate the problem in the effect of system-level technologies on the NTS. Whereas the second and the third groups can be regarded as sensitivity analysis, hybrid simulation scenarios are conducted in the last group. These test cases have a direct relevance to the second supplementary research question in §2.4.2.

Implementation of all the scenarios is straightforward, but tricky on the other hand. For example, vehicle attributes can be easily changed and the corresponding simulation scenario is constructed by replacing the original vehicle GAP for this notional new mobility vehicle. The problem is determining PAV cost since it is linked to the enormous uncertainties associated with design requirements and any enabling technology set—the cost of owning and operating any new system is simply unknown under a great deal of noise effects. As a strategy to handle this difficulty, a unit scenario is performed with PAV cost being treated as an independent variable, with all else being equal. Accordingly, the baseline scenario runs against the PAV (or GAP in this particular case) price indices ranging from 0.75 to 1.25 with 0.05 interval, a multiplication factor of potential operating cost. Completion of this process produces a curve rather than a point as in Figure 89.

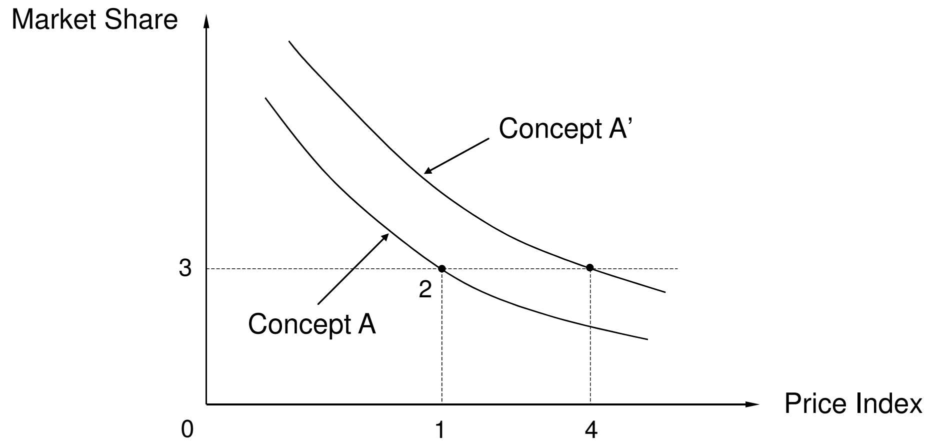
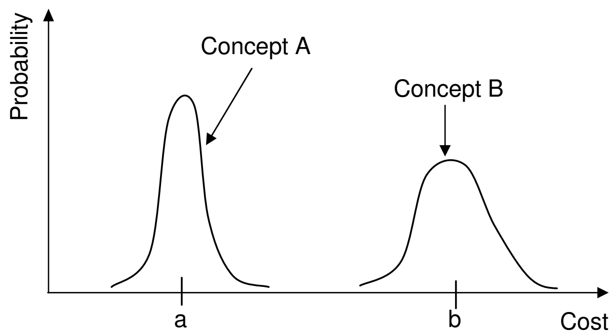


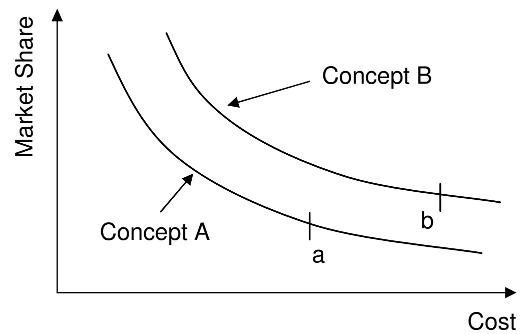
Figure 89: Example of Model Output with Varying Prices

It is noteworthy that this simple strategy partially mitigates limitations of the model by the same token mentioned in the second paragraph of this section. The ATS database as the proxy world does not sufficiently represent general aviation trips done by a wide variety of GA fleets. It would be more sensible to look at the responses of the model behaviors with respect to the perceived GA cost that has the most indefinite nature. On the other hand, adaptation of this strategy throughout all simulation scenarios offers many advantages. First, a flexible concept trade is possible regardless of the nature of the technology or the configuration of PAV. For example, suppose a situation where conventional GAP (A) and a roadable PAV (A') should be considered together. The simulations for both vehicles can be performed independently, resulting in Figure 89 that conceptually represents the two vehicles' estimated market share as a function of the various pricing options. More specifically, assume that GAP can be built at price 1. In this case, business planners can set price target for the new PAV on the price 4 to maintain the same marketability. The area of the rectangular 0123 indicates the total sales at price 1, and if the objective is to maximize it, an analytic study on the pricing strategy can be performed with an regression function of the curve.

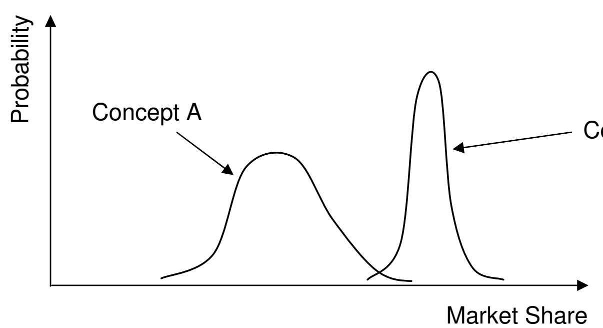
A further extension can be imagined. Engineering and economic analysis tools at vehicle level can be interweaved for the integrated decision-making process. A conceptual illustration is sketched in Figure 90. Under the goal of maximizing the market share, Concept A seems to be superior to Concept B by looking at Figure 90(a). However, if the two distributions are interpreted with the marketability chart on Figure 90(b), the opposite conclusion can be drawn as portrayed in Figure 90(c).



(a) Engineering/Economic Analysis of Vehicles



(b) Concept Trade Chart from ***Mi***



(c) Mapping (a) to (b)

Figure 90: Use of Other Simulation Result with Agent-Based Simulation Result

6.2 *Sensitivity Analysis*

Starting from the baseline (code BSLN), changing certain assumptions allows exploration of sensitivities and answers to "what-if" questions, as changes are revealed by differing agent behaviors within the simulation. The sample size of agents is set to 10 million throughout the study, which will produce approximately ten percent of the real trip amounts over a year in the U.S.

6.2.1 Model Parameters and Socioeconomic Conditions

The first scenario collectively perturbs the level of economic affluence of the nation, identified as the most crucial driver entity. Two sub-scenarios are considered. Scenario $\mathrm{INC}_{\oplus}$ assumes a 25 percent net increase in personal wealth. The public's average income is decreased by the same 25 percent in Scenario $\mathrm{INC}_{\ominus}$. Table 51 compares the overall market share results with that of the baseline.

Table 51: Overall Market Share Result of Scenarios $\mathrm{INC}_{\ominus}$ and $\mathrm{INC}_{\oplus}$

	$\mathrm{INC}_{\ominus}$	BSLN	$\mathrm{INC}_{\oplus}$
Mode CAR	79.52%	75.14%	71.62%
Mode ALN	19.91%	24.11%	27.43%
Mode GAP	0.45%	0.57%	0.68%
Mode GAJ	0.11%	0.19%	0.27%

When the economy in the virtual world was strong, the market share of mode CAR was adversely affected whereas higher speed modes became more popular. As expected, the agents were willing to spend more money during better economies. On the other hand, a bad economy put a chilling effect on the use of modes GAP and GAJ. Mode ALN was also significantly affected as mode CAR picked up the balance. Figure 91 contrasts the effect of the economy in more visual way, where each series of the dots indicates the diminishing market share trend of mode CAR along varying trip distance.

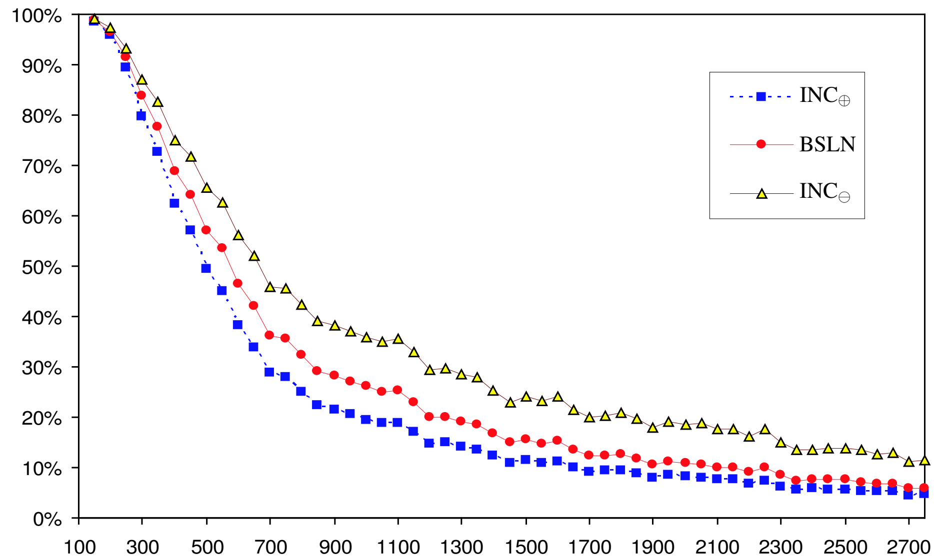


Figure 91: Adoption Rate of Mode CAR by Distance

Besides the market share analysis, the present model is capable of predicting trip volumes. A change in income proportionally influences the monetary part of the mobility budget of each consumer. Hence, the total amount of trips can vary accordingly in response to a certain situation. Table 52 shows the total trip volumes by mode, with the elasticity values calculated using Equation 11. Note that the sign of elasticity becomes positive as net effective price, which is a function of income, is reduced from Scenarios INC$_\ominus$ to INC$_\oplus$.

Table 52: Trip Volumes of Scenarios INC$_\ominus$ and INC$_\oplus$

	INC$_\ominus$	BSLN	INC$_\oplus$	Elasticity
Mode CAR	39,121,883	44,442,527	48,665,578	+0.43
Mode ALN	9,797,283	14,258,793	18,634,662	+1.26
Mode GAP	221,632	335,329	461,272	+1.43
Mode GAJ	56,352	109,985	183,599	+2.31
All Modes	49,197,150	59,146,634	67,945,111	+0.63

It was found that even the volume of automobile trips increased in Scenario INC$_\oplus$, although it was the least elastic to the income changes. The market share changes of general aviation modes were insignificant in terms of the absolute number of trips. But amplified with the total volume changes, modes GAP and GAJ showed the most dramatic boost.

It is commonly known that luxury goods tend to be more elastic, and the agents follow this economic rule without the modeler's intervention—which is a case of an *emergent behavior*.

In Scenarios $\mathrm{INC}_{\ominus}$ and $\mathrm{INC}_{\oplus}$, the net effective prices of all modes were changed at the same time by the same ratio. Three subsequent scenarios are formulated to closely investigate the interactions amongst the modes in response to individual price changes by ± 5 percent, making the price index (px) 0.95 or 1.05. Modes GAP and GAJ are combined for simplicity, and thus the price indices of the two modes are chained. The trip volume changes (i.e., the induced trips) are shown in Table 53. Figure 92 compares the individual elasticity of each mode and total volume due to variation of price index of each mode.

Table 53: Trip Volume Change by Price Index of Each Mode

	Mode CAR		Mode ALN		Mode GA	
	$px = 0.95$	$px = 1.05$	$px = 0.95$	$px = 1.05$	$px = 0.95$	$px = 1.05$
Mode CAR	44,852,823	43,944,223	44,281,489	44,551,936	44,420,709	44,448,519
Mode ALN	14,270,876	14,283,796	15,197,698	13,396,275	14,254,347	14,294,262
Mode GA	446,576	446,280	438,417	457,634	491,610	408,393
All Modes	59,570,275	58,674,299	59,917,604	58,405,845	59,166,666	59,151,174

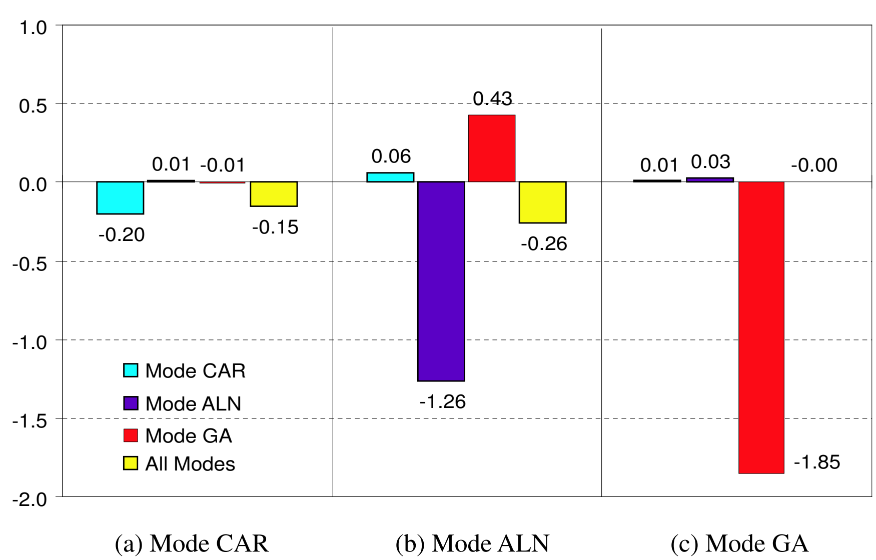


Figure 92: Individual Mode Elasticity

This individual elasticity plot confirms again that mode CAR was the least elastic and mode GA was the most elastic goods of all. Interestingly, the interaction between the different modes was not symmetric. For example, when the price index of mode ALN increased, the volume of mode GA increased as shown in Figure 92(b). This means that modes ALN and GA are substitute goods for the consumers. But an observation from Figure 92(c) reveals that mode ALN is not affected from the price change in mode GA. This partial interchangeability is credited to the effect of the volume inertia. Another intriguing finding is that the price of mode ALN is the most sensitive to the total volume change. If this is true for the real world, travel related businesses (e.g., hotels, tourist attractions, etc) would watch airline ticket prices more closely than gasoline prices as an early indicator of sales. From the modeling perspective *per se*, the finding means that economic assumptions of the provider ALN:FARE are the most critical to the simulation outcome.

The next set of scenarios concerns the modified gravity model of which parameter γ reversely indicates the strength of trip demand towards more remote destinations. The values of γ used in the baseline model are 1.6 and 1.4 for households and enterprise, respectively. These values were increased by 0.2 (code $DST_{\ominus}$) and decreased by the same amount (code $DST_{\oplus}$). In some sense, Scenario $DST_{\ominus}$ represents the past whereas Scenario $DST_{\oplus}$ represents the future, provided that the evolution of the transportation system ultimately reduces the public's perception about travel distance as an impedance function. The resulting volume and market share (M/S) of each mode are presented in Table 54.

Table 54: Simulation Result of Scenarios $DST_{\ominus}$ and $DST_{\oplus}$

	$DST_{\ominus}$		$DST_{\oplus}$		M/S Sensitivity
	Trip Volume	M/S	Trip Volume	M/S	from $DST_{\ominus}$ to $DST_{\oplus}$
Mode CAR	55,561,954	79.82%	34,789,761	69.58%	-12.83%
Mode ALN	13,587,904	19.52%	14,774,994	29.55%	51.39%
Mode GAP	346,110	0.50%	324,056	0.65%	30.35%
Mode GAJ	111,184	0.16%	108,124	0.22%	35.39%
All Modes	69,607,152	100%	49,996,935	100%	0.00%

As shown in the table, the small (relatively) variation in γ value was influential to the simulation outcome. The changes in the mode-wise market shares were reasonable, which explains the total volume expansion in Scenario $DST_{\ominus}$. This implies that the value of γ should have a inverse correlation with the income or the mobility budget in order to be more realistic. The effect of γ on mode ALN can be visualized in Figure 93, where median air trip distance changes from approximately 820 to 770 miles for Scenario $DST_{\ominus}$ and to 940 miles for Scenario $DST_{\oplus}$.

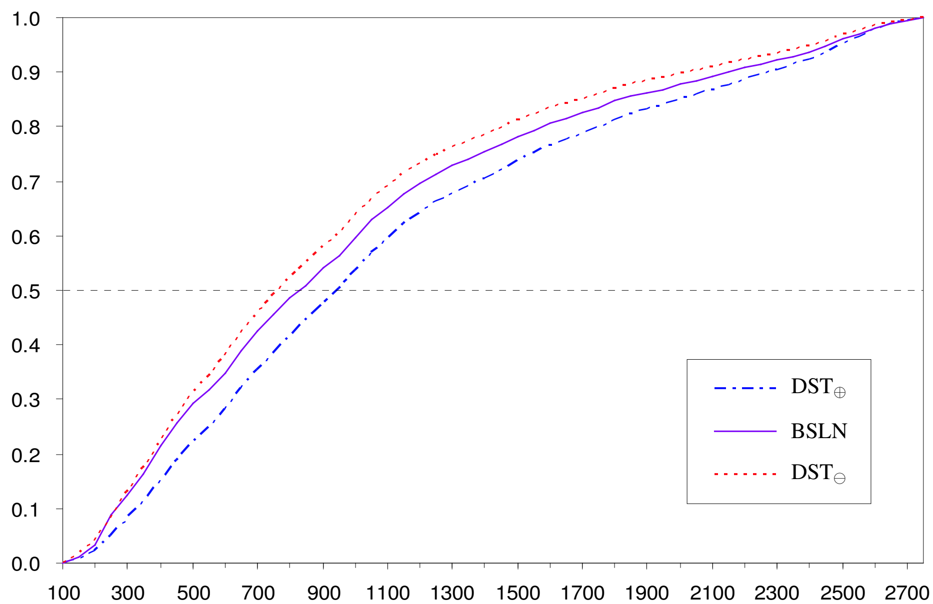


Figure 93: The Effect of γ on Mode ALN

The above scenarios illustrate the point that the corresponding parameters were sensitive to the output of the simulation, thus care must be taken to determine an estimate of each parameter. In some cases, however, the model behaviors went against expectation. The first example revolves around the long-standing trend of Americans to migrate to large metropolitan areas. Scenario $MET_{\oplus}$ accelerates this trend instantaneously, assuming the percent increase in population of +10%, +5%, -5% and -10% for Locales LAR, MED, SML and NON, respectively. The reverse use of these factors constitutes Scenario $MET_{\ominus}$. The O-D matrix should be accordingly updated for each scenario before the experiment, as shown in Table 55.

Table 55: The O-D Matrices for Scenarios $\mathrm{MET}_{\ominus}$ and $\mathrm{MET}_{\oplus}$

$$\underset{\text{for } \mathrm{MET}_{\ominus}}{\begin{pmatrix} 7.15 & 6.72 & 3.67 & 11.03 \\ 5.19 & 3.84 & 2.51 & 9.47 \\ 2.48 & 2.54 & 1.25 & 5.17 \\ 6.85 & 8.79 & 5.08 & 19.14 \end{pmatrix}} \qquad \underset{\text{for } \mathrm{MET}_{\oplus}}{\begin{pmatrix} 10.68 & 9.08 & 4.06 & 11.03 \\ 7.02 & 4.69 & 2.51 & 8.57 \\ 2.74 & 2.54 & 1.02 & 3.83 \\ 6.85 & 7.95 & 3.76 & 12.82 \end{pmatrix}}$$

The simulation result in Table 56 shows that a population shift to the cities did instigate a small-scale migration towards airline trips as more people were now within driving distance of a hub airport. It was expected that movement to rural areas involved a sizable shift to general aviation, allowing more thorough use of rural roads and airports. The result implies, however, that the current general aviation system would not be a reinforcing driver towards a "de-urbanization" of the nation. But it is noteworthy that this interpretation is based on the *ceteris paribus* assumption. For example, if the income profiles of locales were influenced due to migration, the result would be different.

Table 56: Simulation Result of Scenarios $\mathrm{MET}_{\ominus}$ and $\mathrm{MET}_{\oplus}$

	$\mathrm{MET}_{\ominus}$		$\mathrm{MET}_{\oplus}$		Volume Sensitivity
	Trip Volume	M/S	Trip Volume	M/S	from $\mathrm{MET}_{\ominus}$ to $\mathrm{MET}_{\oplus}$
Mode CAR	44,448,016	75.39%	44,487,712	74.90%	0.09%
Mode ALN	14,057,931	23.85%	14,466,125	24.35%	2.90%
Mode GAP	337,051	0.57%	331,659	0.56%	-1.60%
Mode GAJ	111,386	0.19%	111,739	0.19%	0.32%
All Modes	58,954,384	100%	59,397,235	100%	0.75%

Next, the choice parameter α was the subject of the experiment. The values of α for this sensitivity study are 0.075 (code $\mathrm{ALP}_{\ominus}$) and 0.125 ($\mathrm{ALP}_{\oplus}$). Again, the model was robust to such a magnitude of variation as shown in Table 57. It is recommended that α values should be set to smaller than 0.5 so that the Logit-based choice model can avoid the exponential of a very large negative number.[3]

[3]The smallest positive nonzero value in type "double" in Java is 2^{-1074} or 4.9^{-324}.

Table 57: Simulation Result of Scenarios $ALP_{\ominus}$ and $ALP_{\oplus}$

	$ALP_{\ominus}$		$ALP_{\oplus}$		Volume Sensitivity
	Trip Volume	M/S	Trip Volume	M/S	from $ALP_{\ominus}$ to $ALP_{\oplus}$
Mode CAR	44,529,434	75.22%	44,192,844	74.90%	-0.76%
Mode ALN	14,222,187	24.02%	14,368,734	24.35%	1.03%
Mode GAP	338,106	0.57%	331,355	0.56%	-2.00%
Mode GAJ	110,798	0.19%	109,761	0.19%	-0.94%
All Modes	59,200,525	100%	59,002,694	100%	-0.33%

6.2.2 Analysis of Vehicle-level Design Requirements

The scenarios in this group attempt to gauge the leverage effect of vehicle design requirements, which essentially explores the PAV design space in accordance with Hypothesis 1 posed in §2.4.2. The result of this study would facilitate in formulating a viable PAV concept vehicle while bridging a "decision-making" domain and an "engineering" domain, as emphasized in §1.2.1. The prerequisite of the sensitivity study is a series of simulations of the baseline with the varying price index of mode GAP. The simulation result is shown in Figure 94 where the dots represent the market share with respect to the corresponding price index. The regression analysis indicates that a practically perfect logistic relationship

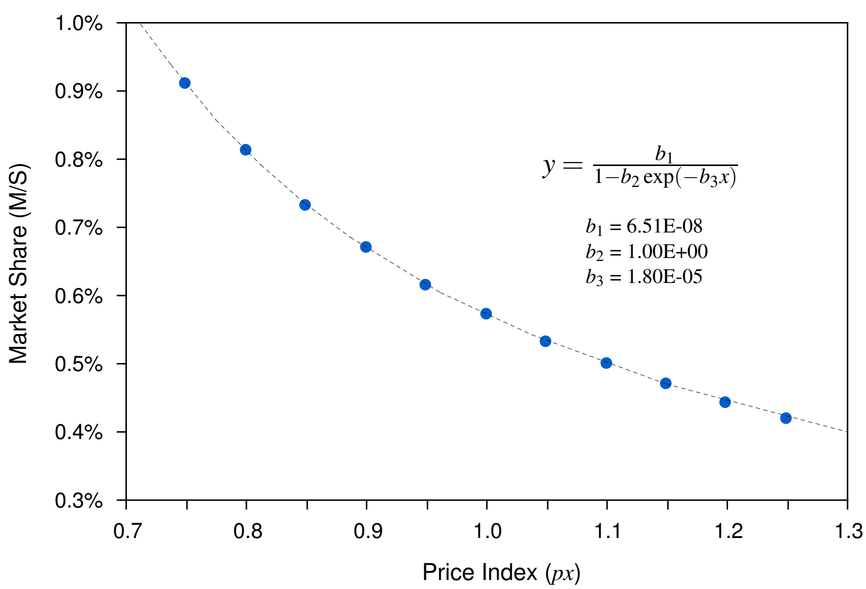


Figure 94: Market Share Variation of Mode GAP

exists. The error sum of squares for this fitted model is merely $SSE = 2.28 \times 10^{-9}$. Similar trends were found throughout all simulation scenarios, which can be considered another case of the emergent behaviors that the present model produces.

Now, mode GAP is being replaced with mode PAV and a selected attribute of mode GAP in Table 29 is changed by ± 25 percent. The first scenario concerns nominal cruise speed through Scenarios $\text{SPD}_{\ominus}$ (V = 120 mph) and $\text{SPD}_{\oplus}$ (V = 200 mph). Obviously, increasing airspeed, all others being equal, gave a benefit in terms of the PAV market share as verified in Figure 95 of which the beauty is to convey the benefit quantitatively.

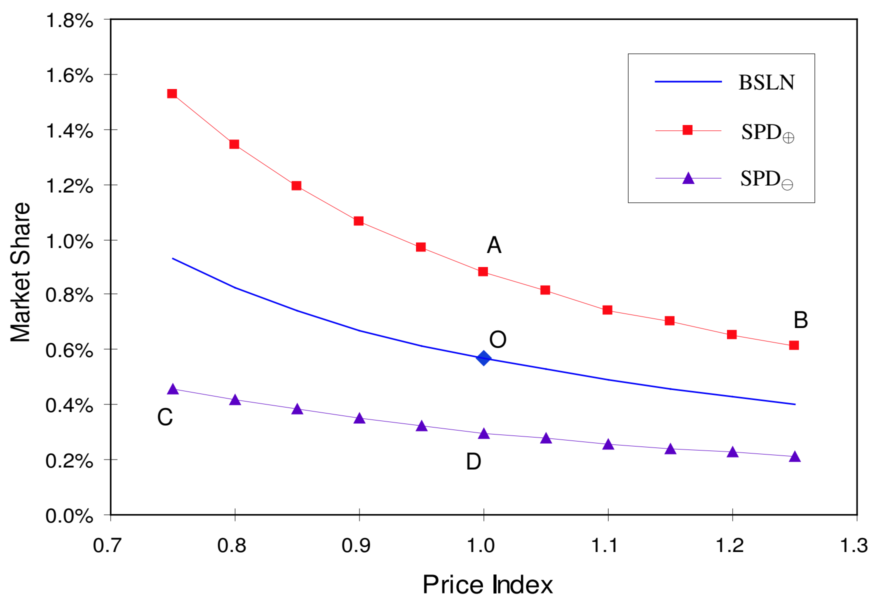


Figure 95: Market Share Variation of Mode GAP in Scenario SPD

Under the assumption that mode GAP achieved the Pareto optimality with respect to economics and performance, a meaningful tradeoff against, for example, the baseline point O can be performed between Points A and B for Scenario $\text{SPD}_{\oplus}$ or between Points C and D for Scenario $\text{SPD}_{\ominus}$. The plot shows that a new PAV with 25% faster than baseline GAP prevails only if this improvement costs less than about 130% of the baseline GAP. On the other hand, it would be an undesirable attempt to design a 25% cheaper PAV at the expense of a 25% speed degradation.

The trip volume results from the speed sensitivity study can also be categorized into different provider type as shown in Table 58. At the baseline speed of 160 mph, the percentage of self-owned vehicle trips approximates to 72% of overall mode GAP use. In Scenario $SPD_{\ominus}$, the percentage of GAP:SELF is approximately 85% whereas in Scenario $SPD_{\oplus}$, it reduced significantly to 62%. This trend of decreasing percentage use of self-owned GAP can be explained by the nature of the customer base. Vehicle owners make decisions not just based on speed performance but also on other non-quantifiable aspects such as joy derived from recreational flying. Hence their behavior is less elastic to changes in speed. In contrast, business travelers who hire pilots are likely to base their mode choice decisions largely on speed alone. Hence, their behavior is seen to be the most elastic to change in speed. In the middle, self-piloted travelers who rent show asymmetric behavior; they lost interest in flying when speed is too low, but they desire faster general aviation as much as owners do.

Table 58: Comparison of Elasticity by Provider Type

Mode GAP	Scenario			Arc Elasticity	
Provider	$SPD_{\ominus}$	BSLN	$SPD_{\oplus}$	$SPD_{\ominus}$→BSLN	BSLN→$SPD_{\oplus}$
GAP:SELF	150,326	242,212	325,726	1.66	1.33
GAP:RENT	3,469	11,153	15,996	4.06	1.62
GAP:HIRE	22,109	81,964	179,938	4.44	3.52

The next scenarios, $PLD_{\ominus}$ and $PLD_{\oplus}$, examined the effect of PAV payload capacity. If other conditions are kept the same, big airplanes incur higher acquisition cost and operations cost. This is not always a bad situation because travel cost per capita can be reduced if a vehicle operates at a full load. The simulation showcased that a down-sizing of passenger seats (including pilot's) from the baseline of four to three resulted in a noticeable decrease in market share, and an increase in passenger seats to five also yields a moderate change, as shown in Figure 96. It can be observed that the size of the gap between the two dot curves is smaller than that of the preceding scenarios, which indicates the leverage effect of speed is important than payload capacity.

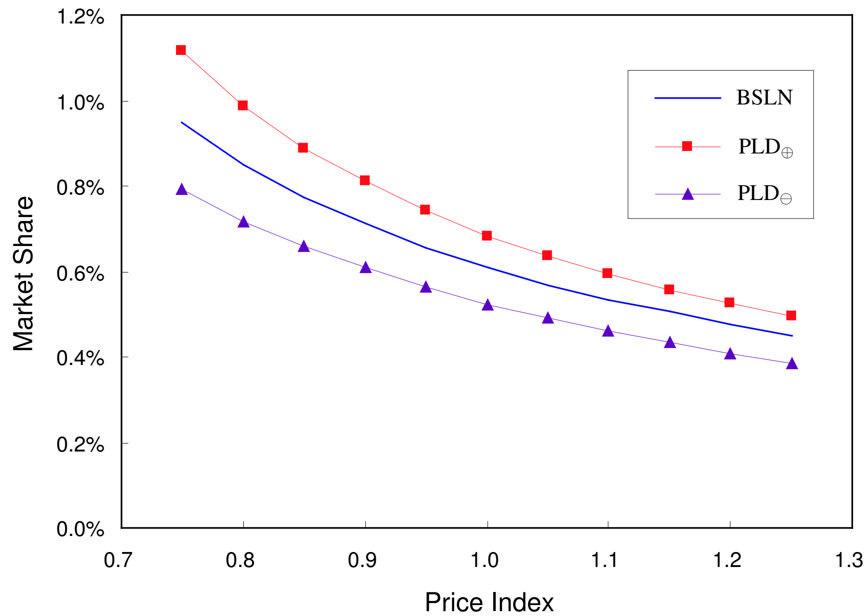


Figure 96: Market Share Variation of Mode GAP in Scenario PLD

Subsequent simulation studies were naturally instigated with varying the passenger seats from one to seven. Figure 97 portrays a decline in the marginal market share as the payload capacity increases (1 passenger = 200 lb) throughout all price indices from 0.75 to 1.25. This chart can be interpreted differently depending on vehicle platform. For example, the curves AB and CD in particular, describe the relationship between changes in the price indices to changes in payload for different platforms. In general, the gross weight is the most influential factor in determining direct/indirect cost within a family of the same platform. It is often accepted that the marginal gross weight of rotorcraft is bigger than that of conventional aircraft. If the curve CD is thought to be a rotorcraft vehicle platform, it can be seen that an increase in payload causes a similarly rapid increase in price index and rapid loss of market share. The curve AB, which is representative of payload efficient fixed-wing concepts, allows for payload increases with a much gradual increase in price index. From these observations, it can be estimated that a two or three-seater rotorcraft is a better pay-off design for the public's need whereas many local optimum solutions exist in the fixed-wing concept.

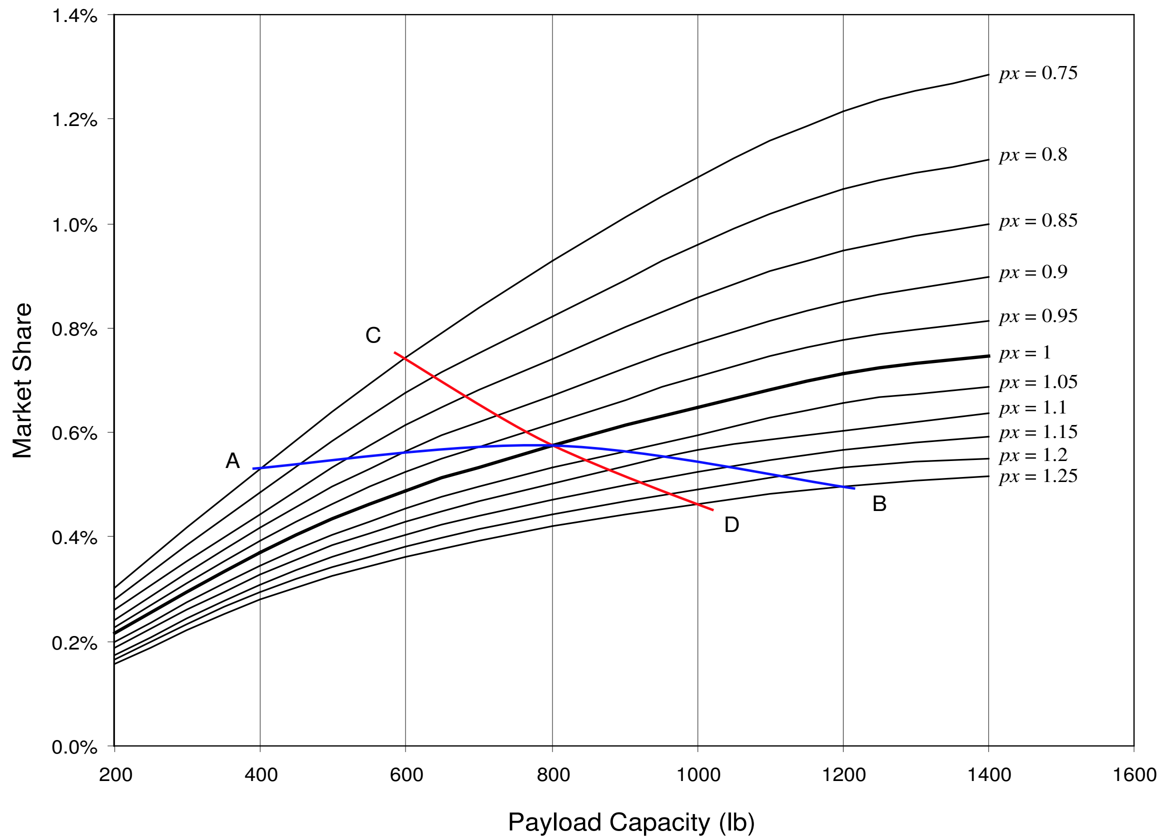


Figure 97: Market Share against Payload Capacity

Scenarios $RNG_{\ominus}$ and $RNG_{\oplus}$ were intended to measure the sensitivity to the changes in refueling range. As opposed to the above design requirements, Figure 98 reveals that changes in refueling range by ±165 ($= 660 \times 25\%$) miles caused nearly indistinguishable differences from the baseline. To understand this counter-intuitive phenomenon, the ratios of the distance-wise market share of mode GAP of Scenarios $INC_{\oplus}$ and $INC_{\ominus}$ to that of Scenario BSNL are portrayed in Figure 99. As shown, GAP trips less than about 400 miles are the most sensitive to any changes from the baseline. PAV with refueling range of 300 miles was simulated to confirm this finding, and the result (marked RNG_{300} in Figure 98) shows a noticeable amount of travelers being lost.

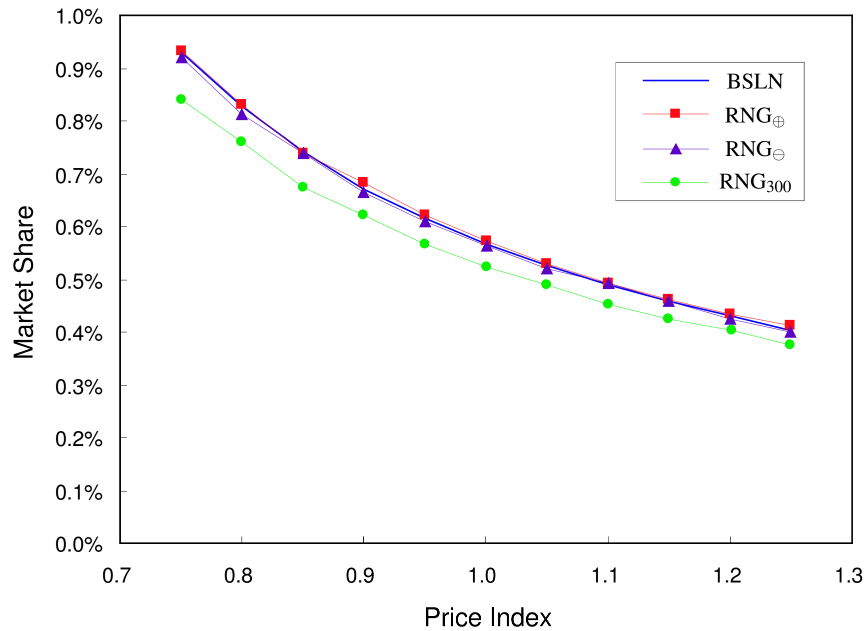


Figure 98: Market Share Variation of Mode GAP

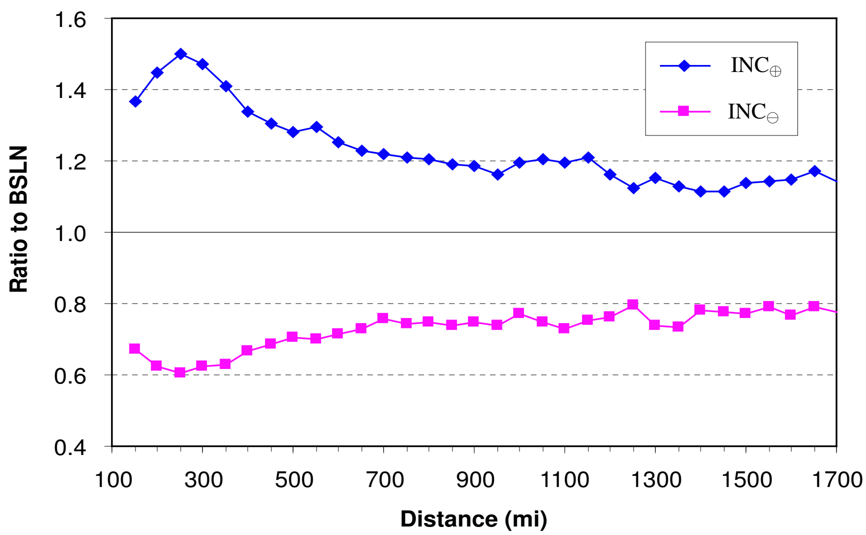


Figure 99: Responsiveness of Mode GAP against Distance

6.2.3 Tradeoff Study for System-level Technologies

Besides the vehicle-centric scenarios examined, ***Mi*** enables formulation of hypothetical situations in which a system-level technology, usually hosting a set of technologies that have far-reaching effect, is implemented in the NTS. Scenario WAIT:ALN$_{\ominus}$ attempts to investigate the potential benefit of improving mode ALN portal efficiency through reducing waiting and standby times at the portals AP:HUB and AP:SML by 25%. The wait-time is increased by the same 25% in Scenario WAIT:ALN$_\oplus$. Table 59 compares the overall market share results of the both scenarios. The resulting volume changes were against expectations in terms of the magnitude, especially for mode ALN. The consumers' complaints about a long line due to enforced security measures made 1.25% loss in the market share. The amount of total enplanement loss, however, was close to 5%. Increased wait-time at air terminals favored use of other modes, but high-speed GA was more responsive. This observation is in line with the recent growth in the use of business jets and the market for fractional shares after the September 11 attack.

Table 59: Simulation Result of Mode ALN Portal Wait Variation

	WAIT:ALN$_\ominus$		WAIT:ALN$_\oplus$		Volume
	Trip Volume	M/S	Trip Volume	M/S	Elasticity
Mode CAR	44,024,216	74.51%	44,833,225	75.71%	0.04
Mode ALN	14,628,303	24.76%	13,922,232	23.51%	-0.10
Mode GAP	325,073	0.55%	346,716	0.59%	0.13
Mode GAJ	106,280	0.18%	115,107	0.19%	0.16
All Modes	59,083,872	100%	59,217,280	100%	0.00

Scenarios WAIT:GAP$_\ominus$ and WAIT:GAP$_\oplus$ are the counterpart of the preceding scenarios, taking a $\pm$25 percent perturbation on wait-time at the portal AP:GAP. As shown in Table 60, few differences can be found except for mode GAP that has a notable decrease.

Table 60: Simulation Result of Mode GAP Portal Wait Variation

	WAIT:GAP$_\ominus$		WAIT:GAP$_\oplus$		Volume
	Trip Volume	M/S	Trip Volume	M/S	Elasticity
Mode CAR	44,437,283	75.11%	44,443,616	75.14%	0.00
Mode ALN	14,256,944	24.10%	14,278,460	24.14%	0.00
Mode GAP	360,117	0.61%	314,817	0.53%	-0.26
Mode GAJ	109,769	0.19%	111,153	0.19%	0.02
All Modes	59,164,113	100%	59,148,046	100%	0.00

The next scenario group concerns requirements of pilot license, which is currently stringent on requirements and expensive to obtain. As a result, only 0.285% of the income earners are qualified to operate the aircraft. Furthermore, the decreasing trend of both private and commercial ratings (See Table 37.) would aggravate the struggle of the general aviation industry. Scenario $LIC_{1/2}$ assumes that this trend is irreversible so the number of pilots drops by 50%. Certainly, there are many on-going technological investments that can overturn this undesirable situation. So, what if advances in avionics, computer and communication technology at both vehicle and system level attract more people? At what point would licensing make a significant effect on the NTS? Scenario LIC_2 assumes a fundamental improvement such that "easy-to-fly" technology doubles the number of licensed pilots. Also, a more aggressive goal is set through Scenarios LIC_3 and LIC_4, making a three-fold and four-fold increase respectively. The results from these scenarios are shown in Figure 100.

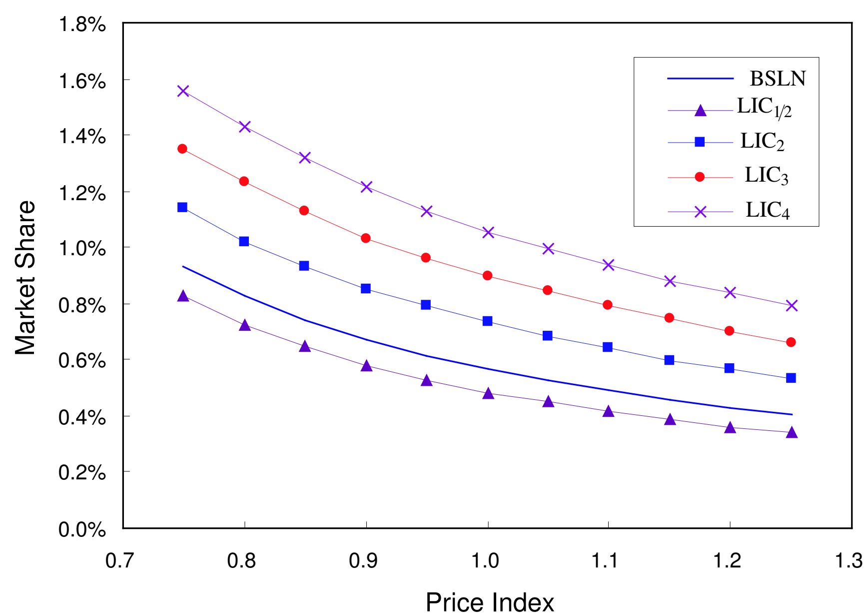


Figure 100: Market Share Variation of Mode GAP

One distinguishing feature is that the curves are placed as if they are formed by parallel shift of the baseline curve whereas other scenarios made a funnel shape with curvature. Also, it is intriguing that a relationship between the number of pilots and the market share is not directly proportional. This is because of interaction between other modes. The advantages of enabling "easy-to-fly" technologies are intuitive, but ***Mi*** describes the effects quantitatively, which was not addressable in the past.

The next scenario treats the effect of weather-resistance technology. In Scenario WTR_1, all pilots have IFR ratings. Scenario WTR_2 assumes that the near-all-weather technology reaches the next level leap. The influence matrix in Table 61 shows the input of the model to implement these conditions. The simulation result, shown in Figure 101, indicates that the benefit from BSLN to WTR_1 is critical whereas that of from WTR_1 to WTR_2 is less significant. This implies that technology goal for WTR_1 is more economical at least for short term goal.

Table 61: Weather Model and Influence Matrix

Scale	0	1	2	3	4	5	6	7	8	9
LoAlt (VFR)	0	0	20	60	100	100	100	100	100	100
WTR_1 (IFR)	0	0	0	0	20	60	100	100	100	100
WTR_2 (SATS)	0	0	0	0	0	20	60	100	100	100

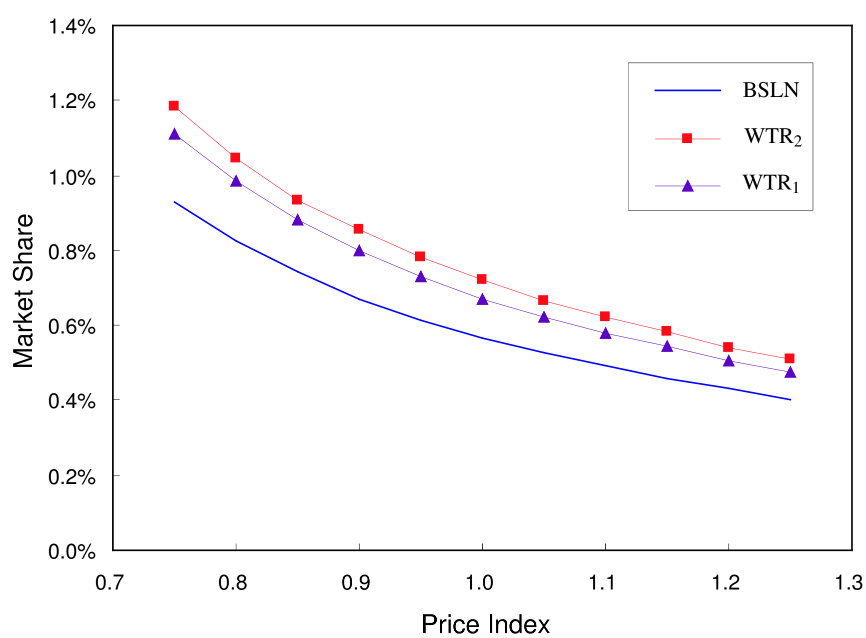


Figure 101: Market Share Variation of Mode GAP

6.3 Case Studies

In this section, two case studies are performed to gauge the leverage effects of hybrid technology portfolios. As explained in Appendix A.2, NASA's initiatives associated with PAV are embodied in the PAVE project and the SATS project. For the first case study, envision a decision maker who desires to evaluate the extent of the effect to which the combination of PAVE and SATS can bring in. Given that both projects are intuitively essential, the decision maker may not be sure which project has a priority or whether they are synergistic. Three specific scenarios are formulated to analyze this case study.

Scenario PAVE consists simply of the replacement of the existing mode GAP with a new mobility vehicle based on NASA's Rural/Regional Next Generation concept. This new concept is portrayed in Figure 102, with its target performance characteristics. For the simulation, adoption of this vehicle forms the Scenario PAVE which is an amalgamation of Scenarios $\mathrm{SPD}_{\oplus}$, $\mathrm{PLD}_{\oplus}$ and $\mathrm{RNG}_{\ominus}$ examined in the previous section.

Cruise Speed: 200 MPH

Range: 500 Miles

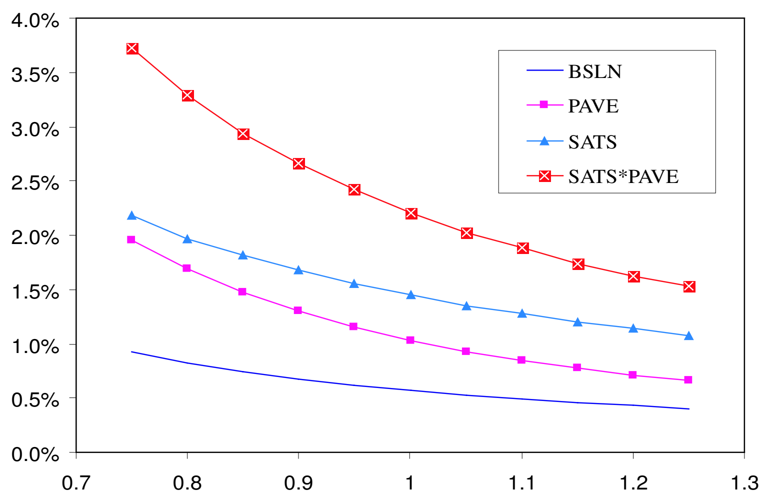
Passenger Seats: 5

Acquisition price: $75,000

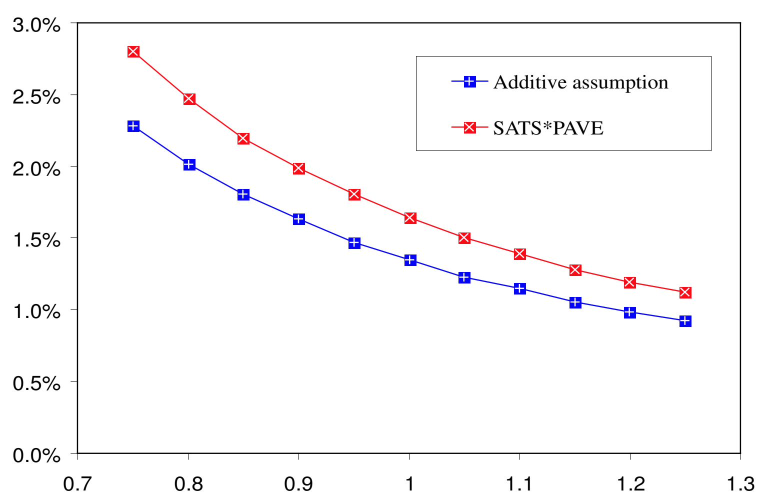
Figure 102: NASA's Tail Fan Concept

In Scenario SATS, it is assumed that a set of enabling technologies is adequately invested and successfully implemented, including the near-all-weather technology (Scenario WTR_2) and the easy-to-fly technology (Scenario LIC_4). These improvements would lower takeoff and landing requirements and thus the currently underutilized airports make usable. To reflect these changes, the access distance of GAP portal in locale NOM was reduced so that entire population resides within 20 miles of GAP portals. Scenario SATS*PAVE is the next logical step and constitutes the third scenario. This scenario imagines that, as the code name implies, Scenarios SATS and PAVE are synthesized together.

The simulation was performed with variation in the price index for technology tradeoff study amongst the scenarios. As shown in Figure 103(a), the gap between Scenarios BSLN and PAVE is diminishing as the price index increases whereas Scenario SATS is generating a rather parallel upward shift. This shape distinction once more delineates the nature of the potential benefit from vehicle-centric technologies and system-level technologies. Also, it was found that both technologies are synergistic as evidenced in Figure 103(b), where the market share increase of SATS*PAVE over the baseline is greater than the simple addition of the two isolated effects. This implies simultaneous funding to both projects would be the best decision before the decision maker.



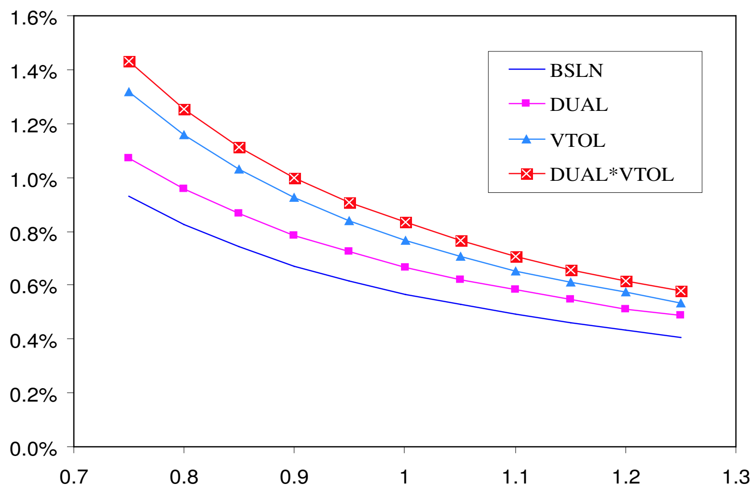
(a) Market Share Change against the Price Index



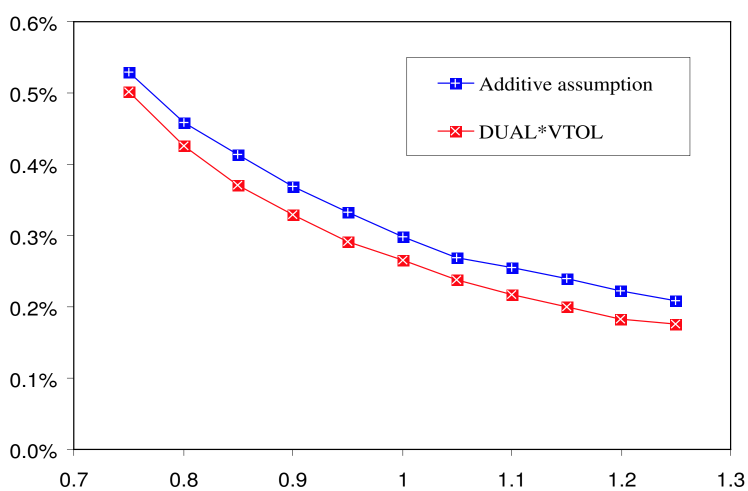
(b) Market Share Expansion from Scenario BSLN

Figure 103: Comparison of Scenarios PAVE and SATS

The next case study involves roadability and vertical takeoff capability of PAV systems. Similar to the first case, three specific scenarios are formulated. In Scenario DUAL, a dual mode PAV is infused in the virtual world, eliminating the need of the secondary ground mode and reducing portal wait time at the origin and destination. Scenario VTOL assumes that a vertical takeoff PAV system switches the baseline mode GAP. This new vehicle is operational with a small takeoff-landing facility similar to heliports and thus access distance is reduced to a tenth. Subsequently, the secondary mode cost is ignored if the access distance is less than a mile. The final Scenario DUAL*VTOL is a combination of the both scenarios. The simulation result is shown in Figure 104.



(a) Market Share Change against the Price Index



(b) Market Share Expansion from Scenario BSLN

Figure 104: Comparison of Scenarios DUAL and VTOL

The leading impression on Figure 104(a) is that the gap between Scenario BSLN and Scenario DUAL*VTOL is smaller than that of Scenario SPD$\oplus$. This means, if other conditions were kept the same, increasing cruise velocity by 40 MPH would be a better engineering decision than making a vehicle either roadable or capable of vertical takeoff and landing, not to mention dual mode VTOL. It was found that the interaction between Scenario DUAL and VTOL was smaller than the additive assumption, as shown in Figure 104(b). These two concept vehicles are not synergistic so only one should be chosen for a cost efficient investment in technology development.

In summation, the behavioral distinctions discovered from the two case studies showcased the fact that there exists a close coupling (positive or negative) of future aviation technologies. Furthermore, the studies exemplify how a rigorous simulation approach can support the decision maker to reach an informed decision, not just based on intuition or expert opinion, but on quantitative analysis that accounts for real world behavior as mimicked by the ABM. This decision-support methodology is a most important contribution of the simulation framework in modeling the NTS.

CHAPTER VII

CONCLUSIONS

Contents

7.1 Recapitulation of the Thesis

The initial motivation of the present research began by recalling the course of evolution of the National Transportation System (NTS) and recognizing the emerging need of a new mobility system. The Personal Air Vehicle (PAV), an instance of a new mobility system, served as an illustration of the limitations on understanding the knowledge boundaries and characterization of the design options space. A key contention then ensued: the PAV design problem, as of today, is not only a decision-making problem, but also a system-of-systems (SoS) problem. Therefore, the decision-making process should be carried out in the context of the whole system—the NTS—to achieve the bold goal of advancing the evolution of transportation.

Under this overarching theme, the literature of relevance was reviewed in the area of aerospace design, decision-making, and transportation system analysis. The qualitative tools surveyed are powerful and important in properly framing overall design problems; the use of utility-focused approaches provide a simple means to analyze design requirements; and large-scale transportation models possess a tremendous capability in tracking microscopic behavior of the computer-generated objects of the investigator's interest. Significant depths are achieved in the respective realms. Nonetheless, they are not sufficient for a firm basis of the decision-making process, and the integrating whole has not been attempted. In other words, these efforts do not fully embrace the SoS perspective. The reason is simple; this is literally a grand challenge—it's not merely an aviation problem, a ground transportation problem, an economics problem, nor a single disciplinary engineering problem. Hence, the capability to examine a new mobility system, e.g. the PAV design problem, in the context of the entire NTS was elusive. In response to these findings, the research questions and hypotheses were formulated to articulate the major thrust of the present research and to guide its execution.

A tangible model that closely imitates the NTS is missing and is required to provide the foundation on which the formulated hypotheses can be tested. The theoretical and methodological foundation was laid in Chapter III. The basic tenet of Agent-Based Modeling/Simulation (ABM/S) was stressed and highlighted as this new approach provides a theoretical and methodological tool to investigate dynamic non-linear behaviors of complex systems. The basic theory of transportation demand was introduced and its limitations were discussed in the face of the ever-changing transportation SoS. Probabilistic choice theory was also introduced, an important theoretical background for implementing the agent's decision-making structure.

With the theoretical foundation set, the point of departure for development of the integrated transportation model was an exploration of the NTS in terms of infrastructure, people, and economic trends to acquire a sound understanding of the target system. The NTS is indeed a complex SoS, both in colloquial and technical senses of the word, stemming from the heterogeneity of constituent elements, the distributed nature of these systems and the presence of deep uncertainty in exploring its future state. Under such circumstances, the usual reductionism strategy would not be complete, or sufficient to tackle the challenge. Therefore, the holistic assumption was adopted; essentially "everything is on the table." The concept of entity was introduced and articulated as the abstract class of transportation artifacts. To organize a multitude of building blocks, four entity descriptors were deduced. This abstraction framework produced the concept of the Transportation Architecture Field (TAF) as a mental model of the NTS. The culmination of the entity-centric abstraction framework crystallized the modeling hypothesis as a strategic blueprint. Then, the tactical blue print was laid out, which guided the construction of a virtual NTS couched in an agent-based model. Verification, Validation & Assurance issues were contemplated and the four legitimacy principles were established to prevent the usual modeling fallacies.

The implementation chapter elaborated how to translate the conceptual virtual world into the computational model with a sufficient level of detail. With a sound understanding of the applicable databases, the scope of the modeling exercise was carefully determined, then details of the model components were described and their implementations were illustrated. The modular structure inherited from the entity-centric abstraction made it possible to represent important relationships among components inside the system with relatively simple processes. From the Census database, a union of aggregated locales was extracted and a virtual transportation environment was derived and constructed. The transportation consumers and the service providers were modeled as agents that apply a set of preprogrammed behavioral rules to achieve their respective goals. The transportation infrastructure and multitude of exogenous entities—disruptors and drivers—in the whole system were also represented without resorting to an extremely complicated structure. The overall mechanism followed the agent-based evolutionary scheme that allows for adaptive and emergent behaviors. All entities modeled were elements of an instance of the virtual TAF, culminated in the unified simulation model ***Mi***. A graphical user interface for ***Mi*** was implemented on the web. The development of a distributed computing environment was implemented to remove the potential bottleneck due to an unmanageable computational burden, which would eclipse the potential usefulness.

The American Travel Survey (ATS) database was used to construct the datum against which the baseline was calibrated. The household trip distribution by primary mode, by travel distance, and by trip motivation was the key subject and a very stringent calibration objective was set. Nevertheless, since the core calculation block and the basic theory was sound, calibration runs went as planned; the results from ***Mi*** were pleasingly close to the historic data without any coercive mechanism, and meaningful reasons were identified for the cases that differed from expectations. A majority of the evidence from simulation indicated that the model is indeed valid. Thus, the final outcome attained the goal—a flexible, scalable, computational model that allows for examination of numerous scenarios which involve the cascade of interrelated effects of aviation technology, infrastructure, and socioeconomic changes throughout the entire system. Further, it was verified that the simplicity revealed by abstraction (at the price of model details) can answer certain questions that would otherwise remain unapproachable if high levels of detail were demanded.

7.2 Contributions and Recommendations

The following paragraphs summarize the key contributions and recommendations from the present research. The contributions can be collected into three major categories: intellectual, experimental, and methodological.

The need for understanding the whole—embracing holism and foregoing reductionism—has been a long-standing theme,[1] yet one that is markedly difficult to conceive and implement for complex systems. The entity-centric abstraction presented in this dissertation enables a working holism and leads to a proper formulation of the appropriate problem. This complete frame of reference provides an un-shifting foundation that is applicable across problem domains and is *the primary intellectual contribution*. (**Hypothesis 2**)

A sound simulation model for holistic treatment of transportation has been obtained. It has been shown to replicate historical data to a sufficient degree (after calibration), yet also capable of predictive power without user intervention (or extrapolation of empirical data). Its object-orientation and modular design make it extendable, as demonstrated by examination of heterogeneous vehicle design concepts quantitatively. This simulation model and the results obtained from it concerning the PAV design problem embody *the primary experimental contribution*. (**Hypothesis 1**)

The entity-centric abstraction and the simulation model, when amalgamated with problem specific data, produces an integrated decision-making framework for transportation problems, *the primary methodological contribution*. Nearly any conceivable combination of transportation resources, economies or policies is admissible and evolvable. The research goal to obtain a method that matches the problem type has been achieved: just as the transportation architecture is a living system, so too is the methodology that models it. (**Hypotheses 1 & 2**)

[1] For example, the Department of Transportation has identified five core principles that indicate a paradigm shift towards decision-making that focuses on "a holistic approach, collaboration and consensus building, being flexible and adaptable, informed and transparent decision making, and innovation." (DOT 2000)

The recommendations also are organized along several lines. The *content of the simulation model* itself can be expanded and refined: to capture important time-variant characteristics in all entity-classes (the MIDAS computational tool will enable representation of agents in a thread), to allow use (where necessary) of physics-based analysis and economic analysis codes, to improve granularity & fidelity in stakeholder dynamics, to incorporate feedback between demand and capacity, to generate optimization landscapes for identification & visualization of local optima, to improve the graphical user interface. The decision model for future transportation systems can be advanced by: including additional stakeholders in framework, obtaining and employing stated-preference data to boost the model credibility and enhance predictive power for decisions, and to creating post-1995 ATS data that is more useful, including a detailed assessment of GA need and the support of multiple calibrations of the model.

The *power of the framework* could be further enhanced and evolved for sophisticated real decision-making problems, provided that all participants from multiple domains—vehicle manufacturers, service providers, customers and policy makers from government agency—share this framework to guide an integrated decision-making process. Within this framework, designers can extract essential technical requirements that allow polishing of vehicle concepts; policy makers can investigate the infrastructure and technology impact of new systems; and business planners can perform an analysis based on their own strategies and market projections.

7.3 Concluding Remarks

The 'grand challenge' of modeling & simulation for the NTS and the subsequent challenge of identifying and achieving the best of the possible architectures are significant and daunting at the same time, perhaps a utopian goal. Yet, this dissertation presented one step forward towards an effective approach for these challenges that does not sacrifice salient features in search of tractability. The example application gives evidence of the potential for modeling a class of innovations, without boundaries that come from stovepipes, towards a destination where nearly any possible alternative is admissible.

This one step is an important one, for history has certainly shown that no single agency, program, or technology alone can solve a SoS type problem. History is also replete with examples of ‘unintended consequences’, in which the careful analysis of the interactions between technology, policy, and economics was absent. Ultimately, then, the ideas contained in this dissertation have the promise for improving future transportation architectures, not through promotion of a single piece of technology or combinations of technologies, but instead through the provision of a new ‘calculus’, a new way of thinking. Just as aircraft designers over the past decades have experienced the paradigm shift from performance to affordability in finding robust designs, today’s engineers and policy makers need to embrace a holistic perspective facilitating collaboration to tackle problems like a reshaping of the NTS.

In the final analysis, however, and despite the best intentions, it is the author’s view that the entire transportation ‘universe’ can never be modeled completely. Even if this task is accomplished (surely with countless further steps, even leaps), the result could be far from the real ‘universe’. Yet, the continued effort to fully integrate all entities is meaningful from a pedagogical point of view. Under these circumstances, the best practices appear to be the considered construction of interfaces to link diverse domains, the inclusion of uncertainty to account for incomplete information across interfaces, and the implementation of programming flexibility to accommodate changes that arise in the living system.

> *I hope that the reader will find many motivating and enlightening ideas in this dissertation. My wish is that this body of work will contribute to the development and further awareness of this new and fascinating field. For, as of now, the future looks bright with the glow of ideas. The challenge is enormous, the collaborations are essential, and so we must get started together.*

APPENDIX A

PERSONAL AIR VEHICLE RELATED EFFORTS

Contents

It is always fascinating to look back at the works of the pioneers. Sometimes this practice proves to be beneficial since such a reflection provides a designer with a solid starting point. The first section presents a brief historical account of various individual attempts at producing PAV. The next section introduces the more formal approaches taken by NASA towards realizing the dream of PAV.

A.1 Individual Attempts

As the public think PAV is a flying car, key efforts of the pioneers of dual mode vehicle are described in this section, where the author's role was merely to organize excellent works from Hall (2001), Bowers (1990), and numerous web sites.

A.1.1 The Past of Flying cars

The dream of flying 'things' has always fascinated human beings. Even before the first flight in 1903, there was an attempt in the 18th century to develop a gliding horse cart. To date, there are nearly 100 patents on file at the U.S. Patent and Trademark Office for various kinds of personal flying cars. (Stiles 1994) Some of these have actually flown, but all have come up short of reaching the goal of becoming the world's first mass-produced PAV.

Curtiss Autoplane

The first serious effort toward producing an automobile that could fly was made by American Curtiss Aeroplane & Motor Company. In 1917, Glenn Curtiss unveiled his first attempt at the Pan-American Aeronautic Exposition held in New York. The three-seat automobile body was built of aluminum. The pilot sat alone at the front of the cabin with two passengers who sat behind him side-by-side. A 100-hp Curtiss OXX airplane engine in the car-like nose drove the rear wheels through a gear box and shaft. The Autoplane had three wings that spanned 40 feet and 6 inches. The car's motor drove a four-bladed propeller at the rear of the car. The Autoplane never truly flew, but it did manage a few short hops. *Specifications: Wingspan 40 feet 6 inches; speed range as an airplane 45–65 mph.*

Waterman Arrowbile

Developed by Waldo Waterman in 1937, the Arrowbile was an all-metal, tailless, hybrid flying automobile. Similar to the Autoplane, its direct predecessor, it had a propeller attached to the rear of the vehicle. The three-wheeled car was powered by an automobile engine, a 100-horsepower Studebaker, behind the two-seat cabin. It was capable of up to

70mph on the ground. No wonder the DMV classified it as a motorcycle. The wings could be detached for storage. A lack of funding ended the project. *Specifications: Span 38 feet; gross weight 2,500 lbs; top speed as an airplane 120 mph.*

Pitcairn Whirlwing Roadable Autogyro

The Pitcairn had been the leading U.S manufacturer of autogyros in the U.S since 1928 and had made many technical contributions to the type. In 1939, Pitcairn introduced its model PA-36 Whirlwing, which looked like a conventional autogyro in that the propeller was in the nose. The engine was buried in the all-metal fuselage behind the cabin and drove the propeller through an extension shaft. One power went to the rotor to pre-spin it prior to takeoff and another went to the rear wheel that drove the Whirlwing, with its rotor blades folded, along the road. The outbreak of World War II ended Pitcairn's work. *Specifications: Rotor 43 feet; length 20 feet 5 inches; speed 122 mph.*

Stout/Spratt Skycar IV

Since 1930, William B. Stout had been trying to develop an easy-to-fly "every man's airplane" through his series of Skycars. He teamed up with George Spratt of the Stout Research Division of ConVair. The Spratt/Stout collabration was built in 1944. The wing was mounted above an elongated auto-like body with a burried engine driving pusher propeller at the rear through an extension shaft. The fixed end fins were used for stability only, not control. The wing was movable since it was pivoted in such a way that it could be banked and pitched, so there was no need for elevators, rudder or ailerons. *Specifications: Span 36 feet; length 21 feet 7 inches; speed 114 mph.*

Fulton Airphibian

Robert Fulton developed the Airphibian in 1946. Instead of adapting a car for flying, Fulton adapted a plane for the road. The wings and tail section of the plane could be removed to accommodate road travel, and the propeller could be stored inside the plane's fuselage. It took only five minutes to convert the plane into a car. The Airphibian was the first flying car to be certified by the Civil Aeronautics Administration, the predecessor of the the Federal Aviation Administration (FAA). It had a 150-horsepower, six-cylinder engine and could fly 120 miles per hour and drive at 50 mph. Despite his success, Fulton couldn't find a reliable financial backer for the Airphibian. *Specifications: Span 36 feet 5 inches; gross weight 2,100 lbs; cruising speed 110 mph.*

ConVairCar

In the 1940s, Consolidated-Vultee developed a two-door sedan equipped with a detachable airplane unit. The novelty of the combination was that the car and air units came separately. ConVair announced a price of $1,500 for the car. The air unit was to be available on a rental basis much like a rental trailer. The ConvAirCar 118 debuted in 1947, and offered one hour of flight and a gas mileage of 45 miles (72 kilometers) per gallon. Plans to market the car ended when it crashed on its third flight. *Specifications: Span 34 feet; speed 113 mph.*

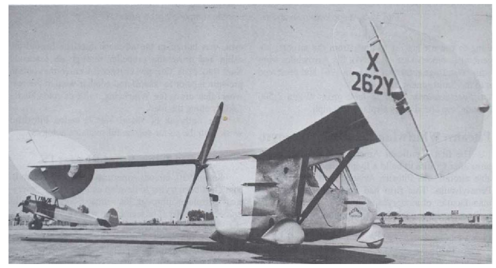
Taylor Aerocar

Inspired by the Airphibian and Robert Fulton, whom he had met years before, Moulton Taylor created perhaps the most well-known and most successful flying car to date. The Aerocar was designed to drive, fly and then drive again without interruption. Taylor covered his car with a fiberglass shell. A 10-foot-long (3-meter) drive shaft connected the engine to a pusher propeller. It cruised at 120 mph in the air and was the second and last roadable aircraft to receive FAA approval. In 1970, Ford Motor Company even considered marketing the vehicle, but the decade's oil crisis dashed those plans. *Specifications: Span 34 feet; gross weight 2,100 lbs; top speed as an airplane 130 mph, as a car 67 mph.*

Bryan Roadable

This was an interesting roadable airplane built by L. Bryan, for which major parts of a standard Ercoupe were used. The basic design was a twin tailboom pusher. The wing could be folded upwards at 90 degrees just outboard of the landing gear and again near the tips so that the wings form a protective square around the propeller. In this airplane, Byran was killed in a crash in 1974. *Specifications: Span 22 feet; length 19 feet; speed 105 mph.*

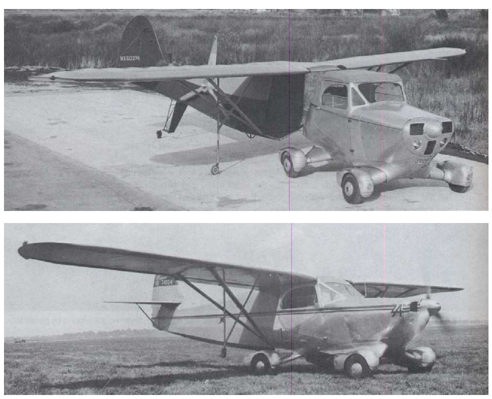
(a) Curtiss Autoplane



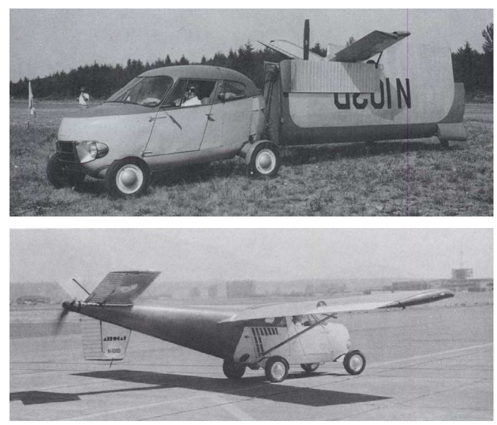
(b) Waterman Arrowbile

(c) Pitcairn Whirlwing

(d) Stout/Spratt Skycar IV



(e) Fulton Airphibian



(f) Taylor Aerocar

(g) ConVairCar

(h) Bryan Roadable

Figure 105: Roadable Personal Air Vehicles (1910s – 1970s)

A.1.2 Recent Efforts

These early pioneers of PAV never managed to develop a viable flying car, and some even died testing their inventions. However, they proved that a car could be built to fly, and inspired a new group of roadable aircraft enthusiasts. With advances in lightweight material, computer modeling and computer-controlled aircraft, the dream is much closer to becoming reality. There is no lack of engineers taking on the challenge to design a new breed of flying cars.

Paul Moller has spent 40 years and millions of dollars developing his Skycar. Moller's latest design, the **Skycar M400**, is designed to take off and land vertically in small spaces. The four-seat Skycar is powered by eight rotary engines that are housed inside four nacelles on the side of the vehicle. The engines lift the craft with 720 horsepower, and then thrust the craft forward. To make the Skycar safe and available to the general public, it will be completely controlled by computers using Global Positioning System (GPS) satellites. In case of an accident, the vehicle will release a parachute and airbags, internally and externally, to cushion the impact of the crash.

MACRO Industries in Huntsville, AL., is developing a flying car that it calls the **SkyRider X2R**. This aero car will be also able to take off and land vertically. SkyRider incorporates the interior design of a 2-seat sports car with the mobility of a helicopter or airplane. MACRO said that the system will be almost fully automatic, but may allow some manual controls. Commands will be entered simply by telling the car what the driver wants it to do. MACRO is planning to demonstrate a working vehicle by 2005. The company is planning to power the vehicle with an enhanced automobile engine to drive the four-ducted fans.

Rafi Yoeli in Israel is developing a family of VTOL vehicles. The **CityHawk** is a prototype of a flying car. The CityHawk is similar to the Skycar and SkyRider in that it also takes off and lands vertically. However, there are some key differences. The CityHawk will be powered by fans that are driven by four internal combustion engines. Much like in the Skycar, this redundancy of engines will allow the vehicle to land even if one of the engines is lost. The CityHawk is about the size of a Chevy Surburban, and will have cruising speeds of 90 to 100 miles per hour.

In 1990, Kenneth Wernicke formed Sky Technologies to develop a small-winged flying car. The **Sky Technologies Aircar** has flown at 200 to 400 mph (322 to 644 kph) and driven at 65 mph (105 kph). It's also small enough to fit into an average parking space.

Recently, Branko Sarh, a senior engineer at McDonnell Douglas Aerospace, has attempted to develop a flying car, called the **Sokol A400**, or Advanced Flying Automobile. Sarh designed a 4-passenger vehicle that would pop out telescoping wings at the push of a button.

LaBiche Aerospace, Inc. has come up with a folding wing concept, called **LaBiche FSC-1**. With a wing span of 32.13 ft, the two sides of the wing are folded in half and stowed in the compartment underneath the cabin when the car is on the ground. The design also has a canard, working as a control and lifting surface to improve aerodynamic capabilities of this sport-car-like vehicle. The canard can be retracted into the compartment in front of the passenger cabin.

Groen Brothers Aviation, Inc. has developed and is certifying with the FAA the world's first turbine powered gyroplane. Powered by a 420 SHP Rolls-Royce Model 250 series gas turbine engine, the **Hawk 4** gyroplane provides USTOL (ultra short takeoff and landing; under sea level standard conditions, as short as 25 feet) capabilities with its patented variable pitch rotor head.

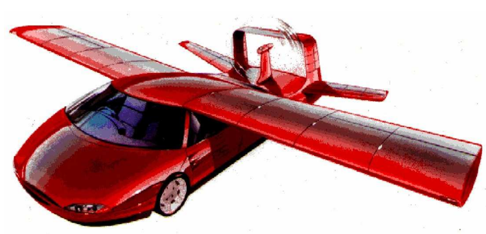
The **CarterCopter** is a vertical takeoff and landing aircraft projected to cruise at 350 MPH at 45,000 feet (200 MPH at sea level). It uses a rotor for vertical takeoff and landing and a small wing for high speed cruise. The company says that the aircraft will demonstrate the speed and efficiency of a fixed wing aircraft and the off-airport abilities of a helicopter through flight tests in the near future.

(a) Moller Skycar M400

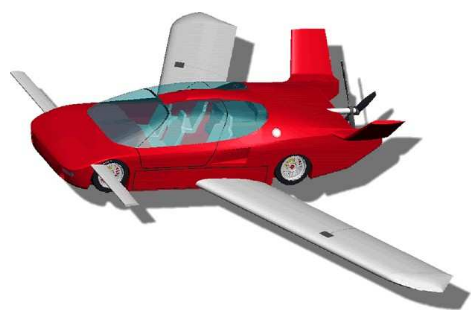
(b) MACRO Skyrider X2R

(c) See-through of Cityhawk

(d) Sky Technologies Aircar



(e) Sokol A400



(f) LaBiche FSC-1

(g) Groen's Hawk 4

(h) CarterCopter

Figure 106: Today's Attempts

A.2 NASA Efforts

Among the many programs and projects within NASA, the Revolutionary Aerospace Systems Concepts (RASC) Program and Small Air Transportation Systems (SATS) Project are closely related to Personal Air Vehicles. The RASC Program aims to develop aerospace systems concepts and technology requirements to enable future civil and possibly military missions and to explore new mission capabilities and discover "What's possible." (NASA LaRC 2002*b*) The initial focus of the RASC Program is developing revolutionary systems concepts including Personal Air Vehicles. The SATS Project is a more infrastructure oriented research task to enhance system capacity which leads to expanded use of general aviation. The following sections introduce two programs that support the ground rules and assumptions of the present research.

A.2.1 Personal Air Vehicle Exploration

Under the framework of the RASC program, NASA PAVE team was formed with recognition that prior efforts related to PAV have jumped into concept development—even prototype development—with minimal justification. Thus, PAVE project aims to build a foundation for future investment as an overall objective of the study. Detailed PAVE study process and objectives are described as follows. (NASA LaRC 2002*a*)

- Establish a foundation
  - Review prior concepts and current relevant technologies.
  - Extract requirements, missions, and constraints.
  - Establish metrics as a basis for comparison.
  - Define potential infrastructure scenarios.
  - Develop baseline vehicles with current technology.
- Explore the design space
  - Define, establish, and integrate synergistic technologies (2015 TRL 6).
  - Develop advanced concepts utilizing physics based methods.

- Compare concepts to reference baselines, each other, and alternate travel modes.

- Determine technology investment approach
  - Show technology sensitivities and gaps for the various mission concepts.
  - Show assumption sensitivities to understand the elasticity of the design space.
  - Present the study results in a highly interactive, intuitive and visual format.

The team categorized concept vehicles into several categories such as reference baselines, advanced single-mode concepts and advanced dual-mode concept. Various configurations, as shown in Figure 107, in each category were analyzed on their design spaces including possible mission, performance requirements and constraints. Also, in order to address accessibility issues, several layers of infrastructures were considered, from CTOL to VTOL compatibility as well as roadability (roadable capability) as shown in Table 62. In addition, comparison metrics were proposed to facilitate the exploration, evaluating baseline vehicles in a unified way.

Table 62: PAV Concept Matrix

	CTOL (2000 ft)	STOL (1000 ft)	SSTOL (500 ft)	VTOL (100 ft)
Single-Runway				
Dual-Taxi				
Dual-Side Street				
Dual-Highway				

Figure 107: Various Personal Air Vehicles under Consideration

Other findings include identification of synergistic technologies in assisting technology investment, and the key guidelines of PAV design are listed below.

- **Safety:**
  - Simple, low complexity systems (non-professional pilots)
  - Very low takeoff and landing speeds
  - Minimum external systems (hanger rash / bump / tamper proof)
- **Environment:**
  - Low noise (close proximity operations)
  - Automotive equivalent emissions
  - Low downwash (ground erosion and foreign object kickup)
- **Cost:** Vehicles must provide a positive ROI compared to value of time
- **Size:** Vehicles must fit into limits imposed by existing infrastructure.

It is expected that follow-on work will provide detailed designs, technologies, and costing as well as greater depth in top-level system benefits in accordance with the design criteria. In summary, the PAVE project has laid a foundation for successful development of future PAV systems in a more organized fashion, which had never been attempted before.

A.2.2 Small Aircraft Transportation System

While the PAVE project is developing both a system and an understanding about its design space, the SATS project is developing enabling airspace technologies and the equivalent of Highways In The Sky (HITS). Initial concept of the SATS program was deeply rooted from the aspiration to revitalize general aviation. (Fallows 2001) Since the 1980s, as briefly discussed in Chapter I, the general aviation industry has gone through a long downturn, making the small airports throughout the nation sit even more idle. The current air transport system relies on hub-and-spoke infrastructure which entails thirty large airports to process more than 70% of the total throughput, which results in big airports, big aircraft, and big irritation. If, somehow, small aircraft can attract many travelers, the benefit would be two-fold, permitting both an offloading of hub-and-spoke infrastructure and congested highway systems. The dream of the SATS project lies here. Once smaller, under-utilized airports become easily accessible with advanced, easy-to-fly and affordable general aviation aircraft, travelers would enjoy reduced door-to-door trip time due to high speed mobility as well as expanded accessibility to ubiquitous portals, even in an on-demand basis. The nation already has enough resources. There are more than 5,400 public use airports and nearly 13,000 private airports including verti-ports as evidenced in Figure 108(b) and 108(c). In examining this vision, NASA leads the SATS program focused on maturing technologies needed.

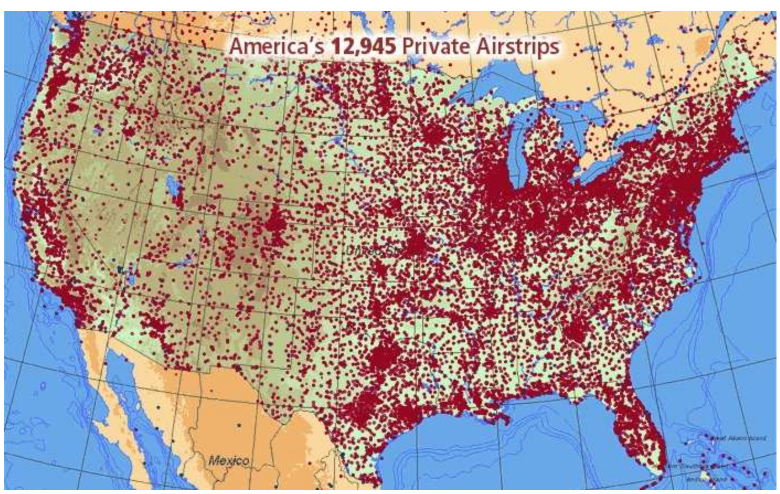
Since the SATS project is a proof-of-concept research, the initial goal is to demonstrate key technologies for precise guided accessibility in small aircraft in near all weather conditions to virtually any airport in non-radar / non-towered airspace. In achieving this goal, the SATS team set the following objectives. (Durham & Creedon 2001)

- Higher volume operation at non-towered / non-radar airports
- Lower landing minimums at minimally equipped landing facilities
- Increase single-pilot crew safety and mission reliability
- En route procedures and systems for integrated fleet operations

The enabling technologies include self-sequencing and separation systems, airborne Internet, software-enabled controls and highway-in-the-sky guidance. The technologies targeted for development are aimed at smaller aircraft used for personal and business transportation missions within the infrastructure of smaller airports throughout the nation. These missions included transportation of goods and travel by individuals and families or groups of business associates. Consequently, the aircraft must be of a similar size-class of typical automobiles and vans used for non-commercial ground transportation. They may be used for on-demand, unscheduled air-taxi transportation of these same user types. Various forms

(a) Large Hub Airports

(b) Small Public Airports



(c) Private Airports

Figure 108: Airports in the U.S.

of shared ownership and usage will likely be the most common means of use. While the aircraft are not specifically designed for commercial operations, the targeted technologies would provide benefits to commuter and major air carrier operations in the hub-and-spoke system as well. For FAA regulatory purposes, SATS technologies are targeted towards aircraft with a maximum take-off weight less than 12,500 pounds.

A.3 Summary

- Individual attempts explored the PAV design space, which resulted in expanding its boundaries and numerous concept vehicles.
- NASA's PAVE project: it is in an appropriate stance to cope with the vast PAV design space with rational ground rules and approach.
- NASA's SATS project: it is focusing on infrastructure and system technologies. This research is vital for any kind of PAV systems in the future.
- Two projects, when combined together, seem to be the state-of-the-art approach to the mobility goal. Hence, it is worthwhile to abide by the ground rules and assumptions learned from both projects.

APPENDIX B

MULTI-CRITERIA DECISION MAKING

Contents

Every decision-making problem involves multiple criteria by nature, are considered simultaneously. This contributes to expanding the knowledge boundary on complex design space in or even before the conceptual design phase. For this reason, it is forecasted that MCDM techniques will be more regularly exercised on practical problems in every domain. (Fishburn & Lavalle 1999, Lootsma 1999) MCDM can be divided into two branches (Miettinen 1999, Triantaphyllou 2000): Multi-Attribute Decision Making (MADM) and Multi-Criteria Optimization (MCO). MADM relates to techniques that aid a decision maker in choosing the best design from a small number of alternatives. MCO is also known as multi-objective optimization or vector optimization, and its task is to present a set of designs that are the most appealing alternatives to a decision maker. The following sections introduce the two branches of MCDM.

B.1 Multi-Attribute Decision Making

MADM concentrates on problems with discrete designs. In these problems the set of decision alternatives are predetermined. Although MADM methods may be widely diverse, many of them have certain aspects in common.

B.1.1 Definitions

Alternatives: Alternatives represent the different choices of action available to the decision maker. Usually, the set of alternatives is assumed to be finite, ranging from several to hundreds. They are supposed to be screened, prioritized and eventually ranked.

Attributes: Each MADM problem is associated with multiple attributes. Attributes represent the different dimensions from which the alternatives can be viewed. Attributes are also referred to as "goals" or "criteria". This is why the terms MADM and MCDM have been used very often to mean the same in the literature. In the context of the present article, however, MADM is always a subset of MCDM.

Incommensurable units: Different attributes may be associated with different units of measure. For instance, in the case of buying a used car, the attributes "cost" and "mileage" may be measured in terms of dollars and thousands of miles, respectively. It is this nature of having to consider different units which makes MADM to be intrinsically hard to solve.

Hybrid measures: Some attributes are evaluated by qualitative statements (i.e., good / fair / bad). In addition, a composite MADM problem can be posed with qualitative and quantitative attributes.

Decision matrix: A MADM problem can be easily expressed in a matrix format. A decision matrix A is an $(M \times N)$ matrix in which the element a_{ij} indicates the performance of alternative A_i when it is evaluated in terms of the decision attribute C_j, (for $i = 1,2,3,\cdots,M$ and $j = 1,2,3,\cdots,N$). It is also assumed that the decision maker has determined the weights of relative performance of the decision criteria (denoted as W_j, for $j = 1,2,3,\cdots,N$). This information is best summarized in Table 63.

Table 63: A Typical Decision Matrix

Criteria / Alternatives	C_1	C_2	$\cdots$	C_N
A_1	a_{11}	a_{12}	$\cdots$	a_{1N}
A_2	a_{21}	a_{22}	$\cdots$	a_{2N}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
A_M	a_{M1}	a_{M2}	$\cdots$	a_{MN}

B.1.2 MADM Methods

Over the years, numerous MADM methods have been published, with each method having its own characteristics. Thus, there are many ways one can classify MADM methods. Some authors classify them according to the type of data, i.e., there can be deterministic, stochastic, or fuzzy MADM methods. However, there may be situations which involve combinations of some or all the above (such as stochastic *and* fuzzy data) data types. Alternately, MADM methods can be classified according to the number of decision makers involved in the decision process. Hence, we have single decision maker MADM methods and group decision making MADM. In Chen & Hwang (1992) deterministic—single decision maker—MADM methods were also classified according to the type of information and the salient features of the information. The Weighted Sum Model (WSM), Analytic Hierarchy Process (AHP), revised AHP, Weighted Product Model (WPM), and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) methods are the more popular methods which are used in practice today and are described in later sections. Finally, it should be stated here that there are many other alternative ways for classifying MADM methods, but, the ones that are summarized above are the most widely used approaches found in the MADM literature.

B.1.2.1 Weighted Sum Model

The Weighted Sum Model (WSM) is probably the most commonly used MADM approach, especially for single dimensional problems. If there are M-alternatives and N-attributes, then the best alternative is the one that satisfies (in the maximization case) the following expression:

$$A^* = \max_i \sum_{j=1}^{N} a_{ij} w_j, \text{ for } i = 1, 2, \cdots, M \tag{76}$$

where A^* is the WSM score of the best alternative, N is the number of attributes, a_{ij} is the actual value of the i-th alternative in terms of the j-th attribute, and W_j is the weight of importance of the j-th attribute.

The assumption that governs this model is the additive utility assumption. That is, the total value of each alternative is equal to the sum of products given as Equation 76. In single-dimensional cases, in which all units are the same (e.g., dollars, feet, seconds), the WSM can be used without difficulty. Nevertheless, difficulty with this method emerges when it is applied to multi-dimensional decision-making problems. Then, in combining different dimensions, and consequently different units, the additive utility assumption is violated and the result is equivalent to adding "apples and oranges". Also, the WSM is incapable of working on concave Pareto fronts, which will be covered in more detail in a later section.

B.1.2.2 Weighted Product Model

The Weighted Product Model (or WPM) is very similar to the WSM. The main difference is that instead of addition, there is multiplication. Each alternative is compared with others by multiplying a number of ratios, one for each attribute. Each ratio is raised to the power equivalent to the relative weight of the corresponding attribute. In general, in order to compare the alternatives A_K and A_L, the following product (Miller & Starr 1969) must be calculated:

$$R(A_k/A_L) = \prod_{j=1}^{N} (a_{Kj}/a_{Lj})^{w_j} \tag{77}$$

where N is the number of attributes, a_{ij} is the actual value of the i-th alternative in terms of the j-th attribute, and w_j is the weight of importance of the j-th attribute. If the term $R(A_k/A_L)$ is greater than one, then alternative A_K is more desirable than alternative A_L (in the maximization case). The best alternative is the one that is better than or at least equal to all the other alternatives. The WPM is sometimes called dimensionless analysis because its structure eliminates any units of measure. Thus, the WPM can be used in single- and multi-dimensional decision-making problems. An alternative approach is one to use only products without ratios. That is, to use the following variant of Equation 77:

$$P(A_k) = \prod_{j=1}^{N} (a_{Kj})^{w_j} \tag{78}$$

B.1.2.3 The Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP) is based on decomposing a complex MADM problem into a system of hierarchies. The final step in the AHP deals with the structure of an $M \times N$ matrix (where M is the number of alternatives and N is the number of attributes). This matrix is constructed by using the relative importance of the alternatives in terms of each attribute. The vector $(a_{i1},\ a_{i2},\ a_{i3},\ \cdot,\ a_{iN})$ for each i is the principal eigenvector of an $N \times N$ reciprocal matrix which is determined by pairwise comparisons of the impact of the M alternatives on the i-th attribute.

Some evidence is presented in Saaty (1980) which supports the technique for eliciting numerical evaluations of qualitative phenomena from experts and decision makers. The entry a_{ij}, in the $M \times N$ matrix, represents the relative value of the alternative A_i when it is considered in terms of attribute C_j. In the original AHP, the sum $\sum_{i=1}^{N} a_{ij}$ is equal to one. According to AHP the best alternative (in the maximization case) is indicated by the following relationship:

$$A^{*}_{AHP} = \max_{i} \sum_{j=1}^{N} a_{ij} w_j, \text{ for } i = 1, 2, \cdots, M \tag{79}$$

The similarity between the WSM (Equation 76) and the AHP (Equation 79) is evident.

The AHP uses relative values instead of actual ones. Thus, it can be used in single- or multi-dimensional decision making problems.

In relation to AHP, Belton & Gear (1983) proposed a revised version of the AHP model. They demonstrated that an inconsistency can occur when the AHP is used. In their example, the indication of the best alternative changes when an identical alternative to one of the non-optimal alternatives is introduced. Saaty (1990) criticized the revised AHP claiming that identical alternatives should not be considered in the decision process. However, Triantaphyllou & Mann (1995) have demonstrated again that similar logical contradictions are possible with the original AHP as well as with the revised AHP even though non-identical alternatives are introduced.

B.1.2.4 The ELECTRE Method

The ELECTRE (for Elimination and Choice Translating Reality; English translation from the French original) method was first introduced in Benayoun, Roy & Sussman (1966). One of the distinguishing features of the ELECTRE method is that it is fundamentally non-compensatory, which means a very bad score on an attribute cannot be compensated by good scores on other attributes. Another original feature is that the ELECTRE method allows for incomparability. The important concepts of the ELECTRE method are "threshold" and "outranking".

The ELECTRE method begins with pairwise comparisons of alternatives under each attribute. Using physical or monetary values $g_i(A_j)$ and $g_i(A_k)$ of the alternatives A_j and A_k respectively, and introducing "threshold" levels for the difference $g_i(A_j) - g_i(A_k)$, the decision maker may declare that she/he is indifferent between the alternatives under consideration, that she/he has a weak or strict preference for one of the two, or that she/he is unable to express any of these preference relations. Therefore, the set of binary relations of alternatives, the so-called "outranking" relations, may be complete or incomplete. Next, the decision maker is requested to assign weights or importance factors to the attributes in order to express their relative importance.

Through a series of consecutive assessments of the outranking relations of the alternatives, ELECTRE elicits the so-called concordance index, defined as the amount of evidence to support the conclusion that A_j outranks A_k, as well as the discordance, the counter-part of concordance index.

Finally, the ELECTRE method yields a whole system of binary outranking relations between the alternatives. Because the system is not necessarily complete, the ELECTRE method is sometimes unable to identify the preferred alternative. It only produces a core of leading alternatives. This method has a more explicit view of alternatives by eliminating less favorable ones, especially convenient while encountering few attributes with a large number of alternatives in a decision making problem (Lootsma, Mensch & Vos 1990).

B.1.2.5 TOPSIS

TOPSIS (the Technique for Order Preference by Similarity to Ideal Solution) was developed by (Hwang & Yoon 1981) as an alternative to the ELECTRE method. The basic concept of this method is that the selected alternative should have the shortest Euclidian distance from the ideal solution and the farthest distance from the negative-ideal solution. TOPSIS assumes that each attribute has a tendency of monotonically increasing or decreasing utility. Therefore, it is easy to locate the ideal and negative-ideal solutions. The Euclidean distance approach is used to evaluate the relative closeness of alternatives to the ideal solution. Thus, the preference order of alternatives is yielded through comparing these relative distances.

B.2 Multi-Criteria Optimization

In contrast to MADM, MCO treats infinite design spaces that are continuous. This necessitates the presentation of mathematical backgrounds as follows.

B.2.1 Fundamentals

Envision an N-dimensional design space $\Theta \subset \mathbb{R}^N$, a design vector $X \in \Theta$, an n-dimensional criterion space $\Omega \subset \mathbb{R}^n$, a criterion vector $Y \in \Omega$, and a mapping $F : X \in \Theta \longmapsto Y \in \Omega$. This is illustrated in Figure 109 as an example taking $N = 3$ and $n = 2$. If the task is optimizing either f_1 or f_2, one simply needs to employ a single objective optimization method. Then the solution would end up to point P or P' in this particular example. Nevertheless, the core of MCO is to minimize both functions at the same time. Mathematically, a general form of an MCO problem is stated as follows:

$$\begin{aligned}&\text{“Minimize”}\ \ F(X) = [f_1(X), f_2(X), \cdots, f_n(X)]^T \\ &\text{Subject to: } G(X) \leq O,\ H(X) = O\end{aligned} \tag{80}$$

The scalar function $f_i(X)$ denotes an i-th criterion that is an element of the criterion vector $F(X)$. The vector function $G(X)$ and $H(X)$ indicate inequality and equality constraints respectively that bound a feasible design space Θ, i.e., $\Theta = \{X | G(X) \leq O,\ H(X) = O\}$. The symbol O simply indicates a zero vector.

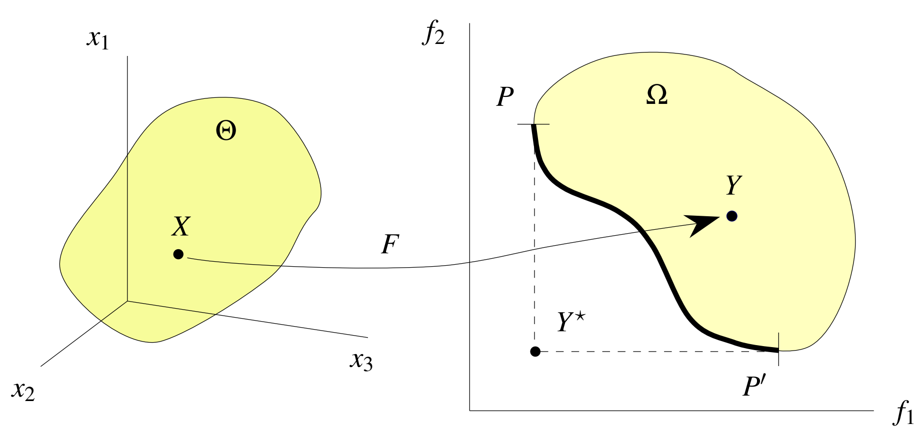


Figure 109: Design Space and Criterion Space

The difficulty comes in due to the nature of $F(X)$. Since $F(X)$ is an n-dimensional vector function, unlike single objective optimization problems, the solution of Problem (80) would be a set of optima rather than a single optimum. For example, any point on the thick curve $\widetilde{PP'} \subset \Omega$ in Figure 109 can be presented as the solution of the MCO problem. This is called a *non-inferior solution*, *nondominated solution*, *efficient point*, or *Pareto optimum*. The mathematical definition of a Pareto optimum is given as follows by Miettinen (1999):

Definition 1. *(Global) Pareto optimum*

> A design vector $X^* \in \Theta$ is called a *Pareto optimum* if there does not exist another design vector $X \in \Theta$ such that $f_i(X) \leq f_i(X^*)$ for all $i = 1, 2, \cdots, n$ and $f_j(X) < f_j(X^*)$ for at least one index j.
>
> A criterion vector $Y^* \in \Omega$ is a *Pareto optimum* if the design vector corresponding to it is a Pareto optimum.

Now, the *Pareto front* $\mathcal{P}$ can be defined as a set of all Pareto optima in Ω. In a two criteria case, for example, it can be visualized by drawing a curve as in Figure 109; in a three criteria case, a surface. A very similar term, a local Pareto optimum, is defined for later use.

Definition 2. *Local Pareto optimum*

> A design vector $X^* \in \Theta$ is called a *local Pareto optimum* if there exists $\delta > 0$ such that X^* is a Pareto optimum in $\Theta \cap B(X^*, \delta)$.
>
> An objective vector $Y^* \in \Omega$ is a *local Pareto optimum* if the design vector corresponding to it is a local Pareto optimum.

It is obvious that X^* being a global Pareto optimum implies that X^* is a local Pareto optimum. The converse does not always hold, which imposes another difficulty in solving MCO problems. Lastly, another useful term is introduced.

Definition 3. *Utopian vector*

> A *utopian vector* $Y^\star \in \mathbb{R}^n$ is defined such that

$Y^{\star} \equiv [\min f_1, \min f_2, \cdots, \min f_n]^T$. Each minimization is required to satisfy the original constraints, i.e., $X \in \Theta$.

The utopian vector $Y^{\star}$ is marked in Figure 109. It can be perceived as a goal or an aspiration point. An MCO problem collapses to a single objective optimization problem when a utopian vector sits on $\mathcal{P}$.

B.2.2 MCO Methods

The objective of MCO methods is to locate the Pareto optima and thus to generate a complete Pareto front $\mathcal{P}$. Numerous methods have been developed in the past few decades to do this task efficiently. Interested readers are referred to Hwang & Masud (1979), Chankong & Haimes (1983) and Ehrgott (2000) for detailed reviews. One cannot judge, however, that a particular method is superior to others. This is mainly because characteristics of MCO problems are too diverse depending on problem-specific situations. It is next to impossible to come up with a generic method that works evenly well for every MCO problem. Therefore, the key criteria in choosing from various methods should be based on practicability. In other words, a user should look into not only whether the method fits her/his specific needs appropriately, but also the difficulty or simplicity of numerical implementation.

In understanding MCO methods, similar to how MADM methods are classified differently by different authors, an absolute way of categorizing MCO methods do not exist. For example, Hwang & Masud (1979) categorized the methods according to the decision maker's participation in the optimization process; namely, no-preference, a posteriori, a priori and interactive methods. Carmichael (1981) did so into three bases: a composite single criterion, a single criterion with constraints and many single criteria. In the present research, the classification is made based on whether a single solution or a set of solutions is generated at a single execution of the method: namely, the one-by-one strategy and the all-at-once strategy.

B.2.2.1 One-by-one Strategy

This strategy begins with switching an MCO problem to a single objective optimization problem by introducing a surrogate function $F^s : X^s \times \alpha \mapsto \mathbb{R}^1$. The new design vector X^s consists of the original design vector X and an extra design vector λ if needed. The parameter vector α is a necessary input to coordinate a search procedure. Now Problem (80) will be converted as follows:

$$\begin{aligned} &\underset{X \in \Theta}{\text{Minimize}} \;\; F^s = F^s(X^s;\alpha) \\ &\text{Subject to: } G^s(X,\lambda;\alpha) \leq O,\; H^s(X,\lambda;\alpha) = O \end{aligned} \tag{81}$$

where G^s and H^s are additional constraints. These extra entities, including λ, may or may not be present to complete the conversion process depending on the method. Now that Problem (80) has changed to a surrogate Problem (81), only a single Pareto optimum would be searched at a single execution if the solution converged successfully. Changing the value of α will entail a new Pareto optimum. Through such a sequential process, the Pareto front $\mathcal{P}$ would be formed by accumulating the Pareto optima. An outline of three basic approaches adopting the one-by-one strategy follows.

• *Weighted Sum Method*

The WSM simply defines F^s as a composite of each criterion.

$$F^s \equiv W \cdot F(X) = \sum_{i=1}^{n} w_i\, f_i(X) \tag{82}$$

Here, $W = [w_1, w_2, \cdots, w_n]$ (usually $\sum w_i = 1$) corresponds to the parameter vector α in Problem (81). By perturbing parameter vector W or weights, each optimization process will produce a different Pareto optimum. The serious drawback of the method is that the method cannot generate complete description of a Pareto front that is not convex. This situation is depicted in Figure 110. No matter how W is altered, the portion of the Pareto front between Y_a^* and Y_b^* can never be obtained.

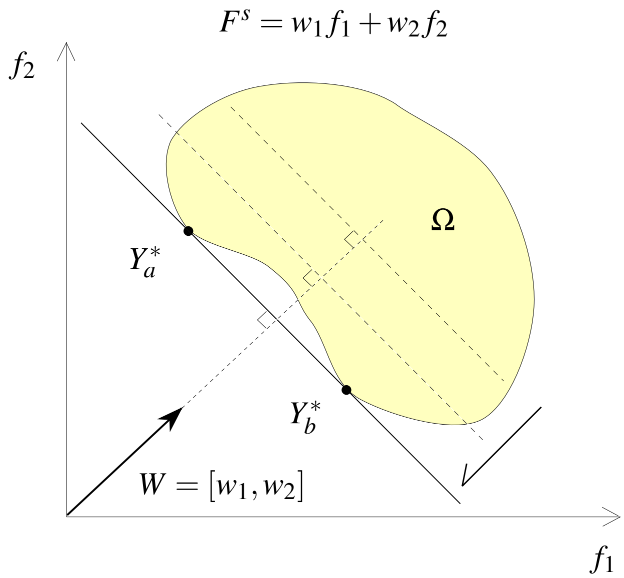


Figure 110: Weighted Sum Method

- ε-*Constraint Method*

The ε-constraint method is also based on a simple idea which has become a routine for single objective optimization process. The most important criterion function is chosen and optimized taking the remaining criterion functions as constraints. Hence, F^s is defined as follows without loss of generality:

$$F^s \equiv f_n(X) \tag{83}$$

The other criteria are incorporated into G^s. This extra constraint vector has $(n–1)$ inequality constraints and they are:

$$f_i(X) - \varepsilon_i \leq 0 \ \ (i = 1, 2, \cdots, n–1) \tag{84}$$

Specific ε_i values (for $i = 1, \cdots, n–1$) need to be determined before the optimization. The optimization process is illustrated in Figure 111 taking $F^s = f_2(X)$. The Pareto optimum Y_0^* would be obtained with given value $\varepsilon_1 = \xi_0$. If the optimization is repeated with $\varepsilon_1 = \xi_1$, the Pareto optimum Y_1^*, which could not be obtained by the previous method, can be searched. The major drawback of this method is that it is time consuming to figure out appropriate numeric values for ε_i, especially when $n = \dim \Omega$ becomes larger.

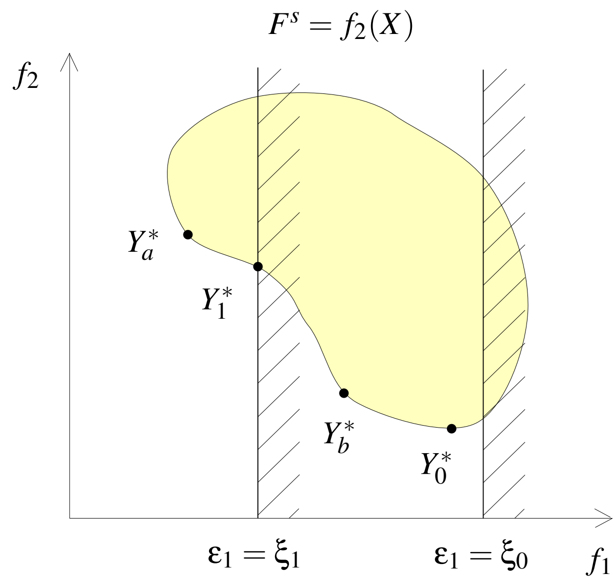


Figure 111: ε-Constraint Method

• *Distance Metric Method*

In this method, F^s is defined as follows:

$$F^s \equiv \|F(X) - \hat{F}\|_p = \left(\sum_{i=1}^{n} |f_i(X) - \hat{f}_i|^p\right)^{\frac{1}{p}} \tag{85}$$

where $\hat{F} = [\hat{f}_1, \hat{f}_2, \cdots, \hat{f}_n]^T$ is a predetermined goal vector. The parameter p is usually set to 1, 2 or infinity. The basic idea behind this method is shown in Figure 112. The function F^s measures the distance between $F(X)$ and $\hat{F}$. If the optimizer minimizes the distance from $\hat{F}_1$, Y_1^* will be obtained. Moving a goal to $\hat{F}_2$ will make the optimizer search for point Y_2^*. However, this method is sensitive to the position of $\hat{F}$. For example, starting from $\hat{F}_3 \in \Omega$ will do not accomplish anything. Furthermore, if $\hat{F}_4$ is chosen, a meaningless point on $\partial\Omega$ will be the final outcome.

B.2.2.2 All-at-once Strategy

Even though various methods adopting the one-by-one strategy have been successfully applied to many applications, there exist a number of weaknesses. First, as the name implies,

$$F^s = \|F - \hat{F}\|_2$$

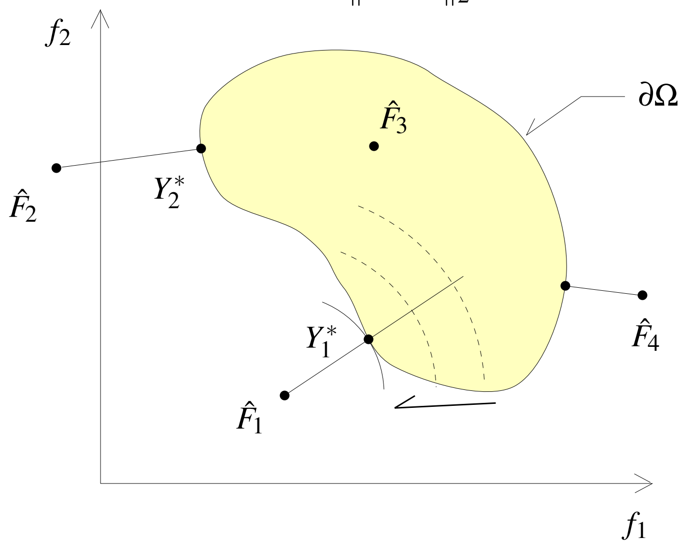
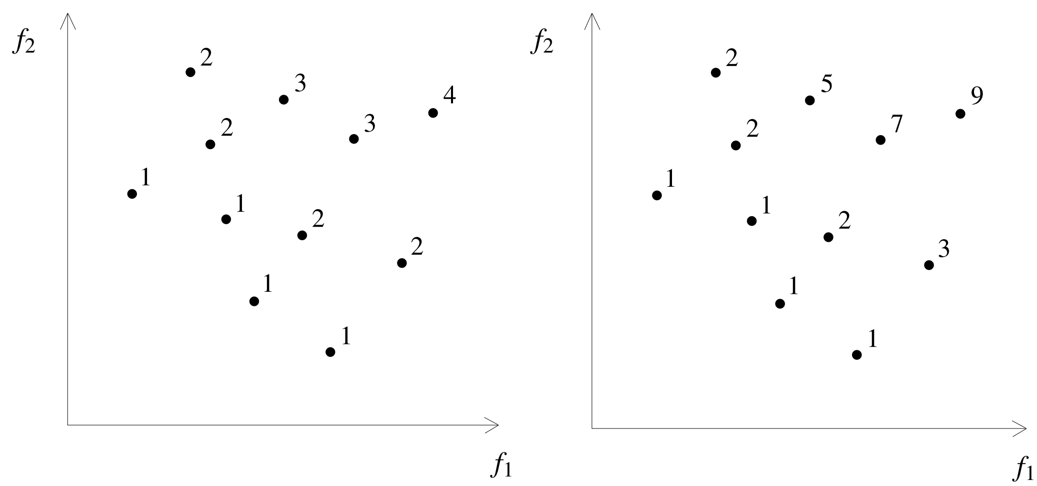


Figure 112: Distant Metric Method

many repetitions are required to obtain the entire Pareto front. Second, some methods are very sensitive to the shape of the Pareto front. Last but not least, in order to do an effective search, the one-by-one strategy requires some degree of a priori knowledge about the criterion space. These limitations can be resolved through the all-at-once strategy also known as Multi-Objective Evolutionary Algorithms (MOEAs) or Multi-Objective Genetic Algorithms (MOGAs). These GA-based MCO techniques have gained much attention over the past decade since they have intriguing concepts and a lot of potential to tackle MCO problems.

Genetic Algorithms (GAs) are sharply distinguished from calculus-based optimization algorithms in that they do not call for analytic information from an objective function. Thus, they can deal with objective functions that need not be differentiable or continuous. This feature of GAs makes them a very versatile optimization method. However, GAs should work on a discrete design space (except the real GAs). Also, unlike calculus-based optimization algorithms, it is difficult to check whether the final outcome is a converged one. The most serious issue of a GA is that it requires much more function calls than any other gradient-based optimization algorithms.

In what follows, two specific ways to combine a GA and MCO will be introduced. One is called nondominated sorting procedure proposed by Goldberg (1989, p. 201) Among the population, the nondominated individuals are ranked 1 then they are removed. The next nondominated individuals are ranked 2 and also removed. This process is repeated as illustrated in Figure 113(a). Fonseca & Fleming (1993) presented a different scheme focusing on each individual, which is depicted in Figure 113(b). If an individual is dominated by R other individuals, then $(R+1)$ is assigned for its rank. Under the Darwinian principle of the survival of the fittest, top-ranking individuals are likely to be chosen to reproduce their offspring for the next generation. A hypothetical snapshot of the GA evolution is shown in Figure 114. While the initial population is scattered randomly in Ω, the individuals of the final generation are gathered near the Pareto front.



(a) Group Ranking (b) Individual Ranking

Figure 113: Population Ranking Methods

GA-based MCO methods rely on aforementioned strengths of GAs. Thus, it is possible to find multiple Pareto optimum in a single execution. Also, one does not need to be concerned with the shape of the Pareto front. More importantly, it is not necessary to have a priori knowledge for a given problem as the knowledge naturally grows from evolutionary process. On the other hand, GA-based MCO methods inherit the disadvantages as well. Due to their stochastic nature, repetitive executions are needed to ensure reliable solutions.

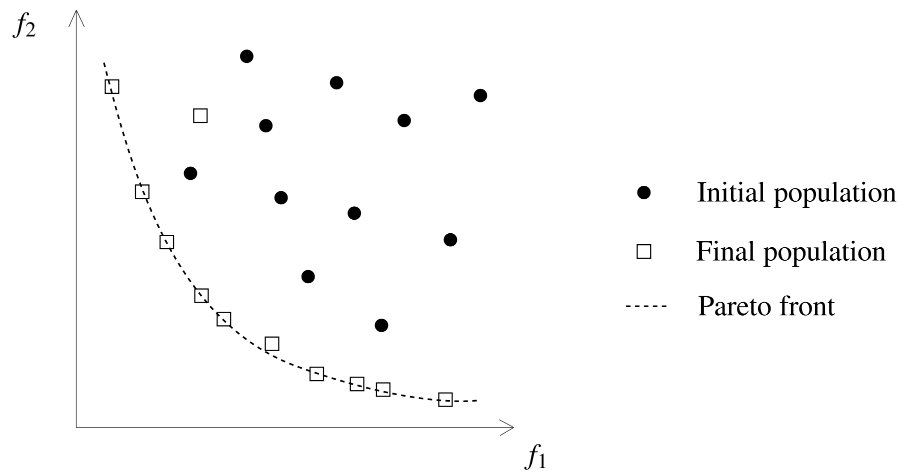


Figure 114: GA Evolution Snapshot

It is again emphasized that GAs require a high number of function calls. In the end, a user often experiences a situation in which any advantages of a GA-based method become quickly diluted due to the huge amount of function calls.

Consequently, the user should be familiar with the advantages and disadvantages from one-by-one and all-at-once strategy so that she/he may choose the most suitable strategy and method. In some occasions, it may be worthwhile to reflect on the potential benefits from combining both strategies wisely for her/his MCO problem.

APPENDIX C

REFERENCE DATA

Tables 64 through 68 show the postprocessing result of the Census 2000 database. Since the scope of the study is on the CONUS, urban areas in HI, AK, PR, etc. are eliminated. For a complete list, refer to `http://www.census.gov/geo/www/ua/uaucinfo.html#lists`.

Table 64: Summary Statistics of the Four Locales

Locale	Statistic	Population	Area (mile^2)	Radius (miles)
	Average	5,521,697	1,462.7	21.05
LAR (N = 15)	Median	3,933,920	1,295.3	20.31
	Std. dev.	4,162,972	697.5	4.90
	Average	1,031,114	400.3	11.03
MED (N = 55)	Median	882,295	369.0	10.84
	Std. dev.	498,762	178.8	2.42
	Average	136,466	73.3	4.56
SML (N = 380)	Median	97,300	53.6	4.13
	Std. dev.	95,377	55.4	1.60
	Average	9,490	6.5	1.30
NOM (N = 3,114)	Median	5,938	4.1	1.14
	Std. dev.	9,001	8.1	0.60

Table 65: Urbanized Areas in Locale LAR

Rank	Name	Population	Area (mile^2)	Radius (miles)
1	New York–Newark, NY–NJ–CT	17,799,861	3,352.6	32.67
2	Los Angeles–Long Beach–Santa Ana, CA	11,789,487	1,667.9	23.04
3	Chicago, IL–IN	8,307,904	2,122.8	25.99
4	Philadelphia, PA–NJ–DE–MD	5,149,079	1,799.5	23.93
5	Miami, FL	4,919,036	1,116.1	18.85
6	Dallas–Fort Worth–Arlington, TX	4,145,659	1,407.0	21.16
7	Boston, MA–NH–RI	4,032,484	1,736.2	23.51
8	Washington, DC–VA–MD	3,933,920	1,156.8	19.19
9	Detroit, MI	3,903,377	1,261.5	20.04
10	Houston, TX	3,822,509	1,295.3	20.31
11	Atlanta, GA	3,499,840	1,962.6	24.99
12	San Francisco–Oakland, CA	3,228,605	526.7	12.95
13	Phoenix–Mesa, AZ	2,907,049	799.0	15.95
14	Seattle, WA	2,712,205	953.6	17.42
15	San Diego, CA	2,674,436	782.3	15.78

Table 66: Urbanized Areas in Locale MED

Rank	Name	Population	Area (mile^2)	Radius (miles)
16	Minneapolis–St. Paul, MN	2,388,593	894.2	16.87
18	St. Louis, MO–IL	2,077,662	829.0	16.24
19	Baltimore, MD	2,076,354	682.7	14.74
20	Tampa–St. Petersburg, FL	2,062,339	802.3	15.98
21	Denver–Aurora, CO	1,984,889	498.8	12.60
22	Cleveland, OH	1,786,647	647.0	14.35
23	Pittsburgh, PA	1,753,136	852.4	16.47
24	Portland, OR–WA	1,583,138	473.9	12.28
25	San Jose, CA	1,538,312	260.1	9.10
26	Riverside–San Bernardino, CA	1,506,816	438.8	11.82
27	Cincinnati, OH–KY–IN	1,503,262	671.8	14.62
28	Virginia Beach, VA	1,394,439	526.8	12.95
29	Sacramento, CA	1,393,498	369.0	10.84
30	Kansas City, MO–KS	1,361,744	584.4	13.64
31	San Antonio, TX	1,327,554	407.6	11.39
32	Las Vegas, NV	1,314,357	285.9	9.54
33	Milwaukee, WI	1,308,913	487.0	12.45
34	Indianapolis, IN	1,218,919	552.9	13.27
35	Providence, RI–MA	1,174,548	503.6	12.66
36	Orlando, FL	1,157,431	453.2	12.01
37	Columbus, OH	1,133,193	397.7	11.25
38	New Orleans, LA	1,009,283	197.8	7.94
39	Buffalo, NY	976,703	366.7	10.80
40	Memphis, TN–MS–AR	972,091	399.8	11.28
41	Austin, TX	901,920	318.1	10.06
42	Bridgeport–Stamford, CT–NY	888,890	465.3	12.17
43	Salt Lake City, UT	887,650	230.7	8.57
44	Jacksonville, FL	882,295	410.5	11.43
45	Louisville, KY–IN	863,582	391.3	11.16
46	Hartford, CT	851,535	469.3	12.22
47	Richmond, VA	818,836	436.8	11.79
48	Charlotte, NC–SC	758,927	434.9	11.77
49	Nashville-Davidson, TN	749,935	430.8	11.71
50	Oklahoma City, OK	747,003	322.4	10.13
51	Tucson, AZ	720,425	291.4	9.63
53	Dayton, OH	703,444	323.6	10.15
54	Rochester, NY	694,396	295.2	9.69
55	El Paso, TX–NM	674,801	219.1	8.35
56	Birmingham, AL	663,615	392.1	11.17
57	Omaha, NE–IA	626,623	226.4	8.49
58	Albuquerque, NM	598,191	224.0	8.44
59	Allentown–Bethlehem, PA–NJ	576,408	289.5	9.60
60	Springfield, MA–CT	573,610	309.0	9.92
61	Akron, OH	570,215	307.8	9.90
62	Sarasota–Bradenton, FL	559,229	270.4	9.28
63	Albany, NY	558,947	284.3	9.51
64	Tulsa, OK	558,329	261.4	9.12
65	Fresno, CA	554,923	138.6	6.64
66	Concord, CA	552,624	176.5	7.49
67	Raleigh, NC	541,527	319.6	10.09
68	Grand Rapids, MI	539,080	257.4	9.05
69	Mission Viejo, CA	533,015	136.9	6.60
70	New Haven, CT	531,314	285.3	9.53
71	McAllen, TX	523,144	313.8	9.99
72	Toledo, OH–MI	503,008	202.3	8.03

Table 67: Urban Areas in Locale SML (Sampled)

Rank	Name	Population	Area (mile^2)	Radius (miles)
73	Baton Rouge, LA	479,019	280.7	9.45
80	Ogden–Layton, UT	417,933	180.1	7.57
87	Flint, MI	365,096	231.1	8.58
94	Spokane, WA–ID	334,858	143.1	6.75
101	Modesto, CA	310,945	86.1	5.23
109	Corpus Christi, TX	293,925	110.3	5.93
116	Fayetteville, NC	276,368	167.1	7.29
123	Greensboro, NC	267,884	135.5	6.57
130	Barnstable Town, MA	243,667	286.2	9.54
138	Eugene, OR	224,049	68.5	4.67
145	Thousand Oaks, CA	210,990	86.2	5.24
152	Victorville–Hesperia–Apple Valley, CA	200,436	124.0	6.28
160	Lorain–Elyria, OH	193,586	87.6	5.28
167	Green Bay, WI	187,316	82.1	5.11
174	Huntington, WV–KY–OH	177,550	106.8	5.83
181	Brownsville, TX	165,776	57.3	4.27
188	Muskegon, MI	154,729	99.6	5.63
195	Deltona, FL	147,713	89.5	5.34
203	Gastonia, NC	141,407	118.8	6.15
210	College Station–Bryan, TX	132,500	49.1	3.96
218	Seaside–Monterey–Marina, CA	125,503	40.6	3.60
225	Vero Beach–Sebastian, FL	120,962	81.5	5.09
234	Tuscaloosa, AL	116,888	76.0	4.92
242	Yakima, WA	112,816	50.4	4.00
249	Harlingen, TX	110,770	59.1	4.34
257	South Lyon–Howell–Brighton, MI	106,139	95.0	5.50
264	Tyler, TX	101,494	57.5	4.28
271	Leesburg–Eustis, FL	97,497	71.0	4.75
278	Jacksonville, NC	95,514	64.1	4.52
285	Grand Junction, CO	92,362	56.0	4.22
292	Vacaville, CA	90,264	25.3	2.84
299	Gainesville, GA	88,680	90.4	5.36
306	Iowa City, IA	85,247	35.8	3.37
313	Charlottesville, VA	81,449	37.5	3.45
320	Terre Haute, IN	79,376	43.2	3.71
329	Johnstown, PA	76,113	42.8	3.69
336	Bay City, MI	74,048	39.8	3.56
343	Temple, TX	71,937	41.3	3.63
350	Hightstown, NJ	69,977	29.9	3.09
358	Avondale, AZ	67,875	29.4	3.06
365	Michigan City, IN–MI	66,199	33.2	3.25
372	Westminster, MD	65,034	53.4	4.12
379	Camarillo, CA	62,798	21.3	2.61
386	Rocky Mount, NC	61,657	40.8	3.60
393	Auburn, AL	60,137	40.0	3.57
400	Bristol, TN–Bristol, VA	58,472	51.1	4.03
407	Madera, CA	58,027	22.6	2.68
414	Blacksburg, VA	57,236	26.3	2.89
421	DeKalb, IL	55,805	17.9	2.39
429	Morristown, TN	54,368	45.4	3.80
436	San Luis Obispo, CA	53,498	14.8	2.17
443	Pittsfield, MA	52,772	33.8	3.28
451	Hazleton, PA	51,746	28.9	3.03
458	Ames, IA	50,726	15.8	2.24
465	Middletown, NY	50,071	28.2	2.99

Table 68: Urban Areas in Locale NOM (Sampled)

Rank	Name	Population	Area (mile^2)	Radius (miles)
467	Midland, MI	49387	30.7	3.13
528	East Stroudsburg, PA	40664	40.6	3.60
586	Hilton Head Island, SC	34400	41.9	3.65
643	Santa Paula, CA	29070	4.2	1.16
700	Seguin, TX	25640	15.6	2.23
760	Newberg, OR	22137	8.5	1.65
819	Galliano, LA	20611	18.2	2.41
876	Corning, NY	18573	7.9	1.59
935	Tamaqua, PA	16915	6.1	1.40
994	Chickasha, OK	15510	9.5	1.74
1052	Coldwater, MI	14293	7.6	1.55
1111	Sylacauga, AL	13291	11.0	1.87
1168	Easton, MD	12503	6.7	1.46
1226	Baraboo, WI	11780	4.7	1.23
1286	Tucson Southeast, AZ	11100	3.4	1.04
1346	Central City–Greenville, KY	10410	9.5	1.74
1403	Lawrenceburg, KY	9899	4.3	1.17
1460	Paris, IL	9376	4.6	1.21
1517	Silsbee, TX	8994	9.6	1.75
1574	Grand Rapids, MN	8525	7.8	1.57
1631	Petoskey, MI	8158	6.5	1.44
1689	Catskill, NY	7812	6.1	1.40
1746	Minot AFB, ND	7489	5.4	1.31
1804	Wilmington (Will County), IL	7107	8.0	1.60
1861	Batesville, MS	6768	4.5	1.20
1918	Cynthiana, KY	6495	3.9	1.12
1975	Escalon, CA	6267	3.8	1.10
2032	Farmville, VA	6029	3.2	1.01
2089	Holdenville, OK	5760	4.9	1.25
2148	Watseka, IL	5590	2.6	0.91
2205	Indiantown, FL	5345	1.4	0.66
2263	Genoa City, WI–IL	5126	4.8	1.24
2320	Chester, IL	4955	3.5	1.06
2378	Mineola, TX	4775	4.6	1.22
2436	Paden City, WV	4618	1.8	0.76
2495	Walnut Grove, GA	4452	6.3	1.41
2554	Perry, MI	4329	3.0	0.98
2613	Fordyce, AR	4182	3.2	1.01
2670	Madisonville, TN	4063	4.0	1.14
2727	Veneta, OR	3946	3.5	1.05
2785	Chincoteague, VA	3819	3.6	1.07
2842	Lyons, NY	3700	3.4	1.03
2901	Aledo, IL	3591	1.4	0.67
2958	Benson, MN	3471	2.6	0.90
3015	Vail, CO	3370	1.9	0.78
3072	Tehama, CA	3261	2.3	0.85
3130	Park Rapids, MN	3175	2.8	0.94
3189	Rogers City, MI	3080	1.8	0.75
3247	Bushnell, FL	3002	2.8	0.95
3305	Neodesha, KS	2913	1.1	0.59
3364	Premont, TX	2837	2.2	0.84
3422	Carson City, MI	2763	3.1	0.99
3481	Cactus, TX	2699	0.6	0.43
3542	Girard, KS	2608	1.5	0.68
3599	Santa Rosa, NM	2540	2.7	0.92

Table 69: Top 20 U.S City Pair Markets (excluding Hawaii)

Rank	City Pair	Distance (mile)	Per Day Pax
1	Boston — New York	183	3,529
2	Los Angeles — New York	2,467	3,496
3	Fort Lauderdale — New York	1,068	3,451
4	New York — Orlando	947	3,319
5*	New York — Washington, DC	217	3,223
6	Chicago — New York	723	3,222
7	Atlanta — New York	756	3,025
8*	Dallas/Fort Worth — Houston	234	2,867
9	New York — San Francisco	2,574	2,643
10*	Los Angeles — Las Vegas	238	2,627
11	New York — West Palm Beach	1,030	2,517
12	Miami — New York	1,097	2,342
13	New York — San Juan	1,603	2,100
14	Los Angeles — San Francisco	341	1,985
15	Los Angeles — Oakland	334	1,914
16	Chicago — Phoenix	1,446	1,894
17	Las Vegas — New York	2,237	1,804
18*	Chicago — Detroit	237	1,743
19	Boston — Washington, DC	400	1,729
20	Los Angeles — Phoenix	366	1,729

Source: http://r2ainc.com/top_us_markets.htm

APPENDIX D

DISTRIBUTED COMPUTING ENVIRONMENT: MIDAS

Contents

The substantial amount of computation is required for execution of ***Mi*** with sufficient level of fidelity. Since single-user/single-platform computing environment cannot afford this, the simulation framework was implemented in a distributed computing environment that was synthesized using the Multi-platform Integrated Development Aid System (MIDAS).

D.1 Background

The nature of the formulated problem of the dissertation is *simulation optimization* (Fu 2001), which requires a large amount of computational resources. The simplest, but not the most economical, way to reduce the execution time to a more manageable level would be to upgrade the existing hardware. An alternate solution is known as *distributed simulation technology* (Fujimoto 2000) that takes advantage of the aggregate computing power with multiple Central Processing Units (CPUs) at each node, connected over a network. It is possible to construct a cluster from standard, inexpensive desktop computers but such an implementation requires expert knowledge of parallel programming and setup. In addition, when the personal computers must process tasks collectively as a cluster, they are no longer useable for the local users.

In order to alleviate this problem, there has been developmental efforts, to use idle processor time on regular desktops in recent years. In such implementations, small clients, sometimes acting as screen savers that include analysis codes for the respective computational tasks, need to be present on distributed desktop computers. These clients connect to a centralized server and download a small amount of data, often called a work unit, to be analyzed. Subsequently, the work unit is executed on the local machine during idle processor time and the result is sent back to the centralized server for further processing. The collective analysis power of potentially millions of computers interconnected through the Internet can result in a formidable supercomputer, but these clients are still limited to their prescribed calculations. There may be a number of bug fixes and enhanced versions of the client available for download, but the task is still the same and new versions have to be updated manually. Furthermore, the computational problems that can be handled are limited to inherently parallel problems that do not require any communications between the clients.

In response to the computing limitations outlined above, the initiative was to develop a user-friendly and generic framework that enables a straightforward initiation of unlimited number of distributed clients on standard office desktop machines without requiring the knowledge of parallel or network programming by aerospace design engineers. The resulting ad-hoc cluster would need to be able to execute any number of computational tasks in an easy-to-use environment without interfering with a local user's needs. Thus, MIDAS is the culmination of the effort to integrate the state-of-the-art in computer science with aerospace systems design applications.

D.2 Development History

Initial research into the matter of distributed clients indicated that a simple client that interfaces with a Network File Server (NFS) or a File Transfer Protocol (FTP) / Hyper Text Transfer Protocol (HTTP) server would be suitable for downloading an analysis code. This solution, however, requires either the use of a shared file system or the definition of protocols and standards for the exchange of executable codes and data. Because such an arrangement needs a number of open shares with full read-and-write access, it is not desirable for security reasons. Moreover, a partial failure of any client or portions of the network will result in a loss of data. Since it was evident that our framework needed to be able to circumvent such issues, Java was chosen as the programming language for MIDAS. Designed as a network language from its conception, the language includes the necessary network Input/Output (I/O) libraries and guarantees platform independence and network security through its built-in implementation of Java Virtual Machine (JVM). Higher level network functionality such as web servers and database servers are either already included in the basic libraries or at the very least have a programming interface that can be utilized. Furthermore, Java includes a functionality called dynamic class loading that allows a JVM to load the code and data during the execution. A security model, which grants access to the dynamically loaded classes, especially in terms of disk or network I/O, ensures that it is not possible for users to inadvertently or maliciously access, delete, or send unauthorized data across the network. This feature, combined with the Remote Method Invocation (RMI) that allows a remote calling of a code, enables the implementation of a master-and-worker type architecture, in which each worker can be triggered to remotely download a piece of code and data from the master, execute it and return the results to the master machine. Nevertheless, the basic RMI lacks a comprehensive security model and does not allow any control over class unloading, which can cause version conflicts when changing task codes.

D.3 Architecture

Further research into this area led to the discovery of an enabling technology called JavaSpaces. It is part of a Java networking package called Jini, which is a basis for enterprise grid computing as advocated by Sun. JavaSpaces is an implementation of a persistent *object* store called a *space*. This space provides a logically shared memory where data can be stored, accessed, and updated in real-time without requiring a physically shared memory. The default implementation of a space exists across a local area network (LAN), and thus it is possible for multiple computational jobs to simultaneously share the same workers without any conflicts. Furthermore, Jini includes several other valuable features. One of the most convenient features is a look-up service in conjunction with a discovery manager. The use of these features allows the self-discovery of different components in a distributed system. It implies that no previous knowledge of computer names or addresses is required in configuring, setting up, utilizing a distributed computing environment. Another essential feature is the built-in transaction manager. It enables the book-keeping of activities across the network and, should any part of an operation fail, the restoration to a state before the failed operation started. While this does not recover the lost time, it does prevent the loss of data. Such functions supported by the transaction manager provide inherent "fault-tolerance" by persevering the Atomicity, Consistency, Isolation, and Durability (ACID) of the distributed environment. These are very important attributes for a distributed architecture that exists across a shared network because they guarantee that nothing is lost due to potential machine, network, or program failures. As a result, the application of JavaSpaces provides a flexible, "scalable" (E. Freeman & Arnold 1999), and reliable framework for creating a system or even systems of distributing applications over a network of master and worker machines. MIDAS adopts a type of master-and-worker style of distributed processing in the context of JavaSpaces technology as shown in Figure 115.

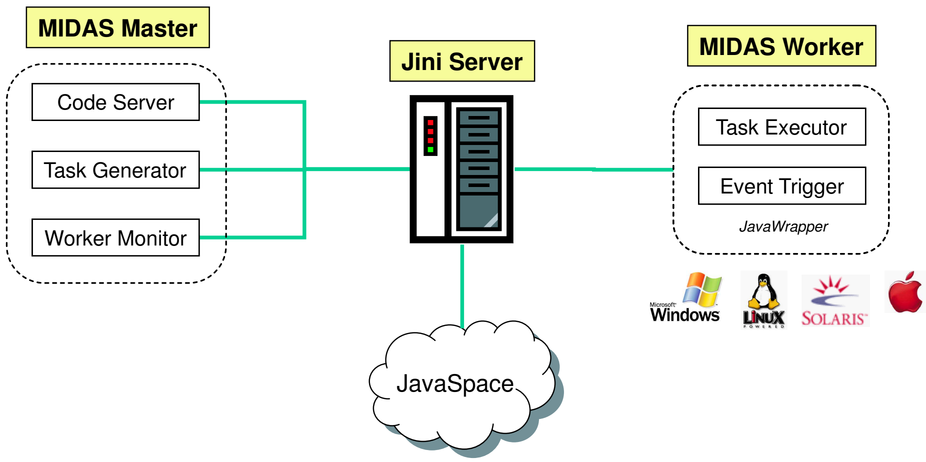


Figure 115: System Level Architecture of MIDAS

Shown here are the workers, generic computation engines, which can search for, receive and execute any tasks placed in the space by the Jini server. Each task is implemented as an "entry" which is a generic object that the JavaSpaces Application Programming Interface (API) uses to store the object in the space. In this context, each task is an object that contains the data as well as the necessary reference executable code to complete the computational job. Through the use of a simple 'demon' or service that runs in the background, the workers wait for entries into the space that match the template of a task. Once a task is found, it is removed from the space and a local copy is created on the worker machine. Subsequently, the worker dynamically loads the remote code into its native JVM and executes the task. Once a worker completes the task, it writes the results back to the space as another instantiation of an entry. Each result does not contain any executable code, but contains any desired number of data objects storing the result values, which include timing information.

The implementation of such a distributed simulation scheme in a space is inherently self-load balancing, because any idle worker would continue to take the tasks and return the results until the space is empty of new tasks. Therefore, there is no need for a special assignment mechanism that actively distributes tasks among workers. The master side is responsible for creating new tasks and collecting the results. This part of the infrastructure consists of modules that relate to user interactions, resource and worker management, and task construction. As long as all the machines on the network have the compatible versions of JVM, such an infrastructure is able to distribute the tasks even amongst a collection of heterogeneous, multi-platform workers. Although only one master machine was used for this research, the JavaSpaces technology allows the interaction between multiple master computers as well as multiple worker computers.

D.4 Implementation

In the initial development phase, the management of the space relied on a commercial product, but the current implementation of the space and other network services is supported by Sun's libraries without the use of proprietary management products. Currently, a designated computer is used as the dedicated provider of the network services, which is operated independently of the master. This has the advantage that all necessary network services will be available continuously and not be dependent on a specific user's computer.

On the master side, three basic modules are implemented. The worker monitor shown in Figure 116 obtains information about the connected workers. The code server provides analysis codes to the space. The task controller generates the tasks, and then collect all results received from each worker. This basic functionality can be further enhanced using a generalized graphical user interface (GUI) which lets the user easily control the system status, submit tasks and avoid the inconvenience of having to change the entire underlying code every time tasks are to be distributed. It is envisioned to create an additional element of the GUI that contains the programming interface information and allows the creation of a code in a simple text editor. Subsequently, a user can easily compile his or her written code by using the GUI on the master machine to make the code immediately available for task execution.

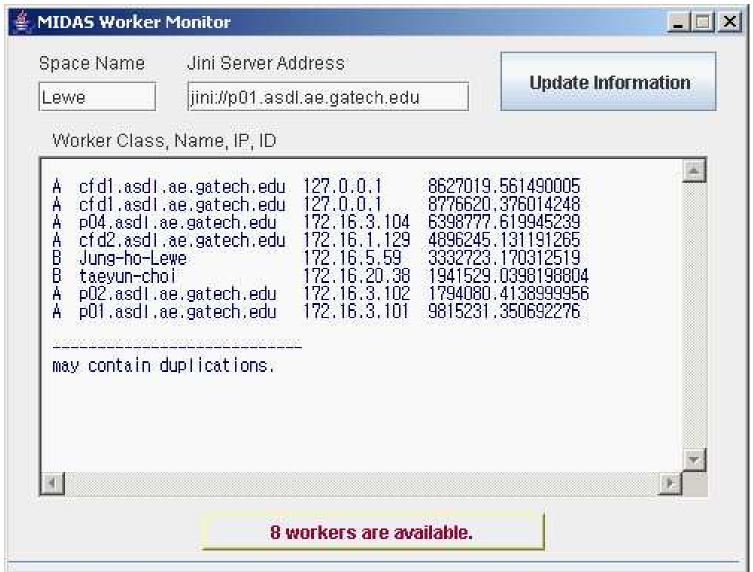


Figure 116: MIDAS Worker Monitor

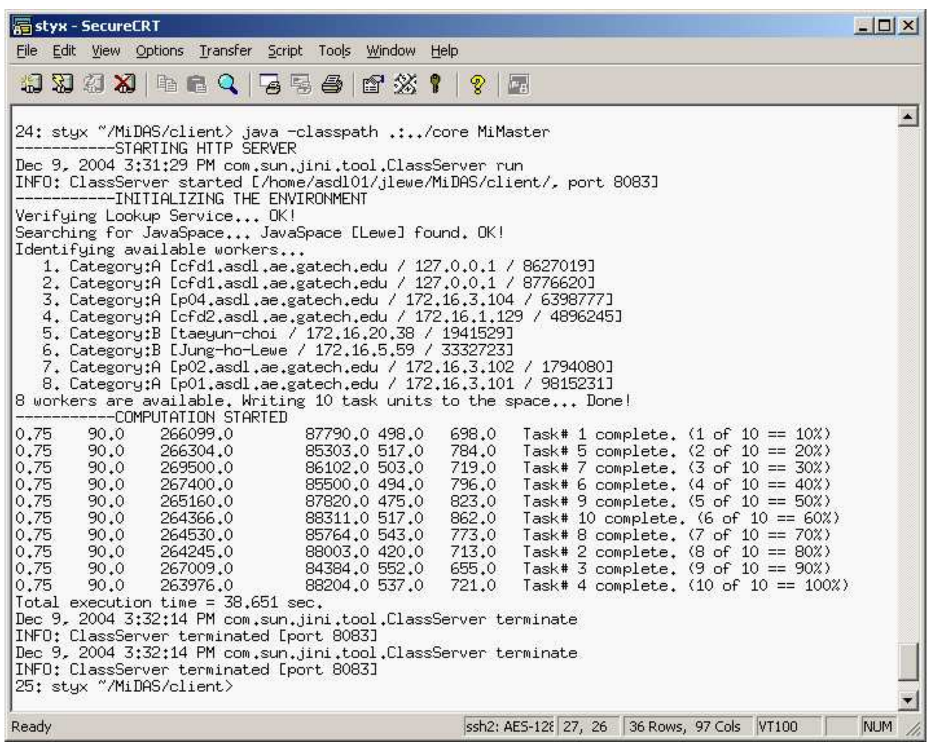
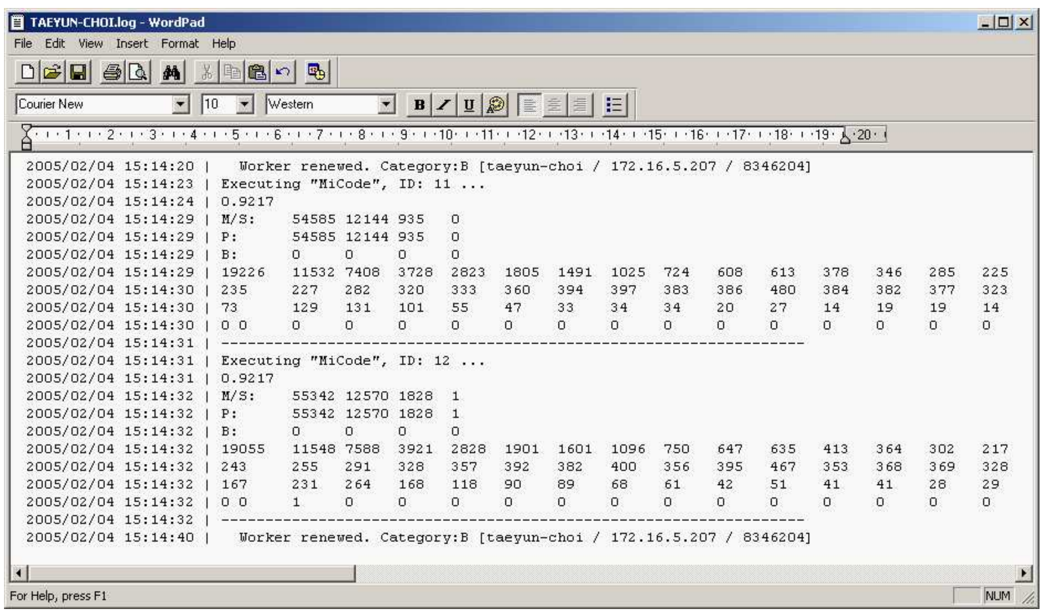


Figure 117: Execution of MIDAS Master

The initial implementation of MIDAS involved the creation of generalized workers that would simply exist as a Java Archive (JAR) containing all the necessary libraries and policy files. This archive acts as the stand-alone executable which can be run using a text console. This simple setup, however, was found to be impractical and problematic for many reasons, and has been improved by employing a non-commercial Java application called the Java Service Wrapper.[1] This powerful application makes it possible to install the worker as a Windows service or daemon processes on UNIX systems, as well as to record a comprehensive log file based on user's needs. After the generic worker connects to the space with the use of the look-up discovery service, it outputs a simple status message about itself, waits for any available tasks, and perform the task as shown in Figure 118.




```
2005/02/04 15:14:20 |   Worker renewed. Category:B [taeyun-choi / 172.16.5.207 / 8346204]
2005/02/04 15:14:23 | Executing "MiCode", ID: 11 ...
2005/02/04 15:14:24 | 0.9217
2005/02/04 15:14:29 | M/S:    54585 12144 935   0
2005/02/04 15:14:29 | P:      54585 12144 935   0
2005/02/04 15:14:29 | B:      0     0     0     0
2005/02/04 15:14:29 | 19226   11532 7408  3728  2823  1805  1491  1025  724   608   613   378   346   285   225
2005/02/04 15:14:30 | 235     227   282   320   333   360   394   397   383   386   480   384   382   377   323
2005/02/04 15:14:30 | 73      129   131   101   55    47    33    34    34    20    27    14    19    19    14
2005/02/04 15:14:30 | 0 0     0     0     0     0     0     0     0     0     0     0     0     0     0     0
2005/02/04 15:14:31 | -------------------------------------------------------------------
2005/02/04 15:14:31 | Executing "MiCode", ID: 12 ...
2005/02/04 15:14:31 | 0.9217
2005/02/04 15:14:32 | M/S:    55342 12570 1828  1
2005/02/04 15:14:32 | P:      55342 12570 1828  1
2005/02/04 15:14:32 | B:      0     0     0     0
2005/02/04 15:14:32 | 19055   11548 7588  3921  2828  1901  1601  1096  750   647   635   413   364   302   217
2005/02/04 15:14:32 | 243     255   291   328   357   392   382   400   356   395   467   353   368   369   328
2005/02/04 15:14:32 | 167     231   264   168   118   90    89    68    61    42    51    41    41    28    29
2005/02/04 15:14:32 | 0 0     1     0     0     0     0     0     0     0     0     0     0     0     0     0
2005/02/04 15:14:32 | -------------------------------------------------------------------
2005/02/04 15:14:40 |   Worker renewed. Category:B [taeyun-choi / 172.16.5.207 / 8346204]
```

Figure 118: MIDAS Worker Log File

D.5 Performance Tests

In order to assess the computational performance and efficiency of our home-grown distributed computing environment created by MIDAS, a number of distributed simulation cases were executed on different clusters over the LAN. Table 70 lists the different type of

[1] Visit http://wrapper.tanukisoftware.org for detailed information.

machines, with varying operating systems, computational capabilities, and JVM versions, which were used for these experiments.

Table 70: Specifications of Machines Used for Performance Test

ID	CPU	Clock Speed	RAM	OS	JVM Version	PI
B_1	Intel Pentium IV	1.6GHz	1GB	Linux	1.4.2	1
B_2	Intel Pentium IV	1.6GHz	1GB	Linux	1.4.2	1
B_3	Intel Pentium IV	1.6GHz	1GB	Linux	1.4.2	1
B_4	Intel Pentium IV	1.6GHz	1GB	Linux	1.4.2	1
$P1$	Intel Pentium III	866MHz	256MB	Windows 2000	1.4.0	1.09
$P2$	Intel Pentium IV	2.0GHz	512MB	Windows 2000	1.4.2	1.59
$P3$	Intel Pentium IV	3.2GHz	1GB	Linux	1.4.2	2.26
$G4_1$	Power PC G4	866MHz	1GB	Mac OS	1.4.2	1.26
$G4_2$	Power PC G4	866MHz	1GB	Mac OS	1.4.2	1.26
$G5$	Power PC G5	1.6GHz	512MB	Mac OS	1.4.2	2.3

Since the performance between the different collections of machines must be equitably compared, a normalized performance index (PI) was introduced. It is defined in such a way that a baseline PI of 1 indicates the computational capability of a single Pentium IV Linux machine (B_i's in Table 70). Furthermore, one computational load (1 λ) is defined as a task that takes 96.4 seconds (1 τ) to complete on the baseline machine. First, it was attempted to measure the relationship between the reduction in the execution time and the number of employed machines. For this purpose, three different tasks were executed over different mixes of B_1 through B_4. Generally, the computational performance of a distributed simulation is measured by the speed-up, which is given as the ratio of the execution time on a single processor to that of n number of processors, i.e., $SU = T(1)/T(n)$. The ideal case is when the speed-up is identical to the number of machines n. In practice, however, the speed-up will be less due to communication bandwidth[2] and latencies[3]. As expected, the result in Figure 119 confirms that the speed-up is linearly proportional to the number

[2]The amount of data that can be sent from one computer to another through a particular connection in a certain amount of time.

[3]The delay in transmitting a message from one computer to another.

of machines. The illustration also shows how the speed-up deteriorates as the number of CPUs increases.

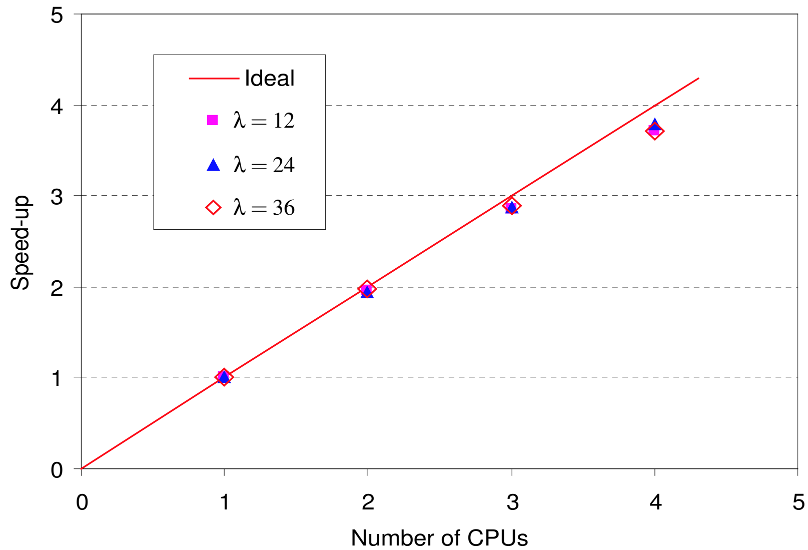


Figure 119: Linear Speed-up vs. Number of CPUs

Second, the computational efficiency of executing distributed simulation over a network of heterogeneous machines was assessed. The purpose of this experiment was to test the effectiveness of MIDAS in creating an efficient multi-platform computing environment. Six such clusters of machines were created, as listed in Table 71.

Table 71: Clusters of Heterogeneous Machines

Cluster ID	Loads (λ)	# of CPUs	B_i	$P1$	$P2$	$P3$	$G4_i$	$G5$
A1	36	6	✓	✓	✓	✓	✓	✓
A2	72	6	✓	✓	✓	✓	✓	✓
B1	36	8	✓ ✓ ✓	✓		✓	✓ ✓	✓
B2	72	8	✓ ✓ ✓	✓		✓	✓ ✓	✓
C1	36	10	✓ ✓ ✓ ✓	✓	✓	✓	✓ ✓	✓
C2	72	10	✓ ✓ ✓ ✓	✓	✓	✓	✓ ✓	✓

As a result, it is possible to compare the computation efficiencies of these six clusters, as shown in Figure 120. The ideal time for each cluster to finish computing a task load (λ) was obtained by dividing the given λ value by its respective aggregate PI value. The solid line in the figure represents such an ideal scenario when all tasks were evenly distributed to each machine within a cluster. Each cluster's actual τ under various loads was then compared against their ideal τ.

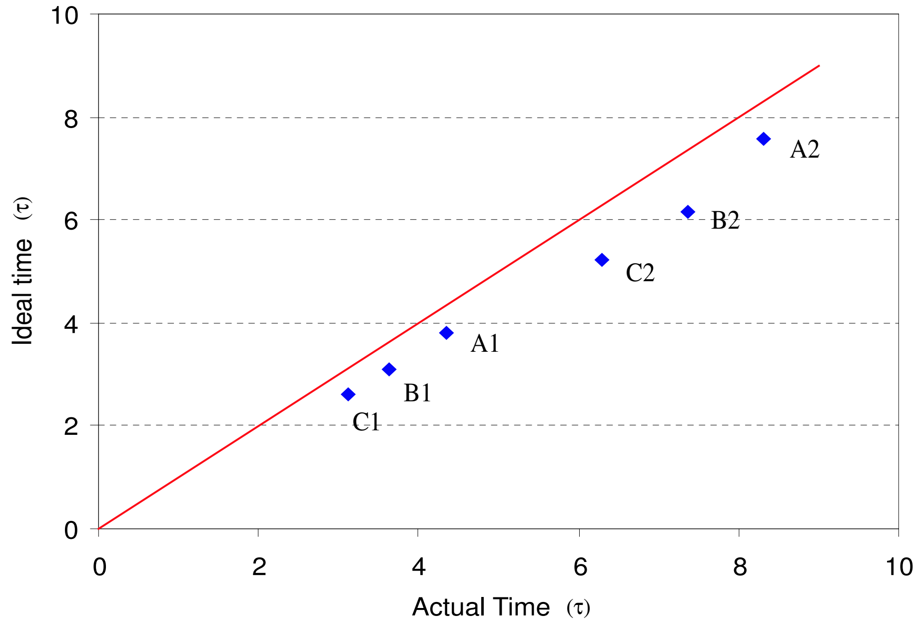


Figure 120: Computational Efficiencies of Various Clusters

As opposed to the first experiment, conducted in identical-platform environments, the degradation in performance is an inevitable price for synthesizing clusters composed of heterogeneous machines. These losses, which occur from different machines completing the same unit-load task at different rates, can be quantified by measuring the horizontal distance between the ideal line and the marked points from Figure 120.

REFERENCES

Adamsen, P. B. (2000), *A Framework for Complex System Development*, CRC Press, Boca Raton, FL.

AIA (2001), 'United states aviation system capacity cannot accommodate future civil aviation demand', Aerospace Industries Association (AIA), as of March 13, 2001, from `http://www.aia-aerospace.org/issues/commission`.

AIAA (1998), *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, American Institute of Aeronautics and Astronautics (AIAA), Reston, VA.

Altshuller, G. (1984), *Creativity As an Exact Science : The Theory of the Solution of Inventive Problems*, Gordon and Breach Science Pub., New York.

Axelrod, R. (1997), *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press, Princeton, NJ.

Bak, P. (1996), *How Nature Works : The Science of Self-organized Criticality*, Copernicus, New York.

Baker, A. P. & Mavris, D. N. (2001), Assessing the simultaneous impact of requirements, vehicle characteristics, and technologies during aircraft design, *in* '39^{th} Aerospace Sciences Meeting and Exhibit (Reno, NV)'.

Barabási, A.-L. (2002), *Linked: The New Science of Networks*, Perseus Pub., Cambridge, MA.

Basu, N., Pryor, R., Quint, T. & Arnold, T. (1996), Aspen: A microsimulation model of the economy, Technical Report SAND-96-2459, Sandia National Laboratories, Albuquerque, NM.

Bellman, K. L. & Landauer, C. (2000), 'Towards an integration science: The influence of richard bellman on our research', *Journal of Mathematical Analysis and Applications* **249**, 3–31.

Belton, V. & Gear, T. (1983), 'On a short-coming of Saaty's method of analytic hierarchies', *Omega* **11**(3), 228–230.

Benayoun, R., Roy, B. & Sussman, N. (1966), 'Manual de reference du programme ELECTRE, note de synthese et formation', No. 25, Direction Scientifique SEMA, Paris, French.

Benenson, I. & Omer, I. (2001), 'Agent-based modeling of residential distribution', Department of Geography and Human Environment, University Tel Aviv, Israel.

Bierlaire, M. (1996), Mathematical Models for Transportation Demand Analysis, PhD thesis, Facultés Universitaires Notre-Dame de la Paix, Namur, Belgium.

Boarnet, M. G. & Chalermpong, S. (2000), New highways, urban development and induced travel, Technical Report UCI-ITS-WP-00-15, Institute of Transportation Studies, Univ. of California, Irvine.

Bonabeau, E. (2002), 'Agent-based modeling: Methods and techniques for simulating human systems', *PNAS* **99**(Suppl. 3), 7280–7287.

Bowers, P. M. (1990), *Unconventional Aircraft*, 2nd edn, Tab Books, Blue Ridge Summit, PA.

BTS (1999*a*), 'The 1995 american travel survey: Microdata file', Data CD-ROM, Bureau of Transportation Statistics.

BTS (1999*b*), *Transportation Statistics Annual Report 1999*, BTS99-03, Bureau of Transportation Statistics, Washington, D.C.

BTS (2000), *The Changing Face of Transportation*, BTS00-007, Bureau of Transportation Statistics, Washington, DC.

BTS (2001), *Airport Activity Statistics of Certified Air Carriers*, BTS01-05, Bureau of Transportation Statistics, Washington, D.C.

BTS (2002*a*), *National Transportation Statistics 2002*, BTS02-08, Bureau of Transportation Statistics, Washington, D.C.

BTS (2002*b*), *Transportation Statistics Annual Report 2001*, BTS02-07, Bureau of Transportation Statistics, Washington, D.C.

Carley, K. M. (1996), Validating computational models, Working paper, Center for Computational Analysis of Social and Organizational Systems, Carnegie Mellon Univ., from http://www.casos.cs.cmu.edu/.

Carmichael, D. G. (1981), *Structural modelling and optimization: a general methodology for engineering and control*, Ellis Horwood Limited, Chichester, England.

Carson, J. & Doyle, J. (2002), 'Complexity and robustness', *Proceedings of the National Academy of Sciences* **99**(Suppl. 1), 2538–2545.

Cederman, L. E. (1997), *Emergent Actors in World Politics*, Prinston Press.

Census Bureau (2003), *Statistical Abstract of the United States: 2003 (123rd edition)*, U.S. Census Bureau, Washington, D.C.

Chakrabarti, R. (2002), An agent based model of corruption, Working paper, Georgia Institute of Technology, as of Feb. 26, 2002, from `http://www.prism.gatech.edu/~rc166/`.

Chankong, V. & Haimes, Y. (1983), *Multiobjective Decision Making: Theory And Methodology*, North-Holland, New York.

Chen, S.-J. & Hwang, C.-L. (1992), *Fuzzy multiple attribute decision making : methods and applications*, Vol. 375 of *Lecture notes in economics and mathematical systems*, Spinger-Verlag, New York.

Clark, C. E. (1961), 'The greatest of a finite set of random variables', *Operations Research* **9**(2), 145–162.

DeLaurentis, D. A., Lim, C. & Kang, T. (2004), 'Approach to characterize mobility through modeling of solution spaces for conceptual aircraft', *Journal of Aircraft* **41**(1), 73–84.

Deutschman, D. H. et al. (1997), 'Scaling from trees to forests: Analysis of a complex simulation model', *Science* **277**, 1688, also available at `http://www.sciencemag.org/feature/data/deutschman/index.htm`.

DOT (2000), Transportation decision making: Policy architecture for the 21st century, Research paper, Department of Transportation, from `http://ostpxweb.dot.gov/policy/Data/PolArch.pdf`.

Downen, T. D. & Hansman, R. J. (2003), 'Identification and analysis of key barriers to the utility of general aviation', *Journal of Aircraft* **40**(2), 232–238.

Downs, A. (1962), 'The law of peak-hour expressway congestion', *Traffic Quaterly* **16**, 393–409.

Drake, H. M., Kenyon, G. C. & Galloway, T. L. (1969), Mission analysis for general aviation in the 1970's, *in* 'AIAA Aircraft Design and Ops Meeting (Los Angeles)'.

Durham, M. H. & Creedon, J. F. (2001), Small aircraft transportation system program, Program plan, NASA LaRC.

E. Freeman, S. H. & Arnold, K. (1999), *JavaSpaces Principles, Patterns, and Practice*, Addison-Wesley, Reading, MA.

Eagan, R. J. et al. (2001), *Theoretical Foundations for Decision Making in Engineering Design*, National Academy Press, Washington, D.C.

Ehrgott, M. (2000), *Multicriteria Optimization*, Vol. 491 of *Lecture notes in economics and mathematical systems*, Springer, Berlin, Germany.

Epstein, J. M. & Axtell, R. (1996), *Growing artifcial societies: Social science from the bottom up*, MIT Press.

FAA (2001), *Aviation Capacity Enhancement Plan 2001*, Office of System Capacity Annual Report, Federal Aviation Administration (FAA).

FAA (2002), National plan of integrated airport systems 2001-2005, Report to congress, Federal Aviation Administration (FAA).

Fallows, J. (2001), ‘Freedom of the skies’, *The Atlantic Monthly* **287**(6), 37–49.

Fishburn, P. & Lavalle, I. (1999), ‘MCDA: Theory, practice and the future’, *Journal of Multi-Criteria Decision Analysis* **8**(1), 1–2.

Fishwick, P. A., Sanderson, J. G. & Wolff, W. F. (1998), ‘A multimodeling basis for across-trophic-level ecosystem modeling: The florida everglades example’, *Transactions of the Society for Computer Simulation* **15**(2), 76–89.

Fonseca, C. M. & Fleming, P. J. (1993), Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization, *in* ‘Genetic Algorithms: Proceedings of the Fifth International Conference’, Morgan Kaufmann, pp. 416–423.

Fu, M. C. (2001), Simulation optimization, *in* ‘2001 Winter Simulation Conference’, pp. 53–61.

Fujimoto, R. (2000), *Parallel and Distributed Simulation Systems*, John Wiley & Sons, New York.

GAMA (2004), ‘General aviation statistical databook 2003’, General Aviation Manufacturers Association, as of April 10, 2004, from `http://www.gama.aero/resources/statisticsl`.

Goldberg, D. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA.

Goldstone, J. A. (1991), *Revolution and Rebellion in the Early Modern World*, University of California Press, Berkeley, CA.

Grosholz, E. (2001), *Cartesian Method and the Problem of Reduction*, Oxford Univ. Press, New York.

Hall, D. (2001), Personal air vehicles, Technical report, David Hall Consulting, Morro Bay, CA.

Hartley, D. S. (1997), Verification & validation in militrary simulations, *in* S. Andradóttir et al., eds., ‘Proceedings of the 1997 Winter Simulation Conference’, pp. 925–932.

Hartvigsen, G. & Levin, S. (1997), ‘Evolution and spatial structure interact to influence plant-herbivore population and community dynamics’, *Proceedings of the Royal Society of London. Series B, Biological sciences* **264**, 1677–1685.

Hecker, J. Z. (2003), ‘Commercial aviation: Issues regarding federal assistance for enhancing air service to small communities’, Testimony Before the Subcommittee on Aviation, Committee on Commerce, Science and Transportation, U.S. Senate.

Helbing, D., Farkas, I. & Vicsek, T. (2000), ‘Simulating dynamical features of escaping panic’, *Nature* **407**, 487–490.

Henrickson, L. & McKelvey, B. (2002), 'Foundations of "new" social science: Institutional legitimacy from philosophy, complexity science, postmodernism, and agent-based modeling', *PNAS* **99**(Suppl. 3), 7288–7295.

Heudin, J.-C., ed. (1998), *Virtual Worlds: First International Conference, VW '98, Paris, France, 2000*, Vol. 1434 of *Lecture notes in computer science*, Springer, New York.

Hitchins, D. K. (2003), *Advanced Systems Thinking, Engineering, and Management*, Artech House, Boston.

Ho, Y.-C., Sreenivas, R. & Vakili, P. (1992), 'Ordinal optimization of discrete event dynamic systems', *Journal of Discrete Event Dynamic Systems* **2**(2), 61–88.

Hoffer, S. et al. (1998), Economic values for evaluation of federal aviation administration investment and regulatory decisions, Technical Report FAA-APO-98-8, Federal Aviation Administration.

Holland, J. H. (1995), *Hidden Order: How Adaptation builds Complexity*, Addison-Wisely, Reading, MA.

Hood, L. (2002), 'Agent based modelling', as of April 18, 2002, from `http://www.brs.gov.au.social_sciences/kyoto/hood2.html`.

Hwang, C.-L. & Masud, A. (1979), *Multiple Objective Decision Making - Methods and Applications*, Vol. 164 of *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, New York.

Hwang, C.-L. & Yoon, K. (1981), *Multiple Attribute Decision Making: Methods and Applications, A State-of-the-Art Survey*, Vol. 186 of *Lecture Notes in Economics and Mathematical Systems*, Spinger-Verlag, New York.

Ilachinski, A. (1997), Irreducible semi-autonomous adaptive combat (ISSAC): An artificial-life approach to land warfare, Technical Report CRM 97-61.10, Center for Naval Analysis.

Ingalls, R. G. (2002), Introduction to simulation, *in* E. Yücesan et al., eds., 'Proceedings of the 2002 Winter Simulation Conference', pp. 7–16.

Johnson, M., McJeon, M. & Szanto, T. (1998), *Simulation Based Acquisition: A New Approach*, Defense Systems Management College Press, Fort Velvoir, VA.

Jones, W. B., Cassady, C. R. & Bowden, R. O. (1999), Developing a standard definition of intermodal transportation, NCIT report, National Center for Intermodal Transportation, Available from `http://www.ie.msstate.edu/ncit/RESEARCH.html`.

Kanafani, A. (1983), *Transportation Demand Analysis*, McGraw-Hill, New York.

Kang, T. et al. (2003), A system dynamics model of the development cycle for future mobility vehicles, *in* '21st System Dynamics Conference (New York, NY)'.

Keating, C. et al. (2003), 'System of systems engineering', *Engineering Management Journal* **15**(3), 36–45.

Koppelman, F. & Wen, C. (2000), 'The paired combination logit model: Properties, estimation and application', *Transportation Research Part B* **34**, 75–89.

Kotonya, G. & Somerville, I. (1998), *Requirement Engineering: Process and Techniques*, John Wiley & Sons, West Sussex, England.

Kruesi, F. E. (1997), 'Departmental guidance for the valuation of travel time in economic analysis', U.S. DoT Memorendum, dated April 9, 1997.

Labbé, M. et al., eds. (1998), *Operations Research and Decision Aid Methodologies in Traffic and Transportation Management*, Vol. 166 of *NATO ASI Series F: Computer and Systems Sciences*, Springer-Verlag, Berlin, Germany.

Lane, D. A. (1993), 'Artificial worlds and economics, part i', *Journal of Evolutionary Economics* **3**(2), 89–107.

Law, A. M. & Kelton, W. D. (2000), *Simulation Modeling and Analysis*, 3 edn, McGraw-Hill, Boston, MA.

Lee, D. B., Klein, L. A. & Camus, G. (1999), 'Induced traffic and induced demand', *Transportation Research Record* **1659**, 68–75.

Lempert, R. J., Popper, S. W. & Bankes, S. C. (2003), *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-term Policy Analysis*, RAND Corporation.

Lewe, J.-H. et al. (2002), An integrated decision-making method to identify design requirements through agent-based simulation for personal air vehicle systems, *in* 'ATIO 20002 Conference (Los Angeles, CA)', AIAA.

Lindsey, C. G. (1998), Aviation weather study, Technical Report NWRA-CR-98-R185, NorthWest Research Associates, Inc., Bellevue, WA.

Litman, T. (2001), 'Generated traffic: Implications for transport planning', *ITE Journal* **71**(4), 38–47.

Lootsma, F. (1999), 'The expeced future of MCDA', *Journal of Multi-Criteria Decision Analysis* **8**(2), 59–60.

Lootsma, F., Mensch, T. & Vos, F. (1990), 'Multi-criteria analysis and budget reallocation in long-term research planning', *European Journal of Operational Research* **47**, 293–305.

Lustick, I. S. (2000), 'Agent-based modelling of collective identity: Testing constructivist theory', *Journal of Artificial Societies and Social Simulation* **3**(1), [online].

Manski, C. E. (2001), 'Daniel McFadden and the econometric analysis of discrete choice', *Scand. J. of Economics* **103**, 217–229.

Mavris, D. N., Bandte, O. & Brewer, J. T. (1995), A method for the identification and assessment of critical technologies needed for an economically viable hsct, *in* '1^{st} AIAA Aircraft Engineering, Technology, and Operations Congress (Los Angeles, CA)'.

Mavris, D. N. & DeLaurentis, D. A. (1995), An integrated approach to military aircraft selection and concept evaluation, *in* '1^{st} AIAA Aircraft Engineering, Technology, and Operations Congress (Los Angeles, CA)'.

Mavris, D. N. & DeLaurentis, D. A. (2000), Methodology for examining the simultaneous impact of requirements, vehicle characteristics, and technologies on military aircraft design, *in* '22^{nd} Congress of the International Council on the Aeronautical Sciences (Harrogate, England)'.

Mavris, D. N. & DeLaurentis, D. A. (2002), Personal air vehicle exploration tool and modeling, Final project report, Georgia Institute of Technology, Atlanta, GA.

Mavris, D. N. & Hayden, W. T. (1996), Formulation of an ippd methodology for the design of a supersonic business jet, *in* '1^{st} World Aviation Congress (Los Angeles, CA)'.

McCarthy, P. S. (2001), *Transportation Economics: Theory and Practice: A Case Study Approach*, Blackwell pub., Malden, MA.

McGill, J. I. & Van Ryzin, G. J. (1999), 'Revenue management: Research overview and prospects', *Transportation Science* **33**(2), 233–256.

McMullin, B. & Varela, F. J. (1997), Rediscovering computational autopoiesis, *in* P. Husbands & I. Harvey, eds., 'Fourth European Conference on Artificial Life (Brington, UK)', MIT Press, Cambridge, MA.

McPhail, C. (1997), Stereotypes of crowds and collective behavior: Looking backward, looking forward, *in* D. E. Miller, M. A. Katovich & S. L. Saxton, eds., 'Constructing Complexity: Symbolic Interaction and Social Forms (Suppl. 3)', JAI Press Inc., Greenwich, CT.

Merelo, J. J., Prieto, A. & Rivas, V. (1997), An agent based model for the study of publicity/consumer dynamics, *in* P. Husbands & I. Harvey, eds., 'Fourth European Conference on Artificial Life (Brington, UK)', MIT Press, Cambridge, MA.

Meyn, L. A. et al. (2004), Preliminary assessment of future operational concepts using the airspace concept evaluation system, *in* '4th AIAA Aviation Technology, Integration, and Operations Forum (Chicago, Illinois)'. AIAA-2004-6508.

Miettinen, K. (1999), *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Boston, MA.

Miller, D. W. & Starr, M. K. (1969), *Executive decisions and operations research*, 2 edn, Prentice-hall, Englewood Cliffs, NJ.

NASA LaRC (2001), 'Small aircraft transportation system (SATS): Wings for all of america's communities', Trifold Broshore, NASA Langley Research Center (NASA LaRC), Hampton, VA.

NASA LaRC (2002*a*), 'Personal air vehicle exploration', NASA Langley Research Center (NASA LaRC), as of May 4, 2002, from `http://rasc.larc.nasa.gov/rasc_new/rasc_fy01_top/PAVE_Top_page.htm`.

NASA LaRC (2002*b*), 'Revolutionary aerospace systems concepts program', NASA Langley Research Center (NASA LaRC), as of Nov 12, 2002, from `http://rasc.larc.nasa.gov/rasc_new/RASC.html`.

Niedringhaus, W. P. (2004), 'The Jet:Wise model of national airspace system evolution', *Simulation* **80**(1), 45–58.

Noland, R. B. (2001), 'Relationship between highway capacity and induced vehicle travel', *Transprotation Research Part A* **35**, 47–72.

Noland, R. B. & Lem, L. L. (2002), 'A review of the evidence for induced travel and changes in transportation and environmental policy in the us and the uk', *Transprotation Research Part D* **7**, 1–26.

Odoni, A. R. et al. (1997), Existing and required modeling capabilityes for evaluating atm systems and concepts, Final Report for NASA's AATT Program under Grant No. NAG2-997, International Center for Air Transportation, MIT, Cambridge, MA.

Pritchett, A. R. et al. (2003), *Applied system simulation : methodologies and applications*, Kluwer Academic Publishers, Boston, chapter 16, pp. 368–379.

Raberto, M. et al. (2001), 'Agent-based simulation of a financial market', *Physica A: Statistical Mechanics and its Applications* **299**(1-2), 319–327.

Reed, W. J. (2003), 'The pareto law of incomes—an explanation and an extension', *Physica A: Statistical Mechanics and its Applications* **319**, 469–486.

Reynolds, C. W. (1987), 'Flocks, herds, and schools: A distributed behavioral model', *Computer Graphics* **21**(4), 25–34.

Richards, J. (1959), 'A flexible growth function for empirical use', *Journal of Experimental Botany* **10**, 290–300.

Rouse, W. B. (2003), 'Engineering complex systems: Implications for research in systems engineering', *IEEE Transactions on Sytstems, Man, and Cybernetics–Part C* **33**(2), 154–156.

Rumbaugh et al. (1991), *Objected-Oriented Modeling and Design*, Prentice Hall, Englewood Cliffs.

Saaty, T. L. (1980), *The Analytic Hierarchy Process*, McGraw-Hill International, New York.

Saaty, T. L. (1990), 'An exposition of the AHP in reply to the paper 'Remarks on the Analytic Hierarchy Process", *Management Science* **36**(3), 259–268.

SACTRA (1994), Trunk roads and the generation of traffic, Technical report, Standing Advisory Committee on Trunk Road Assessment (SACTRA), Department for Transport, London, U.K.

Sargent, R. G. (2000), Verification, validation, and accreditation of simulation models, *in* J. Jones et al., eds., 'Proceedings of the 2000 Winter Simulation Conference', pp. 50–59.

Sargent, R. G. (2001), Some approaches and paradigms for verifying and validating simulation models, *in* B. Peters et al., eds., 'Proceedings of the 2001 Winter Simulation Conference', pp. 106–114.

Scheffer, M. et al. (1995), 'Super-individuals: a simple solution for modelling large populations on an individual basis', *Ecological Modelling* **80**(2-3), 161–170.

Simon, H. A. (1957), *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in a Social Setting*, Wiley, New York.

Srbljinovc, A. & Skunca, O. (2003), 'An introduction to agent based modeling and simulation of social sciences', *Interdisciplinary Descrption of Complex Systems* **1**, 1–8.

Sterman, J. (2000), *Business Dynamics: Systems Thinking and Modeling for a Complex World*, Irwin McGraw-Hill, Boston, MA.

Stettner, M. & Schrage, D. P. (1993), Tiltrotor performance sensitivities for multidisciplinary wing optimization, *in* 'AHS International Specialists' Meeting on Rotorcraft Multidisciplinary Design Optimization (Atlanta, GA)'.

Stiles, P., ed. (1994), *Roadable Aircraft: From Wheels to Wings*, 2nd edn, Custom Creativity, Inc., Melbourne, FL.

Sumpter, D. J. T. (2000), From Bee to Society: An Agent-based Investigation of Honey Bee Colonies, PhD thesis, University of Manchester, Manchester, United Kingdom.

Tesfatsion, L. (2002), 'Agent-based computational economics: Growing economies from the bottom-up', *Artificial Life* **8**(1), 55–82.

Thomke, S. H. (1998), 'Simulation, learning and r&d performance: Evidence from automotive development', *Research Policy* **27**(1), 55–74.

Thurston, L. (1927), 'A law of comparative judgement', *Psychological Review* **34**, 273–286.

Trani, A. A. et al. (2003), 'Integrated model for studying small aircraft transportation system', *Transportation Research Record* **1850**, 1–10.

TRB (1995), *Expanding metropolitan highways: implications for air quality and energy use*, Special Report 245, Transportation Research Board (TRB), National Research Council, National Academy Press, Washington, D.C.

TRB (2002), *Future flight : a review of the small aircraft transportation system concept*, Special Report 263, Transportation Research Board (TRB), National Research Council, National Academy Press, Washington, D.C.

Triantaphyllou, E. (2000), *Multi-Criteria Decision Making Methods: A Comparative Study*, Vol. 44 of *Applied optimization*, Kluwer Academic Publishers, Boston, MA.

Triantaphyllou, E. & Mann, S. H. (1995), 'An evaluation of the AHP and the revised AHP when the eigenvalue method is used under a continuity assumption', *Computers and Industrial Engineering* **2**(1), 35–44.

Turnbull, A. (1999), The typical general aviation aircraft, Contractor Report NASA/CR-1999-209550, NASA LaRC.

Weber, M. et al. (1991), Weather information requirements for terminal air traffic control automation, *in* 'Fourth International Conference on Aviation Weather Systems (Paris, France)', American Meteorological Society.

Weiss, G., ed. (1999), *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, MIT Press, Cambridge, MA.

White, B. B. (1978), 'Empirical tests of the life cycle hypothesis', *The American Economic Review* **68**(4), 547–560.

Wieland, F., Wanke, C. et al. (2002), Modeling the NAS: A grand challenge for the simulation community, *in* 'First International Conference on Grand Challenges for Modeling and Simulation', San Antonio, TX.

Williams, M. L. & Sikora, J. J. (1992), Overview, *in* A. E. Ritchie, ed., 'Simulation Validation Workshop Proceedings', Vol. I, pp. 1–6.

Wilson, A. G. (1967), 'A statistical theory of spatial distribution models', *Transportation Research Part B* **1**, 253–269.

Wooldridge, M. & Jennings, N. R. (1995), 'Intelligent agents: Theory and practice', *The Knowledge Engineering Review* **10**(2), 115–152.

Zowghi, D. (1995), 'Requirements engineering – scope', Joint Research Centre for Advanced Systems Engineering, as of June 28, 2002, from `http://www.jrcase.mq.edu.au/~didar/seweb/scope.html`.

VITA

Jung-Ho Lewe was born to Mr. Jae-Bong Lewe and Ms. Hyung-Moon Lee in Suwon, South Korea. He is married to Ms. Sook-Jin Song and has a son, Eric Jaymin Lewe. His undergraduate degree is in Aerospace engineering from Seoul National University in 1995. His first professional career began at R&D division in Hanwha Corporation where he was an engineer and assistant manager conducting various RDT&E projects with missiles, sounding rockets and satellite propulsion systems. Working in the first Korean national satellite program, he had an opportunity to expand his knowledge background while working with TRW in Redondo Beach, California. After he joined Georgia Tech in 1999, his research interests have spanned personal air vehicles, multi-criteria decision making, transportation systems modeling, agent-based simulations, stochastic optimization methods, and systems engineering focused on a system-of-systems perspective.