Modeling the Interactions between Land Use and Transportation Investments Using Spatiotemporal Analysis Tools

Contract No. BC-851

Prepared for

State Transportation Planner
Systems Planning Office
Florida Department of Transportation
605 Suwannee Street, MS 19
Tallahassee, FL 32399-0450

Prepared in Cooperation with the State of Florida Department of Transportation

by

Fang Zhao and Soon Chung
Lehman Center for Transportation Research
Department of Civil and Environmental Engineering
Florida International University
10555 West Flagler Street
Miami, Florida 33174
(305) 348-3821, (305) 348-2802 (fax)
E-mail: zhaof@fiu.edu

and

Shih-Lung Shaw and Xiaohong Xin
Department of Geography
University of Tennessee
304 Burchfiel Geography Building
Knoxville, TN 37996
Phone: (865) 974-6036, Fax: (865) 974-6025
Email: sshaw@utk.edu

August 2003

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
Supplemental Report for BC851		

4. Title and Subtitle	5. Report Date
MODELING THE INTERACTIONS BETWEEN LAND USE AND TRANSPORTATION INVESTMENTS USING SPATIOTEMPORAL ANALYSIS TOOLS	August 2003
	6. Performing Organization Code
7. Author(s) Fang Zhao, Shih-Lung Shaw, Soon Chung, and Xiaohong Xin	8. Performing Organization Report No.
9. Performing Organization Name and Address Lehman Center for Transportation Research, Department of Civil and Environmental Engineering, Florida International University, Miami, Florida 33199	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. Contract No. BC851
12. Sponsoring Agency Name and Address Research Center, 605 Suwannee Street, MS 30, Tallahassee FL 32399-0450	13. Type of Report and Period Covered Final Report March 2000 – August 2003
	14. Sponsoring Agency Code 99700-3596-119

15. Supplementary Notes

16. Abstract

While it is generally agreed that transportation and land use interact with each other, the feedback mechanism in their relationship has not been well defined at a level of detail that adequately supports travel demand modeling. Most studies to date have been at metropolitan level, thus unable to account for interactions spatially and temporally at smaller geographic scales. Additionally, such studies require a significant amount of historical and geographic data, which are not well supported in the current Geographic Information Systems (GIS). In this study, historical land use and transportation data were collected and converted to GIS data. TransLand, a GIS program that was capable of handling and manipulating spatiotemporal data, was implemented and a framework for identifying the spatial and temporal interactions between transportation and land use was developed. TransLand was implemented in ArcGIS 8.0 environment and is capable of storing temporal spatial data and extracting useful information to support the visualization and analysis of transportation and land interactions. A framework for identifying the spatial and temporal interactions between transportation and land use was developed based on statistical analysis of time series, which was applied to selected corridor areas in Miami-Dade County, Florida. The results of the time-series analysis showed that transportation improvements impacted land uses at varying rates and intensities. Cumulated impact was also measured with time series techniques.

17. Key Word	18. Distribution Statement
Geographic information systems (GIS), spatiotemporal GIS, TransLand, transportation investment, land use, land development, time-series analysis, time lags	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	300	

ACKNOWLEDGEMENTS

We would like to thank the many people who helped in this study by providing the data used in this study. They include Mr. David Korros, Mr. Armando Lopez, Mr. Carlos Gonzales, and Ms. Paola Ramirez of FDOT District 6, Mr. Carlos Roa and Mr. Michael Moore of Miami-Dade County Metropolitan Planning Organization, Mr. Manual Armada of the Miami-Dade County Department of Planning and Zoning, Ms. Krysten A. Laine of the South Florida Water Management District, Mr. Manuel Pernas of the Miami-Dade County Property Appraiser's Office, and Dr. Yun Sool Joo of the Miami-Dade County Building Department. The project team is also grateful to the members of the Florida Model Task Force Land Use and GIS Subcommittee for their technical guidance and advice. Help from Dr. Dimitrios Thomakos from the FIU Economics Department on time series analysis is also gratefully acknowledged. Supports provided by the Department of Geography at The University of Tennessee, Knoxville are also greatly appreciated. Finally, special thanks go to Mr. Shi-Chiang Li, the project manager, for his guidance, support, and help throughout the study.

This project was funded by the Research Center of the Florida Department of Transportation. The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

INTRODUCTION

In the last century, the development of transportation technologies together with transportation networks has been one of most important factors influencing the shape and the size of cities. Automobiles, with their unparalleled mobility, flexibility, and comfort, have expanded the urban size by making long distance trips easier than ever and have resulted in decentralization of employment and urban sprawl.

In response to sprawl, citizens, public interest groups, and state and local governments have begun to develop smart-growth solutions to guide growth into existing areas with mixed use and public transportation options and to conserve farmland and open space. To achieve smart-growth solutions, integrating transportation and land use planning is extremely important. However, as pointed out in the Special Report 245 published by the Transportation Research Board (TRB 1995), "the state of knowledge and modeling practice are not adequate for predicting with certainty the impacts of highway capacity additions." This statement is still true today. In conventional travel demand modeling procedures, interaction or feedback mechanism between transportation and land use is not adequately modeled, including the implications of transportation investments with regard to the patterns of land development, time lags between transportation projects and land development, growth rate of development, and "equilibrium point", or the point at which development levels off for a given level of service of the transportation system. While travel demand models consider impacts of land use on transportation systems (e.g., trip generation is determined based on land use variables), most lack the ability to reflect the implications of transportation investments on the patterns of land development except in a limited way such as considering accessibility measured in interzonal travel time. Additionally, any changes in the transportation system are usually fed into a land use model assuming an immediate response in land use to transportation improvements. However, the responses from land use to transportation improvements usually lag behind and the processes that land use and transportation system influence each other are dynamic and the two co-evolve. This co-evolution process is not captured in traditional transportation models or is over-simplified. The lag effects of transportation improvements on land uses have not been adequately studied. While there have been empirical studies on time lags, e.g., changes in property values or travel behaviors measured in vehicle-miles traveled, these studies are limited in their scopes.

Due to the time lag between transportation improvement construction and the changes to land use stimulated by these projects, the collection and examination of historical data on both transportation projects and land use are needed to gain a better understanding of the interaction between them. In the past, the lack of digital data of historical land-use changes and transportation systems, as well as the lack of the necessary analysis tools, has limited studies on the interaction between land use and transportation to a particular project or a particular area of interest and the spatial, and temporal aspects have not been adequately considered. The last two decades have witnessed great advances in computer technologies, digital mapping, and geographic information systems (GIS), including the development of temporal GIS technologies that are capable of supporting analyses of data that have both a spatial context and a time dimension. These advances offer us a new opportunity to examine the land use and transportation co-evolution at a larger scale and at a more detailed level.

This research has two main goals. One was to develop a temporal GIS to support management of historical data on land use and transportation and to support visualization and analysis of the data for travel model improvement efforts. Another was to design a methodology for analyzing the interactions between land use and transportation. The specific objectives include:

1. Design a temporal GIS data model that supports the storage, retrieval, and analysis of spatiotemporal data;
2. Implement a prototype temporal GIS program that supports the analysis of historical land use and transportation data;
3. Identify availability of historical data on land use and transportation and collect and process such data to build temporal GIS databases;
4. Identify the types of new land developments that occur in transportation project areas; and
5. Study the temporal patterns of land developments, including time lags between transportation improvements and land developments, the rate of development, and length of intensified land developments stimulated by transportation improvements.

In the remainder of this report, the literature on transportation and land use interactions and on temporal GIS is reviewed in Chapter 2. Chapters 3 and 4 discuss the design of a temporal GIS data model and the implementation of the temporal GIS data model, temporal databases, and graphic user interface. Chapter 5 describes time-series analysis as a tool to support the analysis of interactions between land use and transportation. Basic definitions and concepts of time-series analysis are provided. Chapter 6 presents case studies of three corridors and describes the study data, study area selections, methodology, and results. Finally, conclusions and recommendations from this research are provided in Chapters 7 and 8.

LITERATURE REVIEW

The literature reviewed included transportation and land use interactions and advancement in the field of temporal GIS.

Transportation and Land Use Interactions

Highway Impacts on Land Use

Evidence has been well documented that changes in the nation's transportation system have resulted in changes in people's lifestyles, including choice of work and residential locations and travel behavior (Adams 1970, Middleton 1967, Smerk 1967, Smith 1984, Flink 1975, Muller 1986). The availability of inexpensive automobiles and massive construction of expressways have been considered the driving forces behind the exodus of city dwellers to suburbs and relocation and dispersion of employment centers (Moynihan 1960, Muller 1986).

The debate concerning changes to America's urban structure that caused suburban sprawl began 40 years ago and continues today. The issue, however, is much more complex than the relationship strictly between highways, auto ownership, and land use. Scholars have argued that additional, equally important factors also contributed to suburbanization including dependency on property tax and inadequate representation of city governments in the legislature of some states (Meyer 1968), as well as "…cheap credits, favorable mortgage loan terms, accumulations of savings, rapid family formation, the postwar baby boom, favorable tax treatment, a strong preference for home ownership, and the suburbanization of an ever larger number of jobs" (Kain 1970).

Payne-Maxie Consultants (1980) conducted a study to assess the beltways' impacts on land use and urban development. The analysis of growth and economic activity showed that beltways might have had some influence on metropolitan development patterns. Beltway location, mileage, distance from the CBD, or the density of interchanges showed more effects on land use changes than the presence of a beltway alone did.

Several studies were conducted on the effects of highway improvement projects on land use in Texas (Cosby and Buffington 1978a, Cosby 1979, Cosby and Buffington 1978b, Buffington et al. 1985). These studies found increased land development activities in both short-term and long-term, with changes occurring at the highest rate in the short-term, but diminishing in the long-run. A study by Herndon (1980), however, found little impact on land use change and development from a street improvement project and contributed the absence of this impact to the limited opportunities for growth because most of the changes in land use in terms of predominant land use, enforced subdivision restrictions, quality of buildings, and lack of unimproved property occurred before the street improvement began.

Transportation Project Impacts on Property Values

Huang (1994) suggested that highway interchanges and public transit stations positively impacted land value, but the extent and magnitude of the impacts varied. In a review of empirical studies over the past four decades on how transportation facilities affected property values, Ryan (1999) recommended the use of travel time savings as the measure of transportation access. When accurately measured, property values tend to show the theoretically expected relationship with transportation access. Swenson *et al.* (1998) also claimed that there was a strong correlation between transportation infrastructure investments and population, population density, total property values, and growth per square mile in property values. Vadali *et al.* (2001) performed a study on an approximately 9.2 mile segment of the recently reconstructed North Central Expressway (NCE), to evaluate the impact of transportation improvements, and concluded that property values were affected to a greater degree due to traffic-related effects than because of the distance from freeway. Voith (1992) found that homes commanded higher sales prices when they were located in places where highway travel time to the Philadelphia central business district was less, all else being equal.

Location Choice

Location choice theory has played an important role and has contributed theoretically and empirically to regional science. It focuses on the 'where' and 'why there' questions of human activities with respect to both households and businesses and on location patterns. Headicar (1996) considered it fundamental to the evolution of urban settlements that the location of new developments was determined largely by the opportunities for interaction created by the transportation system. Geographically, the pattern of growth had not been uniform, but was biased towards particular types of locations where investments tend to be concentrated:

Kawamura (2001) assessed the difference in location choice patterns based on data from 1981 and 1999 in the six-county Chicago region (Cook, DuPage, Kane, Lake, McHenry, and Will counties) by quantifying changes in the location pattern of various business sectors in relation to transportation facilities. The findings indicated that after controlling non-transportation factors and density of transportation facilities, businesses looked for places closer to freeway ramps as they moved farther away from the central business district (CBD).

Induced Travel

In recent years, the phenomenon often referred to as "induced travel" (also known as "generated traffic") has attracted much attention among transportation planners. Induced travel is the additional travel resulting from improvements to transportation facilities intended to provide greater capacity. This increased capacity will initially result in reduction of congestion on the improved facilities. However, traffic on the new facilities tends to gradually increase until congestion again reaches high levels, while at the same time there is no obvious reduction in congestion on the surrounding roads. The question therefore arises: will new transportation facilities or increased capacity result in more travel, thus consuming all the benefits of the investments in these improvements and resulting in more negative environmental consequences? Many researchers have attempted to answer this question by examining: (1) which factors, in addition to added capacity, may be responsible for the changes in travel demand; (2) whether there is merely redistribution of trips while the total travel (e.g., measured by VMT) remains a constant; (3) whether the change in travel is a result of modal switching; (4) whether the demand actually changes due to increased accessibility; (5) how much increase in demand will be induced by added capacity; and (6) what are the implications of induced travel on land use. Many studies demonstrated that, after factors such as population growth were accounted for, new highway construction did not induce additional travel. Yet, the questions of how much induced travel is caused by added capacity and the mechanism of induced travel remain to be fully understood.

The theory of induced travel is based on microeconomics. Induced travel is a response to the reduced price of travel. As in the case of any commodity, the demand (in this case travel) will increase when the price (travel cost) falls. Added demand may be the result of longer trips made possible by shorter travel time, mode shifts, new trips that have been suppressed by the previous level of congestion, trip redistribution, route reassignment, shift of travel time, etc. (U.K. Department of Transport 1993).

There is a general agreement that induced travel is real. However, there is no clear definition of induced travel or universally accepted units to measure induced demand. Many measures have been introduced (Downs 1962, U.K. Department of Transport 1993, DeCorla-Souza and Cohen 1998, Heanue 1998, Litman 1999, Strathman *et al.* 2000, Barr 2000, Fulton 2000, Abelson and Hensher 2001, Boarnet and Chalermpong 2001). Thus, the extent of induced travel has not been clearly determined.

Modeling the effect of added capacity is complicated by many behavioral or nonbehavioral factors that may contribute to demand. Based on time-series travel data for various types of roadways, the SACTRA report (1994) reported that actual travel volume was 10 percent (in the short term) to 20 percent (in the long term) higher than forecast due to unaccounted induced travel. Johnson and Ceerla (1996) studied California metropolitan areas and found the elasticities to be 0.6 to 0.9 over a period of three years. Cohen (1995) noted that travel time elasticities were generally between 0.0 and –0.1. Noland (2001) estimated a fixed effect cross-sectional time series model and a set of equations by road type. These models suggested that short-term (two-year) elasticity was greater than long-term (five-year) elasticity; and that growth of VMT on freeways and local roads was faster than on arterials. The elasticities given by the simultaneous equation model with a two-year lag were 0.567, 0.267, and 0.509 for interstate, arterial, and collectors, respectively. In another study, Noland and Cowart (2000) attempted to take into consideration the congestion level on roadways based on the notion that only added capacity on congested roads would induce travel, and analyzed annual data on metropolitan level congestion compiled by the Texas Transportation Institute (TTI), urban land area, and lane miles of capacity. After controlling for population size, population density, per capita income, fuel price, and for some of the models the lagged effects (dependent variable lagged by one year as an independent variable), the authors found that the number of lane miles accounted for about 15 percent of annual growth in VMT. However, results showed significant variations between metropolitan areas, which were not well understood. The authors hypothesized that these variations might be partially due to the varying growth rate in lane miles.

Strathman *et al.* (2000) used the cross-sectional approach to analyze the relationship between road capacity and VMT of 48 metropolitan areas by combining the 1995 National Personal Transportation Survey (NPTS) and the TTI mobility study, accounting for the effects of residential location, employment location, and commuting mode choice. The direct effect of a one-percent increase in per capita roadway capacity was estimated to result in a 0.29-percent increase in VMT for the short-run period, when all other variables were controlled. It was estimated that the indirect effect of a ten-percent increase in roadway capacity led to a 0.033-percent increase in VMT as a result of capacity related residential and employment density changes. Marshall (2000) conducted an aggregate analysis to estimate the effect of roadway capacity expansion on induced travel using the TTI's Urban Congestion Study Data Set for 70 U.S. urban areas and found that daily vehicle miles per lane mile were strongly correlated with population, area, and density. Elasticities were estimated to be 0.85 for highways and 0.76 for principal arterials. Land use changes due to the addition of roadway capacity might cause long-term effects (or secondary effects) that lagged behind induced demand. A study by Hansen (1998) pointed to a trend whereby VMT increased with the addition of capacity in terms of lane-miles of new highways. Hansen and Huang (1997) measured elasticity at the metropolitan area level using data of VMT on state highways, state highway lane-miles, population, and per capita income for every urban county in the State of California for the years 1973 – 1990. Controlling for factors such as county population and per capita income, the elasticity of vehicle travel (VMT) with respect to lane miles was determined to be 0.3 to 0.7 for counties and 0.5 to 0.9 in California metropolitan areas for a four to five year period. Fulton *et al.* (2000) found a significant and robust relationship between lane miles and daily VMT in the mid-Atlantic region of the U.S. The model used to estimate both short-run and long-run elasticities had the lag structure employing an exponential distribution, which implied that the effects were assumed to be strongest in the first year and then to decline with time. The elasticity between lane miles and VMT was estimated 0.1 to 0.4 in the short-run, and 0.5 to 0.8 in the long run. The Granger test results showed that lane-mile growth temporally preceded growth in VMT. Barr (2000) concluded that there existed statistically significant relationships between travel time and VMT with elasticities in the range of -0.3 to -0.5, meaning that increases in highway capacity reduced travel time and, in turn, induced travel demand.

Integrated Land Use and Transportation Models

The term "integrated" implies a feedback mechanism between the transportation and land use models. The transportation model deals with forecasting travel demand and determining the adequacy of the supply of transportation services. In almost all currently employed integrated models, the transportation model is a traditional four-step model that consists of trip generation, trip distribution, modal split, and trip assignment. The land use model, on the other hand, is concerned with modeling the demand for and the spatial distribution of employment, residential, shopping, and other activities to allocate the area's residents and workers to specific urban zones. The land use system supplies the transportation system with estimates of the location and volume of travel generators. The travel costs resulting from the equilibrium between transportation demand and supply can be fed back into the residential and employment activity location models, which in turn modify its resident and employment location estimations. This allows transportation system changes to affect land utilization, which in turn feeds back its effects in the form of new levels and locations of traffic generation. The notion of locational accessibility here plays a central role in all currently operational models. As an integral component of such accessibility, travel cost changes become part of the mechanism used to reallocate labor, residents, retail and service activities, and when modeled, freight flows between spatially separated land uses.

Within a number of operational models including the MEPLAN and Kim models, the urban system is modeled as a series of markets, with emphasis placed on clearing a transportation market and one or more other land use markets by solving a suitable set of spatially varying market prices endogenously (i.e., travel costs and site rents). Kim's Chicago model, however, did not consider the lag effects.

Within the less inclusive models, such as ITLUP, which avoid endogenous modeling of non-transportation price mechanisms, equilibrium between the transportation system's demands and supplies can also be brought about, stabilizing the parameters within the residential and employment activity location submodels. Such considerations of equilibrium in urban evolution quickly take us into the area of temporal dynamics. Wegner (1995) suggested that, when considering the rate of change of different phenomena in a spectrum, networks and land uses would be slow to change, population and employment would change rapidly, and work places, housing, goods transportation, and travel would have a medium rate of change. Within the ITLUP, MEPLAN, and Dortmund models, lagged effects play an important role in linking different submodels within the transportation and land use systems both across as well as within a single time period.

More recently, a new generation of microsimulation models of land use and transportation models are emerging. UrbanSim is a simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. With a modular design, it models household location choice, employment location choice, real estate development, interfaces with travel models, and incorporate urban dynamics into the models (Waddell 2001). Urbansim has been motivated by the need to test policies such as those regarding urban growth boundaries at the regional or metropolitan scale, street design, mixing of uses, pedestrian access at the neighborhood or site-specific scale, etc.

TRANSIMS is a disaggregate microsimulation model that simulates household activities and individual trips over a multimodal network. The type of activities and their locations are driven by household and personal demographics such as the age of an individual, the person's income, gender, and employment status. TRANSIMS is designed to give transportation planners accurate, complete information on traffic impacts, congestion, and pollution (LANL 1999). While TRANSIMS is capable to produced detail travel information over a large network in an urban area and accurate air quality information, it does not deal with the land use issues as well as other models (e.g. MASTER and UrbanSim).

Temporal GIS

Temporal Concepts

Time is often conceptualized as a continuum on a linear scale that moves in forward or backward directions (i.e., future, present, and past). This simple conceptualization becomes insufficient when we need to deal with complex real world phenomena. Snodgrass (2000) defines three fundamental temporal data types: instant, interval, and period. He also suggests that there are three kinds of time: user-defined time, valid time, and transaction time. Frank (1998) proposes four different views of temporal taxonomy. The first view is based on the concept of events, which are abstract time points without duration or intervals between two events. The second view is based on the interpretation of processes: linear or cyclic repetitive pattern. The third view depends on the scale of measurements: ordinal time scale and interval scale. The fourth view is based on the order of events: total order, partial order, branching, and multiple perspectives of observing the event orders. Hazelton (1998) also identifies four time metaphors that are found to be in general use, which are linear metaphor, cyclic metaphor, multi-dimensional metaphor, and branching metaphor. These concepts are directly relevant to transportation studies. For example, some transportation projects are implemented with specific cycles and a transportation project could branch into separate projects with different time frames.

Hornsby and Egenhofer (2000), on the other hand, formulate time with an explicit description of change with respect to states of existence and non-existence of identifiable objects. Four types of primitives related to the identity states of objects are described: object existence, non-existing object without history, non-existing object with history, and transition about two identity states of the same object. Therefore, nine change operations are derived from a systematic combination of the first three primitives. Claramunt (1996) discusses changes in terms of events and processes. He points out that there are three main classes of basic spatiotemporal processes: evolution of a single entity, functional relationships between entities, and evolution of spatial structures involving several entities (reconstructing processes). Most real world phenomena are the outcomes of temporal *processes*. A process in turn consists of a set of *events*. An event triggers changes to a real world phenomenon that leads it from one *state* to another state. Temporal databases and spatiotemporal GIS need to represent and manage these states, events, and processes.

Measurement Frameworks and Spatiotemporal Data Representations

Various kinds of changes that occur to transportation and land use are interdependent, and cannot be easily separated. In order to model the interactions between different change patterns (e.g., land use changes and transportation system changes), a framework is needed to take into account the time component (when), the location component (where), and the attribute component (what) associated with these changes and their interrelationships. Sinton (1978) suggests using a measurement framework that includes the three main components (i.e., time, location, and attribute) to serve as the "fixed", the "controlled", and the "measured" parameters, resulting in a total of six possible measurement scenarios. Each of the six measurement scenarios is suitable for a particular type of analysis need. For example, we can "fix" time (e.g., April 1, 2001) and "control" location (e.g., 0.5-mile, 1-mile, and 1.5-mile buffer zones around the interchanges of I-75 in Broward County) to "measure" attributes such as land use types.

There have been many approaches suggested for spatiotemporal data representations. Peuquet (1994) argues that an orthogonal 4-D representation of the (x, y, z) coordinates and the time (t) may be inappropriate for GIS because time and space exhibit important differences in their properties. She sums up some higher-level or derived knowledge of spatiotemporal relationships and patterns and discusses the characteristics of time such as temporal cohesiveness, temporal similarity, temporal continuity, hierarchical organization, and incompleteness. Peuquet and Qian (1995) propose a conceptual TRIAD model to describe geographic information with spatial, temporal, and feature dimensions. Feature-based, location-based, and time-based representations are considered to be complementary elements in the TRIAD model. Erwig (1999) however suggests that spatiotemporal databases are essentially database about moving objects. He attempts to find a general case of geometry that may change in a continuous manner and proposes a new line of research where moving points and moving regions are viewed as three-dimensional (2D space + time) or higher-dimensional entities.

Worboys (1994) defines a spatiotemporal (ST) object as a unified object with both spatial and temporal components. An elemental spatial object (i.e., a point, a line), known as a simplex, is combined with a bitemporal element to form an ordered pair. A finite set of such ST-simplexes satisfying certain properties is then further defined to form an ST-complex on which query algebra is developed. An ST-complex traces changes in discrete steps; therefore, it is unable to represent continuous evolution, but is well suited for processes where mutations occur in sudden jumps. Yeh (1992, 1995) provides a model for highly variable spatiotemporal data using behavioral functions. A behavioral function forms a spatiotemporal object triplet together with a timestamp and spatial data to describe versions of data evolution. The extra information allows the modeling of complex interpolations, making the data less redundant and resolving data deficiency between states.

Yuan (1996) presents a three-domain representation that defines semantic, temporal, and spatial objects in three separate domains. This data representation provides links between the three domains to describe geographic processes and phenomena. It allows the geographic concepts and entities to be represented through dynamical links among the three types of objects from either a layer or an object representation. A major advantage of this presentation is that there is no predefined data schema. In addition, it eliminates the constraints of a linear temporal scale that monitors and analyzes successive states of spatial entities. Claramunt (1996) addresses the semantics representation issue for spatiotemporal geographic information. This research retains the TRIAD data representation framework proposed by Peuquet and Qian (1995) and uses an object-oriented model to define the data structure. A spatial entity is defined by three attribute sets: temporal domain, thematic domain, and spatial domain.

Spatiotemporal Relationships and Spatiotemporal Operators

Temporal relationships among different events and processes are critical to spatiotemporal data analysis. An event can take place before, at the same time instant, or after other relevant events. A process may overlap with another process over a period of time. These temporal relationships are important information for understanding changes over time such as whether traffic flow increased significantly on the Sawgrass Expressway before, during, or after a major residential subdivision was built in Weston, Broward County. Peuquet (1994) defines temporal operators with three distinct classes: (1) metrics (length of the time interval) and topology (before, equal, meets, overlaps, during, starts, ends), (2) Boolean operators (and, or, not), and (3) generalization. Spatial operators include: (1) area generalization, (2) overlay, (3) spatial metric, and (4) topology. She further indicates that object operators are based on cause-and-effect relationships, such as causes/caused-by or becomes/precedes. Faria et al. (1998) suggest that temporal operators can be unary or binary, returning time values, Boolean, or objects. Spatiotemporal operators, which extend the spatial operators by adding a temporal dimension, include: location-temporal, orientation-temporal, and metric-temporal. Additional operators are defined in order to support the formulation of spatiotemporal queries.

Spatial, Temporal, and Spatiotemporal Data Models

A *data model* is a logical representation of data organization and data relationships. There are three types of data models frequently referenced in the literature. *Relational data model* is the most widely used data model in both computer science and GIS fields. It has a strong theoretical foundation and a standard query language (i.e., Structured Query Language or SQL). *Object-*

oriented data model is based on a set of object-oriented concepts (e.g., object, object class, inheritance, encapsulation, and polymorphism) that provides a more intuitive data representation of real world phenomena. It has gained significant attentions in the research community during the last decade. A number of commercial object-oriented database management systems (OODBMS) are available on the market today, although the relational database management systems still dominate the market. A more recent development is the *object-relation data model* that incorporates object-oriented concepts and functions into a relational database framework.

Relational data model has been the choice of implementing spatial data models in most GIS. Due to the limitations of handling coordinates data in a standard relational data model, some commercial GIS software packages also developed a proprietary file format to store the coordinates data and used a relational data model to store the attributes. The coordinate data in the proprietary file system are linked to the attribute data in a relational database through a set of unique identification numbers. The geo-relational data model used in ArcInfo (by Environmental Systems Research Institute, Redlands, CA) is an example. Object-oriented data models also attracted a lot of research interest in the GIS field. A number of commercial GIS software provide object-oriented programming tools for users to customize their GIS applications. However, most GIS packages do not adopt a complete object-oriented data model in their implementation. Due to the extension of relational data model to object-relational data model among several major relational database management system (RDBMS) vendors (e.g., Oracle Spatial, Informix Spatial Data Blade, IBM DB2 Spatial Extender), commercial GIS software vendors have started to adopt the object-relational data model. The Geodatabase data model in ArcGIS 8 is an example of object-relational data model.

GIS data models traditionally focus on the handling of locations and attributes. Time is often treated as a static component in GIS. This is known as the *snapshot* approach since each GIS layer represents the locations and attributes of a particular phenomenon (e.g., land use) at a fixed time point (e.g., March 22, 2001). Computer scientists, on the other hand, have been active on the development of temporal databases in order to better represent the time component in a database. In a relational data model, it is feasible to timestamp tables, records (i.e., tuples) in a table, or attributes in a table in order to keep track of the time component at different levels. Clifford (1982) and Clifford and Tansel (1983) discuss a historical database consisting of a collection of historical relations over the same set of states. A historical relation is viewed as a sequence of relation instances, indexed by valid time, each one representing a different state of the historical relation. Sarda (1990, 1993) describes a Historical Data Model (HDM) to support historical relations that contain either tuples stamped with a time instant or stamped with a time interval. Snodgrass (1995) proposes a conceptual data model – the Bitemporal Conceptual Data Model (BCDM) used as a basis for the definition of the temporal query language TSQL2. Data are time stamped either with sets of time instants or temporal elements.

Steiner (1998) discusses four types of temporal databases differentiated by their abilities of representing temporal information with respect to valid and transaction-time definitions. These four types of temporal databases are snapshot databases, historical databases, rollback databases, and bitemporal databases. He also introduces an object-oriented data model. A non-temporal generic OM model (Norrie 1992, 1993) is generalized into the temporal object data model

(TOM). This means that all aspects of the model including constructs, operations, and constraints have temporal generalization. Furthermore, the temporal dimension applies not only to data but also to metadata. Kafer *et al.* (1990) extend an object data model and describe a temporal data model by mapping time sequences to a complex object data model MAD (molecule-atom data model). A time sequence represents the history of an entity of the real world by preserving the different states of the entity over time in a time-ordered sequence. Kafer and Schoning (1992) further extend a relational data model and use a time interval as timestamp instead of a time sequence. Goralewalla and Ozsu (1993) use the extensibility of the object-oriented database management system, TIGUKAT, to implement a temporal data type. This temporal data type specifies an extensible set of abstract data type (ADT) and a rich set of behaviors to model time. Three different kinds of timestamps are supported: time instants, time intervals, and durations of time. It also includes data types to model discrete, continuous, and dense time.

Armstrong (1988) discusses timestamps of vector and raster GIS data models and describes them using entity, category, and relationship diagrams. For raster-based databases, an attribute history is proposed for each individual cell, thus avoiding the costly storage of whole data layers for each version. For vector-based databases, depending on whether durations are recorded explicitly using from and to dates or implicitly by recording a single timestamp, the method associates interval-stamped attributes with instant-stamped attributed in many-to-many relationships. Wachowicz and Healey (1994) present an object-oriented spatiotemporal model of real-world phenomenon and events. Real-world phenomena are represented as complex versioned objects with geometric, topological and thematic properties. A new instance of an object with a different identifier is created for every version of the object establishing a hierarchical structure for the past, present, and future of the object. Events, on the other hand, are manifestations of actions that invoke update procedures on one or more objects. Time is represented as an independent, linear dimension. This is different from other representations where the time axis is orthogonal to the spatial dimensions.

Langran (1988, 1992) suggests a Space-Time Composite Data Model. It is based on the principle that every line in space and time can be projected to a spatial plane with other lines, thus creating a polygon mesh. Each polygon in this mesh has its own attribute history associated with it. Each new amendment is intersected with the already existing lines, and new polygons are formed with individual histories. The model has been tested with a number of indexing methods. The test results looked promising, but only small data sets were tested. Peuquet and Wentz (1994) suggest that most systems are extensions of either a raster-based or a vector-based GIS. The raster or vector format of the extension depends on the type of queries that an application needs to handle. They observe that time-based representation questions can be answered by proposing a model to capture changes in the environment along a temporal vector. Starting with an initial state, events are recorded in a chain-like fashion in increasing temporal order, with each event associated with a list of all changes or can be triggered when gradual evolution is considered to be significant enough. Peuquet and Duan (1995) propose an Event-based Spatio-Temporal Data Model (ESTDM), which is a TRIAD representation of location-based (raster), feature-based (vector), and time-based components of modeling dynamic geographic objects.

TEMPORAL GIS DATA MODEL DESIGN

The main purpose of this task is to design a temporal geographic information system (GIS) database in order to facilitate analysis of transportation and land use interactions based on historical data. The design considerations focused on the data management, analysis, and display functions of a temporal GIS. We first identified a framework of six scenarios that transportation planners are likely to examine the spatiotemporal relationships between transportation and land use changes. This scenario framework extended the measurement framework proposed by Sinton (1978) and added the capability of examining interactions between two phenomena (i.e., transportation and land use). Based on this scenario framework, we then designed a temporal GIS that integrated the time component into the geodatabase data model of ArcGIS 8 (ESRI, Redlands, CA). The temporal GIS design delivers the following key components:

1) A data model that addresses the issues of spatiotemporal data representations and spatiotemporal relationships for land use-transportation interactions;
2) A set of query and analysis tools that handle time-based, location-based, and attribute-based operations as well as the interactions among the three components for the analysis of land use changes and transportation developments; and
3) A user interface that is designed to facilitate the query and analysis functions mentioned above and to visualize the spatiotemporal change patterns.

The temporal GIS design process involved four recent developments in the GIS community. First, there have been various efforts of developing GIS for Transportation (GIS-T) data models. This project incorporated some object-oriented design concepts from these GIS-T data models to develop a temporal GIS for examining transportation and land use interactions. Second, we used the Unified Modeling Language (UML) to assist in the object-oriented analysis and design process. Third, the geodatabase data model of ArcGIS 8 was chosen as the specific implementation platform for the development of temporal GIS capabilities in the design process. The ArcGIS 8 geodatabase data model is an object-oriented data model implemented with a relational database management system. This object-relational approach is different from the previous georelational data model and brings a physical data model closer to its logical data model (Zeiler 1999). Finally, we used the ArcObjects development environment and a Microsoft Component Object Model (COM) compliant programming language to extend the geodatabase data model in order to develop custom tools in support of the temporal data management and analysis requirements of this project.

Scenarios of Analyzing Spatiotemporal Interactions between Transportation and Land Use

Sinton's measurement framework is based on the concept that location, time and attribute are inseparable parts of measuring a real world entity (Sinton 1978). However, Sinton's measurement framework was designed to measure a single phenomenon, while our focus is to analyze the interactions between two phenomena (e.g., transportation and land use), with each phenomenon having its own location, time and attribute components. The three components therefore must be treated as integrated parts of each phenomenon. In this study, time is used as the organizing unit for examining the interactions between transportation and land use changes. The revised spatiotemporal interactions framework includes six scenarios that keep track of the space (location), time, and attribute components of both transportation and land use phenomena, while adding time into a GIS environment to organize the spatial and attribute changes of each phenomenon over time. Current GIS are well developed in their abilities of performing integrated analysis of location and attribute. This study extends these GIS functions by adding the time dimension into the GIS database design using the ESRI's ArcGIS 8 software.

Each scenario in the revised spatiotemporal interactions framework corresponds to one of the six ways that transportation planners are likely to ask questions about transportation and land use interactions. These six scenarios are described below.

- **Scenario 1** facilitates transportation planners to identify the time frames (i.e., measured component) when a transportation project (i.e., fixed component) might have had an effect on different land use types and their locations (i.e., controlled component). Sample queries under this scenario include:
    - When did different land developments in the region respond to the improvement of accessibility due to the I-595 construction project?
    - What was the temporal evolution of residential land-use changes in response to the I-595 construction?

- **Scenario 2** facilitates transportation planners to identify the land use changes (i.e., measured component) near a transportation improvement project (i.e., fixed component) during different time periods (i.e., controlled component). Sample queries under this scenario include:
    - Where were the new residential land use zones that were created during the five years after the construction of I-595?
    - Where were the areas that experienced greater than 20% land value increases since 1995 due to the construction of the I-595?

- **Scenario 3** facilitates transportation planners to identify the land use changes (i.e., measured component) near different transportation improvement projects (i.e., controlled component) at a selected time (i.e., fixed component). Sample queries under this scenario include:
    - What was the land-use pattern around major transportation projects in 1996?
    - Where were the vacant land parcels within a one-mile zone of an ongoing transportation project in 1996?

- **Scenario 4** facilitates transportation planners to identify the transportation patterns (i.e., measured component) associated with different land use activities (i.e., controlled component) at a selected time point (i.e., fixed component). Sample queries under this scenario include:
    - What were the traffic volumes on major roads in traffic analysis zones with population density higher than 3,000 persons per square mile in 1996?
    - Which road segments in Broward County adjacent to commercial land use in 1998 experienced a level of service of D?

- **Scenario 5** facilitates transportation planners to identify the time frames (i.e., measured component) when a land use pattern (i.e., fixed component) impacted different transportation components (i.e., controlled component). Sample queries under this scenario include:

  - When did major transportation projects take place in traffic analysis zones with an employment density higher than 8000 per square mile?
  - What was the temporal evolution of traffic volumes on major roads in response to a major land development project?

- **Scenario 6** facilitates transportation planners to identify the transportation system changes (i.e., measured component) associated with a land use pattern (i.e., fixed component) during different time periods (i.e., controlled component). Sample queries under this scenario include:
  - Where were the new transportation improvement projects during the five years after the construction of a major shopping mall?
  - Which road segments experienced greater than 20% traffic volume increases in a 3-year period since a traffic analysis zones reached a population density of 2000 per square mile?

GIS Database Design Approach

The above spatiotemporal interactions framework serves as the foundation for the GIS database design. In order to implement a temporal GIS that can support the above six scenarios, we took a GIS database design approach that incorporated: (1) object-oriented GIS-T data models, (2) Unified Modeling Language (UML) for object-oriented analysis and design, (3) a selected commercial object-oriented GIS data model (i.e., the geodatabase data model of ArcGIS 8), and (4) custom programming with ArcObjects.

With the development of an object-oriented GIS data model, it is now feasible to define objects and their relationships that are specific and intuitive to a particular application domain. For example, it is more intuitive and useful for transportation planners to deal with objects such as roads, bridges, transit stations, and traffic accidents instead of the geometric features such as points, lines, and polygons in a GIS. There have been several efforts of developing GIS-T data models. Examples include the National Cooperative Highway Research Program (NCHRP) 20-27(2) project's Linear Referencing Systems (LRS) data model (Vonderohe et al. 1997), the GIS-T Enterprise Data Model (Dueker and Butler 1998, Butler and Dueker 2001), and the Draft Framework Transportation Identification Standard prepared for National Spatial Data Infrastructure (NSDI) by the Federal Geographic Data Committee (FGDC, 2001). More recently, the NCHRP 20-27(3) project further extended the LRS data model to a multimodal, multidimensional data model for transportation location referencing systems that incorporated the time dimension (Adams et al. 2001, Koncz and Adams 2001). In addition, ESRI and a consortium led by the University of California, Santa Barbara jointly developed a Unified Network for Transportation (UNETRANS) data model. The UNETRANS data model is an object-oriented data model that can be used with the ArcGIS 8 for any transportation application development project. Basic transportation object classes and their relationships are defined in the UNETRANS data model. Users can modify the data model to include additional classes, properties and relationships to support their specific applications. This data model is specified in an industry-standard modeling language known as Unified Modeling Language (UML). All of these GIS-T data models offered valuable guidelines to the design of GIS-T databases. The temporal GIS data model developed in this project incorporated various design ideas from these GIS-T data models.

The Unified Modeling Language (UML) was adopted by the Object Management Group and became an industry standard in 1997. UML is a graphic modeling language rather than a programming language. It consists of object-oriented analysis and design notations that can be used for everything from high-level analysis concepts down to very detailed design elements (Richter 1999). A set of diagrams defined with UML notations can describe real world activities and represent their concepts and relationships. UML class diagrams translate real world objects into software entities with attributes, associations, methods, interfaces and dependencies. ArcGIS 8 includes a schema generation utility that can convert UML class diagrams into geodatabase schema. In addition, users can define custom objects with UML and use the ArcGIS 8 code generation utility to add custom object behaviors. In this project, we used the Microsoft Visio 2000 as the computer-assisted software engineering (CASE) tool to define and create various UML diagrams (Visio, 2000).

ArcGIS 8 introduces a new object-oriented data model – the geodatabase data model. In a geodatabase, it is feasible to define objects as they exist in the real world and the relationships between the object classes. For example, instead of using generic geometric features (i.e., points, lines, and polygons) to represent different kinds of real world entities, we now can define a traffic count station feature class or a traffic analysis zone feature class in a geodatabase. We also can define transportation projects as a separate object class (i.e., a semantic object) and then create a relationship class to link each transportation project to its related street segments. Furthermore, users can define attribute domains and validation rules in a geodatabase. The geodatabase data model observes the object-oriented concepts of inheritance, encapsulation and polymorphism. However, its implementation is on a relational database management system (RDBMS). The default RDBMS for storing the geodatabase data model is the Microsoft Access, which is known as a personal geodatabase in ArcGIS 8. With the use of ArcSDE, the geodatabase data model can store data in other commercial RDBMS such as Oracle, SQL Server, Informix, DB2, and Sybase. The object-oriented geodatabase data model provides many new functions that were not available in the georelational data model used in earlier ArcInfo releases. Most importantly, the geodatabase data model of ArcGIS 8 offers an environment for application developers to extend the basic data model for different application domains such as the development of a Transportation GIS data model.

ArcObjects is the development platform of the ArcGIS 8 software and is built on the Microsoft Component Object Model (COM) technology. COM is a client/server architecture of which the server (or object) provides some functionality that a client can use. COM defines a protocol that connects one software component with another in a client/server environment. COM also defines an interface-based programming model that encapsulates the data and methods with each instantiated object behind a well-defined interface. ArcGIS installation includes a set of ArcObjects object model diagrams that use the UML notations to define the ArcObjects classes (with their properties and methods), the interfaces available for each class, and the relationships among the classes. In order to develop a set of query and analysis functions in ArcGIS 8 to support the spatiotemporal study of transportation and land use interactions in this project, we used the geodatabase data model and the ArcObjects with a COM-compliant programming language to develop custom tools for this project. The interface-based programming model of ArcObjects means that all communications between objects are made via their interfaces (Zeiler 2001). An interface determines what requests can be made of an object that has an implementation of the interface. Two ArcObjects classes can have the same interface, but their implementations of the interface may be different. This is known as the polymorphism in an object-oriented approach. In this project, Visual Basic 6 was used as the main programming language to create function-specific dynamic linking libraries (DLLs) that supplement the existing functions in ArcGIS 8 to handle the temporal data.

TEMPORAL GIS DATABASES CONSTRUCTION AND SPATIOTEMPORAL ANALYSIS TOOLS DEVELOPMENT

Land use and transportation interaction is a dynamic process that involves changes over spatial and temporal dimensions between the two systems. Since the 1960s, many theories and models have been proposed to study land use and transportation interaction (e.g., Alonso, 1964; Anas, 1982; Anas and Duann, 1986; Boyce, 1980, 1990; Hansen, 1959; Kim, 1983; Prastacos, 1986; Kim et al., 1989; Hirschman and Henderson, 1990). Giuliano (1995, p. 3) argues that most models “are generally static, partial equilibrium models” even though they employ iterative methods and equilibrium concepts. Without incorporating time explicitly in a model, iterative methods only provide an equilibrium solution that is essentially for a given point in time and is not necessarily based on a valid theoretical foundation. Holding some variables fixed (e.g., treating land use activities as exogenously given variables in a travel demand model), such models are at best partial equilibrium approaches.

Most land use and transportation interaction studies take a *confirmatory* analysis approach that is based on *a priori* theories or models. These studies have helped us establish theoretical foundations and solution methods in our attempts to gain a better understanding of land use and transportation interaction. In the meantime, the literature suggests that little consensus regarding the conclusions can be drawn from the empirical studies (Giuliano, 1995). Land use and transportation system changes take place in a highly dynamic system that involves many forces such as economic development, population growth, and policy decisions. Models based on a confirmatory analysis approach may or may not be able to properly reflect these differences, especially when we need to examine land use and transportation interactions at varying spatial and temporal scales. *Exploratory data analysis* (EDA), on the other hand, takes a speculative and systematic approach to assist us in searching for patterns and processes hidden in the data sets (Openshaw, 1994; Goodchild, 2000). EDA emerged in the 1970s as methods of revealing what would otherwise go unnoticed with the use of standard statistical analysis (Tukey, 1977). EDA does not impose scale-specific assumptions on the analysis procedure; therefore, the patterns and relationships could emerge from the analysis rather than be imposed under a confirmatory data analysis approach.

This project implemented a temporal GIS design that offers exploratory data analysis capabilities for examining land use and transportation interactions at user-specified spatial and temporal scales. This temporal GIS is named "Transland" due to its focus on exploring spatiotemporal interactions between land use and transportation systems.

Creation of Temporal GIS Databases in Transland

This process first created empty geodatabases and then imported the data sets collected by the project team into feature classes. Once the snapshot GIS feature classes were created in geodatabases, a set of Visual Basic programs was used to add time stamps into the snapshot feature classes, to combine snapshot feature classes into space-time composite classes, and to generate time-related object classes. In addition, semantic class for transportation improvement projects was created for the project. A set of UML diagrams was created to help define and illustrate the implementation procedures and their associated Visual Basic forms and programs.

Spatiotemporal Exploratory Analysis Tools Development

Since each study area tends to have its own unique transportation and land use interaction patterns over space and time, spatiotemporal exploratory analysis tools could help search for the interaction patterns hidden in large, multiple data sets. Results derived from these exploratory analysis tools could be used to evaluate the existing model structures or to validate the model parameters for a particular study area. They also could suggest new hypotheses for examining the complex relationships between land use and transportation.

During the temporal GIS data model design stage, we proposed a spatiotemporal interactions framework to keep track of the space, time, and attribute components of transportation and land use phenomena. This spatiotemporal interactions framework consists of six scenarios that can explore land use and transportation interactions according to their spatial, temporal, and attribute data. These six scenarios are based on time, transportation, and land use components. This research takes full advantage of the current ArcGIS capabilities of handling spatial and attribute data and uses the time component to extend the static GIS data model into a temporal GIS data model. In other words, the time component is linked to the space-attribute objects so that users can query and analyze GIS databases according to any combinations of spatial, temporal and attribute components. Under this design, standard GIS functions such as SQL queries and spatial searches can be easily implemented in the Transland user interfaces.

Although ArcGIS provides a powerful set of tools for analyzing spatial and attribute data, it does not offer functions for temporal query and analysis. Since this project extends static snapshot GIS databases into temporal GIS databases, it is necessary to add spatiotemporal query and analysis tools into the user interfaces of Transland. A sequence of temporal query and analysis dialog windows, along with the custom Visual Basic forms and programs, was implemented in Transland. Because the time component is embedded in all transportation and land use GIS databases, these temporal query and analysis functions are integrated with the spatial and query functions available in ArcGIS. These VB programs were compiled into dynamic linking library (DLL) files and added into a custom toolbar in ArcMap as the "Transland Scenario Wizard" icon. Each scenario listed in the spatiotemporal interactions framework involves several steps to carry out the exploratory analysis procedure associated with the scenario. According to the specific scenario chosen by a user, the custom VB programs automatically determine the sequence of dialog windows that will be presented to the user. The user then interactively specifies spatial and temporal extents, along with selected attribute data, to explore land use and transportation interactions. The analysis procedure can be repeated using different spatial, temporal, and attribute specifications under the same scenario or between different scenarios. At the end of each scenario analysis, a report of both animated maps and data tables is presented in the Hypertext Markup Language (HTML) format to help users visualize the change patterns and review the summary data. These HTML files are saved in the default temp directory and are available for users to display at any time until users decide to delete them from the hard disk.

If users only want to visualize temporal changes of various GIS databases without running through the Transland Scenario Wizard, this project also includes an animated visualization tool, written in Java script, to display land use and transportation layers changes over time. Users simply click the "Spatiotemporal Animation" icon in the custom toolbar. The system then displays a dialog window for users to specify the time interval, the map layers, and the display method for the animated visualization. This animation tool also offers various functions for users to adjust the display speed, to pause the animation, and to replay the animated displays.

TIME SERIES ANALYSIS

The methodology utilized in this research is based on time series analysis. Time series is defined as an ordered sequence of values of a variable observed at equispaced time intervals. Time series analysis is designed to describe the dynamic consequences of time series by developing models and forecasting the future of the system based on historical trends. Basic assumption required to model time series in many time series techniques is that the time series is stationary, which means that the mean value of a time series remains the same the over observation period. A stationary process can also be defined as a series without trend, constant variance over time, a constant autocorrelation structure over time, and no seasonality. If the time series is non-stationary, it can be transformed to a stationary process applying one of following techniques: differencing, curve fitting, typically with a straight line, and transformation. A stationary time series can be expressed in terms of the mean and past and present error or innovation vectors. This form of the process is called a moving average (MA) representation. The error should be, by assumption, a white noise, which is a random vector with zero mean. A lag can be defined as a fixed time displacement. If there are observations, $y_1, y_2, \ldots, y_n$, over time, the lag between y_2 and y_7 is 5 (= 7 – 2). In other words, y_7 lags behind y_2 by 5 time units.

Two types of time-series analysis approaches were utilized: univariate models and multivariate models. The main difference between univariate and multivariate models is that in a univariate model there is one endogenous variable whereas a multivariate model has multiple endogenous variables. An endogenous variable is one that is internal to the modeled system, the value of which is the result of solving the system. In contrast, an exogenous variable is one that is external to the system being modeled, and the value of an exogenous variable must be specified externally and input into the system.

Univariate Model

In a dynamic system with pairs of observations (x_t, y_t), both x_t and y_t are often observed at discrete times, therefore x_t and y_t may be discrete series rather than a continuous process. The relationship between x_t and y_t may be represented in a *linear filter* form as shown below:

$$\delta(B)Y_t = \omega(B)X_{t-b}$$

where B is a lag operator and δ and ω are coefficients of variables. The discrete transfer function is obtained as $v(B) = \delta^{-1}(B)\Omega(B)$ and the weights v_0, v_1, v_2, … are called the *impulse response coefficients* of the system.

There are a number of tests for checking model validity and adequacy. For this research, three model diagnostic tests including Q-statistics, ARCH test, and Ramsey's RESET test were introduced. Q-statistics provide test statistics to check for randomness of residuals. Ramsey's RESET test checks for model specification. ARCH test examines the presence of conditional heteroskedasticity. If the model has conditional heteroskedasticity, it is required to be estimated usig the GARCH method.

Multivariate Model

Since transportation improvement projects are determined through a transportation planning decision process and not entirely determined in the system of land use, it should be treated as an exogenous variable while land use variables are endogenous variables. Vector autoregression with an exogenous variable, denoted as VARX(p, s), was developed to model endogenous variables and an exogenous variable. VARX(p, s) may be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_t = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1N} \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N1} & \Phi_{N2} & \cdots & \Phi_{NN} \end{bmatrix}_i \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_{t-i} + \cdots + \sum_{j=0}^{s} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \vdots \\ \Theta_N \end{bmatrix}_j [x]_{t-j} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}_t$$

The VARX order p for endogenous variables can be decided using five order selection criteria: the likelihood ratio (LR), final prediction error (FPE), Akaike information criterion (AIC), Schwarz information criterion (SC), and Hannan-Quinn information criterion (HQ). The VARX order s for an exogenous variable can be determined via trial-and-error.

Three model diagnostic tests including AR root test, Lagrange Multiplier test, and Portmanteau test are necessary. AR root test is to check whether the estimated VARX models are stable. If the stability condition, $\det(I_N - \Phi_1 z - \ldots - \Phi_p z^p) \neq 0$ for $|z| \leq 1$, is satisfied, then the estimated model is said to be stable. The portmanteau test is to check whiteness of the residuals of the estimated model and the overall significance of the residual autocorrelations. The null hypothesis is that no residual autocorrelations exist up to the specified lag. LM test is to check if there is no serial correlation at the specified lag.

In this research, it is of interest to investigate the marginal impact of changes in the exogenous variable. Multiplier analysis allows ones to estimate cumulative impacts of an innovation in an exogenous variable on endogenous variables. In the multiplier analysis, a large estimation uncertainty for the VARX coefficients leads to imprecision in the estimation of cumulative impacts. This practical problem may be cured by imposing constraints on the coefficients, resulting in a restricted system.

CASE STUDIES OF TRANSPORTAION AND LAND USE INTERACTIONS

Background

Added capacity makes the land more attractive to developers by removing "concurrency requirements" and induced developments cause increase of travel demand on the improved road. If the increased travel demand erodes the added capacity away, travel demand will be constrained and, in turn, the development growth will be tempered until an "equilibrium point" is reached. Then the improved roadway is filled up with increased travel demand and needs another improvement.

The series of these actions forms a cycle, which consists of three types of lag: Lag1 representing a time span between the completion of roadway improvement and the submission of building permit applications for new developments; Lag2 being the institutional lags for the building department to review applications and issue permits; and Lag3 being the construction period for new developments. In this study, Lag1 was considered in the modeling procedure.

Study Data and Study Area Selection

For this research, Miami-Dade County was chosen as the study area, mostly for the reason of its excellent availability of digital, and in many cases GIS, data when compared to most of the Florida counties. This study involved the use of historical building permit data and transportation improvement project data. The building permit database was obtained from Miami-Dade County and included building permit records from 1987 to 2001. All records were geocoded by their street addresses. Transportation improvement project information came from the Transportation Improvement Programs (TIPs), prepared every year by the Miami-Dade County Metropolitan Planning Organization. Total twenty-three hard copies of TIP report covering from 1978 to 2000 were obtained from the MPO. The temporal GIS database for these TIP reports was created by geocoding all projects listed in the reports based on the 2000 street network.

Three corridors were selected from the Miami-Dade County for this study based on two criteria: (1) isolation of impacts from other areas and (2) transportation improvements in the corridor (occurrence of growth in population, employment, and development). Three study corridors were selected: Tamiami Trail from SW 112th Avenue to SW 152nd Avenue, Bird Drive from

SW 117th Avenue to SW 157th Avenue, where the street ended, and North Kendall Drive between the Florida Turnpike and Krome Avenue.

Specification of Model Variables

The study employed two land use variables: commercial development and residential development. Commercial and residential developments were represented in terms of the sum of building square footage of applied building permits in a time unit, which was defined as month. Building square footage for commercial and residential development was aggregated as below.

Commercial Developments (*dcom*):

$$dcom_t = \sum_{i=1}^{N} (BLDG\_COM)_{i,t}$$

where N was the number of building permits within a study area in month t and BLDG_COM was the building square footage for commercial use measured in thousand square feet.

Residential Developments (*dres*):

$$dres_t = \sum_{i=1}^{N} (BLDG\_RES)_{i,t}$$

where N was the number of building permits within a study area in month t and BLDG_RES was the building square footage for residential use measured in thousand square feet.

A third variable was lane mile increase, an index to measure transportation improvements in terms of capacity. It was computed as the product of number of lanes and length of the improved section. The first order difference of lane miles, $dlanemile_t$ (= $lanemile_t - lanemile_{t-1}$), was used as a variable in modeling.

Univariate Model Estimations

A univariate model includes one endogenous variable and one exogenous variable, which were the total building square feet from the permit applications in a study area and the transportation improvement index, respectively. The models were estimated by least square method and GARCH method. The models for the study areas were given below:

Tamiami Trail Corridor

$$dtotal(t) = 9.846803 \times dlanemile(t\text{-}2) + 327.8634 \times dlanemile(t\text{-}15) + 0.254607 \times AR(1) + 0.130499 \times MA(4) + 0.260083 \times MA(12) + 11.36979$$

Bird Drive Corridor

$$dtotal(t) = 0.196534 \times dtotal(t\text{-}1) + 0.173770 \times dtotal(t\text{-}15) + 0.280335 \times dtotal(t\text{-}25) + 38.56848 \times dlanemile(t\text{-}23) + 10.89416 + 0.137867 \times \text{dummy@trend}$$

North Kendall Drive Corridor

$$dtotal(t) = 0.082 \times dtotal(t\text{-}10) + 0.085 \times dtotal(t\text{-}12) + 30.38312 \times dlanemile(t\text{-}7)$$

$$+ 75.58657 \times dlanemile(t\text{-}8) + 159.1483 \times dlanemile(t\text{-}11) + 0.256 \times MA(2) + 0.242 \times MA(3) + 21.39$$

where AR stood for autoregression and MA denoted moving average. There was a trend in the development in the Bird Drive corridor, which was accounted for by introducing a trend dummy variable.

The cumulated impacts (long-run effect) of roadway improvements on land use can be calculated as sum of coefficients of exogenous variable weighted by subtracting from one the sum of coefficients of endogenous variable. The cumulated impacts for the Tamiami Trail, Bird Drive, and North Kendall Drive corridors were 339,003, 110,397, and 319,439 square feet, respectively.

The land development lagged behind transportation improvements between two to 23 months, and there was a wide range of the lags depending on the corridors.

VARX Model Estimations

A multivariate VARX model includes one or more endogenous variables and one exogenous variable. In this study, the endogenous variables were the total residential and commercial building square feet from the permit applications in a study area, and the exogenous variable was the transportation improvement index.

The models estimated for each study corridor are shown below.

Tamiami Trail

$$\begin{aligned} dres(t) = \; & 1.379023 \times dcom(t\text{-}4) - 1.327276 \times dcom(t\text{-}6) + 2.321187 \times dcom(t\text{-}16) \\ & - 2.324449 \times dcom(t\text{-}21) + 0.471442 \times dres(t\text{-}1) + 0.011933 \times dres(t\text{-}15) \\ & + 0.010363 \times dres(t\text{-}17) + 0.034389 \times \mathrm{dres}(t\text{-}23) - 0.014000 \times dres(t\text{-}24) \\ & + 12.38801 \times dlanemile(t\text{-}2) - 6.585278 \times dlanemile(t\text{-}3) \\ & + 5.502976 \times dlanemile(t\text{-}4) + 327.0074 \times dlanemile(t\text{-}15) \\ & - 154.2895 \times dlanemile(t\text{-}16) + 4.638388 \end{aligned}$$

Bird Road Corridor

$$\begin{aligned} dres(t) = \; & 0.234462 \times dres(t\text{-}1) + 22.79471 \times dlanemile(t\text{-}18) + 15.76523 \\ & + 0.224561 \times TrendDummy \\ dcom(t) = \; & 0.388299 \times dcom(t\text{-}3) - 0.173641 \times dcom(t\text{-}6) + 0.059454 \times dres(t\text{-}2) \end{aligned}$$

North Kendall Drive Corridor

$$\begin{aligned} dres(t) = \; & 0.096 \times dres(t\text{-}10) + 0.126 \times dres(t\text{-}12) + 0.098 \times dres(t\text{-}18) \\ & + 0.099 \times dres(t\text{-}20) + 0.648 \times dcom(t\text{-}23) + 61.223 \times dlanemile(t\text{-}8) \\ & + 151.979 \times dlanemile(t\text{-}11) \\ dcom(t) = \; & 0.043 \times dres(t\text{-}17) + 16.997 \times dlanemile(t\text{-}7) + 13.804 \times dlanemile(t\text{-}8) \\ & + 3.695 \end{aligned}$$

Based on the results, the cumulative impact from a unit roadway improvement on the residential development was 345,679.5 square feet and zero on the commercial development in the Tamiami

Trail study corridor; 38,945.8 and 7,632.8 square feet for residential and commercial developments in the Bird Drive study corridor; and 349,765.2 and 60,654.9 square feet for residential and commercial developments in the North Kendall Drive study corridor.

The commercial development model for the Bird Drive did not have any significant exogenous variables, but cumulative impact to unit increase of roadway improvement was found. This must be indirect impact caused by roadway improvement through residential development. For instance, roadway improvement affected residential development and, in turn, residential development caused commercial development.

Land use developments began to respond to transportation improvements as soon as two months and as long as 21 months. The VARX model found more lags because univariate models employed total development as an endogenous variable, which was a variable aggregating all types of land use, while VARX model treated both commercial and residential developments as endogenous variables. It was, therefore, possible to track interactions between land use and transportation more closely.

Examination of Traffic Conditions

Development is influenced by transportation supply or accessibility, but is also determined by other factors. One important factor is the congestion level. Congestion not only reduces transportation accessibility, but also prevents further development from taking place through the concurrency requirements. The traffic conditions in two corridors, Tamiami Trail and North Kendall Drive, were examined because they had been state roads and historical traffic data, average annual daily traffic (AADT), were available from the Florida Traffic Information (FTI) CD-ROM.

The estimated v/c ratio for the North Kendall Drive corridor at SW 127th Avenue and SW 137th Avenue were higher than 1.0 before the construction, indicating that these sections had serious congestions. The traffic volume at SW 137th Avenue surpassed that at SW 127th Avenue in 1992, possibly due in part to the concentration of commercial development at the location.

The estimated v/c ratio for Tamiami Trail at SW 122nd Avenue was the only one affected by the road improvement. There was a decrease in the v/c ratio from 1989 to 1994, but after the improvement, the v/c ratio started to grow again. The v/c ratios at SW 137th Avenue increased between 1993 and 1995, reflecting increase of residential developments in the area between SW 137th Avenue and SW 147th Avenue in 1993.

CONCLUSIONS

Land use and transportation interaction has been a research topic for several decades. Many theories and models have been suggested to study this well-known, yet extremely complex, process. Using a confirmatory analysis approach based on prior theories and models certainly has helped us gain some insights into this complex process of land use and transportation interaction. However, empirical studies have suggested that land use and transportation interaction patterns can be highly variable in different geographic areas. Each geographic area tends to have its own unique characteristics that may lead to a different pattern of land use and transportation interactions. The patterns are also likely to vary as we examine them at different spatial and temporal scales. As geographic processes often exhibit properties of both spatial dependency and spatial heterogeneity, the challenge is to identify the spatiotemporal patterns underlying these complex geographic variations.

With recent advancements of GIS technology and research progress on temporal GIS, we are now equipped with better tools to tackle complex geographic processes. This project successfully implemented a temporal GIS, coupled with an exploratory analysis approach, that allow systematic and interactive ways of analyzing land use and transportation interaction among various data sets and at user-selected spatial and temporal scales. Temporal GIS databases implemented in this project makes it feasible for the analysis of spatiotemporal interaction patterns in a more efficient and effective way than the conventional snapshot GIS approach. Extending Sinton's measurement framework into a spatiotemporal interaction framework also provides a systematic means of exploring land use and transportation interactions.

In addition to develop temporal GIS tools to support the study of land use and transportation interactions and to support land use and transportation modeling, historical time series data on land use and transportation have been analyzed for selected corridors in the Miami-Dade County. For corridor level analysis, the geographic scale of land use data in temporal GIS databases should be property to accurately reflect land use changes and the temporal scale should be month.

Both univariate models and multivariate models were developed. Residential development, which was the dominant development, was significantly impacted by roadway improvements and was well expressed in terms of roadway improvements. Larger cumulated impacts on land use were found in the Tamiami Trail and North Kendall Drive corridors than in the Bird Drive corridor.

Land use changes of two principal arterials, Tamiami Trail and North Kendall Drive, responded to roadway improvements faster than that of the Bird Drive corridor. A plausible explanation may be that the traffic congestions in these two corridors were more severe and were constraining the development under the Florida concurrency requirements. Thus increased accessibility resulted in faster response in land use.

It was found that in two principal arterial corridors, Tamiami Trail and North Kendall Drive, the responses from land use changes to roadway improvements were faster than that in the Bird Drive corridor. A plausible explanation may be that the traffic congestions in these two corridors were more severe and were constraining the development because of the Florida concurrency requirements, thus increased accessibility resulted in faster response in land use. Impacts of roadway improvement on land use may continue under the conditions that developments are followed by population and employment growth, that developable land is still available, and that traffic in the corridor does not reach saturation level.

Estimated lag effect for Lag1, which is the time span between the completion of a roadway expansion project and increased building permit applications by developers for new developments, ranged from a few months to one and one half years. The averages of Lag2 and Lag3 were found to be four months and 10 months with standard deviations of four and nine months, respectively. Consequently, it took two to four years for travel demand to respond to road investments in the growing areas in Miami-Dade County.

The historical data available to this study were still limited in the sense that there was only one significant improvement in the transportation system in each of the study corridor. A longer modeling period that includes more than one cycle of transportation improvement-land use development may allow more accurate models to be developed. Additionally, the lack of traffic data also prevented the effects of congestion on land development to be adequately considered. Yet preservation of historical data remains a serious challenge.

RECOMMENDATIONS

Given the state-of-the-art nature of this research topic, there are many follow-up research opportunities to further improve the current achievements of this project. (1) Additional custom application tools can help the spatiotemporal interaction framework used in this study be a more user-friendly system. (2) Integration of existing land use and transportation models into the system will be another possible improvement. (3) The current temporal GIS data model may be further extended to allow representations of different time scales. (4) Innovative approaches to database design and spatiotemporal analysis procedures are needed to achieve a reasonable performance level in working with large files. (5) The spatiotemporal interaction framework will require an extension to facilitate the simultaneous analysis of both vector and raster GIS data such as high-resolution remote sensing images, which is useful to study land use and transportation interaction. (6) It is extremely important for different agencies and even different divisions within the same organization to develop an integrated and shared information system, to share data and to collect historical data. (6) In allocating growth in the long-range planning process, it will be necessary to consider the growth trend in the past, the "need" for commercial development in an area, the congestion level, the availability of developable land, and the lag effect. (8) This study focused on the analysis of land use and transportation interactions. To capture the full market mechanism, it will be necessary to develop more powerful models to take into consideration of other factors such as developable land, land price, housing price, quality of schools, accessibility, etc.

1. INTRODUCTION

In the last century, the development of transportation technologies and transportation networks has been one of the most important factors influencing the shape and the size of cities. Automobiles, with their unparalleled mobility, flexibility, and comfort, have expanded the urban areas by making long distance trips easier than ever and have resulted in decentralization of employment and urban sprawl. *The State of the Cities*, a 2000 report published by the U.S. Department of Housing and Urban Development, depicts our urban areas as expanding at about twice the rate as that of the population growth (U.S. HUD 2000). The result of poorly planned developments, urban sprawl brings many problems such as disappearing natural resources such as water, farm land, and open space, inefficient utilization of infrastructure, crowded schools, rising taxes, increased traffic congestion, and air pollution.

In response to sprawl, citizens, public interest groups, and state and local governments have begun to develop smart-growth solutions to guide growth into existing areas with mixed use and public transportation options and to conserve farmland and open space. To achieve smart-growth solutions, integrating transportation and land use planning is extremely important. With the Clean Air Act Amendments of 1990, the ability of travel demand models to assess the impact of transportation improvement plans on land-use development and air quality has been challenged in the courts (e.g., Citizens for a Better Environment vs. Deukmejian et al. and Sierra Club vs. Metropolitan Transportation Commission et al. in the San Francisco Bay area). Additionally, the Transportation Equity Act for the 21st Century (TEA-21) legislation requires a better understanding of the relationship between transportation and land use and the establishment of a better linkage between transportation investments and desired land uses. However, as pointed out in the Special Report 245 published by the Transportation Research Board (TRB 1995), "the state of knowledge and modeling practice are not adequate for predicting with certainty the impacts of highway capacity additions." This statement is still true today. In conventional travel demand modeling procedures, interaction or feedback mechanism between transportation and land use is not adequately modeled, including the implications of transportation investments with regard to the patterns of land development, time lags between transportation projects and land development, growth rate of development, and "equilibrium point", or the point at which development levels off for a given level of service of the transportation system. While travel demand models consider impacts of land use on transportation systems (e.g., trip generation is determined based on land use variables), most lack the ability to reflect the implications of transportation investments on the patterns of land development except in a limited way such as considering accessibility measured in interzonal travel time. Additionally, any changes in the transportation system are usually fed into a land use model assuming an immediate response in land use to transportation improvements. However, the responses from land use to transportation improvements usually lag behind. Moreover, the processes that land use and transportation systems influence each other are dynamic and the two co-evolve. This co-evolution process is not captured in traditional transportation models but is over-simplified. The lag effects of transportation improvements on land uses have not been adequately studied. While there have been empirical studies on time lags, e.g., changes in property values or travel behaviors measured in vehicle-miles traveled, these studies are limited in their scopes.

Due to the time lag between transportation improvement construction and the changes to land use stimulated by these projects, the collection and examination of historical data on both transportation projects and land use are needed to gain a better understanding of the interaction between them. In the past, the lack of digital data on historical land-use changes and transportation systems, as well as the lack of the necessary analysis tools, has limited studies on the interaction between land use and transportation to a particular project or a particular area of interest and the spatial, and temporal aspects have not been adequately considered. The last two decades have witnessed great advances in computer technologies, digital mapping, and geographic information systems (GIS), including the development of temporal GIS technologies that are capable of supporting analyses of data that have both a spatial context and a time dimension. These advances offer us a new opportunity to examine the land use and transportation co-evolution at a larger scale and at a more detailed level.

This research had two main goals. One was to develop a temporal GIS to support management of historical data on land use and transportation and to support visualization and analysis of the data for travel model improvement efforts. Another was to design a methodology for analyzing the interactions between land use and transportation. The specific objectives included:

1. Design a temporal GIS data model that supports the storage, retrieval, and analysis of spatiotemporal data;
2. Implement a prototype temporal GIS program that supports the analysis of historical land use and transportation data;
3. Identify availability of historical data on land use and transportation and collect and process such data to build temporal GIS databases;
4. Identify the types of new land developments that occur in transportation project areas; and
5. Study the temporal patterns of land developments, including time lags between transportation improvements and land developments, the rate of development, and length of intensified land developments stimulated by transportation improvements.

In order to better understand the relationship between land use and transportation, the concept of *accessibility* needs to be examined first. In broadest terms, accessibility addresses the ability to gain easy access to places and activities of interest, including business, retail, educational, and recreational sites (Woodhull 1992). Accessibility is often measured by the cost of travel in monetary terms, travel time, or a combination of the two. An improvement in a transportation system increases accessibility by better linking or connecting two places, thereby decreasing travel costs.

As the accessibility of an area increases, the land in the area will become more attractive to developers, thereby inducing more activities to locate to the area. The location of activities affects travel patterns, which, in turn, affect the transportation system. Land use change is incremental; a single development project may have an imperceptible impact on the transportation system unless the project is extremely large. Numerous studies have identified links between transportation projects and new land development, particularly around fixed guideway transit stations and highway projects. These studies typically examined the land use changes, property values, residential location choice, and improved accessibility in order to identify the land-use link to transportation projects. This study was focused on the temporal relationship between transportation improvement projects and land use, understanding that accessibility being a key in this relationship.

In the remainder of this report, the literature on transportation and land use interactions and on temporal GIS is reviewed in Chapter 2. Chapters 3 and 4 discuss the design of temporal GIS data models and the implementation of the temporal GIS data models, temporal databases, and graphic user interface. Chapter 5 describes time-series analysis as a tool to support the analysis of interactions between land use and transportation. Basic definitions and concepts of time-series analysis are provided and explained. Chapter 6 presents case studies of three corridors and described the study data, study area selection, methodology, and results. Finally, conclusions and recommendations from this research are provided in Chapter 7 and 8.

2. LITERATURE REVIEW

The link between transportation and land use is complex and controversial. The continuing, contemporary debate began well over forty years ago. While most experts agree that transportation at least have contributed to land use changes in the United States, not all can agree on the degree to which transportation has affected land use. There are also different opinions on what other forces have driven land use changes or transportation investments. The subject has attracted attentions from geographers, economists, urban planners, engineers, social scientists, political scientists, and other professional fields. The diversity of the researchers' background indicates that the question is not a simple one to answer. While the body of literature is vast, this review is limited to work that helps answer specific questions about the relationship between transportation and land use – mainly, the temporal relationship between the two. Assuming such a relationship does exist and the interaction occurs through changes in accessibility, the focus of this study was on how to measure the changes in transportation and land use.

In this chapter, Section 2.1 provides a review of literature on the effects of transportation investments on land development patterns, property values, location choice theories and models, as well as integrated transportation-land use models. Another focus of this research was temporal GIS, a technology useful for storing, retrieving, visualizing, and analyzing historical land use and transportation data. Section 2.2 provides a summary of the literature on temporal databases and spatiotemporal GIS, including the key issues in temporal databases and spatiotemporal GIS.

2.1 Transportation and Land Use

2.1.1 Historical Perspectives

Evidence has been well documented that changes in the nation's transportation system have resulted in changes in people's lifestyles, including choice of work and residential locations and travel behavior. Adams (1970) identifies a four-stage model of intra-metropolitan transport eras and associated growth patterns:

(1) 1800 – 1890: Walking-Horse-car Era;
(2) 1890 – 1921: Electric Streetcar Era;
(3) 1920 – 1945: Recreational Automobile Era; and
(4) 1945 – present: Freeway Era.

At each stage, a particular movement technology and network expansion process shaped a distinctive spatial pattern of urban organization.

A number of historical case studies supported Adams' (1970) contention that electric streetcars contributed to the growth of the metropolitan suburbs and had significant impacts on urban spatial structure in the Electric Streetcar Era (1890 to 1920). The streetcar lines shaped the development patterns of urban cores and nearby suburbs of many American metropolitan cities during this period. Propelled by electric motors, they extended commuting distance, allowing residential development to take place in suburban areas on the outskirts of the city. Population growth followed these lines, and new trolley line extensions invariably led to an increase in land values (Middleton, 1967). It was estimated that about one-quarter of the U.S. population resided in urban and suburban areas during this period (Smerk 1967). Moreover, as rail lines extended into newly forming suburban areas, urban rail ridership increased from 600 million to 15.5 billion trips annually, reflecting the dramatic increases in population of the urban areas. By 1920, approximately half of the U.S. residents lived in urban areas as the number of city residents increased from 11 million in 1880 to nearly 45 million in 1920 (Smith 1984).

With the revolutionary assembly-line manufacturing technology that allowed cars to be produced at lowered costs, automobiles became the major mode of transportation because they offered a faster, more convenient, individualized means of personal transportation. In the early days, automobiles were used primarily by farmers to access local service centers, and for that reason, many paved roadways were built in rural areas. In cities, cars were used for weekend outings, driven on the landscaped parkways that followed scenic waterways. Through the 1920s and 1930s, passenger car registrations in the United States increased explosively and the suburb growth rate, mainly based on residential development, exceeded that of the central cities. Table 2.1 illustrates intra-metropolitan population growth between 1910 and 1960. Flink (1975) found that as early as 1922, 135,000 households in 60 metropolises were completely dependent on motor vehicles. The metropolitan highway network expansion continued through the global war and residential development occurred ever farther away from the downtown cores. In the 1930s and 1940s, population deconcentration was accelerated, as can be seen from Table 2.1.

Table 2.1 Intrametropolitan Population Growth Trends, 1910 – 1960 (Source: Muller 1986)

Decade	Central City Growth Rate	Suburban Growth Rate	Percent Total SMSA[a] Growth in Suburbs	Suburban Growth Per 100 Increase in Central City Population
1910 – 1920	27.7	20.0	28.4	39.6
1920 – 1930	24.3	32.3	40.7	68.5
1930 – 1940	5.6	14.6	59.0	144.0
1940 – 1950	14.7	35.9	59.3	145.9
1950 – 1960	10.7	48.5	76.2	320.3

Notes: U.S. Census of Population.

a. SMSA: Standard Metropolitan Statistical Area, comprised of the central city and the county-level political units of the surrounding suburban ring.

From 1945 to present — the Freeway Era in Adams' four-stage model — the automobile culture was born. Automobiles, previously considered a luxury or recreational diversion, became a necessary transport mode for commuting, shopping, and socializing. Massive production by the automobile industry supplied mobility to urban dwellers. High-speed and limited-access freeway system increased accessibility from the central business district (CBD) to any location (or vice versa) via the expressway network. As increased accessibility extended commuting distance to cover the broad suburban area, metropolitan cities sprawled out into the surrounding, formerly agricultural lands with new residential developments. In the 1950s and 1960s, large cities had encouraged the construction of radial expressways (or beltways, which were originally designed to divert interstate highways from the congested urban core) to maintain continued accessibility to the downtown CBD for the suburban population.

As superior intra-metropolitan accessibility became a spatial good available at any expressway-interchange location, non-residential activities, led by manufacturing and retailing, gravitated toward the intersections of the interstate highway and express beltways and greatly accelerated the deconcentration of these new developments. Muller (1986) claimed that these downtown-like suburban concentrations of retailing, business, and light industry were common features of mini-cities surrounding the central city. Moreover, major new multipurpose activity centers had been emerging in the outer cities. These outer cities provided the totality of urban goods and services to their surrounding populations and became increasingly self-sufficient functional entities. Figure 2.1 illustrates the existence of several new outlying metropolitan-level cores with the boundaries of development according to different historical background.

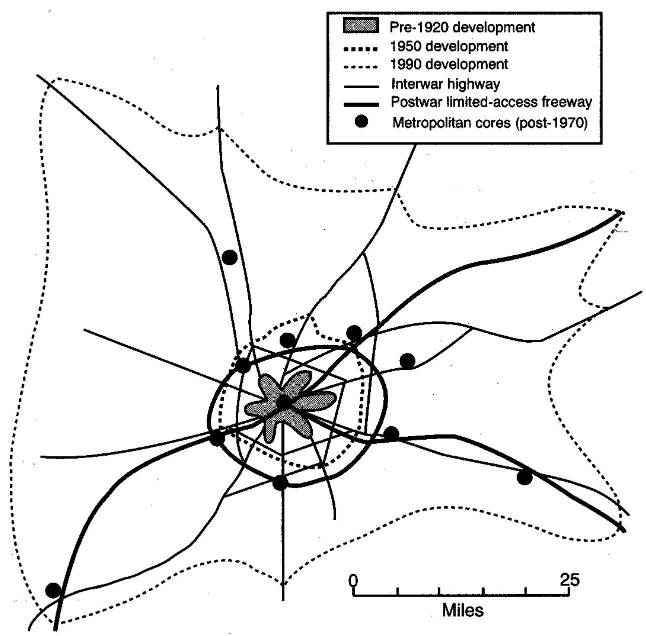


Figure 2.1 Spatial Pattern of Growth in Automobile Oriented Suburbia since 1920 (Source: Muller 1986)

The debate concerning changes to America's urban structure that caused suburban sprawl began 40 years ago and continues today. Moynihan (1960) considered interstate highways, which were under construction at the time of his writing, a driving force that would have a negative impact on land use patterns. He predicted that without proper planning, massive highway construction would draw population, employment, and activity centers out of the central city and into the suburbs, causing the central city to decline. The issue, however, is much more complex than the relationship strictly between highways, auto ownership, and land use. Other scholars have argued that additional, equally important factors also contributed to suburbanization, including dependency on property tax and inadequate representation of city governments in the legislature of some states (Meyer 1968), as well as "…cheap credits, favorable mortgage loan terms, accumulations of savings, rapid family formation, the postwar baby boom, favorable tax treatment, a strong preference for home ownership, and the suburbanization of an ever larger number of jobs" (Kain 1970).

2.1.2 Impacts of Transportation on Land Use

2.1.2.1 Land Use Changes

Wegner (1995) distinguished eight types of major urban subsystems including networks, land use, work places, housing, employment, population, goods transport, and travel, by the rate of change. Wegner suggested that, when considering the rate of change of different phenomena in a spectrum, networks and land uses would be slow to change, population and employment would change rapidly, and work places, housing, goods transportation, and travel would have a medium rate of change, as illustrated in Figure 2.2.

Very Slow	Slow	Intermediate	Fast
Networks Land use	Work place Housing	Goods transportation Travel	Population Employment

Figure 2.2 Rates of Change of Eight Types of Major Urban Subsystems

Payne-Maxie Consultants (1980) conducted a study, commissioned jointly by the U.S. Department of Transportation (USDOT) and the U.S. Department of Housing and Urban Development (USHUD), to assess the impacts of beltways on land use and urban development. The study involved a comparative statistical analysis of 54 metropolitan areas and detailed case studies of eight beltways. Data were collected in seven categories including general economic and demographic information, employment and investment figures, retail trade statistics, commuting information, highway and beltway descriptions, socioeconomic indexes, and residential moving patterns. The database included indicators of several beltway characteristics such as the presence of a beltway, length in miles, number of interchanges, interchange density per mile, percentage of completed beltway construction, age of beltways, and location in terms of both distance from the CBD and political jurisdiction (central city or suburb) where the beltway was located. Two types of statistical analysis techniques were employed to test the beltway influence. The first was analysis of variance (ANOVA) using either two or four sub-groups of the data. The second was multiple regression using beltway characteristics as well as other independent variables in analyzing joint relationships and beltway influence on dependent variables, which were percent changes in urban population, wholesale employment, retail sales, value added by manufacturing, and service employment over a given time period. Since economic cycles and inflation that affected certain variables such as manufacturing capital investment or retail sales had been relatively consistent throughout the country over a given period time, they were not taken into consideration. The analysis of growth and economic activity in 54 cities showed that beltways might have had some influence on metropolitan development patterns. Beltway location, mileage, distance from the CBD, or the density of interchanges showed more effects on land use changes than the presence of a beltway alone did. Significant relationships were found between a beltway's location and the growth of manufacturing, wholesale, and selected service employment in central cities. The influence of beltway location on the workplace of the suburban labor force and that of the length of a beltway on vehicle miles traveled (VMT) per capita were also found to be significant. Other factors affecting VMT per capita included population density, the availability of transit, the geographic size and population of the metropolitan area, per capita income, and a group of highway system descriptors, including total freeway mileage, beltway mileage, and interchanges per mile of beltway. New developments tended to take a place on the urban fringe, where beltways were usually built. It was found that the greatest proportion of suburban residential growth occurred in cities without beltways. Selected beltway studies are summarized in Table 2.2.

Table 2.2 Selected Beltway Study Results

Dependent Variable	Significant Independent Variables
Beltways' Growth Inducing Effects	
Percent change in population SMSA* 1960 – 1970	Percent of beltway located in suburbs (+) Percent change in manufacturing employment (+) City age index (-)
Percent change in population SMSA* 1970 – 1977	Percent change in manufacturing employment (+) City age index (-) Arterial mileage per capita
Percent change in wholesale employment SMSA 1972 – 1977	Distance of beltway from CBD (+) Change in manufacturing employment (+) City age index (-) Principal arterial mileage per 1000 persons (-)
Beltways' Distributional Effects	
Percent changes in central city population, 1960 – 1970 and 1970 – 1977	Percent change in manufacturing employment (+) City age index (-)
Percent change in value added by manufacturing in central city, 1967 – 1972	City age index (-) Beltway mileage (-)
Percent change in retail sales in central city, 1967 – 1972	Percent change central city selected services employment (+) Annual increase per capita income, 1969 – 1974 (+) Principal arterial mileage per 1000 persons, SMSA 1975 (+) Transit vehicles per 1000 persons, 1976 (-)
Percent change in retail sales in central city, 1972 – 1977	Percent change central city population (+)
Percent change in central city wholesale employment, 1972 – 1977	Interchanges per beltway mile (-) Change in central city population (+) Change in central city manufacturing employment (+)
Percent change in central city service employment, 1960 – 1970	Interchanges per beltway mile (-) Beltway distance from CBD (+) Change in central city wholesale employment (+)

* Standard Metropolitan Statistical Area
(+/-): Sign of the regression coefficient

Cosby and Buffington (1978a) conducted research on the effects of a highway improvement project (as opposed to construction of new highways) on land use. The improvement project, taking place in an area of College Station, Texas between July 1972 and April 1974, increased the capacity of State Highway 30 by adding two lanes, a paved shoulder, and a striped median to the existing facilities. The authors analyzed land use changes of both abutting and non-abutting properties on both sides of the highway. The study area was approximately 581 acres and was classified as undeveloped before the improvement. Data were collected for approximately four and one-half years covering the period before construction began up to the end of 1977. The total acreage in each type of land use was estimated for each year, and the types and rates of development before and after the improvements were compared. The study found that after the improvement project was approved, a large increase in multi-family residential development occurred and most of the changes were on lands that abutted Highway 30. It was also shown that commercial and public/governmental use on the abutting land increased. Total improved acreage increased from 13.91 acres in 1968 to 160.68 acres in 1977. Widened and improved Highway 30 was found to be a most important factor influencing land use by providing better access to both abutting and non-abutting properties, therefore making the study area a more likely and desirable place for the developments. In addition to the highway improvements, other causes of development in the area were identified as the growth of Texas A&M University and the proximity of Highway 30 to Texas Avenue, the main street in College Station.

Cosby and Buffington (1978b) conducted another study to examine the impact of improvements on existing roadways in a predominantly residential area of Bryan, Texas. The improvement project upgraded East 29th Street from a two-lane to a four-lane street for an approximately one-mile section. Construction began in 1966 and was completed in 1968. The traffic counts at four locations on East 29th Street in the study area ranged from 1,270 to 1,780 vehicles per day in 1957, which increased to the 5,130 to 5,800 range in 1970, and to the 7,400 to 8,970 range in 1977. Land use data in the study area were collected for 1958, 1965, 1970, and 1977. Approximately 256 acres covered abutting and non-abutting properties on both side of the street. The data were divided into three periods: the "before" period from 1958 to 1965, the "short-term after" period from 1965 to 1970, and the "long-run after" period from 1970 to 1977. The yearly rate of change was calculated for each period. Changes occurred at the highest rate during the "short-term after period" covering the construction years and two years afterwards, but diminished in the "long-run after" period. It was concluded that the increased access and capacity of the street attracted commercial and multi-family residential developments, which was supported by the realtors interviewed. Several reasons behind the decreasing rate of change in the long-run after period were identified as the small amount of unimproved acreage available, a shift in the primary commercial district, and the construction of a loop that diverted traffic away from the area.

Herndon (1979) studied the impacts of roadway improvements in a developed single-family residential area in Houston, Texas. In 1974 and 1975, a 1.05-mile section of the West 43rd Avenue was improved from a two-lane, open drainage road to a four-lane street with curbs and gutters. The study area covered 495 acres of the abutting and non-abutting areas on both side of the improved portion of the street, an area that was approximately 3,600 feet wide and 6,000 feet long. Land use data were collected for 1960, 1965, 1973, 1975, and 1978. Because most of the changes in land use in terms of predominant land use, enforced subdivision restrictions, quality of buildings, and lack of unimproved property occurred before the street improvement began and there were few of these types of changes after the improvement, it was concluded that the West 43rd Avenue improvement project had little impact on land use change and development in the study area.

Buffington et al. (1985) studied impacts of staged construction of two freeways located in Houston, Texas. Most freeways are commonly built in lateral or longitudinal stages. In lateral stages, service roads are constructed and opened to traffic before the main lanes, while in longitudinal stages, the service roads or main lanes are constructed on a freeway section by section. Two freeways, US 290 and US 59, and the one-half mile strip on either side of each of the freeways were selected as the study area. The study strip consisted of four sections and was divided into two parts: the abutting portion, which was within 100 ft of the freeway, and the non-abutting portion, which was the rest of the study strip. The historical land use data from the 1953, 1957, 1962, 1970, 1975, and 1980 records of the Houston City Planning Department and from aerial photographs of the U.S. Department of Agriculture were collected. A regression model was formulated with the percentage of five dominant types of land use, including single residential, multiple residential, commercial, industrial, and undeveloped land uses, as dependent variables. Six binary variables defining location, construction phase, construction type, and freeway capacity change and one continuous variable for average daily traffic were used as explanatory variables (*EV*). To consider interaction among land uses, each dependent variable (*DV*) was expressed as a function of other dependent variable(s) and explanatory variables as shown below.

$$DV_i = \alpha_i + \sum_j \beta_{ij} + DV_j + \sum_k \gamma_{ik} EV_k \qquad \text{for } i \neq j$$

where

i	=	type of land use, where $i = 1, \ldots, 5$;
j	=	type of land use, which is different from i;
k	=	number of explanatory variables, where $k = 1, \ldots, 7$;
EV	=	explanatory variables; and
α, β, γ	=	estimated coefficients.

Based on the regression analysis, it was concluded that single- and multi-family residential uses, as well as industrial uses, were sensitive to staged freeway construction. Residential land use had a significant relationship with freeway construction, while main lanes had a greater impact than service roads. The impacts on commercial and industrial development were similar.

Hirschman (1990) described a methodology to project and evaluate the potential land use impacts of a proposed limited-access highway extension linking the towns of Brockport and Albion to the central business district (CBD) in Rochester, New York. A gravity model of residential location choice was developed with accessibility index scores for subareas. The gravity model also considered other factors such as availability of developable land, net natural increase in population, anticipated economic growth, land use controls, infrastructure, tax rates, and local public services. It was formulated as follows:

$$G_j = G_t \frac{V_j A_j}{\sum_{i=1}^{n} V_i A_i}$$

where

G_j = population growth increment allocated to subregion j;

G_t = total growth projected for the region;
V_j = a product of attributes that interacts with accessibility and affects residential population; and
A_j = a composite weighted travel time between subregion j and all subregions as shown below:

$$A_j = \sum_{i=1}^{n} \frac{E_i}{T_{ij}^a}$$

where

E_i = employment in each subregion;
T_{ij} = the travel time between subregion j and the other subregion; and
a = exponential time-impedance parameter, usually equal to about 2.0 in most applications of this technique.

Employment data from the 1980 Census Urban Transportation Planning Package (CTPP) and zonal highway travel time data from the Rochester regional transportation model was used to calculate the value of A_j for each town in the study area. A survey of businesses inside and outside the corridor and personal interviews were used to evaluate business impacts. The survey included a review of the competitive advantages of the area with and without the highway extension. To assess retail development possibilities, the route 104 Webster Highway corridor northeast of Rochester was chosen for comparison with the Brockport-Albion corridor.

There exists a large body of work on the impacts of transit investments on land uses. However, due to the limited services offered by the Miami-Dade County Metrorail and low transit mode share, this research is focused on highway investments. Only a brief summary of the literature is provided here.

Summarizing literature on the relationship between transit and urban forms, Cervero and Seskin (1995) provided the following insights:

(1) Fixed guideway transit investments themselves did not bring about new growth, but rather redistributed it. Additionally, transit investments rarely induced residential development. Commercial developments occurred more often as a result of efforts to develop areas around transit stations.
(2) Transit investments strengthened the downtown business core by improving accessibility from suburbs to the downtown; however, they encouraged decentralization of activities by creating "subcenters" or "edge cities" surrounding the station areas at the same time.
(3) Bus transit (including busways) and light rail systems generally had a lower impact on land use than heavy rail due to the relatively limited improvement from these modes to accessibility and transit competitiveness against private automobiles.
(4) Transit investments, when combined with strong policies to encourage high-density developments, might produce a positive impact on land use in the station areas by improving accessibility to employment for a larger portion of the population.

Knight and Trygg (1977) argued that there were many factors in addition to rail transit investments that affected land use change and development intensity. Other favorable conditions such as local government land use policies, regional development trends and forces, availability of developable land and physical characteristics of the area played important roles in conjunction with increased accessibility resulting from a rail transit investment. Figure 2.3 depicts some of the factors influencing the land use impacts of transit.

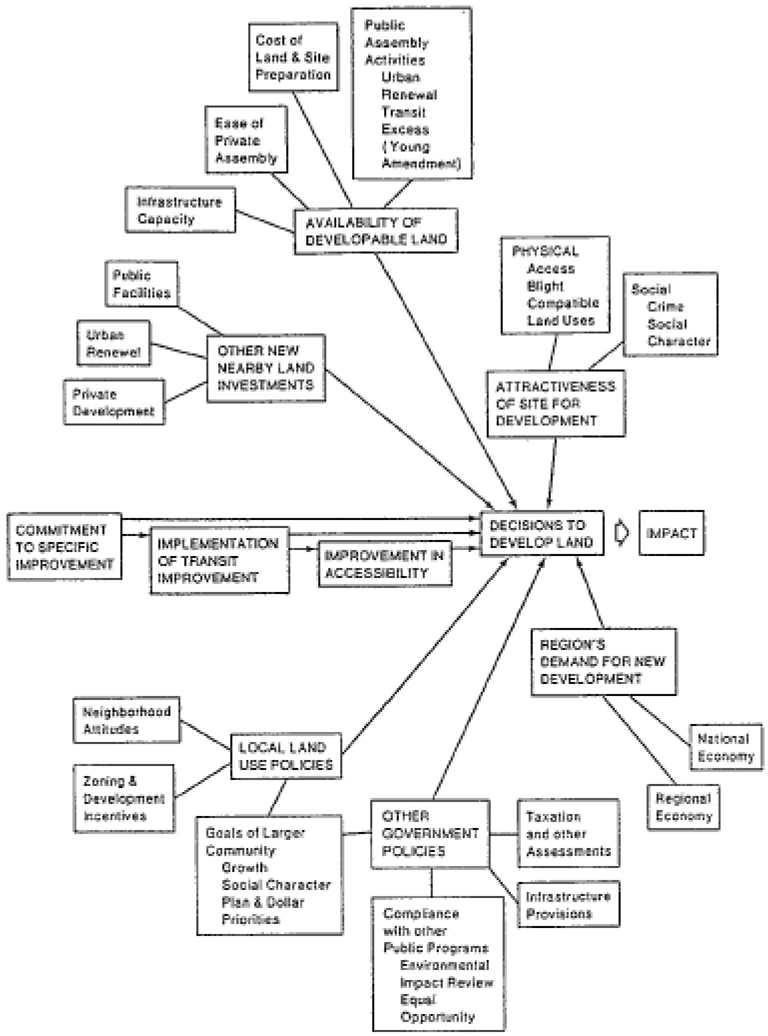


Figure 2.3 Factors Influencing Land-Use Impacts of Transit (Source: Knight and Trygg 1977)

2.1.2.2 Property Values

Improvement in accessibility may increase the attractiveness of an area thus increasing the value of real estate, which may in turn attract more developments. Huang (1994) noted that two techniques had been used to identify the effects of public infrastructure at fixed locations on property values: (1) hedonic price modeling based on cross-sectional data, and (2) analyses of longitudinal data on property value change over time. The author conducted an extensive review of empirical studies on estimation of the effects of transit and highway investments on property values. Tables 2.3 presents a summary by Huang of selected studies related to highway impacts on property values. The author concluded that highway interchanges and public transit stations positively impacted land value, but the extent and magnitude of the impacts varied. This research was focused on the highway impacts therefore literature on transit investments, particularly fixed guideway systems such as light rail, heavy rail, and commuter rail, was not reviewed.

Ryan (1999) provided another review of empirical studies over the past four decades on how transportation facilities affected property values. To explain inconsistent results from those empirical studies, Ryan followed Guiliano's (1988) categorization that divided highway impact studies into two groups: (1) first generation studies conducted in the 1950s and 1960s and (2) second generation studies conducted in the 1970s and 1980s. The first generation studies employed experiment-control methods, while the second generation studies generally used regression analyses. Similarly, transit impact studies were also divided into first and second generation studies: (1) heavy rail and (2) light rail. The author recommended the use of travel time savings as the measure of transportation access. When these savings were accurately measured, property values tended to show the theoretically expected relationship with transportation access. The author noted that measuring accessibility in terms of the distance of a given property from a given transportation facility might not show the relationship correctly. Ryan's (1999) meta-analysis showed that the values of properties located within one mile of highways were affected, while transit impacted the value of properties located within one-third (0.33) of a mile.

Buffington *et al.* (1992) estimated the economic impacts of proposed improvements to U.S. Highway 287 in Wichita Falls and State Highway 199 in northwestern Tarrant County, Texas. This impact study examined traffic-serving businesses and other types of businesses separately through variables such as property values, new development, relocation and employment, municipal tax revenues, and highway users. Effects of the proposed route and/or alternatives were primarily compiled from the literature review. Personal interviews and mail questionnaires were used along with data from the U.S. Bureau of Census, Texas Almanac, Chambers of Commerce, and city offices. Methodologies to estimate each economic impact listed above were described.

Swenson *et al.* (1998) claimed that there was a strong correlation between transportation infrastructure investments and population, population density, total property values, and growth per square mile in property values. The authors used GIS techniques to display and examine the spatial economic growth patterns based on property tax data collected annually between 1987 and 1995 in a nine-county area in the central Iowa region. The spatial pattern of economic changes was linked with the roadway network from TIGER/File data to examine the changes in investments in transportation infrastructure during the same time period. Variables associated with each road segment included total roadway length, changes in length, density of roadway

Table 2.3 Property Value Effects Based on Proximity to Highway (Source: Huang 1994)

Author (Year)	City/Transit System	Property Type	Accessibility	Magnitude of Effect	Extent of Effect
Adkins (1957)	Dallas, TX	Land	Distance bands around Right of Way (ROW)	Highway proximity tends to be associated with higher assessed values and sale prices	4 blocks beyond expressway frontage roads
Cribbins, et al. (1962)	Cumberland, Guilford, & Rowan Counties, NC	Land	Straight-line distance from interchange; Straight-line distance to ROW	No discernable pattern which would have allowed documented value increases to be attributed to highway construction	
Buffington (1964a)	Austin, TX	Unimproved land	Highway corridor versus control area	Highway responsible for premium equal to 163% of original land value, based on hybrid price increase index. Very strong positive effect on abutting land.	Band around highway ROW, ~0.5 mi on either side of ROW
		Subdivided land	Highway corridor versus control area	Highway responsible for discount equal to 13% of original value, based on hybrid price increase index. Effect on abutting land vs. non-abutting land unclear	Band around highway ROW, ~0.5 mi on either side of ROW
Buffington (1964b)	Temple, TX	Land	Highway corridor versus control area	Highway responsible for premium equal to 2562% of original land value, based on hybrid price increase index. Abutting land values increased more than non-abutting land values.	Band around highway ROW
Brown & Michael (1973)	Indianapolis, IN	Land	Distance rings (based on straight-line distance from interchange)	Interchanges had a positive effect on value that decreased with distance	1 mile from interchange
Allen (1981)	Northern VA (Washington, DC area)	SF homes	Noise level exceed 10% of the time various noise measures	House value decreased by $94/decibel	
	Tidewater, VA	SF homes		No significant effect	
Langley (1981)	North Springfield, VA (Washington, DC area)	SF homes	Distance bands (based on straight-line distance to roadway)	Proximity to roadway associated with lower values: -$3000 to -$3500 per house	0 to 343 m from roadway
Palmquist (1982)	Washington State	SF homes	Study area versus control area	12% - 15% higher appreciation in study area versus control area	Study area (up to 1 mi from roadway)
			Noise contours	Up to 7.2% lower appreciation than other study area houses	0 to 600 ft from roadway
Tomasik (1987)	Phoenix, AZ	SF homes	Study area versus local averages	No discernable negative property value effects. Homes within 0.5mi of highway appreciated more than distance houses. But within 0.5 mi. corridor, no correlation between unadjusted value and distance	0 to 0.5 mi. from highway

development, and the number and change in number of interchanges. The authors concluded that high growth urbanized clusters were located at the fringes of a growing metro core city, and that non-urbanized cities located within 20 miles of the metro core grew strongly. Additionally, residential property values declined in the core city while growing dramatically in the suburban fringes. Commercial property values greatly increased in the metro core and even more so in the western suburbs. These findings were consistent with the view that suburbanization, which occurred in most U.S. cities and was encouraged by the development of highways, resulted in the decline of the central city and created multiple activity centers that competed with the traditionally dominant downtown core.

Vadali *et al.* (2001) performed a study on an approximately 9.2-mile segment of the recently reconstructed North Central Expressway (NCE), to evaluate the impact of transportation improvements. Sales data on single-family residential properties between 1979 and 1997 in the study area were geocoded on the Dallas County GIS street maps and the location information was extracted. The entire NCE project consisted of five phases including a pre-planning phase (1979-1983), planning phase (1984-1986), pre-construction phase (1987-1989), construction phase I (1990-1994), and construction phase II (1995-1997). Regression models were developed with the sales data during each project phase between 1979 and 1997, in the following form:

$$Y_j = X_j\beta_j + N_j\delta_j + L_j\gamma_j + D_j\eta_j + U_j$$

where

Y_j = logarithm of the home sales price, j = 1, 2, 3, 4, 5;
X_i = structural attribute variables, j = 1, 2, 3, 4, 5;
N_j = community variables, j = 1, 2, 3, 4, 5;
L_j = location/environmental variables, j = 1, 2, 3, 4, 5;
D_j = time dummy variables;
U_j = error term; and
$\beta, \delta, \gamma, \eta$ = estimated parameters.

Three different sets of location/environmental variables were prepared. The first set included five dummy variables representing each of the sections that were divided based on their distances from the freeway. The second set was defined by a continuous measure of distance from the expressway. The third included sub-location variables based on the area defined by a given distance from the road near the property. After models were developed, three types of hypothesis tests were performed. First, the authors tested the hypothesis of poolability across construction phases for temporal stability of all parameters, which was rejected in all cases at a 0.01 significance level, implying that it was not appropriate to combine any of the phases. Second, sequential (intra-location) stability across project phases was tested. The results showed that property values decreased in the planning period, pre-construction period, and construction period I. Third, the authors tested sectional (inter-location) stability of coefficients for dummy variables representing sections and sub-location variables. Significant differences were found only during construction periods I and II. For the sub-location model, the null hypothesis was rejected in all cases except for the pre-planning phase. From these tests results, the authors concluded that property values were affected to a greater degree due to traffic-related effects than because of the distance from freeway.

In a study of home sale prices during the 1980s in Montgomery County, a suburb of Philadelphia, Voith (1992) found that homes commanded higher sales prices when they were located in places where highway travel time to the Philadelphia central business district was less, all else being equal.

2.1.3 Location Choice

Location choice theory has played an important role and has contributed theoretically and empirically to regional science. It focuses on the "where" and "why there" questions of human activities with respect to both households and businesses and on location patterns. Headicar (1996) considered it fundamental to the evolution of urban settlements that the location of new developments was determined largely by the opportunities for interaction created by the transportation system. These opportunities have increased over time with improvements in transportation technologies and investments in the transportation networks. Geographically, however, the pattern of growth has not been uniform, but biased towards particular types of locations where investments tend to be concentrated:

- Specific corridors rather than roads connecting urban centers (because of economy of scale in the provision of new infrastructure);
- Locations between and around cities and conurbations rather than within them (because of the differences between costs and public acceptability of construction work of new developments between inside and outside of cities and conurbations; also because of the greater share of responsibility held by the federal government for major roads outside cities and conurbations); and
- Specific locations at the edge of individual towns, near the intersection of new inter-urban or by-pass routes and the connecting urban road network (because of the limited opportunities to access the new inter-urban or by-pass routes).

Cascetta *et al.* (2000) introduced a system of models, which consisted of three integrated submodels including a travel demand model, a residential location model, and an activity location model. Figure 2.4 depicts the interactions between each of the sub-models in the system.

The residential and activity location models were estimated based on a behavioral approach consistent with the random utility theory. In the residential location model, the number of people living in zone *o*, *R(o)*, was calculated with the following equation:

$$R(o) = k(o) \cdot \sum_i P^i_{res}(o) \cdot \sum_{d'} E^i_{tot}(d')$$

where

$k(o)$ = the ratio between residents and workers in zone o;
$P^i_{res}(o)$ = the probability that each worker i chose zone o as their residential zone; and
$E^i_{tot}(d)$ = the total number of employed i in zone d.

.

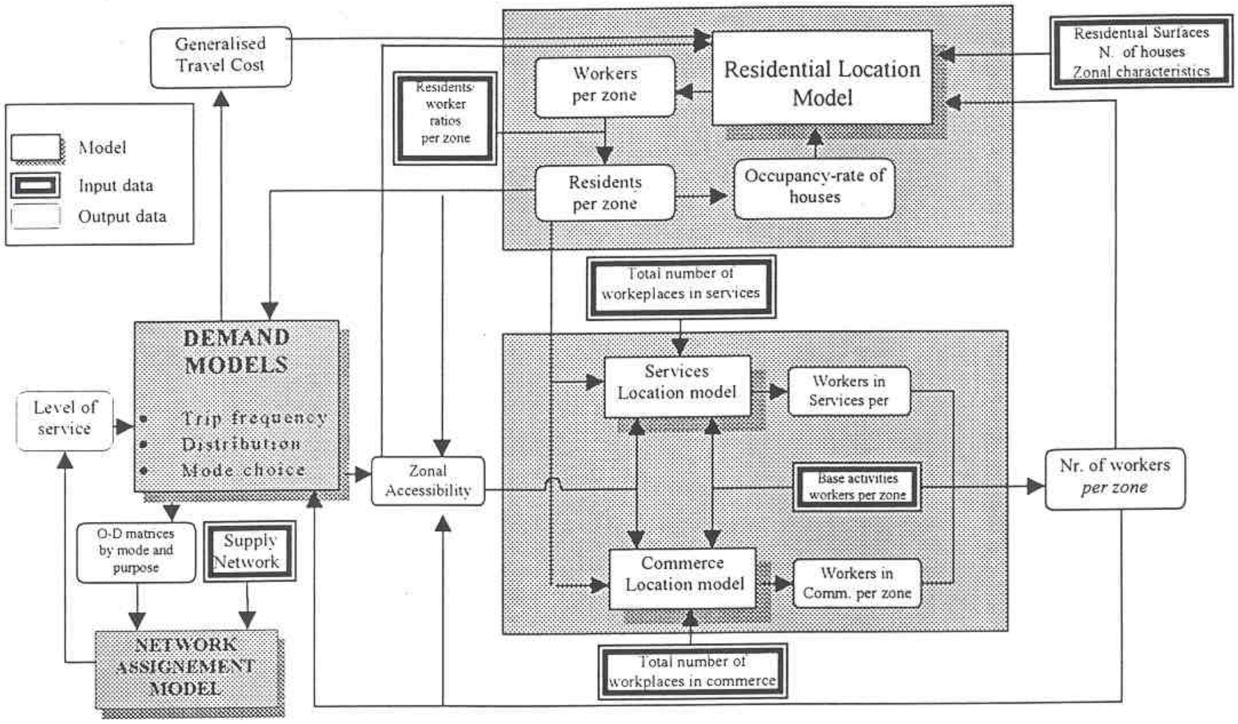


Figure 2.4 Overall Framework of the Models System (Source: Cascetta et al. 2000)

The worker equation was given as follows:

$$P^i_{res}(o) = \sum_d P^i_{res-cond}(o/d) \cdot P^i_{lav}(d)$$

where $P^i_{res\text{-}cond}(o/d)$ was the probability that worker *i* chose to live in zone *o* conditional to working in zone *d,* and $P^i_{Lav}(d)$ was the probability that worker *i* worked in zone *d*.

These probabilities were estimated using following equations.

$$P^i_{Lav}(d) = \frac{E^i_{tot}(d)}{\sum_{d'} E^i_{tot}(d')}$$

$$P^i_{res-cond}(o/d) = \frac{\exp(V^i_{o/d})}{\sum_{o'} \exp(V^i_{o'/d})}$$

where $V^i_{o/d}$ was the systematic utility, which assumed a linear combination of attributes extracted from interviews conducted in the city of Rome, Italy. These attributes included (1) the inclusive value of mode choice between an o-d pair for workplace purpose; (2) a dummy variable representing whether the residential zone and the employment zone belonged to the same zonal area; (3) the housing price in zone *o*, the occupancy rate of houses in zone *o*; (4) the number of houses in zone *o*, which was used as an indicator of the quality of estate of zone *o*; and (5) a dummy variable for prestigious zones.

The activity location models were calibrated based on census data. The probability $P^a(d)$ of locating activity a in zone d is

$$P^a(d) = \frac{\exp(V_d^a)}{\sum_{d'} \exp(V_{d'}^a)}$$

where $V^a{}_d$ was the systematic utility, which was assumed to be a linear combination of the following attributes: the passive accessibility of zone d, residents in zone d, a dummy variable for the city center zone, and the total number of workers in zone d. Among these variables, the passive accessibility (A_d^{pas}) was calculated as follows:

$$A_d^{pas} = \sum_o R(o)^{\gamma_1} \cdot \exp(\gamma_2 \cdot Y_{Other}(o,d))$$

where

$Y_{Other}(o,d)$	=	the inclusive values of mode choice for others purposes (i.e., shopping, personal care, …); and
$\gamma 1,\ \gamma 2$	=	calibrated parameters.

Results of the model tests showed that the errors of residential location ranged from -8% to +8% and the errors of employment location from -4% to +4%. The authors noted that the subdivision of activities (i.e., subdividing activity centers into "wholesale" and "retail") yielded better results.

Kawamura (2001) assessed the difference in location choice patterns based on data from 1981 and 1999 in the six-county Chicago region (Cook, DuPage, Kane, Lake, McHenry, and Will counties) by quantifying changes in the location pattern of various business sectors in relation to transportation facilities. The location and business profile of firms during this period were used to compare proximity to both freeway ramps and transit stations during those two years. The findings indicated that, after controlling non-transportation factors (such as agglomeration effects), factor prices (such as rent, land price, and wage), decentralization, and density of transportation facilities, businesses looked for places closer to freeway ramps as they moved farther away from the central business district (CBD). With the exception of the CBD, most areas within the City of Chicago lost employment, while the western and northwestern suburbs located approximately 20 to 50 kilometers from the CBD made substantial gains in employment. The author noted that the CBD had gained employment in recent years in Chicago, reflecting a revitalization of the city core.

2.1.4 Induced Travel

In recent years, the phenomenon often referred to as "induced travel" (also known as "generated traffic") has attracted much attention among transportation planners. Induced travel is the additional travel resulting from improvements to transportation facilities intended to provide greater capacity. This increased capacity will initially result in reduction of congestion on the improved facilities. However, traffic on the new facilities tends to gradually increase until congestion again reaches high levels, while at the same time there is no obvious reduction in congestion on the surrounding roads. The question therefore arises: will new transportation facilities or increased capacity result in more travel, thus consuming all the benefits of the investments in these improvements and resulting in more negative environmental consequences? Many researchers have attempted to answer this question by examining: (1) which factors, in addition to added capacity, may be responsible for the changes in travel demand; (2) whether there is merely redistribution of trips while the total travel (e.g., measured by VMT) remains a constant; (3) whether the change in travel is a result of modal switching; (4) whether the demand actually changes due to increased accessibility; (5) how much increase in demand will be induced by added capacity; and (6) what are the implications of induced travel on land use. Many studies demonstrate that, after factors such as population growth are accounted for, new highway construction does induce additional travel. Yet, how much induced travel is caused by added capacity and the mechanism of induced travel remain to be fully understood.

The theory of induced travel is based on microeconomics. Induced travel is a response to the reduced price of travel. As in the case of any commodity, the demand (in this case travel) will increase when the price (travel cost) falls. Added demand may be the result of longer trips made possible by shorter travel time, mode shifts, new trips that have been suppressed by the previous level of congestion, etc.

2.1.4.1 Definitions of Induced Travel and Travel Behavior

There is a general agreement that induced travel is real. However, there is no clear definition of induced travel or universally accepted units to measure induced demand. Thus, the extent of induced travel has not been clearly determined. In TRB Special Report 245 it was claimed that "the state of knowledge and modeling practice are not adequate for predicting with certainty the impacts of highway capacity additions" (TRB 1995). Additionally, the report concludes that induced demand effects are "highly dependent on specific circumstances" including location of highway improvement projects, extent of congestion, conditions present for regional growth, and consequences of alternative scenarios.

Dunphy (1998) plotted the congestion indices developed by Lomax and Schrank (1996) of the Texas Transportation Institute, and regressed per capita VMT against the congestion index. If induced travel was real, then areas with lower congestion levels could expect higher levels of VMT. However, the simple regression yielded $R^2 = 0.0062$, indicating no statistically significant linear relationship between VMT and congestion level. Additionally, the regression line had an upward trend with increased congestion levels associated with high levels of driving. Dunphy admitted that this analysis was not scientifically sound, and pointed to the need to better understand the relationship between congestion level and travel behavior, as well as the need for good analytical tools.

Cohen (1995) stated that, on average, induced travel effects in larger areas were greater than for individual facilities. According to the "triple convergence principle" of Downs (1962), three immediate effects of highway expansions were route, time, and mode convergence. The U.K. Department of Transport (1993) identified five major attributes of induced travel when highway capacity was increased: new trips, trip redistribution, changes in mode used to travel (e.g., from transit to automobile), route reassignment, and the time of day of a trip. Abelson and Hensher (2001) argued that the first three attributes (i.e., new trips, trip redistribution, and changes in mode used to travel) resulted in VMT increase on both the improved link(s) and over the entire road network. The other two attributes of induced travel, i.e., route reassignment and time of day when a trip took place, increased traffic on certain routes or at certain times, which was considered as induced travel increased VMT either on the network as a whole or on that part of it where the transportation facility was improved. This finding has been confirmed by many other studies. DeCorla-Souza and Cohen (1998) also defined induced travel through these variables.

In contrast, Strathman *et al.* (2000) defined induced travel as a direct or indirect causal relationship between road capacity increase and increased travel measured by trip frequency or trip distance. The authors claimed that "induced traffic" and "induced travel demand" should be distinguished, the former being the effect in the short-run while the latter the long-run effect. Barr (2000) defined induced travel as any increase in highway system use caused by an addition to highway capacity or other transportation system change that resulted in reduced travel time and/or costs. Barr also claimed that, since shifts in the time of day of a trip did not result in a net increase in highway system use, induced travel did not include these changes, while mode shifts from transit to private automobile contributed to induced travel. Fulton (2000) considered "induced travel" to be the increase in vehicle travel on the roadway above the level that occurred before the capacity addition resulted from additions to roadway capacity. One definition of induced travel by DeCorla-Souza and Cohen (1998) is the "increase in daily vehicle miles of travel, with reference to a specific geographic context, resulting from expansion of highway capacity."

Litman (1999) summarized different types of generated traffic and their nature and impacts on travel and costs, as given in Table 2.4, which included the behavioral aspects of induced travel. The study of behavioral changes is perhaps more complicated than studying the relationship between added capacity and travel volume (e.g., measured by VMT), as more data on travelers may be necessary but often not readily available.

Heanue (1998) claimed that the increase in transportation system capacity did not directly influence travel behavior. However, improved capacity would result in reductions in travel times and costs, and in turn any changes would result in changes in travel decisions. Location, nature, and timing of activity growth and land use changes due to a new transportation facility took the most time to occur. The author concluded that highway capacity expansion interacted with variables such as population, household and employment growth, personal income, auto ownership increases, regional economic growth, and fuel price changes. These variables were determinants of total travel demand, and the major determinants of travel demand were clearly socioeconomic such as growth in population and households, labor force participation and employment, income, auto-ownership, vehicles owned, and licensed drivers. Heanue argued that the induced travel due to new highway capacity could be addressed by considering behavioral and land use change mechanisms.

Dowling and Colman (1998) conducted a survey along with interviews to determine people's travel behaviors in response to changes in travel time. A total of 676 individuals over the age of 16 were randomly selected from the San Francisco and San Diego metropolitan areas. Changing route was the most preferred responses to travel time changes. Other behaviors included changing schedules, consolidating trips, changing modes, and changing destinations. In the survey and interviews, travel time change was provided in increments of 5 minutes for both travel time savings due to increased capacity and travel time increases due to congestion. Based on the survey results, the authors concluded that "current travel forecasting practice probably results in an under-prediction of three to five percent in the number of trips that may be induced by major new highway capacity projects." They also pointed to longer term effects of major highway capacity projects, such as auto ownership, residential, employment, and business location choices.

Table 2.4 Types of Generated Traffic (Source: Litman 1999)

Type of Generated traffic	Category	Time Frame	Travel Impacts	Cost Impacts
Shorter route: Improved road allows drivers to use more direct route	Diverted trip	Short term	Reduction	Reduction
Longer route: Improved road attracts traffic from more direct routes	Diverted trip	Short term	Small increase	Slight increase
Time change: Reduced peak period congestion reduces the need to defer trips to off-peak periods	Diverted trip	Short term	None	Slight increase
Mode shift: existing travel choices Improved traffic flow makes driving relatively more attractive than other modes	Induced vehicle trip	Short term	Increased driving	Moderate to large increase
Mode shift: changes in travel choice Less demand leads to reduced rail and bus service, less suitable conditions for walking and bicycling, and more automobile ownership	Induced vehicle trip	Long term	Increased driving, reduced alternatives	Large increase, reduced equity
Destination change: existing land use Reduced travel costs allow drivers to choose farther destinations	Longer trip	Short term	Increase	Moderate to large increase
Destination change: land use changes Improved access allows land use changes, especially urban fringe development	Longer trip	Long term	Increase driving and auto dependency	Moderate to large increase, equity costs
New Trip: No land use change Improved travel time allows driving to substitute for non-travel activities	Induced trip	Short term	Increase	Large increase
Automobile dependency: Synergetic effects of increased automobile oriented land use and transportation system	Induced trip	Long term	Increase driving, fewer alternatives	Large increase, reduced equity

Table 2.5 Summary of Definitions of Induced Travel

Author (Year)	Broad Definition of Induced Travel	Variables Included in Definition
Downs (1962)	Three immediate effects of highway expansions are route, time, and mode convergence ("triple convergence principle").	route, time, and mode convergence
U.K. Department of Transport (1993)	Five major attributes of induced travel when highway capacity is increased are: new trips, trip redistribution, changes in mode used to travel, route reassignment, and the time of day a trip being made.	new trips, trip redistribution, changes in mode used to travel, route reassignment, and the time of day a trip being made
Cohen (1995)	Induced travel effects for larger areas are greater than for individual facilities	
DeCorla-Souza and Cohen (1998) Abelson and Hensher (2001)	New trips, trip redistribution, and changes in mode used to travel result in VMT increased on both the improved link(s) and over the entire road network. Increased VMT was found either on the network as a whole or on that part of it where the transportation facility was improved as induced travel	new trips, trip redistribution, and changes in mode used to travel, route reassignment, and time of day a trip being made
Heanue (1998)	The major determinants of travel demand are socioeconomic such as growth in population and households, labor force participation and employment, income, auto-ownership, and licensed drivers. Induced travel due to new highway capacity may be addressed by considering behavioral and land use change mechanisms	growth in population and households, labor force participation and employment, income, auto-ownership, vehicles owned, and licensed drivers
Barr (2000)	Any increase in highway system use was caused by addition to highway capacity or other transportation system change which resulted in reduced travel time and/or costs	travel time and/or costs
Fulton (2000)	The increase in vehicle travel on the roadway above the level that occurred before the capacity addition resulted from additions to roadway capacity	
Strathman *et al.* (2000)	Induced traffic is the effect in the short-run while induced travel is the long-run effect	trip frequency or trip distance
Boarnet and Chalermpong (2001)	Induced travel is partially caused by land use changes when they are linked to increases in highway capacity.	land use changes, highway capacity

Boarnet and Chalermpong (2001) suggested that induced travel was partially caused by land use changes when they were linked to increases in highway capacity. They geocoded over 275,000 single-family properties located in Orange County, California for which transactions had been recorded between 1988 and 2000. The sales prices of the homes were regressed against square footage, number of bedrooms, number of bathrooms, lot size, structure age, SAT scores, crime rate, and a dummy variable of year threshold. Investigation of the straight-line distance from a property to the nearest toll road ramp revealed that there was a negative gradient in sale prices. The authors concluded that properties close to the new toll roads commanded higher prices, signaling the willingness of buyers to pay a premium for improved accessibility. This willingness to pay for accessibility influenced land development pattern, which was one of the causes of induced travel. It is unclear, however, if the authors controlled for population growth and land availability, which may also affect the property appreciation.

2.1.4.2 Measuring Induced Travel

Modeling the effect of added capacity is complicated by many behavioral or non-behavioral factors that may contribute to demand. For instance, non-behavioral factors that contribute to changes in the VMT include population, households, labor force participation, income, auto-ownership, licensed drivers, gasoline price, single-family home ownership, and density (Heanue 1998, Kiefer and Mehndiratta 1998). It is clear that some induced travel as a portion of the increase in the VMT in the last several decades was the result of changes in demographics and social trends, as well as changes in urban forms. These changes include the large number of women entering the workforce thus increasing work related travel, reduction in household sizes, increasing automobile ownership, and the deconcentration of urban population that further separates residences from workplaces, resulting in the need for longer commute thus more VMT. By calibrating a regression model using total VMT as the dependant variable and changes in total employment, urban population, median household income, and percentage of suburb-to-suburb flow as the independent variables, Kiefer and Mehndiratta (1998) found that these variables explained 57 percent of the variation in the increase in VMT. They concluded that as the first three trends (female workforce, household size, and automobile ownership) stabilized, they would not play a significant role in the future growth of travel.

Based on time-series travel data for various types of roadways, the SACTRA report (1994) reported that actual travel volume was 10 percent (in the short term) to 20 percent (in the long term) higher than forecast due to unaccounted induced travel. Johnson and Ceerla (1996) also studied California metropolitan areas and found the elasticities to be 0.6 to 0.9 over a period of three years. Cohen (1995) noted that travel time elasticities were generally between –0.1 and 0.0.

Noland (2001) analyzed data at the state level in the U.S. between 1984 and 1996 using a fixed effect cross-sectional time series model and a set of equations by road type. The variables included in the model were state population, per capita income, and cost per energy unit. Other variables such as state licensing rates were tested but were found not significant. The road types considered included controlled access highways, arterial, collector, and local roads, which were further disaggregated by rural and urban classifications. Different models were tested with the dependent variable, VMT per capita, lagging behind the independent variable, lane miles, since behavioral response to increased capacity would take several years. Models were also estimated for each road type individually and simultaneously as a set of equations with zero-, two-, and five-year lags. These models suggested that lane miles contributed to per capita VMT increases, that short-term (two-year) elasticity was greater than long term (five-year) elasticity, and that growth of VMT on freeways and local roads was faster than on arterials. The elasticities given by the simultaneous equation model with two-year lag were 0.567, 0.267, and 0.509 for interstate, arterial, and collectors, respectively. To account for cumulated impact of added capacity, distributed lag models were estimated, in which other dependent variables were assumed to have the same lagged impact as lane miles, and the lagged effect was estimated as an exponential function. The model gave short-term elasticity of 0.128 and long-term elasticity of 0.413.

In another study, Noland and Cowart (2000) analyzed annual data on metropolitan level congestion compiled by the TTI, urban land area, and lane miles of capacity. Instead of using aggregate data (e.g. metropolitan area-wide data), the authors attempted to take into consideration the congestion level on roadways based on the notion that only added capacity on congested roads would induce travel. A cross-sectional time series modeling approach was employed that included fixed effects across both urbanized areas and time. With the fixed effect method, it was not necessary to include all the variables that might have an effect on the dependent variable (Johnston and Dinardo 1997) and the simultaneity bias in the data could be minimized. An instrumental variable (two stage least squares) approach was used to further reduce the simultaneity bias. The general form of the models estimated was:

$$\log(VMT / PC_{it}) = c + \alpha_i + \tau_i + \sum_k \beta^k \log(X_{it}^k) + \lambda \log(LM / PC_{itl}) + \varepsilon_{it}$$

where

VMT / PC_{it}	=	VMT per capita for arterials and freeways in metropolitan area i, for year t;
c	=	constant term;
α_i	=	fixed effect for metropolitan area i, to be estimated;
τ_i	=	fixed effect for year t, to be estimated;
β^k	=	coefficients to be estimated for demographic and other parameters;
X_{it}^{k}	=	value of demographic and other variables for metropolitan area, i, and time, t;
λ	=	coefficient to be estimated for lane mile (LM) parameters;
LM /PC_{it}	=	proxy for cost of travel time (lane mile per capita) by metropolitan area i, for year t; and
ε	=	random error.

In the above model, the lane mile per capita was estimated using the following model:

$$\log(\boldsymbol{LM} / \boldsymbol{PC}_{itl}) = \boldsymbol{c} + \boldsymbol{\alpha}_i + \boldsymbol{\tau}_i + \boldsymbol{kIV}_{it} + \sum_k \boldsymbol{\beta}^k \log(\boldsymbol{X}_{it}^k) + \boldsymbol{\varepsilon}_{it}$$

where IV_{it} is the instrument specified across urban areas and time. After controlling for population size, population density, per capita income, fuel price, and for some of the models the lagged effects (dependent variable lagged by one year as an independent variable), the authors found that the number of lane miles accounted for about 15 percent of annual growth in VMT. However, results showed significant variations between metropolitan areas, which were not well understood. The authors hypothesized that these variations might be partially due to the varying growth rate in lane miles.

The 1995 National Personal Transportation Survey (NPTS) data set has been popularly used in studies to estimate the effect of expanded roadway capacity on induced travel demand. Strathman *et al.* (2000) used the cross-sectional approach to analyze the relationship between road capacity and VMT of 48 metropolitan areas by combining data from the 1995 NPTS and the TTI mobility study, accounting for the effects of residential location, employment location, and commuting mode choice. They contended that VMT was both directly and indirectly related to roadway capacity. The model for estimating the direct effect consisted of four equations with endogenous variables, including commute mode choice, residential density, work place density, and annual VMT, and a wide range of exogenous variables representing individual and household socioeconomic and demographic conditions, factors affecting travel options, and metropolitan characteristics. The direct effect of a one-percent increase in per capita roadway capacity was estimated to result in a 0.29-percent increase in VMT for the short-run period, when all other variables were controlled. It was estimated that the indirect effect of a ten-percent increase in roadway capacity led to a 0.033-percent increase in VMT as a result of capacity related residential and employment density changes. This study also found VMT to be related to vehicle ownership, household socioeconomic characteristics, urban scale, and commuting distance.

Marshall (2000) conducted an aggregate analysis to estimate the effect of roadway capacity expansion on induced travel using the TTI's Urban Congestion Study data set for 70 U.S. urban areas. Regression models for freeways and principal arterials utilized freeway and arterial VMT's per capita as the dependent variables and freeway lane miles per capita, arterial lane miles per capita, and urban area size as the independent variables. The adjusted R^2 for the freeway model was 0.851 and 0.813 for the arterial model, indicating that daily vehicle miles per lane mile were strongly correlated with population, area, and density. Freeway capacity, measured as lane miles per capita, was positively correlated with area and negatively correlated with density. Elasticities were estimated to be 0.85 for highways and 0.76 for principal arterials. Land use changes due to the addition of roadway capacity might cause long-term effects (or secondary effects) that lagged behind induced demand.

Hansen (1998) presented two previous studies performed by Hansen *et al.* (1993) and Hansen and Huang (1997) to estimate capacity elasticity of traffic. The former study estimated the elasticity at the road segment level. Eighteen highway segments with added lanes and increased capacity were selected and annual average traffic count data were collected for each segment between 1960 and 1990. Panel data were used to estimate a fixed-effect model, which related the state highway traffic volumes on the selected segments to the highway capacity, state highway VTM for the entire state system, and increase in highway capacity:

$$\log(Q_{it}) = \alpha_i + \beta \cdot \log(C_{it}) + \gamma \cdot \log(SQ_t) + \lambda \cdot \frac{NC_{it}}{t^{\sigma}} + \varepsilon_{it}$$

where

Q_{it}	=	the traffic volume of segment i in year t (measured from before the beginning or after the completion of the capacity expansion);
C_{it}	=	the capacity (number of lanes) of segment i at time t;
SQ_t	=	vehicle-miles traveled on the California state highway system in year t;
NC_{it}	=	the ratio of capacity added to total capacity for $t > 0$, and zero for $t < 0$;
$\alpha_i, \beta, \gamma, \Psi, \sigma^l$	=	coefficients to be estimated; and
ε_{it}	=	a stochastic error term, drawn from a normal distribution with mean zero.

This study pointed to a trend whereby VMT increased with the addition of capacity in terms of lane-miles of new highways.

Hansen and Huang (1997) measured elasticity at the metropolitan area level using data of VMT on state highways, state highway lane-miles, population, and per capita income for every urban county in the State of California for the years 1973 – 1990. The data were statistically analyzed at the county level and metropolitan level, controlling for factors such as county population and per capita income, together with panel data to estimate a fixed-effect model:

$$\log(VMT_{it}) = \alpha_i + \beta_t + \gamma \cdot POP_{it} + \psi \cdot PCI_{it} + \sum_{l=0}^{L} \omega^l LM_{it-1} + \varepsilon_{it}$$

where

VMT_{it}	=	vehicle-miles travel in area *i* and year *t*;
POP_{it}	=	population in area *i* and year *t*;
PCI_{it}	=	income per capita in area *i* and year *t*;
LM_{it-1}	=	state highway lane-miles in area *i* and year *t-1*;
α_i, β_t, γ, Ψ, ω^l	=	coefficients to be estimated; and
ε_{it}	=	a random variable drawn from a normal distribution with mean zero.

Based on this study, the elasticity of vehicle travel (VMT) with respect to lane miles was determined to be 0.3 to 0.7 for counties and 0.5 to 0.9 in California metropolitan areas for a four to five year period.

Fulton *et al.* (2000) found a significant and robust relationship between lane miles and daily VMT in the mid-Atlantic region of the U.S. The panel data were collected from Maryland, Virginia, North Carolina, and the District of Columbia and included geographic area, population, population density, income per capita, total employment and unemployment rate, and roadway lane miles in different roadway categories at county level. The model developed based on a fixed effect specification approach is shown below:

$$\log(VMT_{it}) = \alpha_i + \beta_t + \sum_k \lambda^k \log(X_{it}{}^k) + \varepsilon_{it}$$

where

VMT_{it}	=	the daily vehicle miles of travel for county *i* in year *t*;
α_i	=	the fixed effect for county *i*, estimated in the analysis;
β_i	=	the fixed effect for year *t*, estimated in the analysis;
X_{it}^k	=	the value of explanatory variable *k* for county *i* and year *t*;
λ_k	=	each of the set of *K* coefficients to be estimated; and
ε_{it}	=	the outcome of a random variable for county *i* in year *t*, assumed to be normally distributed with mean 0.

The model used to estimate both short-run and long-run elasticities had the lag structure employing an exponential distribution, which implied that the effects were strongest in the first year and then declined with time. The elasticity between lane miles and VMT was estimated to be 0.1 to 0.4 in the short-run and 0.5 to 0.8 in the long run. The Granger test results showed that lane-mile growth temporally preceded growth in VMT.

In the context of determining the benefits to users by an improved network, Abelson and Hensher (2001) presented a popular model to estimate induced travel, which assumed a multinomial logit form shown below:

$$\log(T) = \log\alpha + \beta g$$

where:

T = Total travel;

g = Generalized cost; and

α, β = estimators.

To find variation in the elasticity between induced traffic and capacity expansion over time, these authors recommend the use of a time series of cross-section data. Explanatory variables for the above equation included:

- The generalized cost of travel on each route at time period t;
- The population density, car ownership rate, and economic wealth (e.g. real per capita income) as measures of natural growth in traffic in the catchment area;
- The volume-to-capacity ratio on competing routes;
- Any natural barrier variables; and
- Other (dummy) variables that might describe qualitatively the specific infrastructure (e.g. freeway).

While some studies considered the effect of roadway capacity on induced travel, Barr (2000) focused on the effect of travel time changes on induced travel using the cross-sectional disaggregate household data from the 1995 NPTS. Travel time spent by each household was calculated by estimating the inverse of average travel speeds from the reported length and duration of travel. Multivariable statistical tests were performed to estimate the long-term effect of travel time on region-wide VMT after controlling various combinations of demographic, socioeconomic, and land use characteristics. The following variables were included in the model:

- Annual household VMT (miles)
- Inverse value of average daily household travel speed (hour/mile)
- Population density of the census tract in which the household was located (persons/mile^2)
- Annual household income (in dollars)
- Per capita household income (i.e., annual household income divided by household size in dollars)
- Number of members in household
- Number of workers in household
- Median household income of census tract in which the household was located (in dollars)

Fifteen regression models were estimated to explore the relationships between average travel time and VMT for households stratified by: (1) urbanized area; (2) public transportation availability; (3) metropolitan area size; (4) family life cycle; (5) day-of-week of travel; and (6) population density for households in metropolitan areas of one million or more people. It was concluded that there existed statistically significant relationships between travel time and VMT with elasticities in the range of -0.3 to -0.5, meaning that increases in highway capacity reduced travel time and, in turn, induced travel demand. It was also found that elasticities were higher in urbanized areas compared to non-urbanized areas.

Using activity-based approaches to analyze individuals' travel behavior, Fujii and Kitamura (2000) applied a model system to estimate impacts of two hypothetical freeways, Osaka Line and Kobe Line, a circumferential freeway and an extension of the existing freeway, respectively. The parameters of the model system were estimated using one-day activity diary data collected in 1994 in the Osaka-Kobe metropolitan area. A commuter's activity and travel after work in a day were analyzed in terms of variables of time use and trip frequency: (1) the number of trips after work and before returning home for the first time, excluding the return-to-home trip; (2) the total out-of-home activity duration (excluding travel) after work and before returning home for the first time; (3) the increase in travel time due to trips made to engage in out-of-home activities after work and before returning home for the first time; (4) the number of home-based trip chains after returning home for the first time till retiring for the day; and (5) the total amount of time spent at home after returning home for the first time till retiring for the day. The survey questionnaires were distributed to collect one-day activity diary data in the Osaka-Kobe metropolitan area. It was found that the users of an extension of the existing freeway made a greater number of home-based trip chains after returning home than and the users of a circumferential freeway. Increases in the duration of the commute result in reduction of total out-of-home activity duration, travel time due to trips made to engage in out-of-home activities, frequency of home-based trip chains after returning home, and total time spent at home after returning home. The authors concluded that the addition of new capacity did induce new trips.

Elasticities found from previous studies using aggregate level data are summarized in Table 2.6. The short-run period here is typically defined as one year or less and the long-run period ranges two to twenty years.

Table 2.6 Elasticities of Travel Demand with Respect to Lane-miles of Roadway (Source: Strathman et al. 2000)

Study	Time Period	Scale	Elasticities
Hansen and Huang, 1997	Long-run	County	0.6 - 0.7
Fulton, Meszler, Noland, and Thomas, 1999	Long-run	County	0.5 - 0.8
Fulton, Meszler, Noland, and Thomas, 1999	Short-run	County	0.1 - 0.4
Hansen and Huang, 1997	Long-run	Metropolitan	0.9
Noland and Cowart, 1999	Long-run	Metropolitan	0.6 - 1.0
Johnston and Ceerla, 1996	Long-run	Metropolitan	0.6 - 0.9
Marshall 2000	Long-run	Metropolitan	0.76 – 0.85
Noland, 2001	Short-run	States	0.2 - 0.5
Noland, 2001	Long-run	States	0.7 - 1.0

2.1.5 Integrated Land Use and Transportation Models

The term "integrated" implies a feedback mechanism between the transportation and land use models, as illustrated in Figure 2.5. The transportation model forecasts travel demand and determine the adequacy of the supply of transportation services. In almost all currently employed integrated models, the transportation model is a traditional four-step model that consists of trip generation, trip distribution, modal split, and trip assignment. In the future, the traditional models may be replaced by other kinds of transportation models such as activity-based models. The land use model, on the other hand, models the demand for and the spatial distribution of employment, residential, shopping, and other activities to allocate the area's residents and workers to specific urban zones. The land use system supplies the transportation system with estimates of the location and volume of travel generators. The travel costs resulting from the equilibrium between transportation demand and supply may be fed back into the residential and employment activity location models, which in turn modify the resident and employment location estimations. This allows transportation system changes to affect land utilization, which feeds back its effects in the form of new levels and locations of traffic generation. The notion of locational accessibility here plays a central role in all currently operational models. As an integral component of such accessibility, travel cost changes become part of the mechanism used to reallocate labor, residents, retail and service activities, and when modeled, freight flows between spatially separated land uses.

This section reviews several integrated models including ITLUP model, POLIS model, MEPLAN model, Kim's Chicago model, MASTER model, Dortmund model, UrbanSim, and TRANSIM. These models employ different theoretical and/or methodological approaches to representation of temporal and spatial interactions between land use and transportation. The ITLUP model is a Lowry type model. The POLIS model is a mathematical programming based model. The MEPLAN employs spatial input-output analysis. Kim's Chicago model is based on urban economics. The MASTER model is based on microsimulation, and the Dortmund model implements urban dynamics. UrbanSim is a microsimulation model as well as TRANSIM.

Within a number of operational models including the MEPLAN and Kim models, the urban system is modeled as a series of markets, with emphasis placed on clearing a transportation market and one or more other land use markets by solving a suitable set of spatially varying market prices endogenously (i.e., travel costs and site rents). Within the less inclusive models, such as ITLUP, which avoid endogenous modeling of non-transportation price mechanisms, an equilibrium between the transportation system's demands and supplies can also be brought about, stabilizing the parameters within the residential and employment activity location submodels. Such considerations of equilibrium in urban evolution quickly take us into the area of temporal dynamics. Within the ITLUP, MEPLAN, and Dortmund models, lagged effects play an important role in linking different submodels within the transportation and land use systems both across as well as within a single time period.

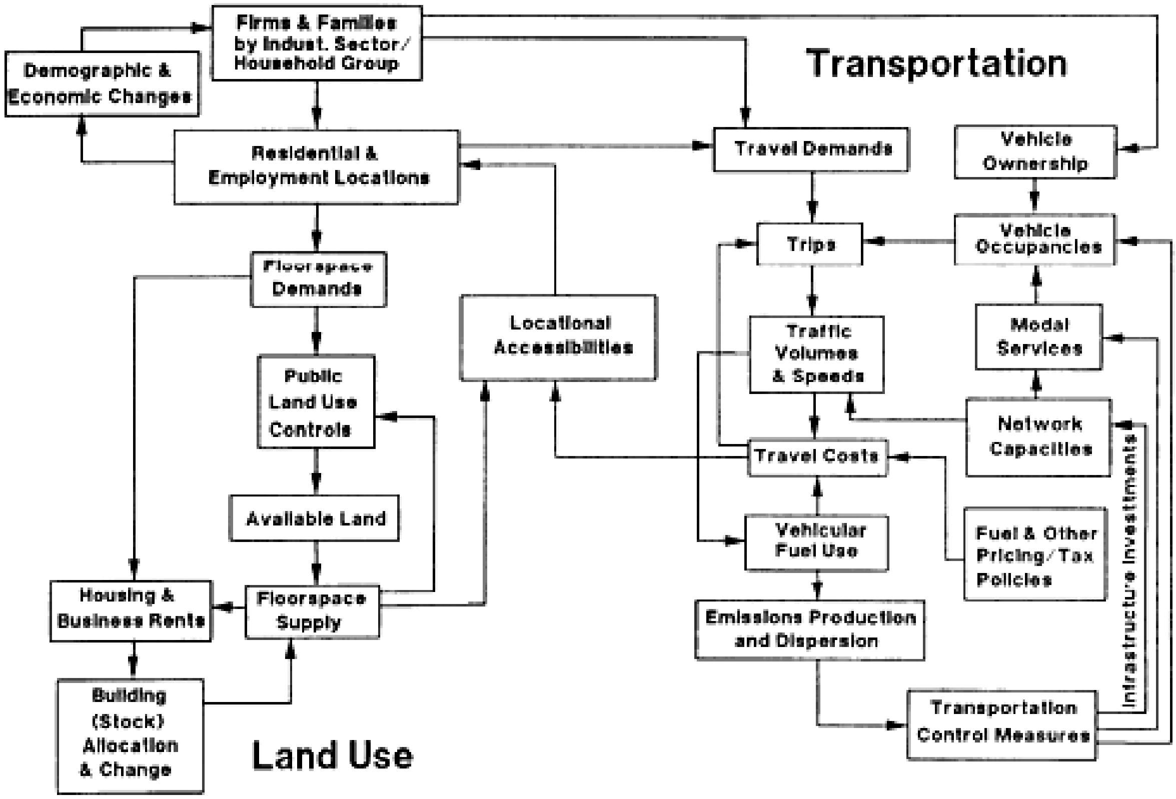


Figure 2.5 Integrated Modeling: General Schematic Flow Chart (Source: Southworth 1995)

Table 2.7 provides a brief summary of 10 land use/transportation models by Wegener (1995). The unified models are those that are tightly integrated while the composite models are those that have submodels, each modeling a particular aspect of land use or transportation.

Table 2.7 Summary and Comparison of Eleven Land Use Models (Source: Wegener 1995)

Model	Subsystems Modeled	Model Theory	Policies Modeled
POLIS *Composite*	Employment Population Housing Land use Travel	Random utility Locational surplus	Land use regulations Transportation improvements
Boyce *Composite*	Employment Population Networks Travel	Random utility General equilibrium	Transportation improvements
KIM *Unified*	Employment Population Networks Goods transport Travel	Random utility Bid rent General equilibrium Input-output	Transportation improvements
METROSIM *Unified*	All subsystems except goods transport	Random utility Bid rent General equilibrium	Transportation improvements Travel cost changes
ITLUP *Composite*	Employment Population Land use Networks Travel	Random utility Network equilibrium	Land use regulations Transportation improvements
TRANUS *Composite*	All subsystems	Random utility Bid rent Network equilibrium Land use equilibrium	Land use regulations Transportation improvements Travel cost changes
LILT *Composite*	All subsystems except goods transport	Random utility Network equilibrium Land use equilibrium	Land use regulations Transportation improvements Travel cost changes
MEPLAN *Composite*	All subsystems	Random utility Network equilibrium Land use equilibrium	Land use regulations Transportation improvements Travel cost changes
IRPUD *Composite*	All subsystems except goods transport	Random utility Network equilibrium Land use equilibrium	Land use regulations Housing programs Transportation improvements Travel cost changes
RUBAN *Unified*	Employment Population Housing Land use	Random utility Bid rent General equilibrium	Land use regulations Transportation improvements

2.1.5.1 DRAM, EMPAL, and ITLUP

Putman (1983) developed the integrated transportation and land-use model package (ITLUP) under a contract with USDOT. This model attempted to balance transportation facility development and land development to achieve market equilibrium between the demand for use of a transportation facility and its characteristics (cost, speed, and capacity) as a source of supply. Putman linked the Disaggregate Residential Allocation Model (DRAM) and the Employment Allocation Model (EMPAL) with the selected submodels of the traditional four-step transportation planning model, including trip distribution, modal choice, and traffic assignment.

A feedback mechanism between DRAM, EMPAL, and the Urban Transportation Planning System (UTPS) model was developed. Using the output of the transportation planning procedure as the input to the land-use planning procedure, the forecasts of spatial distributions of activities produced by the land-use planning procedure became, in turn, the input to a second iteration of the transportation planning procedures. This model has been applied widely in urban areas in this country.

Figure 2.6 illustrates the block diagram of this simulation model package. The entire procedure may be described as follows: First, the various base year descriptions of the existing spatial distributions of activities, along with data on the characteristics of the unloaded base year transportation network, are input into the models. These data are then used to generate a preliminary and probably inflated estimate of trips taking place in the metropolitan area. Given this preliminary estimate of trips, it becomes possible to load future year network so that its characteristics reflect the traffic volumes on the network if there had been no change in the spatial distribution of activities from the base year. These characteristics of the partially congested network, along with the base year spatial distribution data and the forecast region-wide control totals, are then used to generate a trial estimate of the spatial distributions of activities for the projection year. A new estimate of trips will be produced from these spatial distributions. These trips would, in turn, be loaded on the projection year transportation network. The modified characteristics of the more realistically congested transportation networks can then be used to reallocate the projected spatial distribution of activities. This distribution of activities is then compared with the first trial estimate thereof. If there are no significant differences, equilibrium is assumed and the model run is ended. If there are significant differences, new trips are generated and loaded on the networks, and iterations continue. This integrated package of transportation and land-use (location) models thus attempts to eliminate the principal failure of contemporary land-use or transportation studies by explicitly including feedback loops.

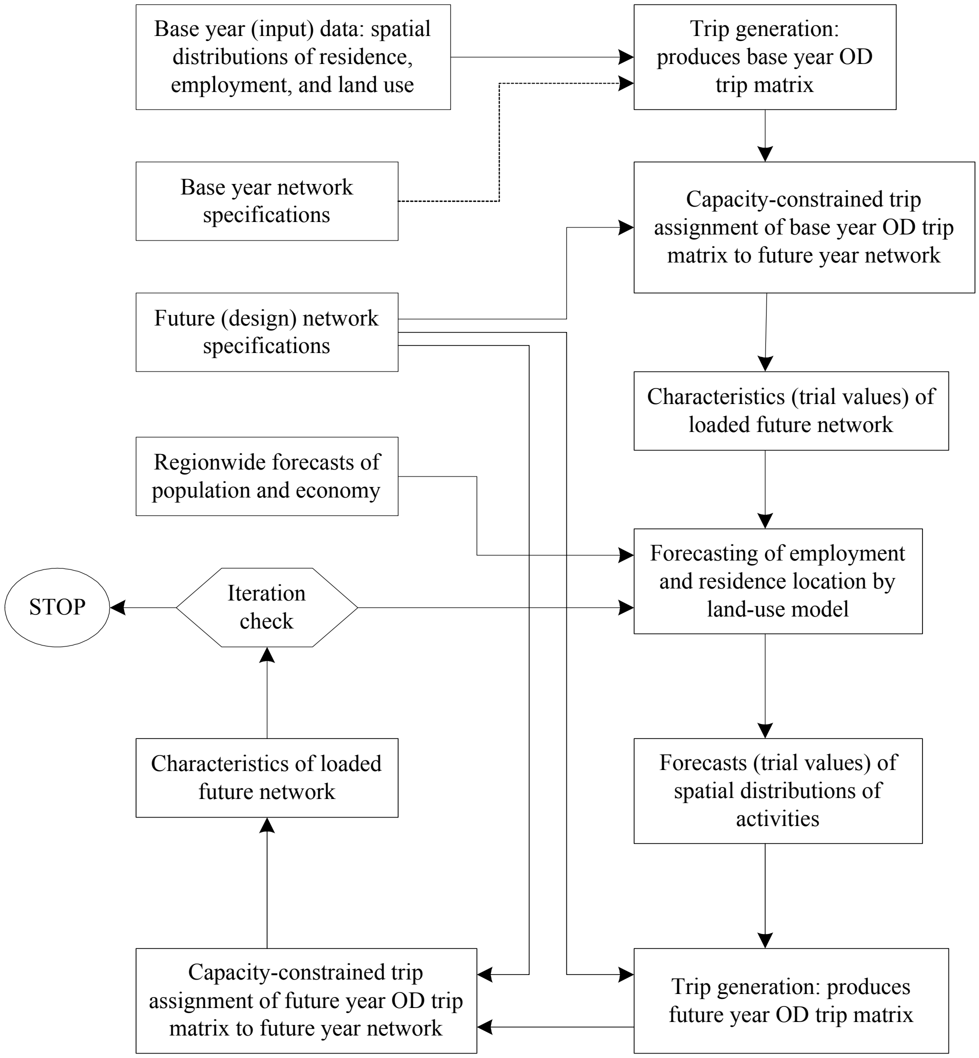


Figure 2.6 Integrated Transportation and Land-Use Model Package (Putman 1983)

2.1.5.2 The POLIS Model

During the 1980s, the Association of Bay Area Governments (ABAG) developed the Projective Optimization Land Use System (POLIS) for San Francisco Bay Area. Prastacos (1985) distinguished the POLIS model from Lowry-type models such as ITLUP with three key aspects: use of microeconomic behavioral principles, formation as a mathematical programming problem, and consideration of job location, basic and non-basic employment, residence selection, and trip making in an integrated manner. The model seeks to maximize jointly the locational benefits associated with multimodal travel to work, retail, and local service sector travel, and, significantly and jointly, the agglomeration benefits accruing to basic-sector employers. The model was subjecting to a set of constraints that require trip production and attraction to be consistent with housing and employment available in a zone and that employment of related sectors be allocated jointly, as well as constraints on limits of housing and employment allocation and exogenous location of employment and housing (policy constraints). POLIS simulates changes in housing, employment, and trip making between two states, the base year and future year. The lagged effects of transportation investments on land use are not modeled.

2.1.5.3 The MEPLAN Model

MEPLAN is a proprietary software package developed by Marcial Echenique and Partners Ltd. in the United Kingdom. Hunt (1997) describes its framework as the interaction between two parallel markets: a land market and a transport market. Behavior in each system is modeled as a response to price or price-like signals (including travel disutility). Each market moves towards equilibrium, but a complete equilibrium is not reached because there are time lags in the system, which are caused by the fact that building stock and transportation infrastructure cannot be changed instantaneously and the fact that information exchange is not perfect.

Time lags are modeled in MEPLAN by ordering the model operations as shown in Figure 2.7. At time *n*, the land use model provides input to the transportation model, which produces network equilibrium. The state of the land use and transportation systems then becomes the input to the land use model at time *n+1*, which utilizes the network conditions at time *n* in allocating population and employment, reflecting the time lag between changes in the transportation system and the response from the land market. The transportation system will in turn respond to changes in land use and produce a new set of network conditions.

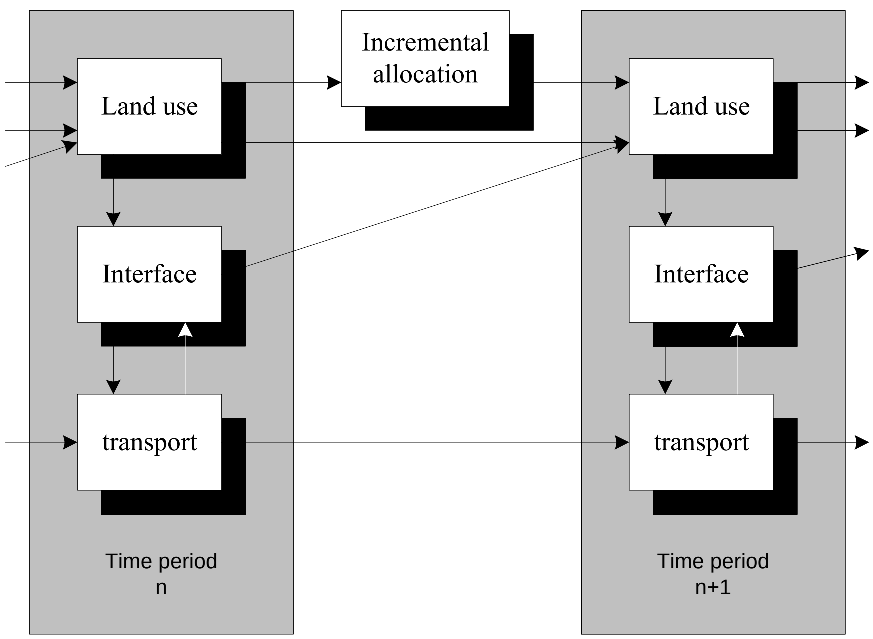


Figure 2.7 Ordering of Model Operations (Source: Hunt 1997)

In addition to time lags between transportation and land-use systems, the lags between activity location and land or building stock and between transportation supply and demand are also modeled in MEPLAN.

Hunt (1997) points out that while MEPLAN requires little data to run for forecast purposes, calibrating MEPLAN is complex and demanding, as "an extremely large and rich set of observed data would be required in order to perform a 'full and complete' calibration of all model components over several periods in a typical application." Particularly, for each time period to be modeled, model data must be available. However, the effort would be worthwhile because its ability to incorporate the interactions of different components in land use and transportation markets, and policy alternatives and effects.

2.1.5.4 Kim's Chicago Model

Kim's Integrated Urban Systems Model for Chicago integrates general urban system equilibrium with probabilistic spatial interaction, combined transportation-facility location models, and concept of equilibrated demand and supply over networks (Southworth 1995). Kim's model is based on urban economic theories and takes into consideration "land use and density, shipment route and mode choice with network congestion" (Kim 1989). The model seeks to minimize the objective function that is the sum of the total costs of moving commodities out of the urban system, and the total land plus rental costs of all zones, commodities, and production techniques used in the urban area. Different intensities of land use may be specified for different production technology, allowing for the production of certain commodities using less land (as in the case of the service industry). In addition to a network with a balanced demand and supply, the activities are allocated spatially in such a way that the marginal cost of producing a commodity in a zone, which includes the capital and land costs, and shipping it to another zone using a given mode is the same as that of producing the same commodity in the second zone. All the variables are exogenously supplied.

Kim has calibrated the model using various and extensive data for the Chicago region. Calibration routines also exist, which makes it an operational model. Its main contribution is the demonstration of the possibility of "bringing important aspects of urban economic theory into intersectoral, spatial-interaction-based discrete choice models in order to move towards more comprehensive urban modeling frameworks" (Southworth 1995). However, the lagged effects are not considered in Kim's model. The equilibrium solution produced by the model will only prescribe an ideal world when information exchange is perfect and responses between land use and transportation are instantaneous, which is far from reality.

2.1.5.5 The MASTER Model

The Micro-Analytical Simulation of Transport, Employment and Residence (MASTER) model, developed in the United Kingdom by Mackett (1990a, 1990b), is a microsimulation model. Microsimulation models are computer models that simulate behaviors of individuals of a representative population and draw conclusions that apply to higher levels of aggregation. These models distinguish themselves from aggregate models in that the explanatory variables reflect the characteristics of individuals and their decision-making processes, whereas in an aggregate model the explanatory variables represent the collective properties of the objects (such as population groups or households of different types) being modeled. Aggregate models therefore model the collective effects of individual behaviors but cannot explain them, thus aggregate models are unable to deal with certain policy issues. Microsimulation models are relatively easy to understand and implement because the decision-making process may be modeled by generating random numbers and using these numbers to make choices given necessary probability distributions of decision choices. As microsimulation models often simulation a process over time, the lagged effects can also be incorporated easily into the models.

The MASTER model is an integrated land use-transportation model that operates at the household level. Population growth and household structure are modeled based on the lifecycle including birth, aging, death, marriage, divorce, and migration. Choice of residential location is based on the weighted function of generalized travel to work cost for the head of household (only work trips are modeled), while housing type choice is based on household size and composition. Supply of housing and jobs are exogenous data input into the model, and vacancy of housing is tracked for each zone. Household members' choice of jobs, employment and becoming unemployed, retirement, education level, sex, social group of head of household, job vacancies, and salary ranges are all modeled in MASTER.

The transportation processes modeled include acquisition of driver's license, auto ownership, car availability, and work trip mode choice, all of which are functions of age, sex, household income, household composition, and travel costs. Traffic assignment is not handled in the model. While incorporating a traffic assignment routine seems to be straightforward, Southworth (1995) questions Mackett's suggestion that 1% sample of households is necessary for model calibration. Instead, Southworth suggested that up to a 100% sample might be necessary if trips are to be assigned to specific routes.

2.1.5.6 The Dortmund Model

Another model that utilizes the microsimulation technique is the Dortmund model developed by Wegener (Wegener 1986) for Dortmund, Germany. Southworth (1995) considers the Dortmund model as one of the most advanced implementations of a multistaged urban land use-transportation system dynamic at the time.

The Dortmund model is the intermediate level model in a three-model hierarchy, which also includes a macroanalytic model that simulates changes in employment by industrial sectors and population by age, sex, interregional migration rates, and nationality in each of the 34 regional labor markets. It is a finer grained model (with smaller zones) that allocates construction to each statistical tract within a zone. The Dortmund model, being a mesoscopic model, inputs employment and population from the macroanalytic model, and simulates changes in the urban system with seven submodels that deal with (1) car ownership and travel; (2) aging of people, households, dwellings, and workplaces; (3) relocation of firms, redundancies, and new jobs; (4) nonresidential construction and demolition; (5) residential construction, rehabilitation, and demolition: (6) change of jobs; and (7) change of residence.

As has been discussed in Section 2.2, different subsystems in an urban system respond to changes in the system at different rates, ranging from fast (e.g., mode choice and route selection) to medium (e.g., home, job, firm relocation) to slow (e.g., activity centers and transportation network). To simulate these urban dynamics, the Dortmund model explicitly maintains the separation of the land use and transportation submodels and limits the ability of the transportation model to relocate work trips (to satisfy the double constrained distribution model) by directly taking input from the land use models on household and job relocations, thus forcing the land-use changes to lag behind changes in travel costs. The model simulation runs in two-year cycles for up to a 30-year planning horizon and allows a "perception delay" of one year to take effect. The transportation model runs at the beginning and end of the two-year cycle, implying that land use changes are not visible to the transportation model within the cycle. New housing construction does not occur until three cycles (6 years) later after changes in the transportation system.

The spatial distribution of activities may be changed in the Dortmund model through aging, exogenously with input of "historical events" such as closing of large industrial factories, and endogenously through changes in accessibility.

2.1.5.7 UrbanSim

Developed by the Urban Planning Department and the Computer Science Department at the University of Washington, UrbanSim is a simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. UrbanSim is a public domain software package, and is intended for use by metropolitan planning organizations and other planning organizations to interface existing travel models with new land use forecasting and analysis capabilities. The University of Washington has received three new grants from the National Science Foundation to continue the development of UrbanSim in September 2001 (UrbanSim 2001).

The development of UrbanSim has been motivated by the need of planning organizations to test policies that deal with environmental, sociological, and economic concerns (UrbanSim 2001). Examples of such policy issues include preserving prime agricultural lands, forests, wetlands, and open space, redevelopment, infill, and inner-city decline. Possible strategies developed based on such policies may range from urban growth boundaries at the regional or metropolitan scale to street design, mixing of uses, and pedestrian access at the neighborhood or site-specific scale.

UrbanSim is a discrete choice model based on the random utility theory. With its modular design, it models household location choice, employment location choice, real estate development, interfaces with travel models, and incorporate urban dynamics into the models. Household are classified in a disaggregate manner by income, persons, workers, and child, employment is classified into 10-20 sectors, and real estate into 24 development types. Real estate measures include acres, units, and floor space. Real estate prices are also modeled. The model operates at a geographic scale of 150-meter grid cells (Waddell 2001).

UrbanSim has eight core models (Waddell 2001):

- The Demographic Transition Model simulates births and deaths in the population of households given control totals. Distribution of income groups, age, size, and presence or absence of children may also be specified. Iterative proportional fitting (Beckman et al, 1995) is used to the addition or deletion of households. Newly created households are added to the household list to be assigned to housing units by the Household Location Choice Model later.

- The Economic Transition Model simulates job creation and loss with control total and distribution of business sectors specified.
- The Household Mobility Model simulates household decisions as where to move based on probabilities determined from historical data.
- The Employment Mobility Model determines which jobs will move from their current locations during a particular year using a similar approach to the Household Mobility Model.
- The Household Location Model chooses a location for each household in the housing list. To do this, a list of vacant housing is maintained and a multinomial logit model calibrated to observed data is used to select a housing from a random sample of the vacant housing units, which are described by attributes of the housing in the grid cell (price, density, age), neighborhood characteristics (land use mix, density, average property values, local accessibility to retail), and regional accessibility to jobs.
- The Employment Location Model is responsible for determining a location for each job that has no location. Alternatives are selected through random sampling, which are described by real estate characteristics in the grid cell (price, type of space, density, age), neighborhood characteristics (average land values, land use mix, employment in each other sector), and regional accessibility to population.
- The Real Estate Development Model simulates developers' choices about location and type of construction to undertake. Each year, the model iterates over all grid cells on which development is allowed and creates a list of possible transition alternatives (representing different development types), including the alternative of not developing. The probability for each alternative being chosen is calculated with a multinomial logit model. Variables in the developer model include characteristics of the grid cell (current development, policy constraints, land and improvement value), characteristics of the site location (proximity to highways, arterials, existing development, and recent development), and regional accessibility to population.
- The Land Price Model simulates land prices of each grid cell as the characteristics of locations change over time based on urban economic theory. The model is calibrated from historical data using a hedonic regression to include the effects of site, neighborhood, accessibility, vacancy rates, and policy effects on land prices. The model variables are similar to those in the Development Model.

The model requires input on population and employment estimates, regional economic forecasts, transportation system plans, land use plans, and land development policies such as density constraints, environmental constraints, and development impact fees. The user is allowed to create "scenarios" as input to UrbanSim by specifying alternative forecasts of population and employment, land-use policy assumptions, transportation infrastructure assumptions, etc. The model then provides output regarding future year distributions of population, households by type (e.g. income, age of head, household size, presence of children, and housing type), units of housing by type, businesses by type (e.g. industry and number of employees), land use by type (user-specified), square footage of nonresidential space by type, densities of development by type of land use, and prices of land and improvements by land use (UrbanSim 2001).

UrbanSim has been validated for the Eugene-Springfield, Oregon area (population 375,823) using data from 1980 to 1994, and has since been applied in Honolulu and Salt Lake City. Other metropolitan areas are beginning to utilize it as well. The UrbanSim software is distributed as Open Source software under the GNU General Public License, which allows anyone to use, modify and redistribute the source code at no cost. The source code of UrbanSim is available at www.urbansim.org.

2.1.5.8 TRANSIM

TRANSIMS is part of the Travel Model Improvement Program (TMIP) sponsored by the U.S. Department of Transportation (USDOT) and the Environmental Protection Agency; it was developed by the Los Alamos National Laboratory. The lab has conducted a major TRANSIMS study in Albuquerque and completed a microsimulation of vehicle traffic patterns in Dallas, Texas and Portland, Oregon. Six cities selected by the U.S. DOT that represent diversity in growth and population area expect to test TRANSIMS. TEA-21 has set aside $25 million in funding for the completion and deployment of TRANSIMS.

TRANSIMS is a disaggregate microsimulation model that simulates household activities and individual trips over a multimodal network. The type of activities and their locations are driven by household and personal demographics such as the age of an individual, the person's income, gender, and employment status. TRANSIMS is designed to give transportation planners accurate, complete information on traffic impacts, congestion, and pollution (LANL 1999).

Briefly, TRANSIMS has five main modules (LANL 1999):

- The Population Synthetic Generator module creates synthetic population figures that represent every household and person in a given metropolitan region. Households and persons are described by their demographics such as the age, income, gender, and presence of children in households.
- The Activity Generator module generates a list of activities for each member of the synthetic population. Each activity is specified by its type, priority, starting and ending time preferences, a preferred mode of transportation, a vehicle preference if appropriate, a list of possible locations for the activity, and a list of other participants if the activity is shared. The set of activities for each household is based on the household demographics, network congestion, and land use and employment data.
- The Route Planner module is responsible for generating routes for travelers. The routes between different locations are constrained by the network and by preferences of individual travelers.
- The Traffic Microsimulator module simulates the movement and interactions of individual travelers in the transportation system of a metropolitan region based on their trip plans provided by the Route Planner. Intermodal travel, vehicle occupancy, trip chaining, and vehicles with different operating characteristics can all be simulated. Driver behaviors, including acceleration, deceleration, turning, changing lanes, passing, and responding to other vehicles, signs, and signals, can also be simulated. The Traffic Microsimulator uses a cellular automata approach to provide the computational speed necessary to simulate an entire region at the individual traveler level. With this approach, each link in the transportation network is divided into a finite number of cells and vehicle locations in the

cells are determined at each time step of the simulation. Because of the sheer number of travelers and the level of detail in the microsimulation, a TRANSIM runs requires the use of multiple CPUs where available.

- The Emissions Estimator module translates traveler behavior into emissions of nitrogen oxides, hydrocarbons, and carbon monoxide, as well as energy consumption and carbon dioxide emissions. Different types of vehicles and different operating states may be modeled in terms of their tailpipe emissions.

While TRANSIMS is capable to produced detail travel information over a large network in an urban area and accurate air quality information, it does not deal with the land use issues as well as other models (e.g. MASTER and UrbanSim). Therefore, it will not be able to explain the impacts of land-use policy on transportation or test the effects of transportation investment decisions on land use.

2.2 Temporal Geographic Information System

Time is an inseparable component of all changes that take place in the real world. The transportation system evolves over time as construction and maintenance of planned improvements change the transportation network. Land use patterns also change slowly but constantly due to the varying needs of human activities over time. Furthermore, various kinds of changes that occur to transportation and land use are interdependent, and cannot be easily separated. In order to model the interactions between different change patterns (e.g., land use changes and transportation system changes), a framework is needed that takes into account the time component (when), the location component (where), and the attribute component (what) associated with these changes and their interrelationships. In addition, when dealing with a real world problem, the semantic level referring to the meanings of objects to be analyzed becomes important.

Sinton (1978) suggests using a measurement framework that includes the three main components (i.e., time, location, and attribute) to serve as the "fixed", the "controlled", and the "measured" parameters, resulting in a total of six possible measurement scenarios (see Table 2.8). Each of the six measurement scenarios is suitable for a particular type of analysis need. For example, we can "fix" time (e.g., April 1, 2001) and "control" location (e.g., 0.5-mile, 1-mile, and 1.5-mile buffer zones around the interchanges of I-75 in Broward County) to "measure" attributes such as land use types. Application of this measurement (scenario #3 in Table 2.8) can provide land use information within the buffer zones around the interchanges of I-75 in Broward County on the selected date. Alternatively, by fixing attributes (e.g., average daily traffic count higher than 8,000) and controlling time (e.g., Year 1990 through Year 2000), we can "measure" locations (e.g., locations of highway segments or traffic count stations). This example corresponds to the scenario 6 in Table 2.8.

Table 2.8 Measurement Framework Proposed by Sinton (1978)

	Fixed Component	Controlled Component	Measured Component
Scenario 1	Location	Time	Attribute
Scenario 2	Location	Attribute	Time
Scenario 3	Time	Location	Attribute
Scenario 4	Time	Attribute	Location
Scenario 5	Attribute	Location	Time
Scenario 6	Attribute	Time	Location

Sinton's measurement framework is also relevant to the development of GIS applications. The current commercial GIS software products are built around the measurement scenarios #3 and #4. That is, the users build their GIS databases by fixing the time component while controlling and measuring the location and the attribute components. In a raster GIS, a user normally controls the spatial resolution level (i.e., grid cell size) and measures the representative attribute value for each grid cell observed at a fixed point in time. Using a satellite imagery to derive land-use classification is a typical example of this measurement scenario. On the other hand, a vector GIS often controls the attributes (e.g., land-use types) and measures the locations (i.e., the polygon boundaries representing different land-use types) for data observed at a fixed point in time. In either case, time is treated as a fixed component in order to represent the other two components in a GIS environment. It is feasible to repeat the same measurement process under the scenarios #3 and #4 to derive measurements for different points in time (e.g., Years of 1990, 1995, and 2000). These repeated measurements result in the "snapshots" of location and attribute information.

The treatment of time as a fixed component in a GIS environment reflects the traditional focus of GIS on location and attribute at the expense of time. In addition, this simplistic approach avoids many challenging (and mostly unresolved) issues of handling the complex interdependency between the space dimension and the time dimension. However, the "snapshot" approach is insufficient to meet the requirements of preserving historical data and supporting the modeling and analysis of complex interactions between land-use changes and transportation system developments. An innovative and practical approach of handling spatiotemporal interactions is needed to support the needs of transportation modeling and planning. With the rapid advances in both computer science and GIS, we have seen an increasing level of research on temporal databases and spatiotemporal GIS. This review provides a summary of the relevant literature on the issues of integrating the time dimension within a GIS environment. Since most GIS use a database management system (DBMS) to help organize and manage the data, this literature review effort went beyond the GIS publications to also include the relevant literature of temporal databases.

This section first introduces a framework that organizes the large amount of literature related to different research issues of temporal GIS (see Section 2.2.1 below). Selected publications under each major research issue are discussed in Section 2.2.2 to provide a review of the approaches proposed to pursue each research issue. Section 2.2.3 then discusses the relevance of previous studies to this research. In addition, a list of references identified from this literature review task on temporal GIS is included in Appendix A as a supplement to the discussions here.

2.2.1 An Organizational Framework of Temporal and Spatiotemporal Research Issues

Integration of the time component into a GIS environment introduces many challenging research issues. These research issues cover a wide range of topics and have been studied in different disciplines. The computer scientists, for example, have been conducting research on temporal databases for a number of years with an emphasis on a better means of handling the time dimension in a database environment. On the other hand, geographers and researchers in GIS-related fields have been paying more attention to the interaction between and integration of the space and the time dimensions in a GIS environment. There has also been an increasing trend towards cross-referencing research efforts among the various disciplines interested in spatiotemporal integration.

Table 2.9 presents a framework that organizes the various research issues related to the development of a spatiotemporal GIS identified in this literature review task. At the highest level, the proposed framework organizes the relevant literature into two broad categories: conceptual issues and implementation issues. *Conceptual* issues focus on the publications related to the topics such as temporal and spatiotemporal concepts, data models, and data access. *Implementation* issues cover the challenges of developing prototype implementations of temporal databases and spatiotemporal GIS and the implementations using a simulation approach.

Table 2.9 Organizational Framework of the Research Issues Identified in the Literature Review

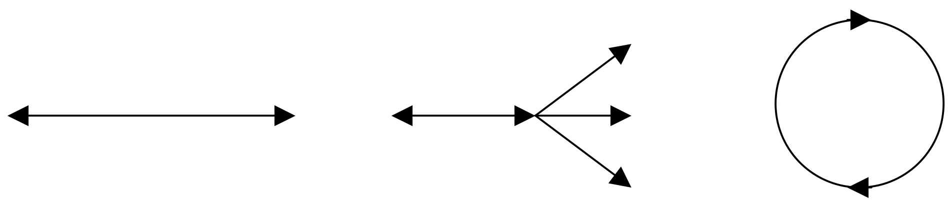
I. Conceptual issues:
A. Temporal and spatiotemporal concepts
1. Temporal definitions
2. Spatiotemporal data representations
3. Spatiotemporal relationships
4. Spatiotemporal database frameworks
B. Data models
1. Spatial data models
2. Temporal data models
3. Spatiotemporal data models
C. Data access
1. Temporal/Spatiotemporal data indexing methods
2. Temporal/Spatiotemporal query languages
3. User interface design/Visualization
II. Implementation issues:
A. Prototype-based implementations
1. TRIAD
2. SIIASA
3. ATLSS
4. Wang (1998)
5. Raza (1998)
6. Batty and Xie (1994a, 1994b)
7. LUCAS
8. Beller, et al. (1991)
9. GAEA
10. QUEST

11. TOM
B. Simulation-based implementations
1. Theodoridis, et al. (1999)

2.2.2 Conceptual Issues

2.2.2.1 Temporal and Spatiotemporal Concepts

Time is often conceptualized as a continuum on a linear scale that moves in forward or backward directions (i.e., future, present, and past). This simple conceptualization becomes insufficient when we deal with real world phenomena that can branch into different scenarios (e.g., three transportation improvement plans for 2010 proposed in year 2000) or have a cyclical temporal pattern (e.g., daily or weekly traffic patterns). Figure 9.1 shows the three possible temporal forms. In addition, time can be recorded as *world time* (i.e., when an event took place in the real world) and *database time* (i.e., when an event is entered into a database). The distinction between world time and database time is critical when we need to evaluate, for example, if a travel demand analysis was performed before a future land use plan was entered into the database although the land use change had not occurred in the real world.



(a) linear (b) branch (c) cyclical

Figure 2.8 Three Possible Forms of Temporal Patterns

Computer scientists often treat time and space in the same way. Tsotras (1997) suggests that each individual piece of data is associated with space and time dimensions in a multi-dimensional space. There are three space dimensions, which are x-dimension, y-dimension, and z-dimension. There are also two time dimensions, namely a valid-time (or world-time) dimension and a transaction-time (or database-time) dimension. This view, however, is too simple to model many complex temporal phenomena.

Most real world phenomena are the outcomes of temporal *processes*. A process in turn consists of a set of *events*. An event triggers changes to a real world phenomenon that leads it from one *state* to another state. Temporal databases and spatiotemporal GIS need to represent and manage these states, events, and processes. One challenge is whether phenomena should be represented as a continuous process or as discrete events. Another important consideration *is temporal resolution*. When a real world phenomenon is evaluated based on a time instant versus time duration, the focus will be on different aspects of the phenomenon. An event is normally associated with a time instant. A list of time-sequenced events, on the other hand, can reflect a process that lasts for a given period. The time intervals (i.e., temporal resolution) used to collect data associated with the events can therefore have an impact on the analysis of a temporal process.

The temporal relationship among different events and processes is another important consideration. An event takes place before, at the same time instant, or after other relevant events. A process may overlap with another process over a period of time. These temporal relationships are important information for understanding changes over time such as whether traffic flow increased significantly on the Sawgrass Expressway before, during, or after a major residential subdivision was built in Weston, Broward County. This example further indicates that if the time component is brought into a GIS environment, the spatiotemporal relationships that involve the interactions among the time, location, and attribute of different processes must also be addressed.

The above discussions indicate many challenging conceptual issues that must be addressed and understood before a spatiotemporal GIS is developed. A review of some key references in the literature on these conceptual issues is provided below.

(1) Temporal definitions

Snodgrass (2000) defines three fundamental temporal data types: instant, interval, and period. A period is an anchored duration of time (e.g., March 5th through May 12th, 2001) compared to an interval simply indicating a length of time (e.g., 14 months). He also suggests that there are three kinds of time: user-defined time, valid time, and transaction time. In addition, he identifies three basic kinds of time-oriented statements: current, sequenced, and non-sequenced. A non-sequenced statement ignores the time element.

Frank (1998) proposes four different views of temporal taxonomy. The first view is based on the concept of events, which are abstract time points without duration or intervals between two events. The second view is based on the interpretation of processes: linear or cyclic repetitive pattern. The third view depends on the scale of measurements: ordinal time scale and interval scale. The fourth view is based on the order of events: total order, partial order, branching, and multiple perspectives of observing the event orders.

Hazelton (1998) uses temporal metaphors to describe a language for communicating about time. Four metaphors are found to be in general use: linear metaphor, cyclic metaphor, multi-dimensional metaphor, and branching metaphor.

Hornsby and Egenhofer (2000) formulate time with an explicit description of change with respect to states of existence and non-existence of identifiable objects. Four types of primitives related to the identity states of objects are described: object existence, non-existing object without history, non-existing object with history, and transition about two identity states of the same object. Therefore, nine change operations are derived from a systematic combination of the first three primitives.

Claramunt (1996) discusses changes in terms of events and processes. He points out that there are three main classes of basic spatiotemporal processes: evolution of a single entity, functional relationships between entities, and evolution of spatial structures involving several entities (reconstructing processes).

(2) Spatiotemporal data representations

Sinton (1978) suggests that geographic information consists of three components: attribute (of phenomena or objects being observed), location (of the phenomena), and time (of the observation). He argues that without all three components present and a record of the precision and reliability of the observation, no geographic data should be entered into an information system.

Effenberg (1992) indicates that the semantics of time can be incorporated into data models using three approaches: as a parameter, as a property, or as a dimension. Under the parameter approach, time is used as a control argument while other variables are investigated. Time also can be treated as a property of an entity. Alternatively, time can be considered in the same way as spatial dimensions. Each of the approaches is likely to result in a different type of data model.

Peuquet (1994) argues that an orthogonal 4-D representation of the (x, y, z) coordinates and the time (t) may be inappropriate for GIS because time and space exhibit important differences in their properties. She sums up some higher-level or derived knowledge of spatiotemporal relationships and patterns and discusses the characteristics of time such as temporal cohesiveness, temporal similarity, temporal continuity, hierarchical organization, and incompleteness.

Peuquet and Qian (1995) propose a conceptual TRIAD model to describe geographic information with spatial, temporal, and feature dimensions. Feature-based, location-based, and time-based representations are considered to be complementary elements in the TRIAD model. Unlike traditional spatial data model that are designed to be self-contained, all three components of the TRIAD model are designed to be cooperative and interdependent. This dimensional integration makes all dimensions of the data accessible to the user and enables the user to observe and analyze data from varying dimensional perspectives within a single representation.

Yuan (1996) presents a three-domain representation that defines semantic, temporal, and spatial objects in three separate domains. This data representation provides links between the three domains to describe geographic processes and phenomena. It allows the geographic concepts and entities to be represented through dynamical links among the three types of objects from either a layer or an object representation. A major advantage of this presentation is that there is no predefined data schema. In addition, it eliminates the constraints of a linear temporal scale that monitors and analyzes successive states of spatial entities.

Claramunt (1996) addresses the semantics representation issue for spatiotemporal geographic information. This research retains the TRIAD data representation framework proposed by Peuquet and Qian (1995) and uses an object-oriented model to define the data structure. A spatial entity is defined by three attribute sets: temporal domain, thematic domain, and spatial domain.

Faria *et al.* (1998) define spatiotemporal objects to encapsulate three basic components: conventional, spatial, and temporal. Each component is a complex object by itself.

(3) Spatiotemporal relationships

Peuquet (1994) defines temporal operators with three distinct classes: (1) metrics (length of the time interval) and topology (before, equal, meets, overlaps, during, starts, ends), (2) Boolean operators (and, or, not), and (3) generalization. Spatial operators include: (1) area generalization, (2) overlay, (3) spatial metric, and (4) topology. She further indicates that object operators are based on cause-and-effect relationships, such as causes/caused-by or becomes/precedes.

Faria et al. (1998) suggest that temporal operators can be unary or binary, returning time values, Boolean, or objects. Spatiotemporal operators, which extend the spatial operators by adding a temporal dimension, include: location-temporal, orientation-temporal, and metric-temporal. Additional operators are defined in order to support the formulation of spatiotemporal queries.

Erwig (1999) examines spatial and temporal integration. He indicates that spatiotemporal databases are essentially database about moving objects. He attempts to find a general case of geometry that may change in a continuous manner and proposes a new line of research where moving points and moving regions are viewed as three-dimensional (2D space + time) or higher-dimensional entities. This proposed concept could be integrated as base (attribute) data types into relational, object-oriented, or other DBMS data models.

(4) Spatiotemporal database frameworks

Paton *et al.* (2000) argue that it is easier to propose a spatiotemporal model than to implement it and emphasizes that researchers should be slow to propose models that they are not prepared to prototype. It is unlikely that scalable working prototypes can be developed in the short to medium term that support vector and raster data, valid and transaction time, spatial and temporal indeterminacy, moving objects, and constraints programming. The authors also suggest that spatiotemporal DBMS can be designed to support DBMS functionalities over spatial and temporal data orthogonally or synergistically. This can be called an extensible approach, which extends a non-spatial database to support spatial applications by requiring an extension of the set of supported data types only. By contrast, extending a snapshot spatial database to support temporal applications requires not only extending the set of supported data types, but also making provisions for all data model types to be associated with a history (i.e., everything changes!). In other words, spatial extensions can be easier than temporal extensions. An architectural framework is proposed in which DBMS is viewed as comprising programming language interfaces, query processing components, and a services manager.

Manuel and Moreira (1997) focus on the implementation, evaluation, and optimization of spatiotemporal database operations. They propose to use C and C++ programming languages, the Informix Universal Server, and O2 system to tackle two challenges. One is to define sets of workloads that simulate a rich set of scenarios representative of a variety of real world applications. The other is to define a common schema for the representation of results of the experiments.

Steiner (1998) discusses four types of temporal databases differentiated by their ability of representing temporal information with respect to valid and transaction-time definitions. These four types of temporal databases are snapshot databases, historical databases, rollback databases, and bitemporal databases.

2.2.2.2 Data Model

A *data model* is a logical representation of data organization and data relationships. There are three types of data models frequently referenced in the literature. *Relational data model* is the most widely used data model in both computer science and GIS fields. It has a strong theoretical foundation and a standard query language (i.e., Structured Query Language or SQL). *Object-oriented data model* is based on a set of object-oriented concepts (e.g., object, object class, inheritance, encapsulation, and polymorphism) that provides a more intuitive data representation of real world phenomena. It has gained a significant attention in the research community during the last decade. A number of commercial object-oriented database management systems (OODBMS) are available on the market today, although the relational database management systems still dominate the market. A more recent development is the *object-relation data model* that incorporates object-oriented concepts and functions into a relational database framework.

Relational data model has been the choice of implementing spatial data models in most GIS. Due to the limitations of handling coordinates data in a standard relational data model, some commercial GIS software packages also developed a proprietary file format to store the coordinate data and used a relational data model to store the attributes. The coordinates data in the proprietary file system are linked to the attribute data in a relational database through a set of unique identification numbers. The geo-relational data model used in ArcInfo (by Environmental Systems Research Institute, Redlands, CA) is an example. Object-oriented data models also attracted a lot of research interest in the GIS field. A number of commercial GIS software provides object-oriented programming tools for users to customize their GIS applications. However, most GIS software does not adopt a complete object-oriented data model in their implementation. In recent years, due to the extension of the relational data model to the object-relational data model among several major relational database management system (RDBMS) vendors (e.g., Oracle Spatial, Informix Spatial Data Blade, IBM DB2 Spatial Extender), commercial GIS software vendors also have started to adopt the object-relational data model. The Geodatabase data model in ArcGIS 8 is an example of the object-relational data model.

GIS data models traditionally focus on the handling of locations and attributes. Time is often treated as a static component in GIS. This is known as the *snapshot* approach since each GIS layer represents the locations and attributes of a particular phenomenon (e.g., land use) at a fixed time point (e.g., March 22, 2001). Computer scientists have been active on the development of temporal databases in order to better represent the time component in a database. In a relational data model, it is feasible to timestamp tables, records (i.e., tuples[1]) in a table, or attributes in a table in order to keep track of the time component at different levels. In an object-oriented data model, we may timestamp individual objects. These principles developed in temporal database

[1] Tuple: a fixed size collection of elements

research could be extended to spatiotemporal data models. However, when the time and space components and their interactions must be considered together, new challenges arise in developing such data models. A review of some key publications on these research issues is provided below.

(1) Spatial data models

- <u>Relational data model</u>

  Hadzilacos and Tryfona (1996) describe a GeoRelational data model (GRDM) as a relational data model because it incorporates five modeling constructs in GIS applications. These five constructs are layers, relations, virtual layers, object classes, and integrity constraints. A layer corresponds to a thematic data layer that describes geographic properties and relates key geometric features (such as points, lines, and polygons) equivalently. A relationship is used to represent non-geographic entities. A virtual layer is used to represent information computed from existing layers. The computations are defined as overlay, buffer, reclassification, and attribute derivation. Object classes are equivalent to spatial joins, and are defined on one or more layers. Integrity constraints are basically the first-order normalized form. These are used to process queries by selecting object classes that satisfy some specified constraints.

  Gadia (1993) assumes that a set of spatial features is a subset of a universal region. This leads to the closure property for the operators of union, intersection, subtraction, and complementation among the records in a relationship. Different from the GRDM, Gadia conceptualizes a region as an aggregate of simple geometries. In this model, a function is defined to map a spatial region to the attribute domain. Spatial operators are defined over regions to return a set of points that satisfy the conditions expressed in the operation. One limitation is the large amount of data redundancy.

- <u>Object-relational data models</u>

  ArcGIS 8's geodatabase data model is called an object-relational data model since it applies object-oriented data modeling concepts and functions to a relational database environment (Zeiler 1999). This data model permits users to define the features in their GIS databases with their natural behavior and allow any sort of relationship to be defined among features.

- Object-oriented data models

  A Bi-level Object-oriented GIS data model is presented in (Choi and Luk 1992). It facilitates data modeling at two levels: geographic and geometric. There are three main constructs at each level: object, object class, and functions. Objects are instances of object classes and are divided into primitive and non-primitive objects. Primitive objects are the built-in data types of a system. Non-primitive objects are those derived data types such as paths, physical regions, or themes at the geographical level, and points, lines, and polygons at the geometric level. Functions are used for data manipulation at both geographic and geometric levels. They include semantic retrieval (tools), set (union, distinct), aggregate (minimum, maximum, average), and superfunctions (shortest path algorithmic implementation). Geometric functions are categorized into five groups: overlap, containment, component, geometric object computation, and arithmetic computation.

(2) Temporal data models

- Relational data model with timestamp at the tuple level

  Clifford (1982), Clifford and Tansel (1983) describe time as an "instant" and discuss a historical database consisting of a collection of historical relations over the same set of states. A historical relation is viewed as a sequence of relation instances, indexed by valid time, each one representing a different state of the historical relation. All tuples are assumed to appear in all states of the relation. A complete relation contains a tuple in each state for every entity that appears in a relation at any state of the database history. Therefore, this model is highly redundant and Clifford (1983) does not even propose to implement the method.

  Jones *et al.* (1979) use time intervals in a data model, named as LEGOL 2.0, with two implicit time attributes - START and STOP. The two time attributes represent a closed valid-time interval. A similar model, Temporal Relational Model, is introduced in (Navathe and Ahmed 1988, Navathe 1989, Navathe and Ahmed 1993). However, this model distinguishes synchronous and asynchronous time-varying attributes.

  Sarda (1990, 1993) describes a Historical Data Model (HDM) to support historical relations that contain either tuples stamped with a time instant or stamped with a time interval. The data model enables users to input temporal attributes with time instances and query with intervals.

  Snodgrass (1995) proposes a conceptual data model – the Bitemporal Conceptual Data Model (BCDM) used as a basis for the definition of the temporal query language TSQL2. Data are time stamped either with sets of time instants or temporal elements. The BCDM is a unifying model since it can be mapped to several existing bitemporal data models. It serves as an appropriate basis for expressing time-varying data.

- Relational data model with timestamp at the attribute level

  Clifford and Tansel (1985) support historical relations with a time interval as timestamps. A historical relation may have four types of attributes: atomic attributes, triplet-values attributes (time-interval plus an atomic value), set-valued attributes (set of atomic values), and set-triplet-valued attributes (set of triplets).

  Dealing with vertical and horizontal temporal anomalies is discussed in Gadia and Yueng (1985), Gadia (1986, 1988), and Gadia et al. 1993. Vertical temporal anomaly refers to the fact that the history of a real world entity is spread over several tuples. There is no way to omit this in temporal data models within First Normal Form (1NF). Timestamps at the tuple level can be used to address this issue. Horizontal temporal anomaly refers to the problem that attributes of the same relation change their values at different time instants. One way to deal with this is by decomposing a temporal relation into relations in time normal form (TNF).

- Object-oriented data model

  Steiner (1998) introduces an object-oriented data model. A non-temporal generic OM model (Norrie 1992, 1993) is generalized into the temporal object data model (TOM). This means that all aspects of the model including constructs, operations, and constraints have temporal generalization. Furthermore, the temporal dimension applies not only to data but also to metadata. It introduces a new form of object timestamp. Instead of adding validity time period to objects in a form of special attributes that denote when a specific state of an object was valid, the notion of temporal object identifiers is introduced in TOM that timestamp objects with their entire time period of existence. TOM is a generic temporal object-oriented data model. Further, the membership associations between objects and the collections to which they belong are time stamped. With regard to the operational functions, TOM supports a full temporal algebra, a temporal query language, and temporal constraints. Additionally, metadata can have temporal properties that allow the modeling of role, association, and constraint lifespan.

- Extensible temporal data model

  Kafer *et al.* (1990) extend an object data model and describe a temporal data model that overcomes the vertical anomaly by mapping time sequences to a complex object data model MAD (molecule-atom data model). A time sequence represents the history of an entity of the real world by preserving the different states of the entity over time in a time-ordered sequence.

  Kafer and Schoning (1992) further extend a relational data model and use a time interval as a timestamp instead of using time sequences. They proposes a direct extension of the MAD data model to become the TMAD. The main difference between the MAD and the TMAD is that a valid-time interval is used in TMAD instead of a time instant in MAD to timestamp the atoms.

  Goralewalla and Ozsu (1993) use the extensibility of the object-oriented database management system, TIGUKAT, to implement a temporal data type. This temporal data type specifies an extensible set of abstract data type (ADT) and a rich set of behaviors to model time. Three different kinds of timestamps are supported: time instants, time intervals, and durations of time. It also includes types to model discrete, continuous, and dense time.

Su *et al.* (1991) extend the object-oriented data model OSAM * with time. In OSAM */T, objects are time stamped with a time interval. Rose and Segev (1991) and Rose (1993) present another approach using time sequences to extend a basic object-oriented data model with new types.

(3) Spatiotemporal data models

- Relational data model

Clifford (1983) discusses temporal information that is stored in flat files or tables in a relational data model. This method simply assigns a timestamp to each GIS relational table. Then a duplicate, but updated, copy is created and treated as the current state.

Snodgrass and Ahn (1985) propose to keep all related objects (records) together in a chorological order. This clustering method makes it easier for the application to find relevant subsequent records when trying to calculate the period of a particular state.

Lum *et al.* (1984) discuss a chaining method to deal with tuple level database design. The data model separates the time-variable and non-time-variable attributes that uses as little memory as possible and allows fast access to the historical data.

Stickler *et al.* (1992) examine the procedures needed to maintain records of the changes in spatial attributes for land parcels through time in ArcInfo. A solution is to construct a "shadow" history table associated with the intermediate link table. In other words, at any time T_n, the link table contains information at the current state (e.g., the mapping between polygons and land parcels) and each record also contains information about land parcels that existed at previous periods, T_{n-1}, T_{n-2}, etc., including time stamps indicating when a state came into existence. Thus the history table can be traversed backwards through time to extract information about previous states of the land parcel. The polygon table contains spatial information about the boundaries of all land parcels that arise through the fragmentation and concatenation process.

Hunter and Williamson (1990) propose to timestamp parcels by their date of creation and date of cessation. They argue that it is not practical to store full layers of geographical information for different time periods in digital cadastral databases. They suggest a system that keeps a graphics file of current parcels for day-to-day use while archiving historical spatial data into a separate file. Reference to the information is kept in files that store aspatial information via multiple versioned copies of the same parcel record.

Armstrong (1988) discusses timestamps of vector and raster GIS data models and describes them using entity, category, and relationship diagrams. For raster-based databases, an attribute history is proposed for each individual cell, thus avoiding the costly storage of whole data layers for each version. For vector-based databases, depending on whether durations are recorded explicitly using from and to dates or implicitly by recording a single timestamp, the method associates interval-stamped attributes with instant-stamped attributed in many-to-many relationships.

Raafat *et al.* (1991) extend the relational database of an image database system to handle both transaction and valid time intervals by adding new algebraic operations to the standard relational operations in order to manipulate the temporal dimension effectively. Beller (1991) presents a grid-based model of a temporal GIS. This article defines an "event" as an object with spatial and temporal extents. A temporal GIS is defined as a collection of Temporal Map Sets (TMS) and Neighborhoods with some operations. An event is a binary TMS where each cell is designated as either belonging to the event or not. A major shortcoming of this approach is its high data redundancy.

Raafat *et al.* (1994) recognize that changes affect both spatial and thematic attributes in a GIS and propose a relational model for accessing spatial and temporal topologies. In their model, a geographical entity goes through a series of historical states of various durations caused by mutations until it loses its essential property, which is normally a user-defined object identifier.

- Object-oriented data models

Wachowicz and Healey (1994) present an object-oriented spatiotemporal model of real-world phenomenon and events. Real-world phenomena are represented as complex versioned objects with geometric, topological and thematic properties. A new instance of an object with a different identifier is created for every version of the object establishing a hierarchical structure for the past, present, and future of the object. Events, on the other hand, are manifestations of actions that invoke update procedures on one or more objects. Time is represented as an independent, linear dimension. This is different from other representations where the time axis is orthogonal to the spatial dimensions.

Bonfatti and Monari (1994) describe an integrated approach to modeling both geographical structures and phenomena. They argue that cross-references between objects to express relationships are ambiguous; therefore, better means are needed to characterize object structure and behavior. Their proposed solution is the use of complex objects comprising of several components to express structure and relationships. In addition, laws that determine the possible states that an object can have are implemented to describe the behavior of the components. Spatiotemporal processes can be modeled easily with the framework by attaching timestamps to objects and expressing motions as laws for complex objects.

Worboys (1994) defines a spatiotemporal (ST) object as a unified object with both spatial and temporal components. An elemental spatial object (i.e., a point, a line), known as a simplex, is combined with a bitemporal element to form an ordered pair. A finite set of such ST-simplexes satisfying certain properties is then further defined to form an ST-complex on which query algebra is developed. An ST-complex traces changes in discrete steps; therefore, it is unable to represent continuous evolution, but is well suited for processes where mutations occur in sudden jumps.

Rojas-Vega and Kemp (1995) describe a structure for distributed, multi-media spatial applications and develops the Structure and Interface Definition Language (SIDL). To achieve full encapsulation needed by the distributed nature of the spatial database, the basic object type has a structural part and an interface part. The structure part defines an object identifier, conventional attributes, an object component grammar, and conceptual relationships, while the interface part contains methods operating on the object. With these two parts, complex object structures can be built to fully model real-world entities and their interactions. Time is introduced by separate objects that can be attached to time-varying components, with the use of separate objects for different models of time.

Yeh (1992, 1995) provides a model for highly variable spatiotemporal data using behavioral functions. A behavioral function forms a spatiotemporal object triplet together with a timestamp and spatial data to describe versions of data evolution. The extra information allows the modeling of complex interpolations, making the data less redundant and resolving data deficiency between states.

Hamre (1994, 1995) proposes a model based on a four-dimensional space consisting of points, lines, surfaces, volumes, and temporal volumes. The four classes of objects are defined in a class-hierarchy where subtypes of the four classes are designed to meet various needs. The implementation shows it is possible to integrate both vector and raster data along with non-spatial data in one spatiotemporal data model.

- Other approaches

Langran (1988, 1992) suggests a Space-Time Composite Data Model. It is based on the principle that every line in space and time can be projected to a spatial plane with other lines, thus creating a polygon mesh. Each polygon in this mesh has its own attribute history associated with it. Each new amendment is intersected with the already existing lines, and new polygons are formed with individual histories. The model has been tested with a number of indexing methods. The test results looked promising, but only small data sets were tested.

Peuquet and Wentz (1994) suggest that most systems are extensions of either a raster-based or a vector-based GIS. The raster or vector format of the extension depends on the type of queries that an application needs to handle. They observe that time-based representation questions can be answered by proposing a model to capture changes in the environment along a temporal vector. Starting with an initial state, events are recorded in a chain-like fashion in increasing temporal order, with each event associated with a list of all changes or can be triggered when gradual evolution is considered to be significant enough.

Peuquet and Duan (1995) propose an Event-based Spatio-Temporal Data Model (ESTDM), which is a triad representation of location-based (raster), feature-based (vector), and time-based components of modeling dynamic geographic objects.

Price (1989) creates an abstract data type (ADT) that keeps track of changes to the property lot sizes. A list of original lots and the subsequent changes to the lot size are stored. A nil value in lot size indicates that the lot no longer exists. Although this method does not address the attribute data temporality, it represents a departure from the versioning approach that requires extensive use of data archives.

Koeppel and Ahlmer (1993) propose two techniques for the integration of temporal data into Automated Mapping/Facilities Management (AM/FM) systems. Dynamic segmentation built on a topological data structure and associated temporal event tables can help keep track of changes. Data redundancy is avoided by not storing explicit topological information. The other technique uses change detection matrices in a spreadsheet to record differences between two time periods. Each axis represents a time period. Standard matrix algebra operations are used to derive change detection matrices for multiple time periods.

Rowe and Stonebraker (1987), Stonebraker (1986), and Stonebraker and Rowe (1987) use POSTGRES, which is an object-relational model based on a relational model including certain object-oriented features. These object-oriented features include additional built-in data types, ability to define abstract data types, support for inheritance, support for shared objects, capability of supporting rules, and functions for manipulating data. However, the main reasons of exploring POSTGRES as a suitable platform for a spatiotemporal GIS are its built-in capabilities for manipulating transaction time and valid time, its support for geometric data types and associated spatial operators, and its ability to manage "large object" types that can include bitmaps for raster data. The POSTGRES DBMS maintains historical data by saving deleted or modified data in a separate relation. This enables queries to be executed on past states of the database. The database maintains a record of all the tuples that have been updated along with approximate time stamps, which can subsequently be queried using the standard query language. Versioning capability in POSTGRES enables existing relations to be versioned at specific, user-determined points in time. A related facility enables versions to be created from "snapshots" of a relation, in which case it is merely a historical record of an entity at a particular point in time. Experience has indicated that relational level versioning is not suitable for most GIS applications.

Parent *et al.* (2000) propose a spatiotemporal modeling approach at the conceptual level, called MADS (Modeling of Application Data with Spatio-temporal features). MADS currently supports spatiotemporal features for applications related to land management or utility networks. MADS is designed in cooperation with GIS database designers and it is truly an object-relational conceptual model. MADS is the first one to explicitly obey the orthogonal principle of adding space and time to data structures. Spatiotemporal features may be associated to objects, attributes, and relationships. The space and time combination has an immediate desirable effect (i.e., MADS naturally supports modeling of moving objects). The spatial features of MADS support both the discrete and the continuous views of space. The temporal features of MADS are mainly characterized by the fact that no constraint is enforced over timeframes of related facts.

Moreover, MADS allows the specification of relationships that have space and/or time-related semantics. Thus, topological relationships may be defined as named schema elements, and bear attributes and methods. Similarly, synchronization relationships can be defined to constrain the lifespan of related objects. Transition and generation relationships are included in the MADS to meet user requirements.

Erwig (1999) presents a new approach to temporal data modeling based on a very general notion of temporal object. This paper proposes to embed temporal objects in a database through the use of abstract data type (ADT). This author focuses on object representations and establishes a representation hierarchy, concludings: "Spatio-temporal objects are special cases of temporal objects." The author expresses a wide range of integration opinions for temporal and spatial objects into spatiotemporal objects in relational databases and compares their expressiveness.. The investigation of the relative expressiveness of the different models gives a clear picture of the relationships between existing data models. In particular, it argues that, compared with the traditional approaches of temporal databases, the ADT approach is more versatile and offers much more control over temporal behavior, even for linearly constrained objects.

Kang and Choy (1995) point out that the relational data model is well formulated based on mathematical concepts of relations.. However, it is not suitable for representing a complex hierarchical structure, which is an important characteristic of most geographic objects. On the other hand, object-oriented data model can naturally represent a complex hierarchical structure, but there is a difficulty in sharing data with the relational data model that is currently used in most commercial GIS software. Therefore, this paper presents a hierarchical structuring method of geographic information using object grouping. This method supports various concepts of the object-oriented data model, with guaranteed compatibility with relational data model and SQL.

Yearsley and Worboys (1995) consider the modeling and design of a system that handles planar spatial objects and provides a rich topological representation and temporal reference. This paper aims at implementing a similar representation like a geometric layer and extends it through the addition of temporal information. It describes how abstract spatial data types can be constructed over the geometric layer that provides explicit topological information. The paper also considers steps toward a higher-level topological data model (spatiotemporal object layer). It then proposes that deductive database technology provides an appropriate paradigm to implement such a model.

The deductive database used in the project is Eclipse. Designed to be a fusion of object-oriented and deductive technologies, it is based upon an earlier system that is intended to be a platform on which next generation knowledge/database management could be built. .

2.2.2.3 Data Access

Another important group of research issues of spatiotemporal GIS is related to data access. A spatiotemporal database often has a large data volume. In order to provide a reasonable performance level when accessing spatiotemporal data, it is necessary to design appropriate indexing methods. Query languages are another important consideration of data access. As we move from attribute-based databases to spatial, temporal, and spatiotemporal databases, there arise additional requirements on the query languages. Finally, a GIS user must be able to interact with the system through user interfaces and to visualize the spatiotemporal data. This brings up the issues related to user interface design and visualization. A review of some key publications on these research issues is provided below.

(1) Temporal/Spatiotemporal data indexing methods

- Spatial indexing

  Lu and Ooi (1993) show the evolution of spatial index structures (see Figure 2.9). A solid arrow indicates a relationship between a new structure and the original structures that it is based upon. A dashed arrow indicates a relationship between a new structure and the structures from which the techniques used in the new structure originated, even though some were proposed independently. The indexes in the diagram are classified into four groups based on their base structures: binary trees, B-trees, hashing, and space-filling methods.

  During the design of a spatial index, the following needs to be minimized:
  (a) The area of covering rectangles maintained in internal nodes,
  (b) The overlaps between covering rectangles for indexes developed based on the overlapping native space indexing approach,
  (c) The number of objects being duplicated for indexes developed based on the non-overlapping native space indexing approach, and
  (d) The directory size and its height.

  There is no straightforward solution to fulfill all the above conditions. The fulfillment of the above conditions by an index can generally ensure its efficiency. In addition, the design of an index needs to take the computation complexity into consideration, although this is a less dominant factor considering the increasing computational power of today's systems. Other factors that affect the performance include buffer design, buffer replacement strategies, space allocation on disks, and concurrency control methods.

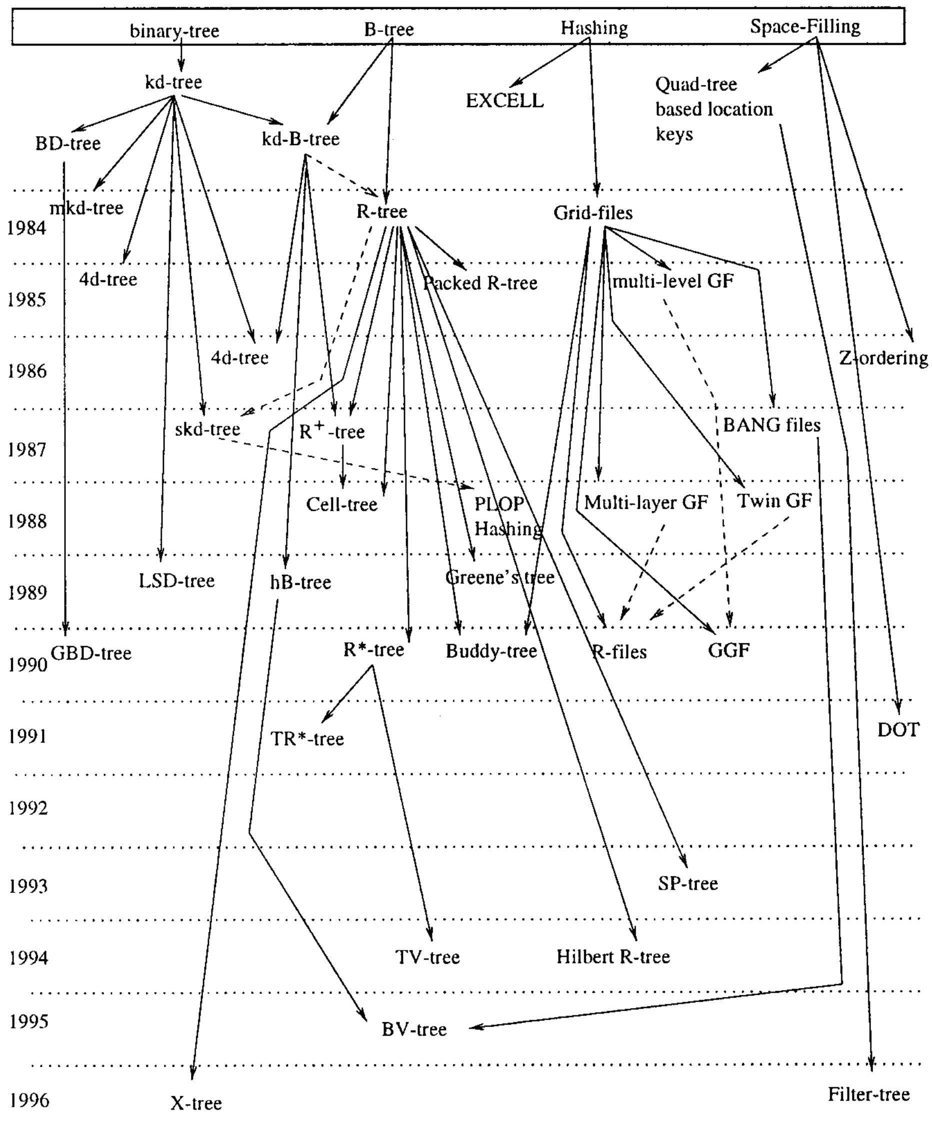


Figure 2.9 Evolution of Spatial Indexing (Source: Bertino et al., 1997)

- Temporal indexing

Bertino et al. (1997) survey a number of promising temporal indexes. Many of these indexes are proposed either for valid time or transaction time in a database. Researchers only began to work on indexing in a bitemporal databases recently. For transaction time databases, the TSB-tree approach is very efficient since it manages to keep the volume of I/O access low and uses tight bounding intervals to support fast search. However, it cannot handle disjointed intervals that may be present in the valid time databases. Direct application of B-trees (such as the AP-tree) includes indexing a single time point (starting or ending), which is efficient in terms of storage space, but is inefficient for any search that involves interval because no information of the actual data space in the "child node" is captured for pruning the search space. Hence, a simple time-slice search requires the scanning of a large portion of leaf nodes.

Spatial indexes such as the R-tree can be used for indexing both transaction times and valid times. R-tree can be used to index temporal data as line segments or points. As indicated by the experiments noted above, the performance of lines indexed by the R-tree is not as ideal as that of the TP-tree. However, if the lines are mapped into points, its efficiency should become comparable to that of the TP-tree.

- Spatiotemporal indexing

  Vazirgiannis *et al.* (1998) present a straightforward approach to index spatiotemporal data and treats time axis as just another dimension (like the spatial dimensions). A traditional indexing method, like R-tree or Quadtree, then can be used with these dimensions. This approach uses space that is proportional to the number of changes. However, this approach will lead to extensive object overlap and thus deteriorate the query performance.

  Langran (1988a) evaluates a simple form of spatiotemporal indexing, which is called partitioning. She argues that most temporal GIS applications will feature either spatial or temporal "dimensional dominance," and therefore proposes space and time-dominant partitioning approaches. The space partition creates partitions according to the data requirements, and then creates temporal layers of these cells. The temporal partition accumulates temporal data first and then organizes them into spatial cells.

  Easterfield *et al.* (1991) refute Langran's idea and argue that there is little need for temporally indexing spatiotemporal data. Because of the relatively long periods that data are valid, the effective reduction in the data volume by indexing for temporal query would not be sufficient to warrant its use.

  Hazelton *et al.* (1992) confirms the difficulty of indexing spatiotemporal data by highlighting problems with data access, database size, and disk storage schema.

  Xu *et al.* (1990) investigate the R-tree for spatiotemporal indexing. Improvements on the R-tree indexing methods can be made for handling spatiotemporal information in image sequences. The simplest method is to create an R-tree for each image and link them in temporal order. However, a space saving technique, called MR-Tree, can be used if the R-tree indexing consequent images are allowed to reference parts of the R-tree structure of the original image. An alternative approach is to use a single, modified R-tree structure, the RT-tree, that stores temporal interval information in each node with start and end dates denoting image recording time.

---

[1] Child node: an item of a tree referred to by a parent item
[2] Leaf node: a terminal or "bottom" item of a tree, i.e., an item with no child

Kolovson and Stonebraker (1991) create Segment Indexes to reference both interval and point data in a single index. Aspects of the Segment are merged with R-tree features in an attempt to improve the performance and efficiency of spatial indexing.

Kolovson (1993) continues the work by outlining two modified spatial indexing techniques that can be applied to spatiotemporal data. Multi-dimensional segment indexes handle historical data represented by time intervals, with time treated as a separate dimension.

(2) Temporal/Spatiotemporal query languages

- Spatial database query

GEOQL (Ooi 1988), SAND (Aref and Samet 1991), Spatial SQL (Egenhofer 1994), GEO-Kernel (Wolf 1989), and PSQL (Roussoppolous et al. 1988) are attribute-based Structured Query Languages (SQLs) of a relational data model. A spatial attribute is added to the list of non-spatial attributes, every tuple has a single value corresponding to the spatial attribute. The value of the spatial attribute could be of a spatial data type such as point, line, and region.

Gadia (1993b) proposes SpaSQL. In this tuple-based approach the spatial attribute is hierarchically subdivided based on the values of non-spatial attributes. For example, a geographic region such as a state can be subdivided into constituent counties or into soil zones. This can be represented as two relations with multi-valued non-spatial attributes. Referencing the same geographic region using multiple spatial subdivision schemes is referred to as restructuring.

Egenhofer (1991, 1992) presents abilities of displaying query results in a graphic form, which are beyond graphical query capabilities provided by other spatial SQL. In Egenhofer (1992), a Graphical Presentation Language (GPL) specifies graphical properties such as color, pattern, and symbols for displaying spatial objects. In addition, users can use five graphic operations in specifying and modifying the display.

- Temporal database query

Sadeghi *et al*. (1988) suggest a temporal extension of the relational query language called DEAL. DEAL has nested queries, conditional statements, loops, and function definitions that allow recursive queries. This makes it more powerful than traditional query languages, such as SQL. Both metric (longer, shorter, equal) and relationship (before, after, meets, etc.) operators are defined to handle temporal elements. Its similarity to spatial relationship operators and its expressive power makes this language a possible candidate for extensions of handling spatiotemporal query.

---

[1]DEAL: a relational language with deductions, functions, and recursions

Snodgrass (2000) proposed to develop temporal extensions to SQL the temporal database community. In response to his proposal, a virtual committee was formed to design extensions to the 1992 editions of the SQL standard. The outcome is now known as TSQL2.

- Spatiotemporal database query

  Qian and Peuquet (1998) propose a visual spatial query language for geographic information systems. This study intends to develop a language to construct sophisticated spatial and temporal data queries through a visual process of relevant function blocks or modules. Each module is essentially a grey box that encapsulates a fixed array of inputs/outputs and functions, while at the same time offers a set of "switches" and "knobs" through which users can fine-tune various data operations.

(3) User interface design/Visualization

Zhao and Brien (1999) describe the design of a prototype GIS-based tool for visualizing the interaction between land use and transportation. This tool provides many functions to facilitate the understanding of important land use and transportation issues and dissemination of information through the Internet. The main purposes of the project are to provide the planning professionals, the public officials, and the general public with a tool to access information regarding land uses and transportation improvements and to assist them in making land use-transportation decisions. This research identifies the relevant variables for visualization, the applicable technologies, and issues related to the development of visualization tools, including data availability, data collection, data conversion, integration of travel demand models and land use models.

Blaser and Egenhofer (2000) pursue an approach that is based on freehand sketching as part of a comprehensive prototype implementation of Spatial-Query-by-Sketch, which provides feature-based and relation-based spatial similarity retrieval. The method presented in the paper is unique because it allows users to freely specify each object's properties, particularly shapes, object types, locations, orientations, and spatial relations among the objects. This system seeks to match the topological, metrical, and the directional configuration of a sketch with that of similar sets of objects in the database. This approach simplifies the query formulation by shifting the responsibility of formulating a correct query from the user to the system. Spatial-Query-by-Sketch has been implemented in C++ and C using Microsoft's MFC classes for the user interface. It runs on Windows 95/98/NT and under Virtual PC - the Window's simulation on Mac OS. This paper also reviews the approaches to the use of query-by-example metaphor, which allows users to specify a sample of what they are looking for. Some approaches focus on image properties such as color, shape, or hue. Other approaches consider the spatial location of homogenous image areas, while a third group approaches favors object-oriented query models that use templates of abstract geometric elements. A metaphor adopted in the method is based on the analysis of sketched objects and their spatial relations.

Kraak and Maeachren (1994) review different aspects of visualizing dynamic processes. They define a temporal map as "a representation or abstraction of changes in geographic reality: a tool (that is visual digital or tactile) for presenting geographical information whose locational and/or attribute components change over time." In addition, they detail six dynamic visual variables for use in dynamic maps to complement the existing static ones: (1) display date - the time at which some display change is initiated, (2) duration - the time between two separate states, (3) frequency - the number of identifiable states per unit time, (4) order - the sequence of frames, (5) rate of change - the number of identifiable states per unit time for each of a sequence of frames, and (6) synchronization - the temporal correspondence of two or more time series. They also present several earlier classification attempts for temporal maps.

Koussoulakou and Kraak (1992) distinguish between three distinct methods of displaying processes: (1) static maps, with the temporal component transcribed graphically by means of variables, (2) series of static maps of progressive time slices, and (3) animated maps where change is observed through real movement on the map itself. Evaluations of the different methods indicate that there is a significant statistical difference in response time among the methods, with the animated maps being the preferred method. This leads to a conclusion that the utility of animated maps in spatiotemporal visualization is a viable option, while observing that it may not be the best choice.

Holmberg (1994) concentrates on giving users the ability of communicating with the display in a language form. The paper extends the Classical Map Language (CML) with dynamic variables to handle dynamic features on a map. The standard visual variables in CML (e.g. size, value, color, textures, grain, shape and orientation), which may be associated with complex map symbols made up from the basic point, line and area symbols, are complemented with dynamic properties of movement oscillation, pulsation and rotation to form the Dynamic Map Language (DML).

Trepied (1995) discusses language feasibility. This paper states that, to GIS end users, non-textual languages may prove to be the most useful for querying purposes. The alternatives, such as SQL, are considered either too difficult or poorly adapted to the user. The paper outlines the features and limitations of different non-textual languages. To provide a solution to the limitations, he proposes the use of dynamic icons that can represent temporal point, line and region type objects and can change shape, size, and color. They also can move, disappear, or be combined with other icons to fully depict the temporal processes in the environment. Each icon consists of a sign, some text describing the spatial-temporal object it represents and its temporal locations, a frame with boundary lines and interior, a color to make object type identification easier, and a screen location that assists in the determination of topological relationships between objects. This approach can lead to intuitive database schema/query expression/result visualization, uniform handling of spatial and temporal objects, and visualization of the dynamics aspects of geographical processes.

Slocum *et al.* (1993) discuss the development of a prototype visualization system. They review three approaches that are currently in use: (1) use separate software packages for individual tasks; (2) use a development kit to develop a library of tools; and (3) use a single specific exploration package. Unfortunately, none of these approaches provides satisfactory solutions. The paper proposes to add new features to support different data and map types, design animation and display multiple frames of animation on the screen. Comparisons of data sets can be done with graphical, statistical, and tabular displays of the data sets. In order to achieve the best results, the system, the GUI, and the set of tools are tested according to a participatory design, which means that potential users provide feedback throughout the development process.

2.2.3 Implementation Issues

2.2.3.1 Prototype-Based Implementation

- TRIAD

  Peuquet and Qian (1995) and Peuquet and Duan (1995) use an object-oriented approach to build TRIAD model. The conceptual representation under this approach consists of attributes, objects, and relationships. Entities have attributes, or properties, associated with them. An object is any type of data element that serves as the basis of the representation. Furthermore, an object is defined on the basis of the data it encapsulates. In other words, it includes attributes and the specification of operations that can be performed it. These operations may include inheritance that allows the definition of a new object class to inherit from existing objects. In the TRIAD model, there are three independent, but interrelated, representations – feature, location, and time. A feature is a thing that exists and can be seen or is purely conceptual. A feature serves as the basic "object" of the feature-based organization. Locations possess attributes that can vary with time. Events are used to track the changes over time. There are two types of events: (1) a change occurs at a time instant that is related to a specific feature or set of features, and (2) a change occurs at a time instant that is related to a specific location or set of locations. The data model of organizing changes around events is called Event-based Spatiotemporal Data Model (ESTDM). The ESTDM in its simple form stores event lists and their associated changes for a thematic domain. The ESDTM data structure consists of a header, a base amp that defines the initial world state for the entire geographical area at t = 0, and an event list of individual event entries. The prototype ESTDM is implemented in C programming language. An ESTDM-formatted file that represents the spatiotemporal dynamics of a single thematic domain for a specific geographic area is equivalent to a single thematic layer.

- SIIASA

  Yates and Crissman (1993) report the Spatial Information Infrastructure for Asian Studies in Australia (SIIASA), which is a project funded by the Australia Research Council for 1992 through 1994. The China GIS project started in 1992 and was the early major beneficiary of the SIIASA grant. This project's spatial database was designed from the outset to incorporate a temporal dimension. Therefore, a challenge of the project was to devise a means to maintain relationships among changing elements over time and to store the temporal variations in the spatial data economically in a non-redundant database. As the SIIASA project is required to support multiple GIS software platforms, it was also necessary to develop a general solution that could be implemented in a variety of different commercial software such as MGE, ArcInfo, and MapInfo. The solution developed for the China GIS project involves two basic approaches. First, for time-based spatial data, a space-time composite data model is adopted. Second, for time-based attribute data, a versioning method is employed. Changes in the attributes pertaining to a spatial feature will cause a new row to be created in the attribute table. Both the old version and new versions of the attributes are stored with approximate timestamps.

- ATLSS

Duke-Sylvester and Gross (1999) present an agent-based or individual-based model (Across Tropic Level System Simulation or ATLSS) that allows for variations in the state and behavior of the basic objects interacting within the model. Modeling each individual as a separate entity allows for spatially explicit components to be included so that the individuals can interact with a heterogeneous landscape and each other. Integrating spatial data into an agent-based system requires a significant level of GIS functionality be incorporated into the modeling system. This approach may seem redundant and costly, but current GIS systems do not offer a framework for building dynamic agent-based models. One of the goals of ATLSS project is to investigate the relative hydrologic scenarios over a thirty-year planning horizon. The ATLSS includes component models, each of which represents different biotic components of the Everglades system in Florida. These component models are linked together to form a multi-model. Each model within the ATLSS needs to be integrated at different levels. The level of aggregation is a function of many considerations, including computational efficiency, availability of empirical data to support a model, and the level of spatial and temporal resolution needed from the model. Each agent has its own separate set of state variable and behaviors. A collection of C++ classes is designed to handle different facets of spatial data. Objects from other classes provide an interface to several input and output devices.

The ATLSS also provides GIS-type functions that can be applied by the agents on the spatial data. Each spatial data set is associated with metadata that describes the spatial features. A set of classes has also been developed to coordinate each metadata set with the appropriate spatial data set. All data used in the ATLASS are raster data. Providing spatial data to the model means that a set of classes store the data and provide a single, common, and flexible interface. In the ATLASS project, these classes are referred to as the Landscape Classes. Input and output tasks are carried out with the IODevice class objects. These objects encapsulate the details of translating spatial data between the outside world and the Landscape class objects. The temporal aspect of spatial data sets is important but not dealt well in the project. All of the ATLASS models are spatially explicit. The individuals or functional groups are distributed over the landscape. The model does not incorporate spatial distribution in the same manner. However, the ALTASS uses the same data structure for both the spatial data set and the internal structure of a model. This design takes advantage of the C++ templates, which allow construction of generic classes that the basic structure of the class is described, but details concerning the type of data they manage is defined at the compile time.

One of the most significant uses of the Landscape classes is as a means of communicating between the component models. This allows several different models, representing different aspects of the system, to be joined together. All forms of communication with spatial data between models are carried out using the Landscape classes. Using the same set of tools for communicating information between models simplifies and unifies the models. The same set of cell referencing and geo-referencing tools can be applied to all parts of the modeling system. The same basic tools for manipulating the size and shape of spatially data also can be applied universally.

- Wang (1998)

Wang (1998) presents the design and implementation of a temporal GIS prototype that supports transportation applications. The prototype is built with the ArcInfo ODE (Open Development Environment) as the back-end GIS server and Visual C++ as the front-end user interface. The ArcInfo handles all spatial manipulations. These are accomplished by a number of stored procedures written in Arc Macro Language (AML). ArcInfo ODE behaves as a bridge between the user interface and the temporal GIS database. Each control in the user interface retrieves data from the database or loads data into the database through ODE. The prototype has three functions: data loading, data update, and temporal spatial queries. A concept about schema conversion is introduced by revealing that how a conceptual semantic schema can be converted into a relational database. However, the thesis does not further explore how the concept is implemented.

- Raza (1998)

Raza and Kainz (1998) discuss a temporal GIS (TGIS) that is implemented on top of the relational data model. Tuple-level versioning method is used to provide lower storage costs. The system is developed in ArcInfo and ArcView. It identifies three fundamental components of real world objects: location (spatial), attribute (aspatial), and time (temporal). The paper calls it spatio-temporal-attribute object (STAO). The spatial component can further be divided into two parts, which include geometry and topology. No further subdivision is considered for the temporal component. These STAO can be decomposed into spatiotemporal object (STO) and attribute-temporal object (ATO). In a land use change application, land use (attribute) is a vital concept; therefore, it is an essential property and other properties are considered to be non-essential. Based on OO concepts, a change in an essential property (land use) means birth of a new STAO or death of an existing STAO. On the other hand, a change in non-essential property triggers a new version of the same STAO.

A change model depicted in the paper consists of three layers: object, properties, and change. Each layer is classified into two components. Object layer has object and version, properties layer has essential and non-essential properties, and change layer consists of essential and non-essential changes. Incorporating the changes in STO is more complex than incorporating the changes in ATO. Graphic display of a STO is one of the fundamental requirements of any generic GIS. Incorporating the changes in STO and then display them is a complex process.

This study chooses a spatiotemporal composite model. Each change causes the changed portion of the coverage to break from its parent object and to become a discrete object with its own history. Two time dimensions (i.e., world time and database time) are attached to objects (STAO), atoms (STO and ATO), event, and evidence data types. In a relational database environment, three design alternatives (table-level versioning, attribute-level versioning, and tuple-level versioning) have been proposed for the temporal implementation in GIS. Object-oriented concepts have been incorporated into a relational data model. It is possible to create, edit, and delete STAOs in a relational data model. Tuple-level versioning has been adopted to implement a spatiotemporal model in ArcInfo. The event and evidence data types help us better understand the land use changes.

- Batty and Xie (1994a, 1994b)

Batty and Xie (1994a, 1994b) published two related papers. The first paper discusses strategies of linking models to GIS and the continuum between loose-coupling and strong-coupling. The authors present a strongly-coupled GIS-model system with models essentially embedded within the GIS. However, many modeling modules exist outside the GIS, thus it is arguably a loose-coupling GIS. Models and data are linked to GIS with the entire interface being driven through the GIS, the external modules being quite invisible to the user. These new functions are based on a large number of distinct external computer programs (more than 60 C codes and 150 macros), and these modules depend on the GIS to function. In this sense the model-based GIS can be argued as strongly-loosed coupling. The models in the papers are based on the monocentric theory of population density in contemporary urban economics and were applied to seven counties in western part of New York State. The interface between the models and their data is through the ArcPlot module in ArcInfo, which is organized as a menu-driven user interface to mimic the modeling process: data selection, data analysis, model selection, calibration, and prediction.

The second paper covers issues about the framework of embedding urban models within a GIS. Most proprietary GISs are built around representations, not simulations. GIS is therefore a storage and display medium for spatial data. If processes are to be modeled in whatever geographical domain, such modeling must be achieved outside of GIS. The coupling between GIS and the models discussed in these two papers establishes two distinct foci -- an emphasis on graphics for selection and evaluation, and an insight into data and modeling.

- LUCAS

Berry and Flamm (1994) present a Unix-based Land-Use Change Analysis System (LUCAS). The map layers used in LUCAS are derived from remotely sensed images, census, ownership maps, and output from econometric models. The structure adopted by LUCAS consists of three subject modules linked by a common database. These modules are socioeconomic model module (transportation networks, land cover, population density, etc.), landscape-change model module, and impacts model module. LUCAS is a computer-based application specially designed to integrate current and forthcoming information for (1) providing a multidisciplinary modeling environment to address research questions of land use and its impacts, (2) applying adaptive management research in order to address management questions concerning landscape impact assessment, and (3) designing a tool for workstations supporting the Unix, X-windows, and Motif user libraries. Geographic Resources Analysis Support System (GRASS) was the chosen GIS because it is a public-domain package available for many workstation environments. GRASS also is easy to be integrated into the LUCAS because it offers a series of map manipulation libraries with the source code and well-defined programming interface.

- Beller et al. (1991)

Beller et al. (1991) describe a prototype system which is an object-oriented tool set interfacing with a conventional GIS (Genemap) that stores raw satellite-derived vegetation index data with temperature and precipitation information for a selected study area. Additional capabilities provided by the system include: (1) temporal database management built on the Temporal Map Sets (TMS) concept; (2) temporal interpolation methods to provide continuity with a TMS even though only a limited number of slices may be available (i.e., the ability to transform existing TMSs into new ones); (3) use of animation to visualize temporal data with the "event" concept to enable investigation of causal relationships between objects; and (4) provision of export/import routines to interface with external statistical and modeling packages.

- GAEA

Hachem *et al.* (1992) present GAEA, an object-oriented scientific database management system for global change research. The system includes sophisticated analytical tools integrated within the database manager, and an operator set that is interactively extensible by the users. Specifically, GAEA includes spatial and temporal analysis capabilities, visual queries, distributivity due to database size, and location data. Data objects possess three dimensions: spatial, temporal and type. Type defines the operators available for the particular object. The components of the database are GUI, Spatio-Temporal Object Database (STODB), Query/Analysis Processor, the set of Current Objects of Interest, the Entry Processor (converting input data into the native format of the system), Entry Operators (describing the transformations taken during the conversion process), the Operator Object Base

(OOB, containing available operators for scientific analysis), and the Operator Editor (creating new or updating existing operators in the OOB).

- QUEST

Muntz *et al.* (1995) report on the QUEST (QUEries over Space and Time) project. QUEST is a prototype system built to serve as a testbed for "validating various techniques and demonstrating the feasibility and benefits of building information systems for atmospheric and earth science databases." System components include a GUI, a database manager, a visualization manager implemented in IDL for plotting, animating and analyzing data, and a query manager. LDL++ is a deductive database system chosen to manage complex spatiotemporal queries because of its extensibility, rapid prototyping and advanced query processing abilities. It provides additional reasoning power to that found in POSTGRES, which is used as a "preprocessor" to LDL++ queries. LDL++ is also preferred over traditional temporal query language since the latter normally require data to be stored in their own native format as opposed to the flat relational table used in POSTGRES. LDL++ offers a new dimension for queries using a rule-based approach. The Event Pattern Language (EPL), which utilizes event tables derived from the original cyclone relation, allows nested and recursive queries to further enhance analytical power.

- TOM

Steiner (1998) describes Temporal Object-oriented data Model (TOM). TOM is developed from TimeCenter technology and is generalized from non-temporal generic OM model. OM is a semantically rich object data model and used to investigate temporal object-oriented databases. TOM is generic and supports a full temporal algebra, a temporal query language, and temporal constraints. Additionally, metadata also can have temporal properties.

2.2.3.2 Simulation-Based Implementation

Theodoridis *et al.* (1999) implement an efficient benchmark environment for spatiotemporal access method, including modules for: (1) generating synthetic datasets, (2) storing datasets, (3) collecting and running access structures, and (4) visualizing experimental results. Data repository modules generate a collection of synthetic data that would simulate a variety of real life scenario. This paper discusses the parameters to be considered by a generator for data type describing temporal evolution of spatial objects. The authors propose an algorithm, called "Generate_Spatio_Temporal_Data" (GSTD), which generates sets of moving point or rectangular data that follow an extended set of distributions.

2.2.4 Summary

The literature clearly indicates the challenging nature of developing a complete spatiotemporal GIS. Although various approaches have been examined in the literature for a wide range of research issues, there is no single operational spatiotemporal GIS implementation that addresses all of these challenging issues. In order to examine the land use-transportation interactions under a spatiotemporal context for this research, a feasible and practical approach is necessary to allow the storage of the temporal component in an operational GIS environment such that spatiotemporal patterns may be managed, queried, analyzed, and visualized.

There are four major approaches to incorporating time into a GIS environment. The first approach is to use a *snapshot data model* that is based on conventional GIS software. It allows us to use the *Date* data type in a standard relational database as timestamps to keep track of the time instant or a time period of which a phenomenon existed. The second approach is the *space-time composite data model*. It creates a composite GIS layer that combines all changes ever existed and records the time elements associated with these changes. If all spatial features in the composite GIS layer existed for a single period of time, we can simply use the Date data type to record the begin time and end time. In reality, many of these spatial features could experience several changes during various time periods. This variable-length time list violates the normalization properties of a conventional relational data model. A third approach is the *event-based data model*. It organizes the data based on a time line and keeps track of the event lists that are associated with the changes to spatial features as well as to attributes. These event lists can be stored in a relational database as an Abstract Data Type, which is an extension of the conventional relational model that allows the handling of variable-length attributes. The fourth approach is the *three-domain data model* that incorporates the concept of semantic objects defined for a specific application domain. A semantic object can be linked to spatial features and their time components.

At the implementation level, the choices include a relation data model, an object-relational data model, and an object-oriented data model. The literature indicates the advantages and disadvantages of each data model for tackling various issues related to the implementation of a spatiotemporal GIS. Traditionally, most GIS software products are built around a relational database management system. This approach takes the advantages of the well-developed relational data model theories such as normalization and structured query language (SQL). On the other hand, it also suffers from the limitations imposed by the relational data model, especially for handling of spatial and temporal data. It is evident from the literature that object-oriented approach has been gaining a strong interest in the research community for representing spatiotemporal data. However, there exist some challenges that must be resolved before a fully functional object-oriented GIS will become available to the user community. A compromise position appears to be the object-relational approach that has been actively pursued and developed by both database management systems and GIS software vendors. This approach extends the conventional relational approach to incorporate the concepts and functions of object-oriented data modeling.

3. TEMPORAL GIS DATA MODEL DESIGN

The main purpose of this task was to design a temporal geographic information system (GIS) database in order to facilitate analysis of transportation and land use interactions based on historical data. The design considerations focused on the data management, analysis, and display functions of a temporal GIS. We first identified a framework of six scenarios that transportation planners are likely to examine the spatiotemporal relationships between transportation and land use changes. This scenario framework extended the measurement framework proposed by Sinton (1978) and added the capability of examining interactions between two phenomena (i.e., transportation and land use). Based on this scenario framework, we then designed a temporal GIS that integrated the time component into the geodatabase data model of ArcGIS 8 (ESRI, Redlands, CA). The temporal GIS design delivers the following key components:

1) A data model that addresses the issues of spatiotemporal data representations and spatiotemporal relationships for land use-transportation interactions;
2) A set of query and analysis tools that handle time-based, location-based, and attribute-based operations as well as the interactions among the three components for the analysis of land use changes and transportation developments; and
3) A user interface that is designed to facilitate the query and analysis functions mentioned above and to visualize the spatiotemporal change patterns.

The temporal GIS design process involved four recent developments in the GIS community. First, there have been various efforts of developing GIS for Transportation (GIS-T) data models. This project incorporated some object-oriented design concepts from these GIS-T data models to develop a temporal GIS for examining transportation and land use interactions. Second, we used the Unified Modeling Language (UML) to assist in the object-oriented analysis and design process. Third, the geodatabase data model of ArcGIS 8 was chosen as the specific implementation platform for the development of temporal GIS capabilities in the design process. The ArcGIS 8 geodatabase data model is an object-oriented data model implemented with a relational database management system. This object-relational approach is different from the previous georelational data model and brings a physical data model closer to its logical data model (Zeiler 1999). Finally, we used the ArcObjects development environment and a Microsoft Component Object Model (COM) compliant programming language to extend the geodatabase data model in order to develop custom tools in support of the temporal data management and analysis requirements of this project.

This chapter is organized into three sections. Section 3.1 discusses the six scenarios of analyzing transportation and land use interactions over time. Section 3.2 describes the GIS database design approach taken in this project. Section 3.3 presents the specifications of a temporal GIS data model design.

3.1 Six Scenarios of Analyzing Spatiotemporal Interactions between Transportation and Land Use

Sinton's measurement framework (see Table 2.9) is based on the concept that location, time and attribute are inseparable parts of measuring a real world entity (Sinton 1978). However, Sinton's measurement framework was designed to measure a single phenomenon, while our focus is to analyze the interactions between two phenomena (e.g., transportation and land use), with each phenomenon having its own location, time and attribute components. The three components therefore must be treated as integrated parts of each phenomenon. In this study, time is used as the organizing unit for examining the interactions between transportation and land use changes. The revised spatiotemporal interactions framework shown in Table 3.1 keeps track of the space (location), time, and attribute components of both transportation and land use phenomena, while adding time into a GIS environment to organize the spatial and attribute changes of each phenomenon over time. Current GIS are well developed in their abilities of performing integrated analysis of location and attribute. This study extends these GIS functions by adding the time dimension into the GIS database design using the ESRI's ArcGIS 8 software.

Table 3.1 Scenario Framework of Transportation-Land Use Interactions

	Fixed Component	Controlled Component	Measured Component
Scenario 1	Transportation (space-attribute-time)	Land use (space-attribute-time)	Time
Scenario 2	Transportation (space-attribute-time)	Time	Land use (space-attribute-time)
Scenario 3	Time	Transportation (space-attribute-time)	Land use (space-attribute-time)
Scenario 4	Time	Land use (space-attribute-time)	Transportation (space-attribute-time)
Scenario 5	Land use (space-attribute-time)	Transportation (space-attribute-time)	Time
Scenario 6	Land use (space-attribute-time)	Time	Transportation (space-attribute-time)

Each scenario in the above spatiotemporal interactions framework corresponds to one of the six ways that transportation planners are likely to ask questions about transportation and land use interactions. These six scenarios are described below.

- **Scenario 1** facilitates transportation planners to identify the time frames (i.e., measured component) when a transportation project (i.e., fixed component) might have had an effect on different land use types and their locations (i.e., controlled component). Sample queries under this scenario include:
  - When did different land developments in the region respond to the improvement of accessibility due to the I-595 construction project?
  - What was the temporal evolution of residential land-use changes in response to the I-595 construction?

  To support this scenario, the GIS graphic user interface (GUI) needs to assist the users with the following steps:

  Step 1: The user selects a transportation project. The system displays the location, temporal, and attribute information of the transportation project.

  Step 2: The user specifies a spatial extent (i.e., a search distance) and a land use change type (e.g., changes to residential land use type) around the selected transportation project. The default spatial extent is the entire study area.

Step 3: The system reports the years when the user-specified land use change took place within the search extent for all of the years with valid land use data. The system can also retrieve user-specified land use characteristics within the specified spatial extent and compute the rates of changes (if applicable) for the entire time period with valid land use data. The locational and temporal land use changes can be presented as: 1) maps showing when and where land-use changes appeared in the region, 2) charts and tables of the computed statistics, and 3) files of the statistics for further analyses.

Step 4: Steps 2 and 3 can be repeated for different user-specified spatial extents and land use characteristics.

- **Scenario 2** facilitates transportation planners to identify the land use changes (i.e., measured component) near a transportation improvement project (i.e., fixed component) during different time periods (i.e., controlled component). Sample queries under this scenario include:
  - Where were the new residential land use zones that were created during the five years after the construction of I-595?
  - Where were the areas that experienced greater than 20% land value increases since 1995 due to the construction of the I-595?

  To support this scenario, the GIS graphic user interface (GUI) needs to assist the users with the following steps:

  Step 1: The user selects a transportation project. The system displays the location, temporal and attribute information of the transportation project.

  Step 2: The user specifies a time period for analysis.

  Step 3: The system computes the distances between the selected transportation project and each land use data reporting zones. It then reports the land use data according to the time period specified in Step 2. Users can use the data to examine the spatial and temporal changes of land use patterns.

  Step 4: Steps 2 and 3 can be repeated for different user-specified time periods.

- **Scenario 3** facilitates transportation planners to identify the land use changes (i.e., measured component) near different transportation improvement projects (i.e., controlled component) at a selected time (i.e., fixed component). Sample queries under this scenario include:
  - What was the land-use pattern around major transportation projects in 1996?
  - Where were the vacant land parcels within a one-mile zone of an ongoing transportation project in 1996?

  To support this scenario, the GIS graphic user interface (GUI) assists the users with the following steps:

  Step 1: The user selects a time point (e.g., 1996).

  Step 2: The system automatically identifies transportation projects that are related to the time point specified in Step 1. The user selects transportation projects for analysis.

  Step 3: For each user-selected transportation project, the system computes the distances between the selected transportation project and each land use data reporting zones and reports the land use data according to distance ranges for the user-specified time point.

Step 4: Steps 2 and 3 can be repeated for different user-specified transportation projects.

❖ **Scenario 4** facilitates transportation planners to identify the transportation patterns (i.e., measured component) associated with different land use activities (i.e., controlled component) at a selected time point (i.e., fixed component). Sample queries under this scenario include:

- o What were the traffic volumes on major roads in traffic analysis zones with population density higher than 3,000 persons per square mile in 1996?
- o Which road segments in Broward County adjacent to commercial land use in 1998 experienced a level of service of D?

To support this scenario, the GIS graphic user interface (GUI) assists the users with the following steps:

Step 1: The user selects a time point.

Step 2: The system automatically identifies the land use data related to the selected time point specified in Step 1. The user selects land-use types for analysis.

Step 3: For each user-selected land use type, the system computes the distances between the selected land use zones and each transportation project. It then reports the transportation data according to distance ranges at the user-specified time point for further analysis.

Step 4: Steps 2 and 3 can be repeated for different user-specified land use types.

❖ **Scenario 5** facilitates transportation planners to identify the time frames (i.e., measured component) when a land use pattern (i.e., fixed component) impacted different transportation components (i.e., controlled component). Sample queries under this scenario include:

- o When did major transportation projects take place in traffic analysis zones with an employment density higher than 8000 per square mile?
- o What was the temporal evolution of traffic volumes on major roads in response to a major land development project?

The GIS graphic user interface (GUI) supports this scenario with the following steps:

Step 1: The user selects a land use characteristic (e.g., TAZs with an employment density of more than 8000 per square mile). The system displays the location of the selected land use characteristic.

Step 2: The user specifies a spatial extent (i.e., a search distance) and a transportation characteristic (e.g., transportation project) around the selected land use location(s). If the user does not specify a spatial extent, the default will be the entire study area.

Step 3: The system then reports the years when the user-specified transportation characteristic took place within the search extent for all of the years with valid transportation data. The system also can retrieve the specified transportation characteristics within the specified spatial extent and compute the rates of changes (if applicable) for the entire time period with valid transportation data. The locational and temporal transportation system changes can be presented as: 1) maps showing when and where land-use changes appeared in the region(s), 2) tables of the computed statistics, and 3) files of the statistics for further analyses.

Step 4: Steps 2 and 3 can be repeated for different user-specified spatial extents and transportation characteristics.

- **Scenario 6** facilitates transportation planners to identify the transportation system changes (i.e., measured component) associated with a land use pattern (i.e., fixed component) during different time periods (i.e., controlled component). Sample queries under this scenario include:
  - Where were the new transportation improvement projects during the five years after the construction of a major shopping mall?
  - Which road segments experienced greater than 20% traffic volume increases in a 3-year period since a traffic analysis zones reached a population density of 2000 per square mile?

The GIS graphic user interface (GUI) supports this scenario with the following steps:

Step 1: The user selects a land use characteristic. The system displays the location of the selected land use characteristic.

Step 2: The user specifies a time period for analysis.

Step 3: The system computes the distances between the selected land use location(s) and each transportation feature. The system then reports the transportation data according to distance ranges for the time period specified in Step 2. Users can use the data to examine the spatial and temporal changes of transportation patterns.

Step 4: Steps 2 and 3 can be repeated for different user-specified time periods.

3.2 GIS Database Design Approach

The above scenario framework serves as the foundation for the GIS database design. In order to implement a temporal GIS that can support the above six scenarios, we took a GIS database design approach that incorporated: (1) object-oriented GIS-T data models, (2) Unified Modeling Language (UML) for object-oriented analysis and design, (3) a selected commercial object-oriented GIS data model (i.e., the geodatabase data model of ArcGIS 8), and (4) custom programming with ArcObjects. This section provides background information of these four GIS database design approaches.

3.2.1 Object-Oriented GIS-T Data Models

GIS have been widely used in the transportation community. With the development of an object-oriented GIS data model, it is now feasible to define objects and their relationships that are specific and intuitive to a particular application domain. For example, it is more intuitive and useful for transportation planners to deal with objects such as roads, bridges, transit stations and traffic accidents instead of the geometric features such as points, lines and polygons in a GIS. There have been several efforts of developing GIS-T data models. Examples include the National Cooperative Highway Research Program (NCHRP) 20-27(2) project's Linear Referencing Systems (LRS) data model (Vonderohe et al. 1997), the GIS-T Enterprise Data

Model (Dueker and Butler 1998, Butler and Dueker 2001), and the Draft Framework Transportation Identification Standard prepared for National Spatial Data Infrastructure (NSDI) by the Federal Geographic Data Committee (FGDC, 2001). More recently, the NCHRP 20-27(3) project further extended the LRS data model to a multimodal, multidimensional data model for transportation location referencing systems that incorporated the time dimension (Adams et al. 2001, Koncz and Adams 2001). In addition, ESRI and a consortium led by the University of California, Santa Barbara jointly developed a Unified Network for Transportation (UNETRANS) data model. The UNETRANS data model is an object-oriented data model that can be used with the ArcGIS 8 by any transportation application development project. Basic transportation object classes and their relationships are defined in the UNETRANS data model. Users can modify the data model to include additional classes, properties and relationships to support their specific applications. This data model is specified in an industry-standard modeling language known as Unified Modeling Language (UML). All of these GIS-T data models offered valuable guidelines to the design of GIS-T databases. The temporal GIS data model developed in this project incorporated various design ideas from these GIS-T data models.

3.2.2 Unified Modeling Language (UML)

Object-oriented analysis and design (OOA&D) allows software developers to decompose a complex real world system into a set of diagrams to represent its processes, classes and relationships. These diagrams serve as useful information in a database design process. A number of OOA&D methods have been proposed in the literature (e.g., Chen 1991, Jacobson 1992, Booch 1993). These OOA&D methods are collectively known as the computer-assisted software engineering (CASE) tools. One problem with the various OOA&D methods was the different notations used in the diagrams to represent a real world system. In 1997, the Unified Modeling Language (UML) was adopted by the Object Management Group and became an industry standard.

UML is a graphic modeling language rather than a programming language. It consists of object-oriented analysis and design notations that can be used for everything from high-level analysis concepts down to very detailed design elements (Richter 1999). A set of diagrams defined with UML notations can describe real world activities and represent their concepts and relationships. Among the various UML diagrams, class diagrams are the most critical to this project. UML class diagrams translate real world objects into software entities with attributes, associations, methods, interfaces and dependencies. ArcGIS 8 includes a schema generation utility that can convert UML class diagrams into geodatabase schema. In addition, users can define custom objects with UML and use the ArcGIS 8 code generation utility to add custom object behaviors. In this project, we used the Microsoft Visio 2000 as the CASE tool to define and create UML diagrams (Visio, 2000).

Two examples, which are based on the UNETRANS data model, are provided below to illustrate some basic UML notations. Figure 3.1 shows a class hierarchy of different types of intersections and stations. It starts with the top box (ESRI Classes: SimpleJunctionFeature), which is an ArcGIS 8 predefined feature class of simple junction points in a geometric network. From this ESRI class, we can define a subclass of TransportJunction that inherits the properties and behaviors of its parent class (i.e., ESRI SimpleJunctionFeature class) along with additional attributes specific to the TransportJunction class (i.e., TransportJunctionID, Capacity, Impedance, Description, etc.). Intersection and Station in turn are two subclasses that inherit from the TransportJunction class. Again, additional attributes can be defined for each of these subclasses. This inheritance hierarchy allows us to define our project-specific classes based on a parent ESRI class. Furthermore, we can use a UML association notation to define subtypes of the Intersection class and the Station class. A subtype partitions the instances of a class (e.g., Station class) into a number of subtypes (e.g., BusStation, TrainStation, TransferPoint) that have the same properties and behaviors of its associated class.

Figure 3.2 shows a UML example of relationships between classes. In this example, a Road segment can be associated with 0 or more (0..*) traffic count stations while a TrafficCountStation can be associated with only one (1) road segment. The possible cardinality (also known as multiplicity) at each end of an association between two classes may be one and only one (1), zero or one (0..1), zero or more (0..*), one or more (1..*), or a specific number (n). ArcGIS 8 code generation utility recognizes these UML notations and uses the notations to automatically generate a geodatabase schema with one-to-many (1:M), many-to-one (M:1) or many-to-many (M:N) relationships defined in the UML diagrams.

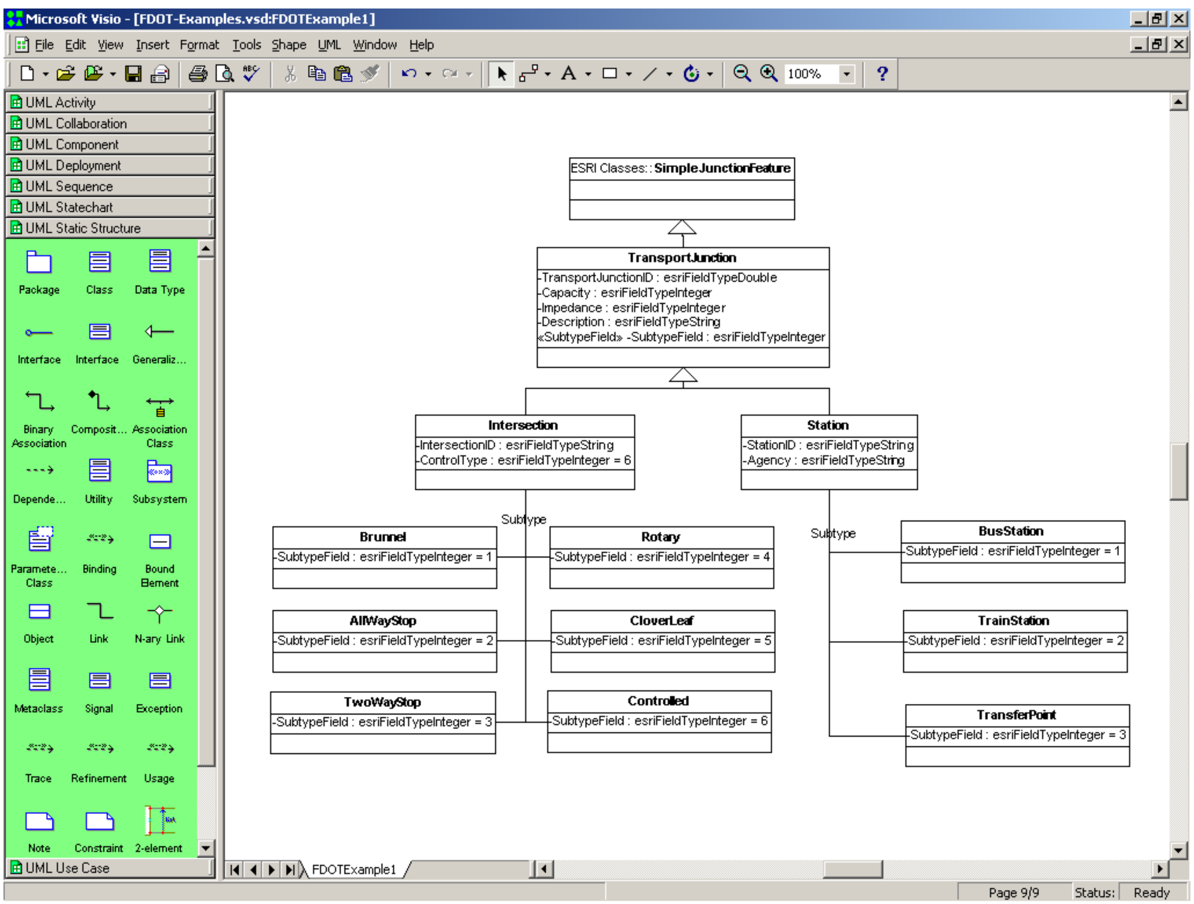


Figure 3.1 UML Example of Classes

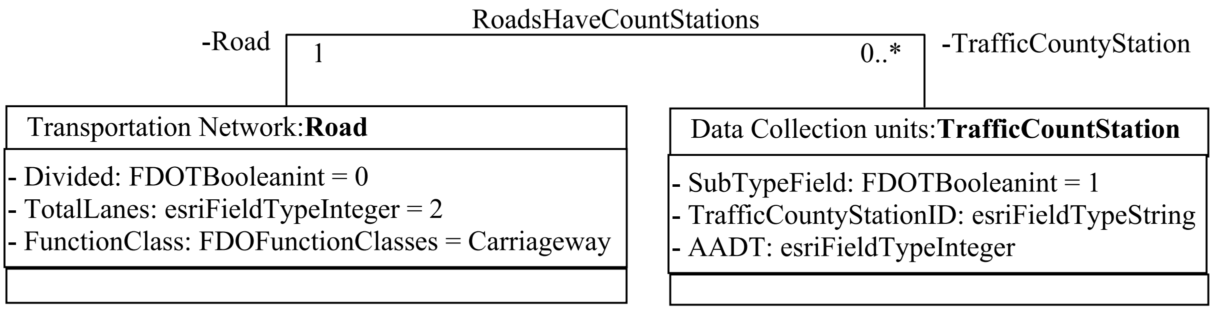


Figure 3.2 UML Example of Relationships

The classes can be accessed through interfaces with custom programs, which will be discussed in the next section. A single class may expose one or multiple interfaces. An interface is a collection of related properties and methods and is separated from the implementation of the classes. In a traditional object-oriented environment like Avenue, there is a one-to-one correspondence between the way we interact with the class and the implementation of the class. With ArcObjects as a COM development platform, these two parts are separated. Therefore, a class can provide multiple interfaces to manipulate the class (Figure 3.3). The UML refinement notation is used to express the relationships between a class and its interfaces.

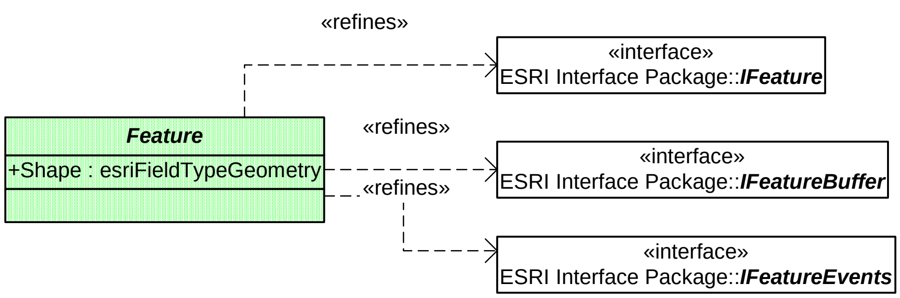


Figure 3.3 Feature Class and Its Multiple Interfaces

3.2.3 ArcGIS Geodatabase Data Model

ArcGIS 8 introduces a new object-oriented data model – the geodatabase data model. In a geodatabase, it is feasible to define objects as they exist in the real world and the relationships between the object classes. For example, instead of using generic geometric features (i.e., points, lines and polygons) to represent different kinds of real world entities, we now can define a traffic count station feature class or a traffic analysis zone feature class in a geodatabase. We also can define transportation projects as a separate object class (i.e., a semantic object) and then create a relationship class to link each transportation project to its related street segments. Furthermore, users can define attribute domains and validation rules in a geodatabase.

The geodatabase data model observes the object-oriented concepts of inheritance, encapsulation and polymorphism. However, its implementation is on a relational database management system (RDBMS). The default RDBMS for storing the geodatabase data model is the Microsoft Access, which is known as a personal geodatabase in ArcGIS 8. With the use of ArcSDE, the geodatabase data model can store data in other commercial RDBMS such as Oracle, SQL Server, Informix, DB2, and Sybase. The object-oriented geodatabase data model provides many new functions that were not available in the georelational data model used in earlier ArcInfo releases. Most importantly, the geodatabase data model of ArcGIS 8 offers an environment for application developers to extend the basic data model for different application domains such as the development of a Transportation GIS data model. Additional information about the geodatabase data model can be found in two introductory books: Modeling Our World: The ESRI Guide to Geodatabase Design (Zeiler 1999) and Building A Geodatabase (MacDonald 1999).

3.2.4 Custom Programming with ArcObjects

ArcObjects is the development platform of the ArcGIS 8 products. Since ArcObjects is built on the Microsoft Component Object Model (COM) technology, it is possible to extend ArcObjects using any COM-compliant development languages (Zeiler 2001). COM is a client/server architecture of which the server (or object) provides some functionality that a client can use. COM defines a protocol that connects one software component with another in a client/server environment. COM also defines an interface-based programming model that encapsulates the data and methods with each instantiated object behind a well-defined interface. ArcGIS installation includes a set of ArcObjects object model diagrams that use the UML notations to define the ArcObjects classes (with their properties and methods), the interfaces available for each class, and the relationships among the classes. With this architecture, it is feasible for users to extend any part of the ArcObjects architecture as the ESRI developers do.

In order to develop a set of query and analysis functions in ArcGIS 8 to support the spatiotemporal study of transportation and land use interactions in this project, we used the geodatabase data model and the ArcObjects with a COM-compliant programming language to develop custom tools for this project. The interface-based programming model of ArcObjects means that all communications between objects are made via their interfaces (Zeiler 2001). An interface determines what requests can be made of an object that has an implementation of the interface. Two ArcObjects classes can have the same interface, but their implementations of the interface may be different. This is known as the polymorphism in an object-oriented approach. In this project, Visual Basic 6 (VB 6) was used as the main programming language to create function-specific dynamic linking libraries (DLLs) that supplement the existing functions in ArcGIS 8 to handle the temporal data.

3.3 Temporal GIS Data Model Design with UML

Conventional GIS are based on a static approach of data representation. Each GIS map layer is treated as a snapshot that captures the spatial pattern of a selected theme at a given time (e.g., land use of 1998). The time component usually is not recorded in a snapshot GIS database. This approach of storing geographic data in snapshot GIS layers without explicit treatment of time introduces many shortcomings when we need to examine changes of spatial and attribute data over time. For example, a GIS overlay function must be repeatedly performed on pairs of snapshot land use layers in order to find out the land use change pattern from 1980 to 2000 (see Figure 3.4). This is an inefficient and cumbersome way of analyzing land use change patterns over time.

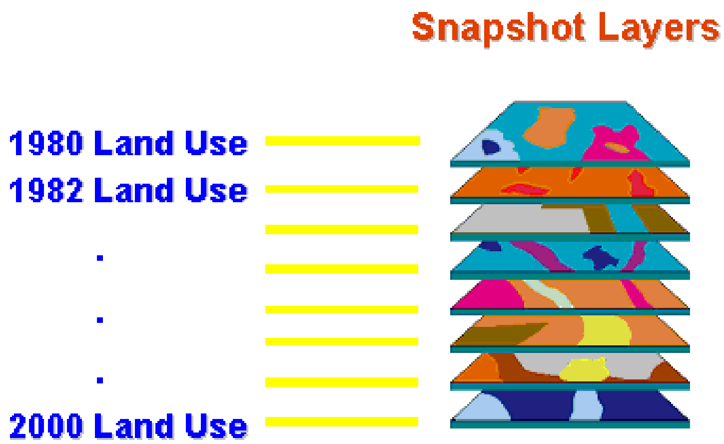
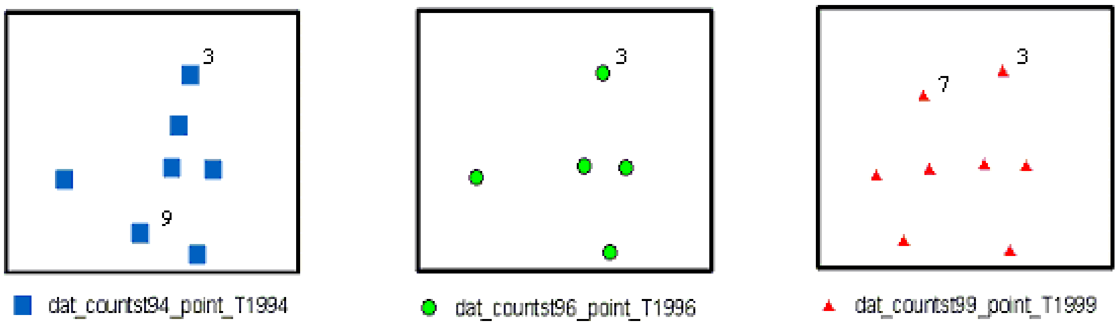
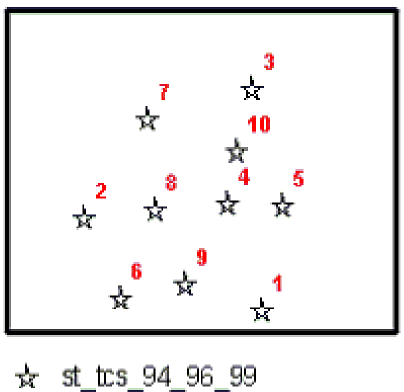


Figure 3.4 An Example of Snapshot GIS Layers

It becomes even more challenging when we need to take into account the time lag effect of interactions between land use and transportation system changes. Certain land use changes (e.g., opening of a major shopping mall) may have an immediate major impact on traffic volumes in their surrounding areas, while other land use changes (e.g., a downtown redevelopment project) tend to induce changes in traffic patterns over a longer time period. Similarly, different transportation projects (e.g., construction of a beltway in a metropolitan area, expansion of an existing major highway, or building of a light rail system) are likely to have impacts on land use patters over different time horizons. In these cases, transportation analysts must measure the changes of one phenomenon (e.g., land use) introduced by the changes of another phenomenon (e.g., transportation) over an extended period of time. Current GIS based on a snapshot approach do not provide adequate supports for such spatiotemporal query and analysis needs.

Change is a key concept of temporal GIS. Worboys (1998) suggests that a temporal information system must include a history of changes in the application world. The space-time composite data model is an attempt to improve the snapshot data model by combining the separate GIS layers representing different time points into a single GIS layer of all changes that have occurred over all time periods (Langran and Chrisman 1988, Langran 1992). In a space-time composite GIS layer, each record represents the smallest geographic unit that experienced changes in history and is assigned with time stamps to keep track of its changes over time. For example, Figure 3.5 shows a space-time composite layer that combines the traffic count stations of 1994, 1996 and 1999. The time stamps indicate that traffic count station #3 existed in all three years and traffic count station #7 existed in 1999 only. For line and polygon features, snapshot GIS layers of each theme for different years are again combined to derive the space-time composite layers (Figure 3.6). The resulting space-time composite databases are able to answer questions such as "Which locations experienced land use changes between t(i) and t(j)?" via a simple query of "Select all land use polygons where the land use type at t(i) is not equal to the land use type at t(j)."





OBJECTID	T1994	T1996	T1999
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	1
6	0	0	1
7	0	0	1
8	0	0	1
9	1	0	0
10	1	0	0

Figure 3.5 Space-Time Composite Point GIS Layers

Yuan (1996, 1999), on the other hand, proposes a conceptual structure of three-domain data model that consists of semantic domain, temporal domain, spatial domain, and domain links. The semantic domain defines real world entities with unique identifiers throughout the study duration. The temporal domain stores each time instance as a unique object, while the spatial domain is based on the space-time composite data model to derive a set of common spatial features with unique identifiers. Domain links are used to record the links among semantic, temporal, and spatial objects with their unique identifiers. This project adopts and extends several temporal GIS approaches reported in the literature to design and implement a temporal GIS for exploratory analysis of land use and transportation interactions.

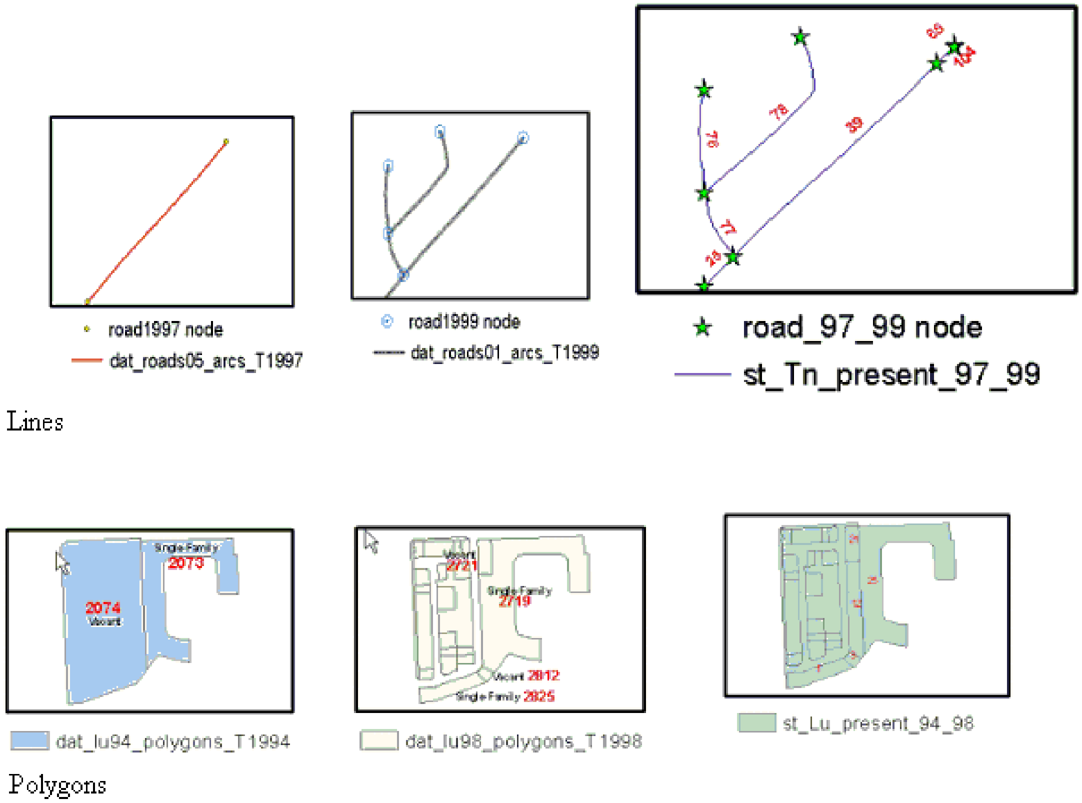


Figure 3.6 Space-Time Composite Line and Polygon GIS Layers

Figure 3.7 presents a unified modeling language (UML) diagram of the overall design framework of this project. This temporal GIS design is named “Transland” because of its capability of supporting spatiotemporal analysis of transportation-landuse interactions as defined in the six scenarios above. The system design is presented as related components specified in UML notations using the Microsoft Visio 2000 software (Visio, 2000).

The "snapshot component" block in Figure 3.7 stores and manages the snapshot GIS data sets using the built-in functions available in ArcGIS 8. These snapshot data sets are organized into transportation-related data sets and land use-related data sets, respectively. Time associated with these snapshot data sets is often implied in the file name or provided through a metadata document. Explicit time stamps normally are not included in the data fields of these datasets. Therefore, the first step is to add time stamps into the snapshot data sets (as indicated by the t(i) and t(i+j) loops in Figure 3.7).

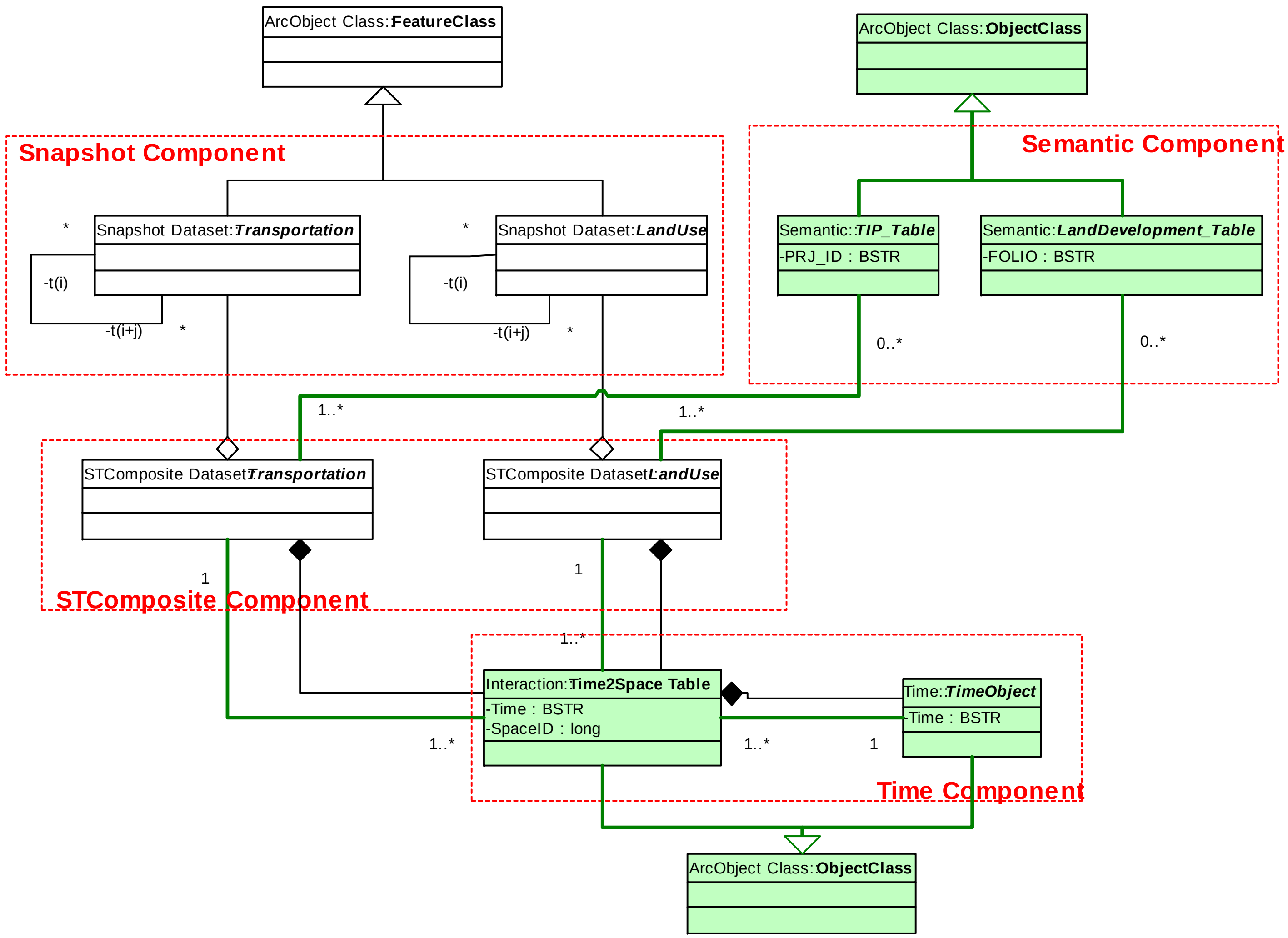


Figure 3.7 UML Diagram of the Transland Design Framework

The "space-time (ST) composite component" block in Figure 3.7 consists of a set of custom Visual Basic (VB) programs to generate space-time composite GIS databases from the snapshot GIS data sets. Each space-time composite GIS database combines all snapshot GIS layers of a particular theme (e.g., land use) into a single GIS database that consists of spatial and attribute changes over time. In addition, the custom VB programs also automatically create a Time object class to record all unique time instances associated with the snapshot GIS layers, as well as a Time2Space object class to associate the unique identifier of each geographic unit in a space-time composite database with all time instances that exist in the individual snapshot layers. The Time object class and the Time2Space object class make up the "time component" block identified in Figure 3.7.

The "semantic component" block in Figure 3.7 is designed to handle entities that are associated with different locations and attributes at different time instances. For example, a transportation improvement project (TIP) often consists of multiple phases. Each phase may be associated with works on different highway locations and may have different attribute values (e.g., budget amount spent at each phase). To facilitate spatiotemporal analysis of processes such as transportation improvement projects and land development projects, we also define Semantic object classes that allow linkages of a single semantic object instance to multiple locations and attributes over time.

The Transland system design is built upon the ArcGIS geodatabase data model and the ArcObjects development environment to create additional functions and tools for handling temporal data and spatiotemporal analysis of transportation-landuse interactions. In general, the Transland system consists of four groups of related objects: ESRI class package, ESRI interface package, Transland class package, and Transland custom tool package.

3.3.1 ESRI Class and Interface Packages

The ESRI classes and interfaces come with the ArcGIS installation as part of the ArcObjects development environment. The top-level object class in the ESRI Class Package is the *Row* Class (see the simplified UML diagram in Figure 3.8). *Row* is a basic concept in the relational data model. The ESRI class package builds around the row concept since its implementation is based on a relational database management system. However, it also extends the relational data model to manage object and feature classes. An *Object Class* is a table that consists of rows representing the entities (Zeiler 2001). A *Feature Class* is an Object Class whose objects are spatial features with geometry. The geometry is stored under a “shape” data field defined as the esriFieldTypeGeometry in a relational table. In ArcGIS, the geometry types include Point, Multipoint, Polyline and Ploygon. The ability of storing geometry in a relational database is one characteristic of the ArcGIS geodatabase data model that moves it from a relational data model into an object-relational data model. The *Object* and *Feature Classes* are critical to the creation of custom features in this project. All custom classes in the FDOT Transland database design are created from these two ESRI classes.

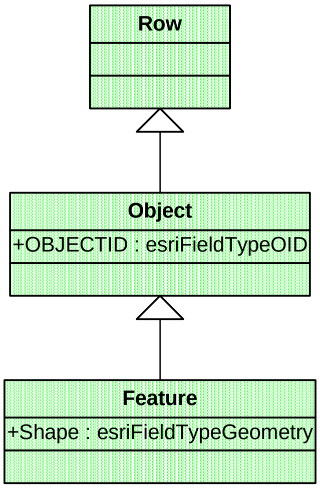


Figure 3.8 ESRI Row, Object and Feature Classes

Interfaces are the behavior of classes. As mentioned before, COM is an interface-based programming model that encapsulates the data and methods with each instantiated object behind a well-defined interface. In other words, COM-compliant custom programs access the properties and methods of various objects through the interfaces implemented with specific object classes. COM interfaces therefore are different from the common use of the word of interface in the context of user interfaces. The ESRI ArcObjects data model provides a rich set of interfaces for programmers to access and manipulate various ArcObjects classes in different modules. For this project, the most relevant modules are ArcCatalog, ArcMap, Geodatabase, and Geometry. By convention, interface names start with a letter of *I*. For example, *IRow, IObject, and IFeature* are interfaces for the *Row* class, the *Object* class, and the *Feature* Class, respectively. All COM interfaces derive from the *IUnknown* interface, and all COM objects must implement this interface (Zeiler 2001). Like the ESRI object classes, an ESRI interface can inherit from another ESRI interface. For example, Figure 3.9 shows that the *IRowEdit* interface inherits from the *IUnknown* interface. The *IFeatureEdit* interface in turn inherits from the *IRowEdit* interface. This figure also illustrates a standard UML notation that includes three sections in each box. The top section shows the name (e.g., *IUnknown*). The middle section indicates the properties (e.g., SubtypeCode is a property of the *IRowSubtypes* interface class), while the bottom section lists the methods available to an interface (e.g., get_Fields, get_Value and put_Value are the methods available through the *IRowBuffer* interface to read the data fields, to read the data values, and to write the data values, respectively.)

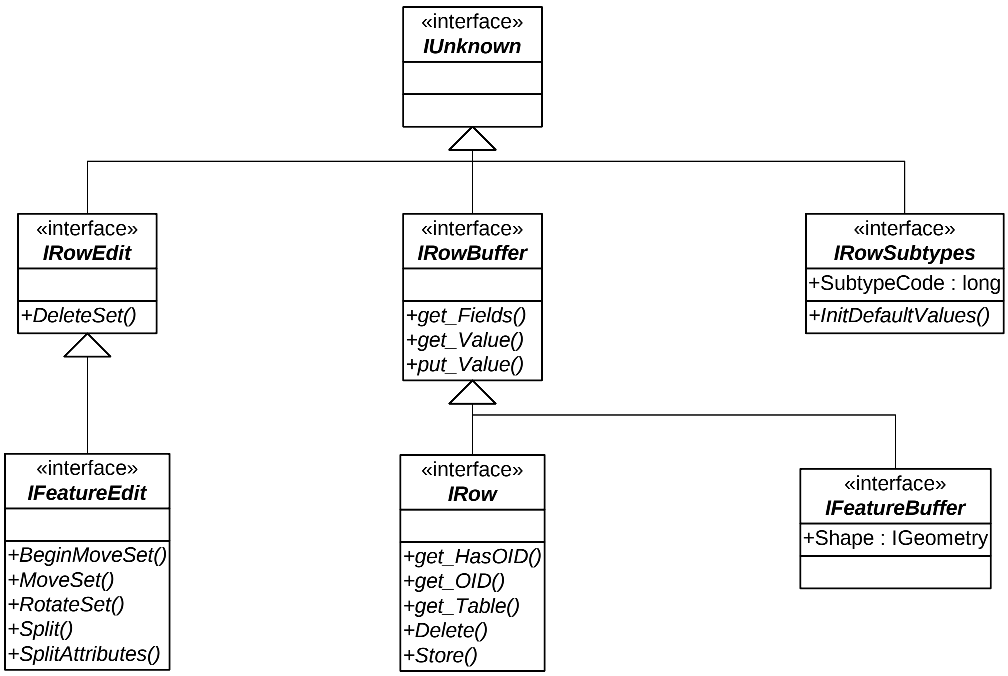


Figure 3.9 Partial ESRI Interface Package

3.3.2 Transland Class Package

The Transland Class package is designed to support the spatiotemporal GIS needs arising from the six scenarios discussed in Section 3.1. This package includes four main sub-packages - FDOT dataset package, STComposite package, Time package, and Semantic package. Brief descriptions of these sub-packages are provided below.

3.3.2.1 FDOT dataset package

The FDOT dataset package is created to hold all source data sets collected for this project. It is further subdivided into the TransportData sub-package and the LanduseData sub-package. Figure 3.10 illustrates the TransportData package that covers point data sets (e.g., traffic count stations), line data sets (e.g., street networks), geometric network data sets (for complex network features and network analysis), and polygon data sets (e.g., TAZs). These data sets can be ESRI coverages or shape files converted into a geodatabase.

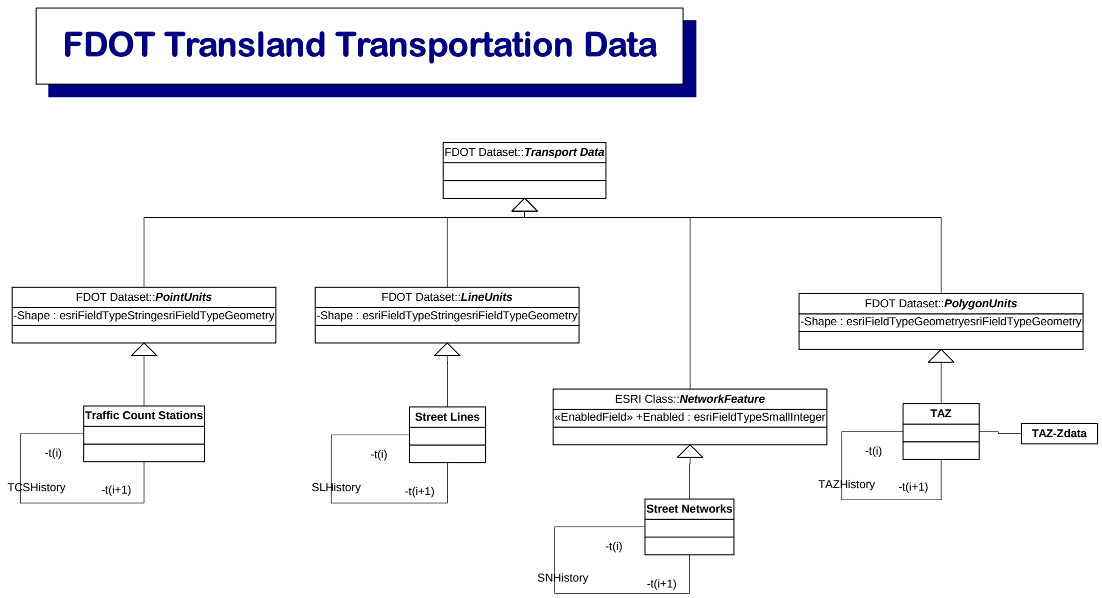


Figure 3.10 FDOT TransportData Package

3.3.2.2 STComposite Package

The space-time composite data model (STComposite) is an approach of combining snapshot GIS data sets into a composite data layer that minimizes spatial data redundancy while keeping track of the temporal changes associated with individual spatial elements. We defined three STComposite classes (i.e., PointSTComposite, LineSTComposite and PolygonSTComposite) to handle points, lines and polygons, respectively (see Figure 3.11). Creation of these STComposite GIS layers is based on the existing spatial relationships (via the ESRI IRelationalOperator interface) and topological operators (via the ESRI ItopologicalOperator interface) available in ArcObjects (Zeiler 1999, 2000).

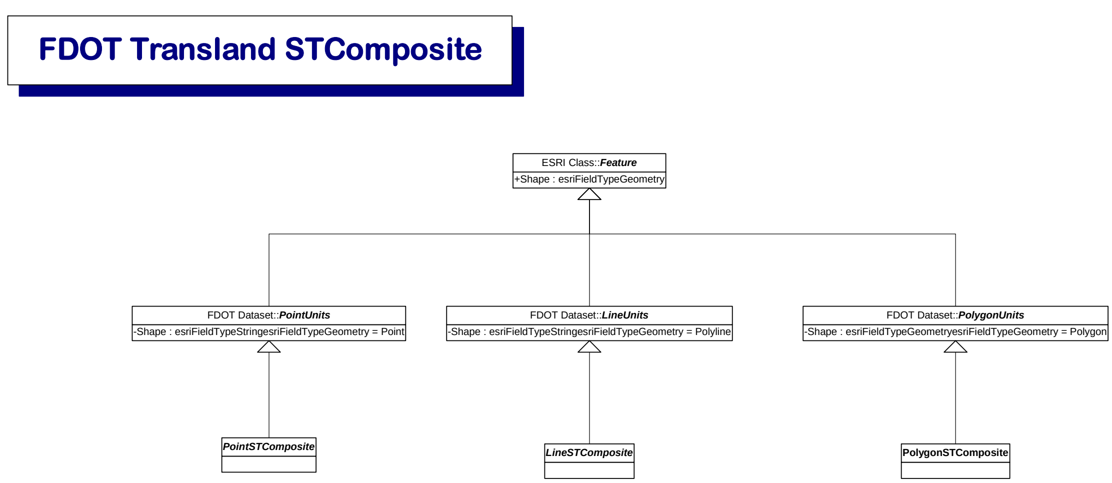


Figure 3.11 STComposite Package

3.3.2.3 Time Package

Time is an inseparable component to any spatial and attribute data changes. One major challenge of temporal GIS database design is where and how we should store the temporal information in relation to the spatial and attribute data. The commercial GIS products available on the market have developed efficient ways of integrating spatial and attribute data for various spatial analysis needs. Our approach is to take full advantage of the existing GIS software capabilities while extending them into the temporal domain. In the Transland database design, time object classes are created to store temporal data explicitly. These time objects in turn are associated with their corresponding feature/attribute data derived from the STComposite layers. This design approach makes it feasible to access the temporal, spatial and attribute data individually as well as to easily access their associated data in other domains. For example, users could first use the existing GIS functions to select features of interest based on locations and/or attributes. Once these features are identified, their associated temporal data can be easily retrieved for spatiotempooral analysis. Users also could use the time objects defined in the Transland to quickly find the features of interest at a given time instant or during a specific time interval. Existing GIS functions then could be used to display these features or to perform additional analysis of these features. Associations between the time objects and their corresponding features are maintained through the unique IDs generated in the STComposite layers. The same time object can be associated with different types of features (e.g., traffic count stations, TAZs, street segments, property parcels). This allows users to access all relevant data from the time objects. As a result, it is possible to examine the interactions between different data sets through the time domain.

The literature review suggests that a temporal GIS database should include both valid time (i.e., when an event actually occurs in the real world) and transaction time (i.e., when a transaction takes place in the database). Both of them are included in the design of Time package. However, most data sets collected for the project do not include information on transaction time. The system implementation therefore focused on valid time only. Transaction time class could be implemented in the future when needs arise. Three different valid time object classes are defined in this project. The TimeInstant class is for single time points. The TimeInterval class is defined by a start time instant and an end time instant. The TimeEventList class can consist of multiple time instants and/or multiple time intervals.

3.3.2.4 Semantic Package

The Semantic package is designed to deal with application-specific entities that do not fit well with the conventional GIS layer concept. For example, a transportation improvement project or a land use development is best represented as an object by itself that is associated with different spatial and attribute data over time. Transportation planners often need to retrieve the temporal, spatial and attribute information related to a transportation improvement project or a land development project. The most intuitive way of performing such tasks is to access them through the project names or the project IDs. Semantic object classes defined in the Transland facilitate these tasks. There are two initial types of semantic classes included in the Transland, which are TransportSemantic and LanduseSemantic (see Figure 3.12). They inherit from the ESRI Object class. It is anticipated that transportation improvement projects could be represented as points, lines or polygons. Therefore, the TIPSemantic class consists of three subclasses: TIPPoint, TIPLine and TIPPolygon. Additional semantic classes could be defined in the Transland to accommodate future user needs.

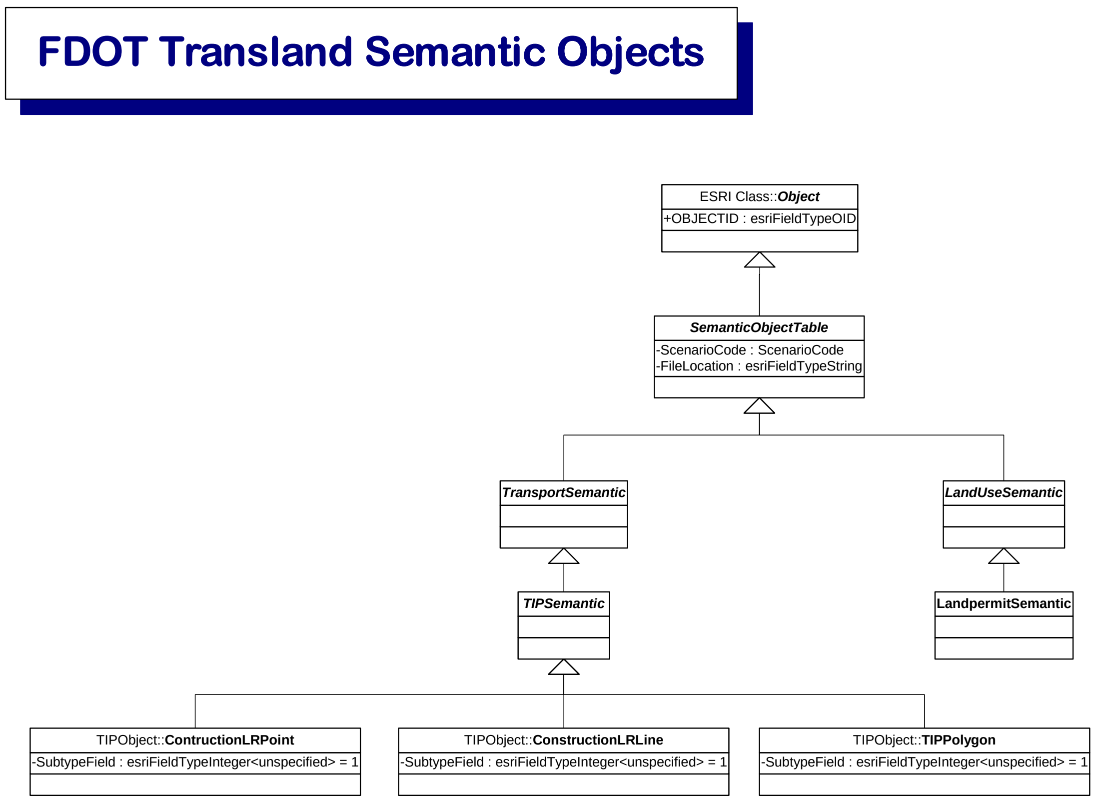


Figure 3.12 Semantic Package

3.3.3 Transland Custom Tool Package

Several custom tools are needed to manage and create the FDOT datasets, the STComposite layers, the semantic objects and the time objects as discussed above. First of all, source data sets collected for this project are available in different formats (e.g., ArcInfo coverage, ArcView shapefile, dBase files, and text files). It is also common to encounter a problem that the same feature type (e.g., TAZs) of different years consist of different data fields and/or different data field names. Custom tools therefore must be developed to consolidate these source data sets and convert them into a geodatabase stored in the Microsoft Access format. Secondly, source data sets can include point, line, or polygon features. In order to create STComposite layers from these different geometry types, we need custom tools to derive the spatial, attribute and temporal changes of point data, line data and polygon data. Furthermore, custom tools are required to organize and store the temporal data derived from the STComposite creation process into the Time object classes.

In addition to the custom tools required to create and manage temporal GIS databases, we also need to provide application tools to support spatiotemporal analysis based on the six scenarios presented in Section 3.1. ArcObjects with Visual Basic are chosen to serve as the main development environment for these custom tools. In many cases, dynamic linking libraries (DLLs) are created for these custom tools. These DLLs are then integrated into the graphic user interfaces of ArcMap as an additional toolbox for users to access. Some of these custom tools also include dialog windows to offer user friendly interfaces.

4. TEMPORAL GIS DATABASES CONSTRUCTION AND SPATIOTEMPORAL ANALYSIS TOOLS DEVELOPMENT

Land use and transportation interaction is a dynamic process that involves changes over spatial and temporal dimensions between the two systems. Changes in land use systems can modify the travel demand patterns and induce changes in transportation systems. Transportation system evolution, on the other hand, creates new accessibility levels that encourage changes in land use patterns. Since the 1960s, many theories and models have been used to study land use and transportation interaction (e.g., Alonso, 1964; Anas, 1982; Anas and Duann, 1986; Boyce, 1980, 1990; Hansen, 1959; Kim, 1983; Prastacos, 1986; Kim et al., 1989; Hirschman and Henderson, 1990). Giuliano (1995, p. 3) argues that most models "are generally static, partial equilibrium models" even though they employ iterative methods and equilibrium concepts. Without incorporating time explicitly in a model, iterative methods only provide an equilibrium solution that is essentially for a given point in time and is not necessarily based on a valid theoretical foundation. Holding some variables fixed (e.g., treating land use activities as exogenously given variables in a travel demand model), such models are at best partial equilibrium approaches.

Most land use and transportation interaction studies take a *confirmatory* analysis approach that is based on *a priori* theories or models. These studies have helped us establish theoretical bases and solution methods in our attempts to gain a better understanding of land use and transportation interaction. In the meantime, the literature suggests that little consensus regarding the conclusions can be drawn from the empirical studies (Giuliano, 1995). Land use and transportation system changes take place in a highly dynamic system that involves many forces such as economic development, population growth, and policy decisions. Models based on a confirmatory analysis approach may or may not be able to properly reflect these differences, especially when we need to examine land use and transportation interactions at varying spatial and temporal scales. *Exploratory data analysis* (EDA), on the other hand, takes a speculative and systematic approach to assist us in searching for patterns and processes hidden in the data sets (Openshaw, 1994; Goodchild, 2000). EDA emerged in the 1970s as methods of revealing what would otherwise go unnoticed with the use of standard statistical analysis (Tukey, 1977). EDA does not impose scale-specific assumptions on the analysis procedure; therefore, the patterns and relationships could emerge from the analysis rather than be imposed under a confirmatory data analysis approach. In recent years, use of EDA with GIS has been actively pursued in spatial data mining and spatial knowledge discovery (e.g., Miller and Han, 2001, Ladner et al., 2002).

This chapter presents the construction of a temporal GIS design that offers exploratory data analysis capabilities for examining land use and transportation interactions at user-specified spatial and temporal scales. This temporal GIS is named "Transland" due to its focus on exploring spatiotemporal interactions between land use and transportation systems. This chapter is organized as follows. Section 4.1 discusses the procedures and custom computer programs involved in the creation of various feature and object classes for Transland. It also describes the organization of ArcGIS geodatabases constructed for the Transland. Section 4.2 offers detailed information of the design and implementation of the six spatiotemporal interaction scenarios discussed in Chapter 3. In addition, this section provides information on the implementation of user interfaces. These custom user interfaces provide analysis tools for transportation planners to construct temporal GIS databases and to explore spatiotemporal interactions of land use and transportation systems.

4.1 Creation of Temporal GIS Databases in Transland

The previous chapter presents the design of a temporal GIS and describes the various object classes and custom tools required by this project. As we move from the design stage to the database construction stage, specific procedures and programs must be developed to implement the design framework. Figure 4.1 shows a UML sequential diagram that illustrates the sequence of operations for the implementation of various object classes discussed in Chapter 3. This process first manually creates empty geodatabases and then imports the data sets collected by the project team into feature classes. Once the snapshot GIS feature classes are created in geodatabases, a set of Visual Basic programs is used to add time stamps into the snapshot feature classes (i.e., TimeStampedSnapshot program), to combine snapshot feature classes into space-time composite classes (i.e., GenerateSTComposite program), and to generate time and time2space object classes (i.e., GenerateTime program). In addition, semantic class for transportation improvement projects is created for the project. More detailed information on the creation of various feature and object classes is provided below.

4.1.1. Creation of time-stamped snapshot feature classes

In order to generate a space-time composite class from the snapshot feature classes, we first must append time to the data fields in the snapshot feature classes. This is necessary because the snapshot feature classes of the same theme (e.g., land use feature classes of various years) are likely to include identical data fields (e.g., land use type). When these data fields in different snapshot feature classes are combined into a single space-time composite feature class, it is critical to be able to identify the time associated with each data field. The time-stamped snapshot feature classes are named with a prefix of "t_' and are stored in the same directory as the snapshot features classes. Figure 4.2 shows the flow of a Visual Basic program to create time-stamped snapshot feature classes.

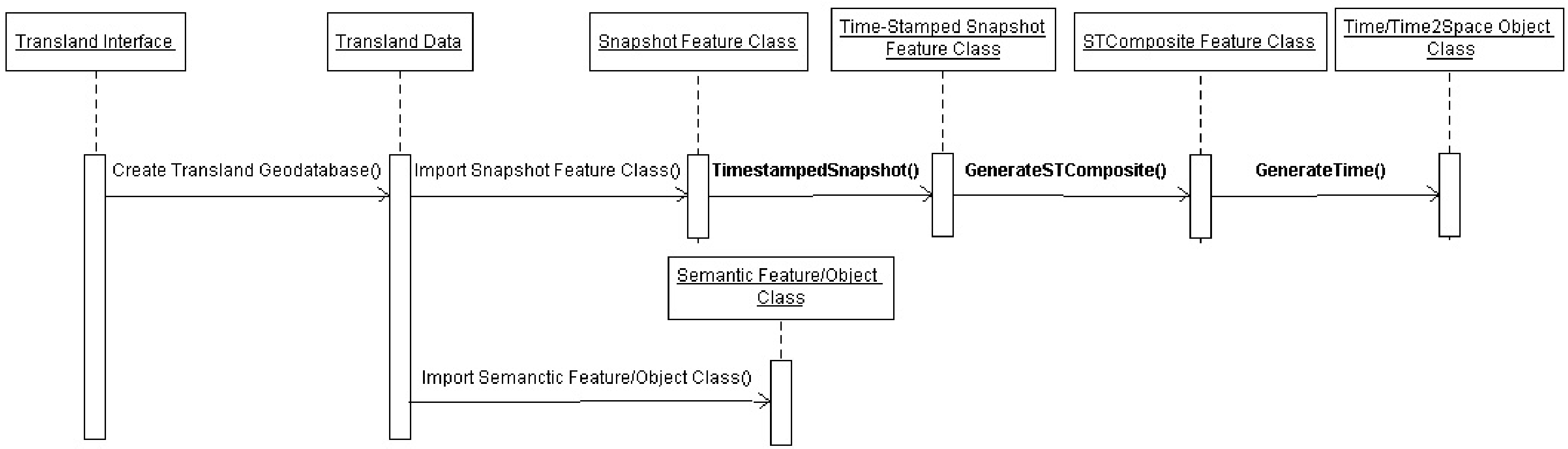


Figure 4.1 Transland classes represented in a UML sequential diagram

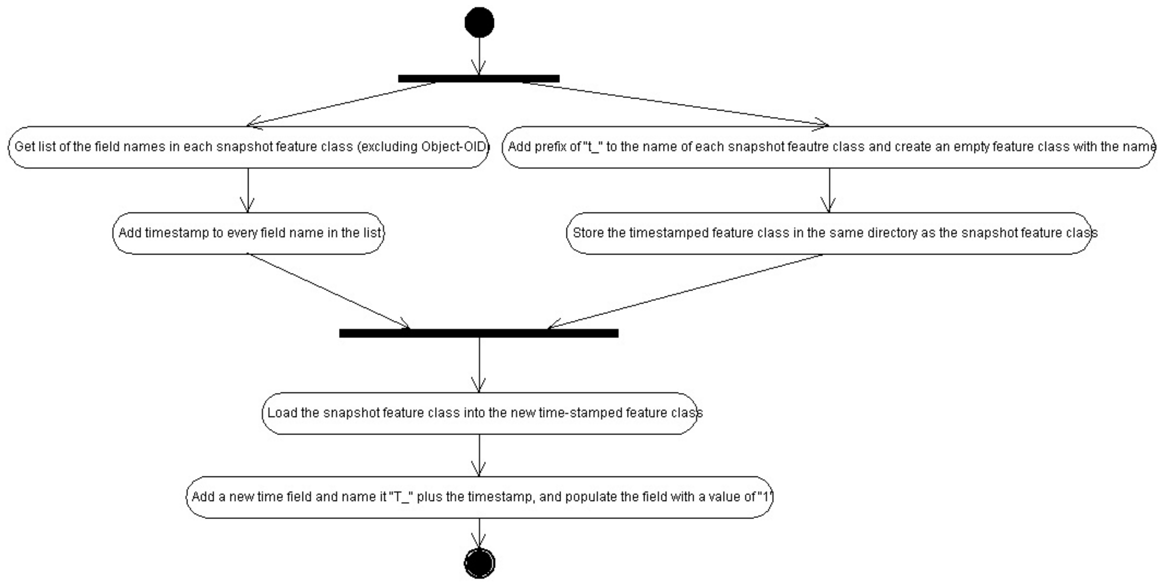


Figure 4.2 Flow of VB Program to Create Time-Stamped Snapshot Feature Classes

Table 4.1 lists the collection of Visual Basic class modules, standard modules, and forms that comprise the VB project for the creation of time-stamped snapshot feature classes. Figure 4.3 shows an example of two modules that are related to the ICommand implementation. ICommand is an interface available in the ArcObjects. The time-stamped snapshot feature class implements the ICommand with its own specific behavior. This custom class is implemented as a button tool that is added into the ArcMap graphic user interface (GUI). In this example, the basSettings module defines global variables and includes common functions. The *DataLoader* function in a class module that calls both the *CreateAccessFile* and the *CalcTime* functions to complete its operation.

4.1.2. Creation of space-time composite feature classes

As indicated in Figures 3.5 and 3.6 of the previous chapter, different geometric features (i.e., points, lines, and polygons) require different procedures of creating the space-time composite (STComposite) feature classes. For example, the three maps shown in Figure 3.5 represent the snapshot GIS layers of traffic count stations in three different years (i.e., 1994, 1996 and 1999). In this example, traffic count station #3 existed in all three years, while traffic count station #9 existed in 1994 only. The STComposite feature class table includes all traffic count stations that existed in 1994, 1996 and 1999. A data value of “1” in the table indicates the existence of a feature at a particular time and a “0” means a feature did not exist at that particular time. As a result, the traffic count station #3 has a value of 1 for T1994, T1996 and T1999, while the traffic count station #9 has a value of 1 for T1994 and 0 for T1996 and T1999.

Table 4.1 Modules and Forms Included in the VB Project for Time-Stamped Feature Class

Class Module Name	Class Module Description
Data_Prepare.cls	Class module that implements ICommand and initializes frmPrepare.frm with its OnClick event
Application_Status.cls	Class module that implements IApplication and IStatusMessage
IApplicationStatus.cls	Class module that defines an abstract class which sets a value for *Application*
IStatusMessage.cls	Class module that defines a subroutine for displaying a message

Module Name	Module Description
Basic_Setting.bas	Basic module that dimensions global variables
Database_Utility.bas	Module that contains utility functions for managing geodatabases

Form Name	Form Description
Prepare_Form.frm	Form for selecting features classes
frmResources.frm	Form defining an icon representing dataPrepare.dll when it is added to ArcMap GUI as a button

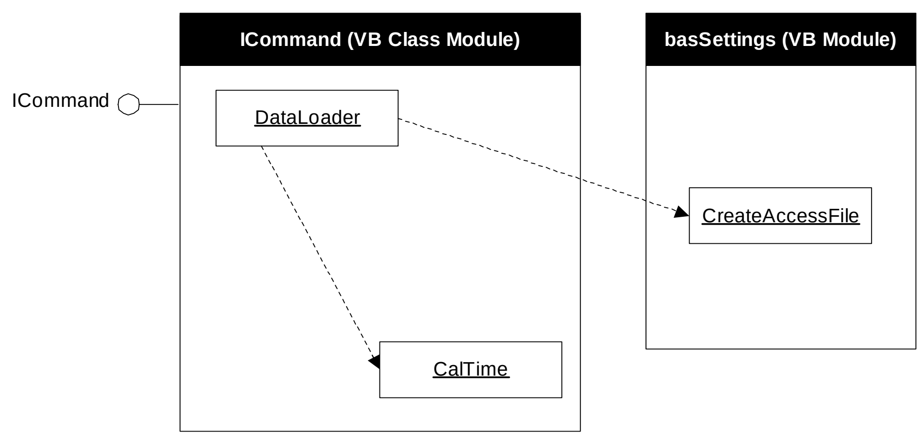


Figure 4.3 ICommand Implementation of Time-Stamped Snapshot Feature Classes

Line and polygon features are more complex than point features to create space-time composite feature classes. When we combine multiple snapshot line/polygon feature classes to generate an STComposite feature class, existing lines/polygons may be split or merged to form new features. For example, the top three maps in Figure 3.6 shows that combining the road map of 1997 with the road map of 1999 results in an STComposite feature class of new line features. Similarly, polygon features could be split or merged over time as illustrated by the bottom three maps in Figure 3.6. This raises a critical issue of maintaining changes of an object identity over time.

Table 4.2 shows an example of STComposite feature class table that includes six objects with their unique object identifiers (OIDs), which are 1, 2, 3, 4, 5 and 6. Since the STComposite feature class is derived by combining the snapshot features classes of 1990, 1994 and 1998, the original OIDs (i.e., OID1990, OID1994, and OID1998) are kept in the final STComposite feature class table. The first five objects in the STComposite feature class table reflect the change history of a land parcel with an OID of 1074 in 1990 that was split into two parcels of 2074 and 2073 in 1998. Parcel 2074 in turn was split into four parcels (2721, 2719, 2812, and 2825) by 1998 and parcel 2073 changed its OID to 2719. The last object (i.e., OID = 6) in the STComposite feature class table indicates that this parcel did not exist in 1990 and 1994 and was created with an OID of 2936 by 1998.

Table 4.2 History of Object Identity (OID) Changes in an STComposite Feature Class

OID	OID1990	LU1990	T1990	OID1994	LU1994	T1994	OID1998	LU1998	T1998
1	1074	Vacant	1	2074	Vacant	1	2721	Vacant	1
2	1074	Vacant	1	2074	Vacant	1	2716	Single-Family	1
3	1074	Vacant	1	2074	Vacant	1	2812	Vacant	1
4	1074	Vacant	1	2074	Vacant	1	2825	Single-Family	1
5	1074	Vacant	1	2073	Single-Family	1	2719	Single-Family	1
6	Null	Null	0	Null	Null	0	2936	Commercial	1

While the above table is feasible to keep track of changes of object identities through time, it becomes inefficient when we need to query data for multiple years. For instance, in order to find out if feature OID 1 existed in all three snapshot years, we must formulate a query statement with three cursors referencing T1990, T1994, and T1998. This query statement can slow down the search process significantly when dealing with large data sets. To overcome this shortcoming, the Transland design also creates TimeObject class and Time2Space class as shown in the design framework of Figure 3.7. The TimeObject class table (see Table 4.3) stores the time instances of all snapshot GIS layers, while the Time2Space class table (See Table 4.4) keeps track of the relationships between all snapshot time instances and all unique spatial features in the STComposite feature class. Both TimeObject class table and Time2Space class table are generated simultaneously with the STComposite feature class via a set of custom Visual Basic programs developed for this project. The TimeObject and Time2Space object classes together facilitate queries of spatial features over time. When they are linked with the STComposite feature class, it is also feasible to query spatial and attribute changes over time.

The above procedures of generating STComposite feature classes and time-related object classes are automated via a set of custom VB programs with ArcObjects. Figure 4.4 outlines the relationships among various VB modules and subroutines. Further discussions of the key VB programs (i.e., *DoSTComposite*, *Layer2STComposite*, *CompositePointFeatures, SplitPolylineFeature,* and *SplitPolygonFeature*) are provided below.

Table 4.3 TimeObject Class Table

TimeOID	Time
1	T1990
2	T1994
3	T1998

Table 4.4 TimeObject Class Table

Time	OID
T1990	1
T1990	2
T1990	3
T1990	4
T1990	5
T1994	1
T1994	2
T1994	3
T1994	4
T1998	5
T1998	1
T1998	2
T1998	3
T1998	4
T1998	5
T1998	6

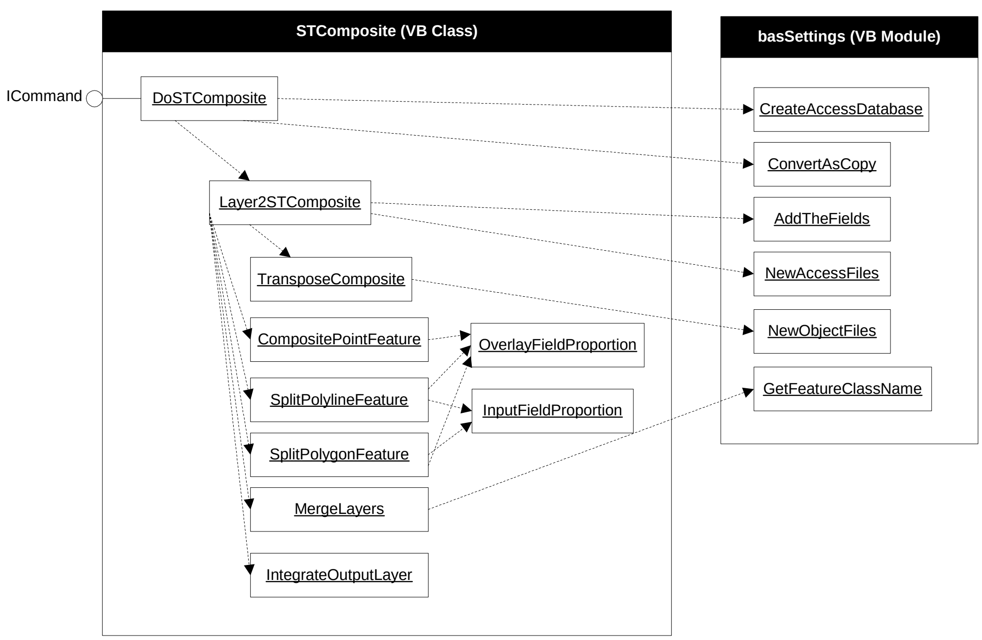


Figure 4.4 Visual Basic Modules and Subroutines for the Creation of STComposite Feature Classes and Time-Related Object Classes

The *DoSTComposite* VB program combines multiple time-stamped snapshot feature classes into an STComposite feature class. Figure 4.5 illustrates the activity diagram of this VB program. It selects the first two time-stamped snapshot feature classes to generate an STComposite feature class. The program then checks if there are additional time-stamped snapshot feature classes. If yes, it generates a new STComposite feature class by combining the current STComposite feature class with the next time-stamped snapshot feature class. It stops when all time-stamped snapshot feature classes have been processed. The program then creates TimeObject and Time2Space object classes from the final STComposite feature class.

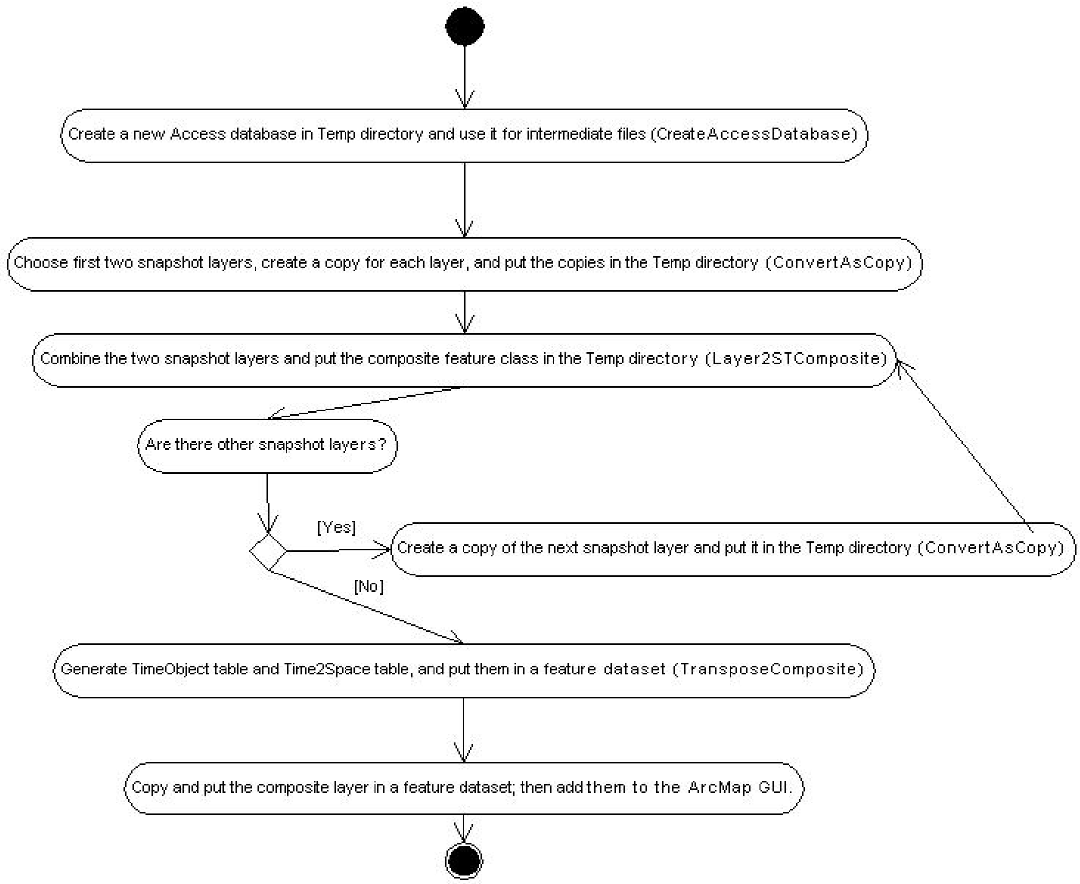


Figure 4.5 Activity Diagram of *DoSTComposite* (VB subroutine Names Are Indicated in Parentheses)

The *Layer2STComposite* subroutine is called from the *DoSTComposite* program to combine two snapshot feature classes into an STComposite feature class. Figure 4.6 is the activity diagram of *Layer2STComposite* subroutine. This subroutine starts with the creation of an empty STComposite feature class with data fields copied from the first two time-stamped snapshot feature classes. It then selects each feature in the first time-stamped snapshot feature class and determines if it intersects with any features in the second time-stamped snapshot feature class. Intersection is defined here as the overlapped parts that have the same dimension as the input feature classes. For example, an intersected feature between two line segments (1-dimensional geometry) must be a line segment. Therefore, disjoint or touched features are not considered as intersected features according to the definition of intersection. Figure 4.7 illustrates examples that are considered as an intersection or a non-intersection for polylines and polygons. Depending on the geometry type (i.e., points, polylines, or polygons) of the input time-stamped snapshot feature classes, the *Layer2STComposite* subroutine calls the respective subroutine of *CompositePointFeatures*, *SplitPolylineFeatures*, or *SplitPolygonFeatures* to generate an STComposite feature class. The *Layer2STComposite* subroutine stops when all time-stamped snapshot feature classes have been processed.

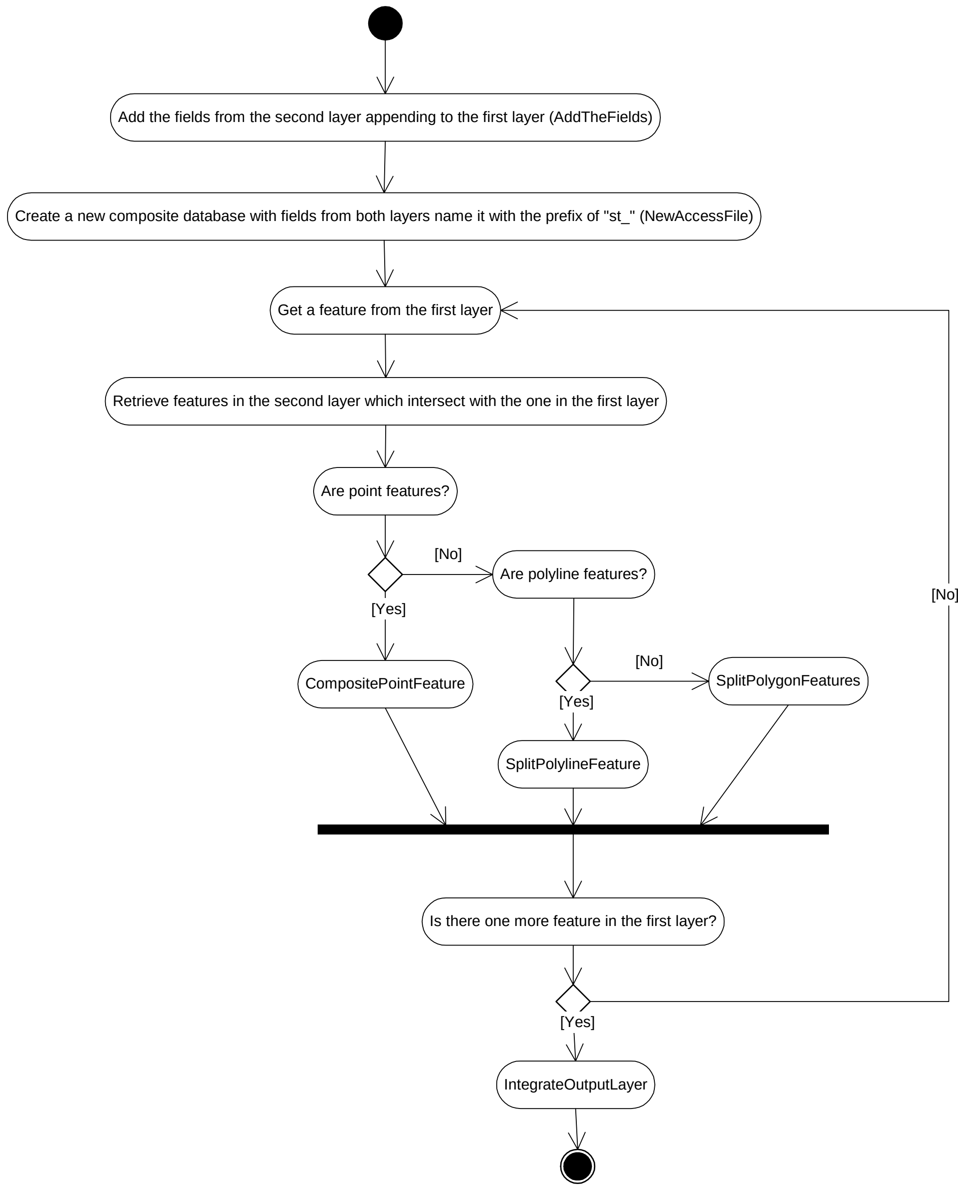


Figure 4.6 Activity Diagram of *Layer2STComposite* (VB Subroutine Names Are Indicated in Parentheses)

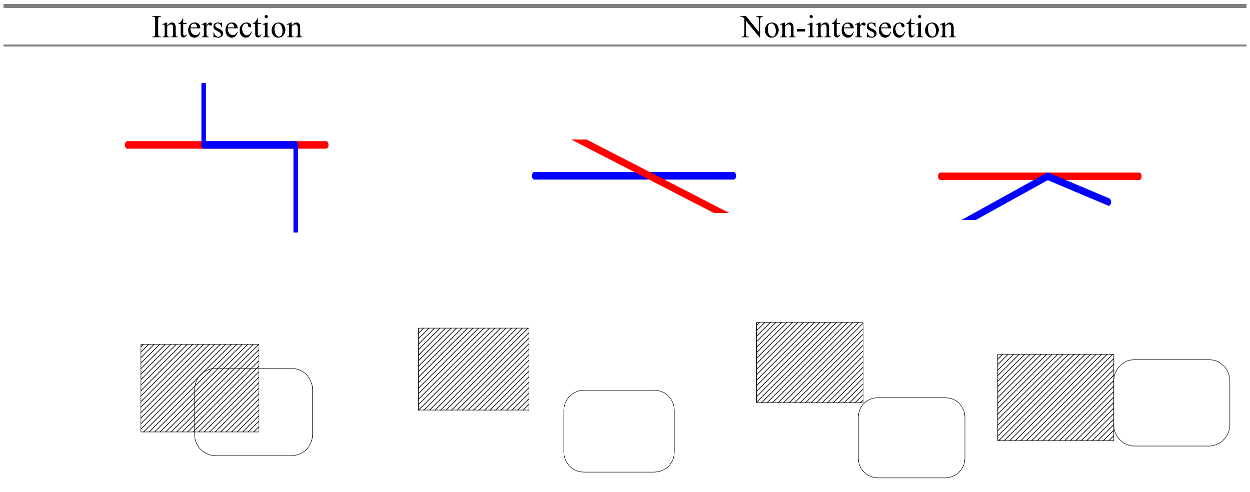


Figure 4.7 Examples of Intersection and Non-Interaction Cases of Polylines and Polygons

When the input snapshot feature classes are points, the process of combining them into an STComposite is straightforward. The *CompositePointFeatures* subroutine first identifies the intersected point features in the input snapshot feature classes and adds them into the STComposite feature class (see Figure 4.8). It then updates the attribute data fields of the STComposite feature class. Finally, it merges the other point features in the snapshot feature classes into the STComposite feature class.

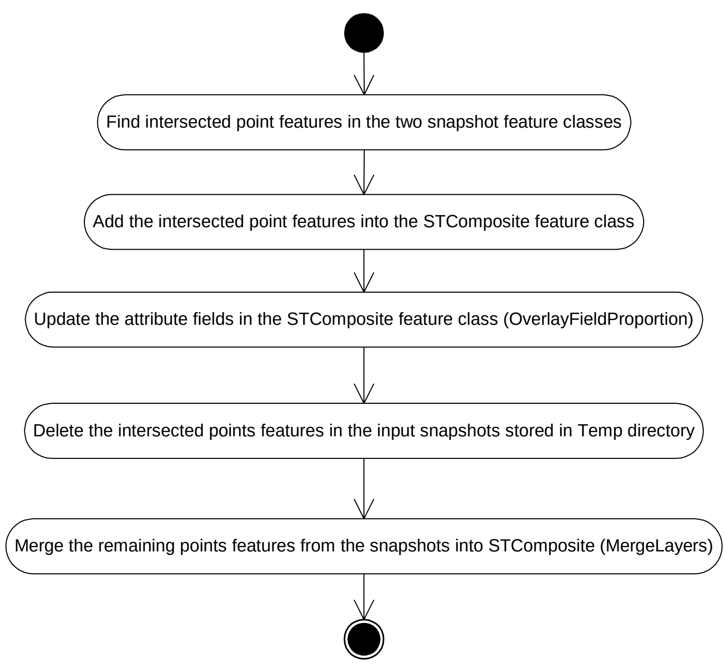


Figure 4.8 Activity Diagram of *CompositePointFeatures* (Subroutine Names Are Indicated in Parentheses)

When the input snapshot feature classes are polylines or polygons, the process first splits the intersected features between the two input snapshot feature classes. It then updates the input and the overlay features in terms of their spatial features and nonspatial attributes. After all input features have been processed, features are merged to create an STComposite feature class. *SplitPolylineFeatures* and *SplitPolygonFeatures* VB subroutines are developed to handle these polyline and polygon snapshot feature classes.

Table 4.5 presents a summary of the VB class modules, standard modules, and forms included in the STComposite VB project.

Table 4.5 Visual Basic Programs for Creating STComposite Feature Classes

Class Module Name	Class Module Description
Stcomposite.cls	Class module that implements ICommand and initializes frmComposite.frm with its OnClick event
ApplicationStatus.cls	Class module that implements IApplication and IStatusMessage
IApplicationStatus.cls	Class module that defines an abstract class which sets a value for *Application*
IStatusMessage.cls	Class module that defines a subroutine for displaying a message
Module Name	Module Description
basSettings.bas	Basic module that dimensions global variables
DataConvert.bas	Module that includes subroutines for converting data between personal geodatabases, coverages, shape files, and SDE
dbUtil.bas	Module that contains utility functions for managing geodatabases
stringParseUtil.bas	Module that contains utility functions for string parsing.
Form Name	Form Description
FrmComposite.frm	Form for selecting features classes and providing STComposite-related information
frmResources.frm	Form defining an icon representing dataPrepare.dll when it is added to ArcMap GUI as a button

4.1.1 Creation of Transland Personal Geodatabases

ArcGIS 8 includes a geodatabase data model to organize and manage geographic information. Its personal geodatabase is based on the Microsoft Access. A geodatabase allows users to organize datasets into different classes (e.g., feature datasets, feature classes, object classes, and relational classes) in a single Access file. In Transland, a geodatabase is created for each theme (e.g., traffic analysis zones). Within each geodatabase, a feature dataset with an identical name of its geodatabase is created to store all snapshot feature classes of the specific theme. Each feature dataset has a defined geographic reference that applies to all feature classes within it.

The data sets collected by the project team include S*treets* (Street Network), *TAZ* (Traffic Analysis Zone), *LOS* (Level of Service), *AADT* (Annual Average Daily Traffic), and *TIP* (Transportation Improvement Project) for transportation features, along with *LU* (Land Use), *TRACT* (Census Tract), *Permits* (Building Permit), and *Property* (Property Parcel) for land use related data. These data sets are available in either ArcView 3 shapefile format or in dBase format. A summary of these data sets is given below.

(1) AADT (Annual Average Daily Traffic) data set includes a space-time composite feature class and annual AADT data from 1970 through 2000;
(2) TAZ (Traffic Analysis Zone) data sets include 1992 and 2000 snapshot feature classes;
(3) Street (Street Network) data sets include 1996, 1997, 1999, and 2000 snapshot feature classes;
(4) LOS (Level of Service) data set includes 1999 snapshot feature class;
(5) Tract (Census Tracts) data sets include 1980, 1990, and 2000 snapshot feature classes;
(6) LU (Land Use) data sets include 1994 and 1998 snapshot feature classes;
(7) Property (Land Property) data sets include 1998 point feature class and 2001 polygon feature class. They cannot be directly used to generate an STComposite feature class due to their different geometric representations;
(8) Tax (Property Tax) data sets include dBase files for 1992, 1996, 1997, 1998, 1999, and 2001. The 2001 Tax data can be associated with 2001 Property data via the "Folio" data field;
(9) TIP (Transportation Improvement Projects) and Permit (Building Permits) data sets are treated as semantic classes in the Transland. They include data for multiple years. TIP and Permit object classes are associated with the Street and the Property feature classes. These associations can be used to identify the spatial locations of a project over time; and
(10) Zdata1 and Zdata2 datasets from FSUTMS.

There are three important notes regarding the geodatabase creation process. First, this project uses the snapshot data sets of TAZ, Street, Tract, and LU to create STComposite feature classes by running the custom VB programs discussed above. The data set for LOS is available for the year of 1999 only; therefore, it is excluded in the process of creating STComposite feature classes. Second, the AADT (Annual Average Daily Traffic) data include a point shapefile that includes traffic count stations with traffic count data from 1970 to 1999. This AADT shape file is converted into geodatabase format and is treated as an STComposite point feature class. Third, TIP and Permit data sets are considered in this project as semantic classes. Semantic objects are maintained in geodatabases as pairs of a feature class and an object class. The feature class identifies the locations of semantic objects, while the object class stores attribute data related to the semantic objects. The TIP object class is associated with the TIP feature class based on the foreign key of Project_ID. The building permit data set is an attribute data file. It is converted into an object class in geodatabase. This object class is then related to its spatial counterpart of the 1999 property parcel feature class through the unique property folio numbers.

After all data sets collected by the project team had been converted into proper geodatabases, we applied the custom Visual Basic programs to generate time-stamped snapshot feature classes, STComposite feature classes, and their related TimeObject and Time2Space object classes. The organization of resulting geodatabases, feature datasets, feature classes, and object classes of the Transland project is listed below.

Transland Data Folder:

- TAZ geodatabase (taz.mdb)
    - TAZ feature dataset

        - TAZ snapshot feature class
        - TAZ time-stamped snapshot feature class
        - TAZ STComposite feature class
    - TAZ TimeObject class
    - TAZ Time2Space object class
- LOS geodatabase (los.mdb)
    - LOS feature dataset
        - LOS snapshot feature class
- AADT geodatabase (aadt.mdb)
    - AADT feature dataset
        - AADT snapshot feature class
        - AADT time-stamped snapshot feature class
        - AADT STComposite feature class
    - AADT TimeObject class
    - AADT Time2Space object class
    - AADT relationship class
- LU geodatabase (lu.mdb)
    - LU feature dataset
        - LU snapshot feature class
        - LU time-stamped snapshot feature class
        - LU STComposite feature class
    - LU TimeObject class
    - LU Time2Space object class
- TRACT geodatabase (tract.mdb)
    - TRACT feature dataset
        - TRACT snapshot feature class
        - TRACT time-stamped snapshot feature class
        - TRACT STComposite feature class
    - TRACT TimeObject class
    - TRACT Time2Space object class
- Semantic geodatabase (semantic.mdb)
    - TIP feature class
    - TIP_Table object class
    - LandDevelopment feature class
    - LandDevelopment_Table object class

Figure 4.9 illustrates the geodatabase organization as shown in the ArcCatalog graphic user interface. Each theme (e.g., CENSUS.mdb for census tracts) has its own geodatabase. Within each geodatabase, a feature dataset with an identical name of its geodatabase is created to store all snapshot feature classes of the specific theme (e.g., pop1980, pop1990, and pop2000), along with the time-stamped snapshot feature classes (e.g., t_pop_T1980, t_pop_T1990, and t_pop_T2000) and the STComposite feature class (e.g., st_census_80_90_1). Each feature dataset has a defined geographic reference that applies to all feature classes within it. In addition, a TimeObject class (e.g., Time_OIDst_census_80_90_00) and a Time2Space object class (e.g., Time_st_census_80_90_00) are created at the same level as the feature dataset in a geodatabase. For the semantic geodatabase (i.e., SEMANTIC.mdb), each semantic class has a spatial feature class (e.g., tip) and its related object class (e.g., tip_table).

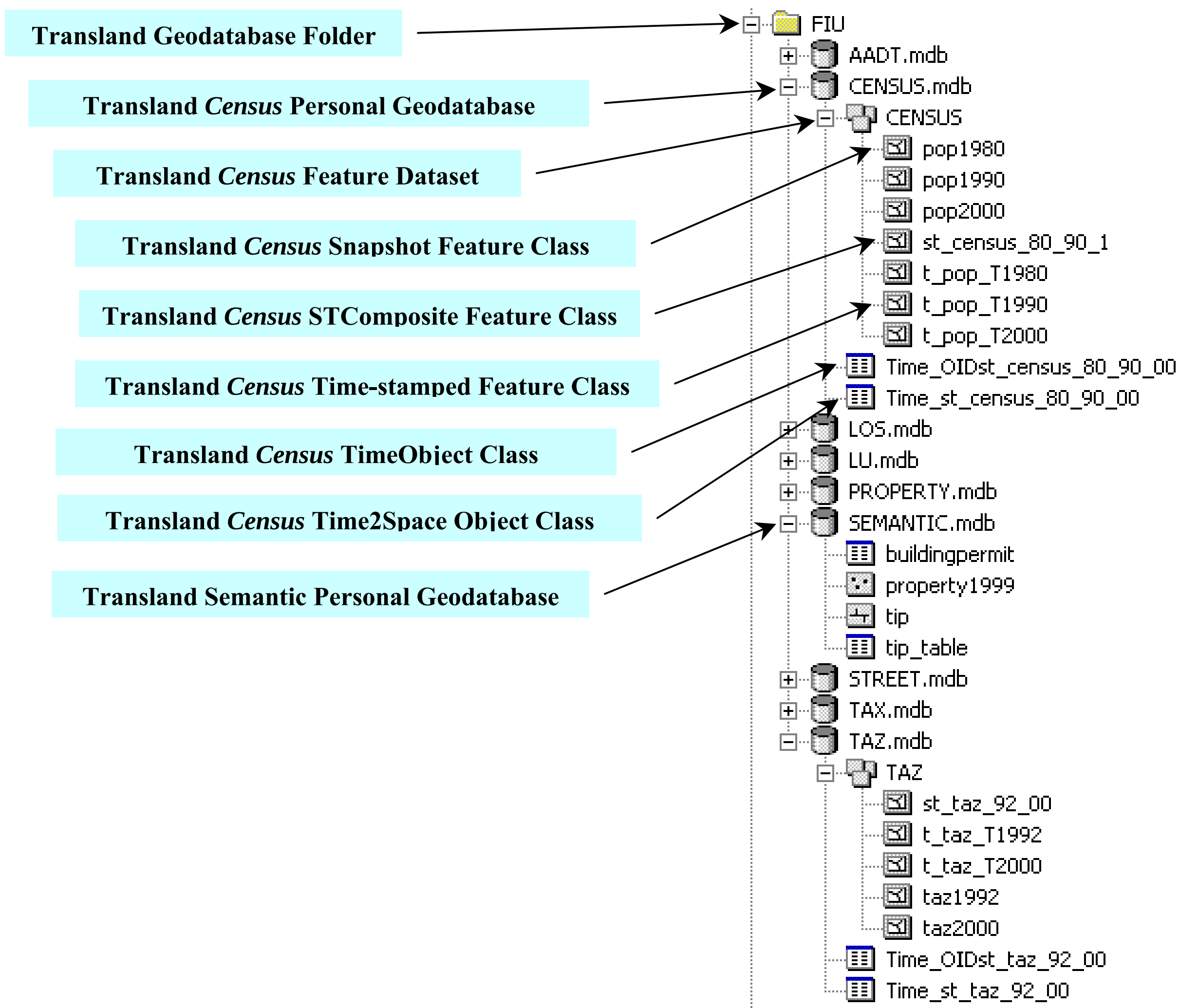


Figure 4.9 Transland Geodatabase Organization as Shown in ArcCatalog

Users could encounter two problems when loading data into ArcGIS geodatabases. First, when a new record is added to feature or object classes of more than 9500 records in ArcMap, an error message will indicate that the file is locked. This is due to the parameter setting in the MaxLocksPerFile registry entry for Microsoft Jet Engine. To fix this problem, run Regedit and find jet 4.0 in the path of *HKEY_LOCAL_MACHINE*, *software*, *Microsoft*, *jet*, *4.0*, and *engines*. Then, modify the MaxLocksPerFile parameter setting to a number larger than the number of records in your table. Second, if the procedure of creating time-stamped snapshot feature classes fails to execute, make sure that you don't have any data fields added after the "Shape_Length" field. This restriction is due to the setup in the custom DataLoader subroutine that does not permit any data fields following the Shape_Length field in a feature class table.

4.2 Spatiotemporal Exploratory Analysis Tools Development

There have been many models suggested in the literature for studying transportation and land use interactions (see literature review in Chapter 1). Most of these models take a *confirmatory* analysis approach that is based on *a priori* theories. These studies have helped establish theoretical foundations and solution methods to gain a better understanding of land use and transportation interaction. In the meantime, empirical studies also report the existence of significant spatial and temporal variations between different geographic areas and at different spatial and temporal scales.

Land use and transportation system changes take place in a highly dynamic system that involves many forces such as economic development, population growth, and policy decisions. Models built on a confirmatory analysis approach are often based on a pre-specified spatial and temporal scale; therefore, they are less flexible of examining the dynamic nature of various forces acting at different spatial and temporal scales on land use and transportation interaction. *Exploratory data analysis* (EDA), on the other hand, takes a speculative and systematic approach to assist us in searching for patterns and processes hidden in the data sets. Since each study area tends to have its own unique transportation and land use interaction patterns over space and time, spatiotemporal exploratory analysis tools could help search for the interaction patterns hidden in large, multiple data sets. Results derived from these exploratory analysis tools could be used to evaluate the existing model structures or to validate the model parameters for a particular study area. They also could suggest new hypotheses for examining the complex relationships between land use and transportation.

Like most other spatial processes, land use and transportation interaction involves time (when), location (where), and attribute (what) that are interrelated with each other. Sinton (1978) proposed a measurement framework to treat location, time, and attribute as “fixed”, “controlled”, and “measured” components when performing measurements (see Table 2.9). However, Sinton’s measurement framework was designed to measure a single phenomenon. When our focus is to analyze the interactions between two phenomena (e.g., transportation and land use), each phenomenon has its own location, time, and attribute components. We must treat the three components as integrated parts of each phenomenon and develop an approach of analyzing the interactions between the phenomena.

In Chapter 3, we proposed a spatiotemporal interaction framework to keep track of the space, time, and attribute components of transportation and land use phenomena (see Table 3.1). This spatiotemporal interaction framework consists of six scenarios that can explore land use and transportation interaction according to spatial, temporal, and attribute data. Each scenario in this framework corresponds to one of the six ways that transportation planners may ask questions about transportation and land use interactions. A list of sample questions for the six scenarios is provided in Table 4.6. For example, scenario 2 allows users to “fix" locations and attributes of transportation features (e.g., select a half-mile zone around the I-595) and “control" time (e.g., 1995 to 2000) in order to “measure” the land use changes in the user-specified area over time. For more detailed discussions of these six scenarios, please refer to Section 3.1.

Table 4.6 Sample Questions of the Six Scenarios for Exploratory Analysis of Land Use and Transportation Interaction

Scenarios	Sample Question
1	When did major land developments take place within a one-mile zone of I-595?
2	Where were the areas that experienced greater than 20% land value increases between 1995 and 2000 within a half-mile zone of I-595?
3	Where were the vacant land parcels within a 1-mile zone of ongoing transportation projects in 1996?
4	Which street segments that are adjacent to commercial land use in the study area had a traffic volume/capacity ratio greater than 1 in 1998?
5	When did traffic volumes on major streets increase by more than 30% in those traffic analysis zones reaching an employment density of 10,000 per square mile?
6	What were the traffic volumes on major streets in each census tract with an annual population growth rate over 5% from 1990 to 2000?

4.2.1 Design of the Six Scenarios under the Spatiotemporal Interaction Framework

The six scenarios presented in Table 3.1 are based on time, transportation, and land use components. This project takes full advantage of the current ArcGIS capabilities of handling spatial and attribute data and uses the time component to extend the static GIS data model into a temporal GIS data model. In other words, the time component is linked to the space-attribute objects so that users can query and analyze GIS databases according to any combinations of spatial, temporal and attribute components. Under this design, standard GIS functions such as SQL queries and spatial searches can be easily implemented in the Transland user interfaces. Figure 4.10 shows a dialog window, which is implemented as part of the Transland user interfaces, for users to query the spatial and/or attribute data. In this dialog window, users can select specific features in either transportation or land use feature classes through a combination of SQL attribute queries (see Figure 4.10a) and interactive spatial selection tools (see Figure 4.10b).

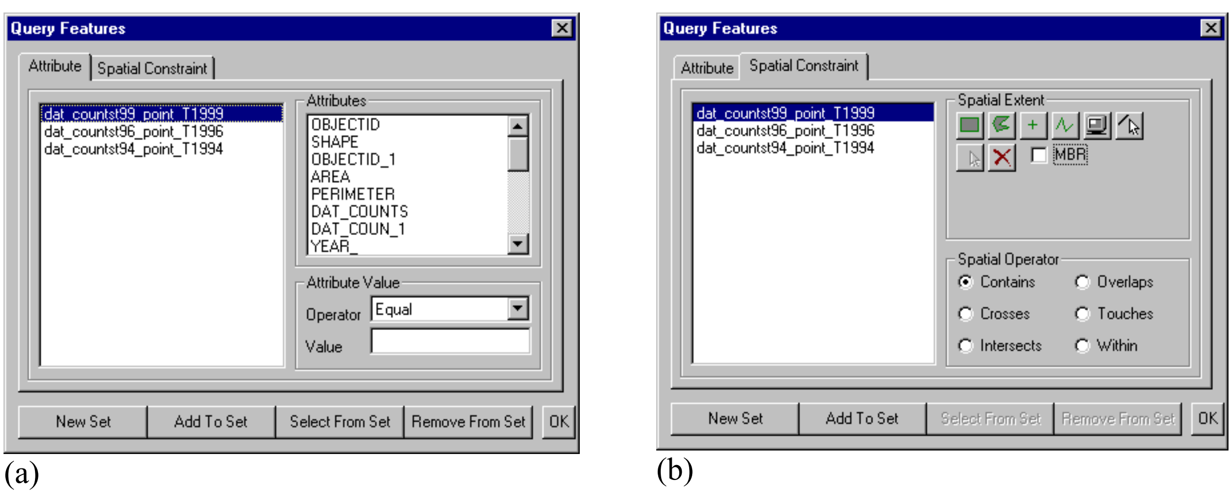
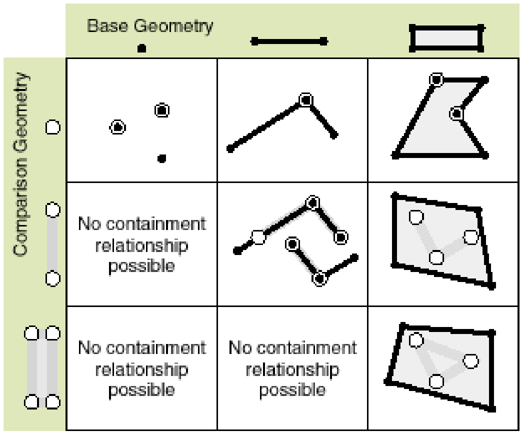


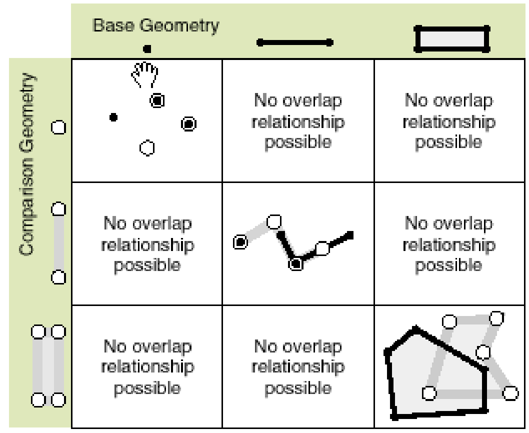
Figure 4.10 Transland Space-Attribute Query Dialog Window: (a) Attribute Query, and (b) Spatial Constraint

Query functions available through this dialog window are based on the ArcMap functions of *Select By Attributes* and *Select by Location*. Users can click the buttons of *New Set, Add To Set, Select From Set,* and *Remove From Set* in the dialog window to perform multiple selections from a list of feature classes in order to select the specific features of interest. The *Spatial Constraint* (i.e. *Selection By Location*) allows users to interactively define a spatial extent of search area and then choose a specific spatial operator to carry out the spatial search. Figure 4.11 illustrates the six spatial operators implemented in this dialog window. For instance, the "Contain" spatial operator will find all features in a selected GIS layer (i.e., the comparison geometry indicated in Figure 4.11) that are completely contained within the user-specified spatial extent (i.e., the base geometry indicated in Figure 4.11). The user-specified spatial extent is defined interactively by drawing a polygon, a line, or a point on the display screen.

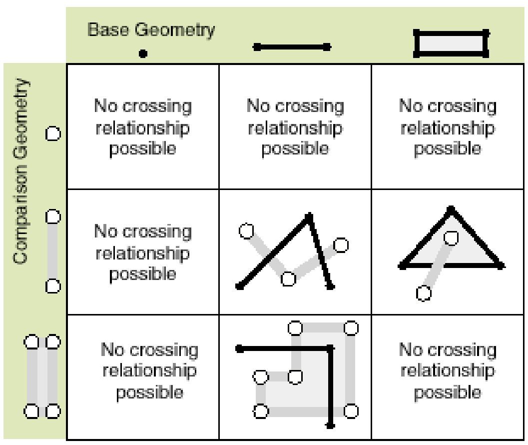
Although ArcGIS provides a powerful set of tools for analyzing spatial and attribute data, it does not offer functions for temporal query and analysis. Since this project extends static snapshot GIS databases into temporal GIS databases (see Section 4.1 above), it is feasible to add temporal query and analysis tools into the user interfaces of Transland. Figure 4.12 shows the temporal query and analysis dialog windows implemented in Transland. The dialog window on the left-hand side in Figure 4.12 handles individual time instants associated with the snapshot feature classes, while the dialog window on the right-hand side is designed to process user-specified time intervals. Because the time component is embedded in all transportation and land use GIS databases, these temporal query and analysis functions are integrated with the spatial and query functions available in ArcGIS.



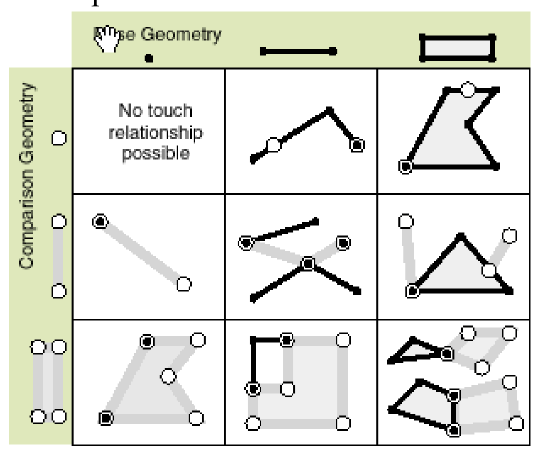
Contain



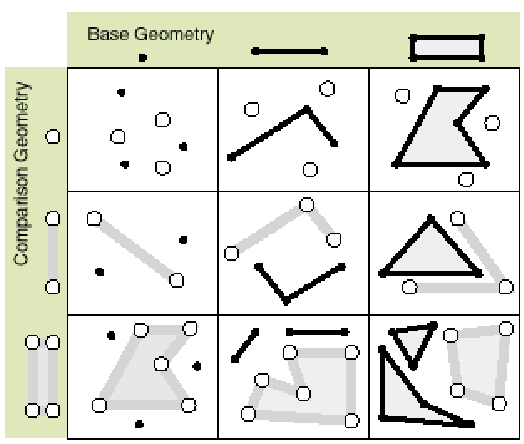
Overlap



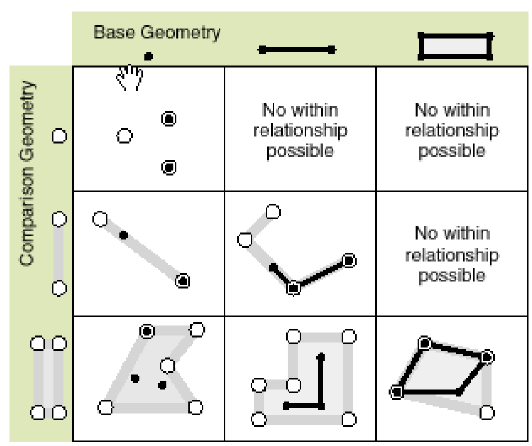
Cross



Touch

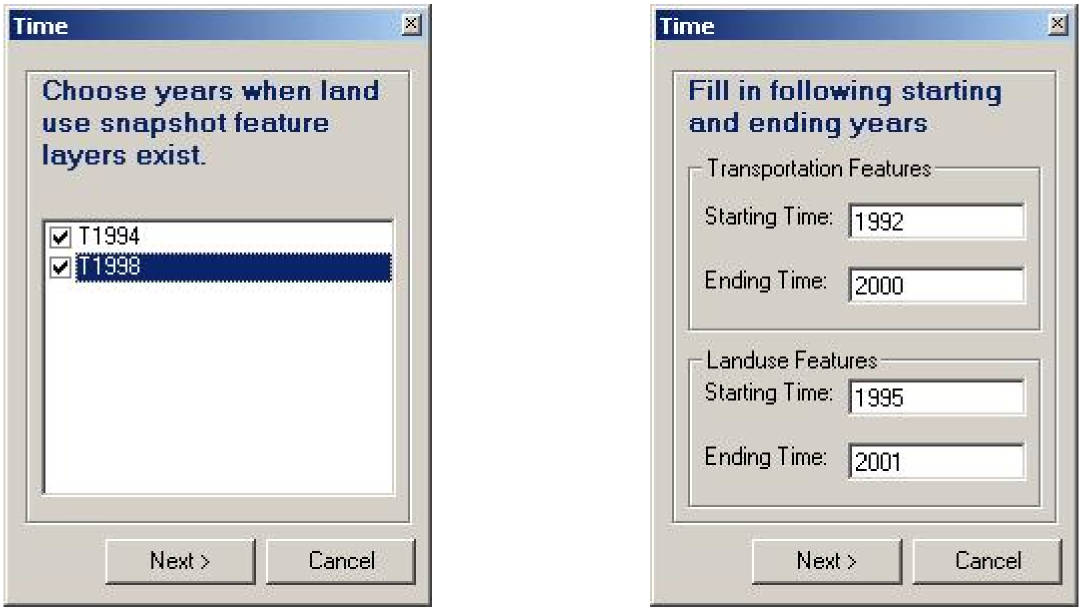


Disjoint



Within

Figure 4.11 Illustration of the Six Spatial Operators Included in Transland Spatial Constraints User Interface (Source: Zeiler, 2001)



(a) (b)

Figure 4.12 Transland Temporal Query and Analysis Dialog Windows: (a) Time Instants, and (b) Time Intervals

4.2.2 Steps of Processing the Six Scenarios in the Spatiotemporal Interaction Framework

Each scenario listed in the spatiotemporal interaction framework of Table 3.1 involves several steps to carry out the exploratory analysis procedure associated with the scenario. This section describes the step-by-step procedures to carry out these analysis tasks. In addition, a UML collaboration diagram, which illustrates the flow and the messages passed from one step to the next step, are presented for each scenario. Again, a set of custom VB programs is developed in this project to implement these spatiotemporal exploratory analysis tools. Like other tools developed in this project, the VB programs are compiled as dynamic linking library (DLL) files and added into a custom toolbar in ArcMap as the "Transland Scenario Wizard" icon shown in Figure 4.13 below. This toolbar provides the entry point for users to access all spatiotemporal exploratory tools available for analyzing land use and transportation interactions.

Figure 4.13 Custom Toolbar of Transland in ArcMap

With land use and transportation temporal GIS layers loaded and displayed in ArcMap, a click of the "Transland Scenario Wizard" icon in the custom toolbar launches the main Transland Scenario dialog window shown in Figure 4.14. This main dialog window provides a description of each scenario and allows users to choose a specific scenario to perform spatiotemporal exploratory analysis.

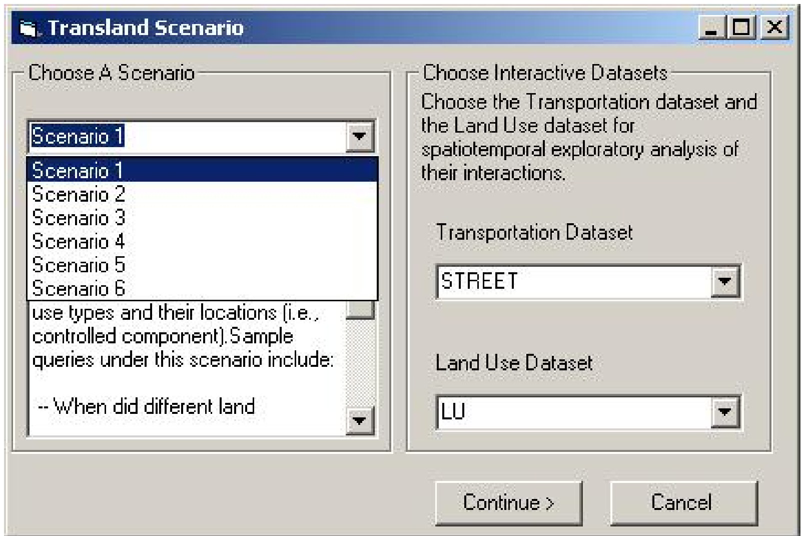


Figure 4.14 Main Dialog Window of Transland Scenario Wizard

Scenario 1 is designed to identify the time frame (i.e., measured component) that transportation systems (i.e., fixed component) have on different land use types and their locations (i.e., controlled component). Table 4.7 lists a sample question and the steps involved in carrying out spatiotemporal exploratory analysis for scenario 1. Figure 4.15 presents a UML collaboration diagram showing the flow and user actions associated with the steps. In this example, we would like to find out "*When did major land developments take place within a one-mile zone of I-595?*" Users first choose scenario 1 from the list presented in the main dialog window (see Figure 4.16a). Next, users are presented with choices of either selecting specific TIP semantic objects or selecting specific transportation features as the fixed component (see Figure 4.16b). This choice is reflected as the two branches in the collaboration diagram (see Figure 4.15).

Table 4.7 Analysis Steps of Transland Scenario 1

Scenario 1 example: When did major land developments take place within a one-mile zone of I-595?	
Steps	Transland Classes Related to Each Step
Choose scenario 1	List of all six scenarios
Choose TIPs or transportation features	A list of TIP semantic class and a list of all transportation STComposite feature classes included in ArcMap
Specify spatial extent and/or attribute query	Selected TIP semantic classes or selected features from transportation STComposite feature classes
Specify land use data fields to be included in the analysis	Land use STComposite and snapshot feature classes within the constraints specified in step 3
Generate final display and summary report	Land use snapshot, STComposite, TimeObject and Time2Space classes

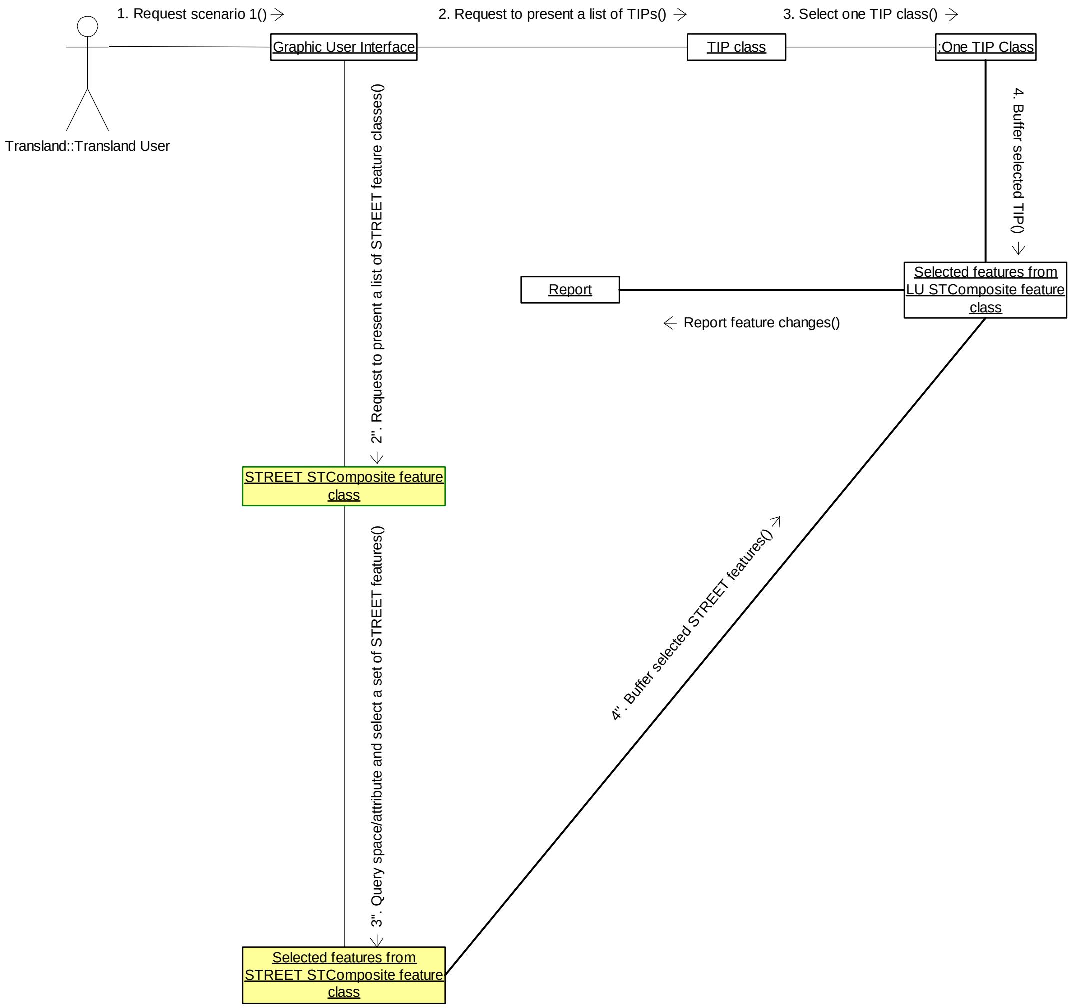
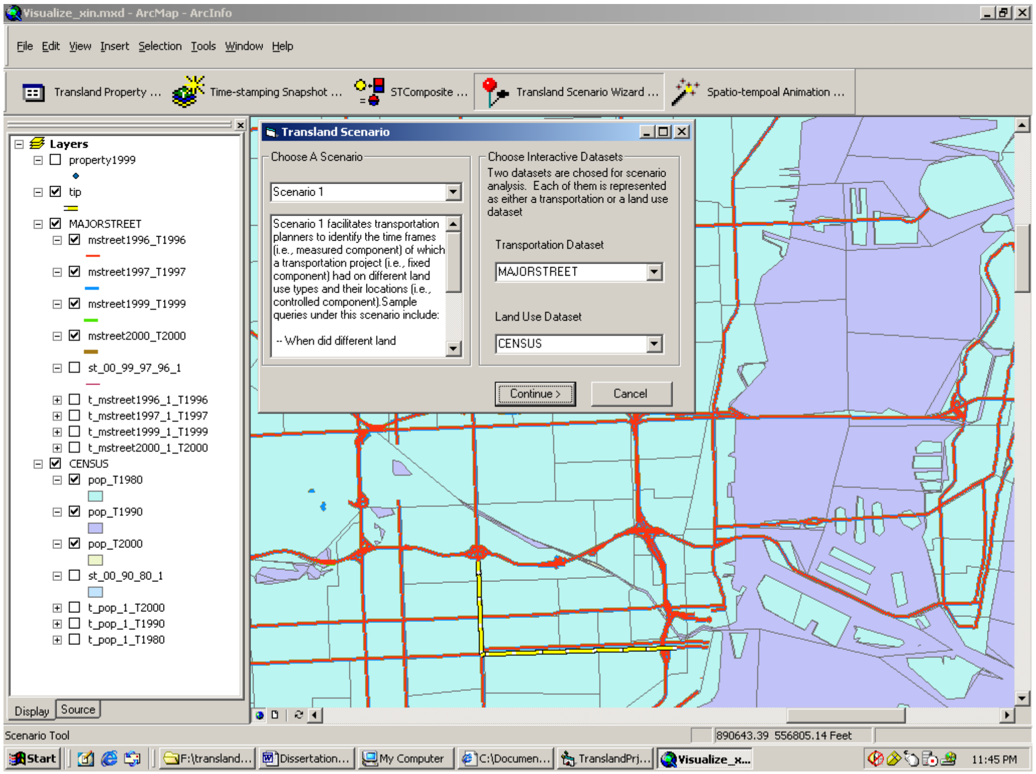


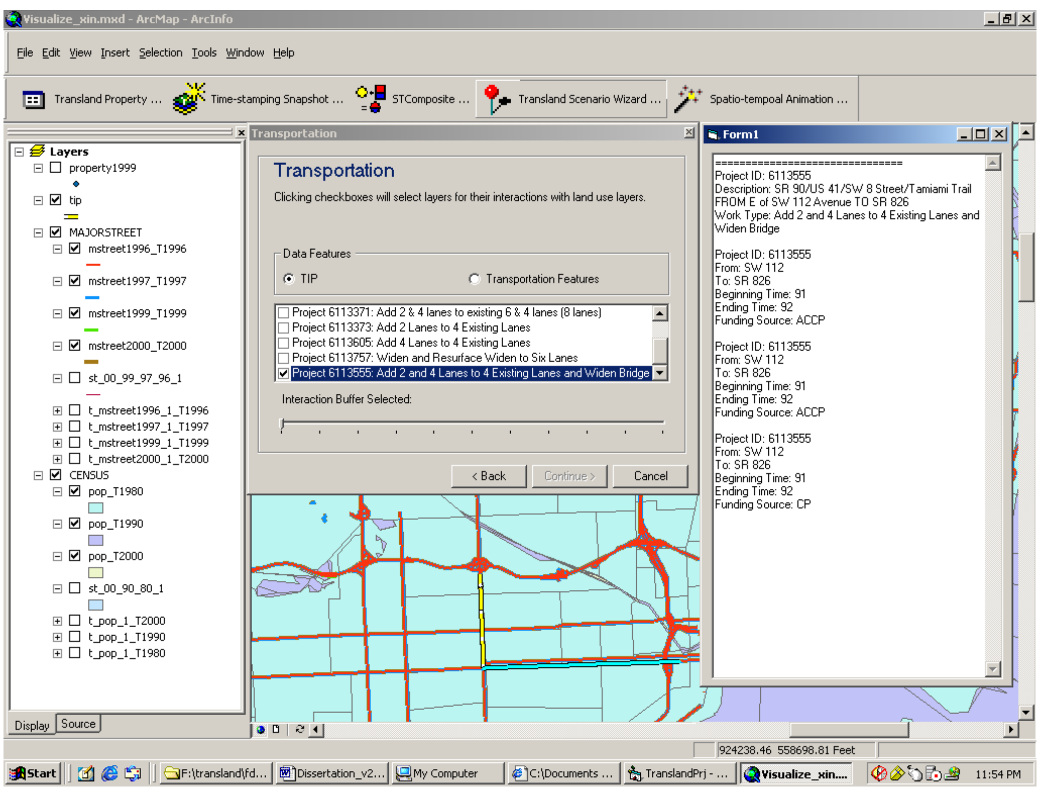
Figure 4.15 UML Collaboration Diagram of Scenario 1

If users choose TIP semantic objects, the next dialog window allows them to specify a buffer distance around the selected TIP features for searching land use features (see Figure 4.16c). If users choose transportation features, they are provided with a different dialog window to interactively define attribute queries and the spatial extent along with a selected spatial operator (see Figure 4.16d). Once specific transportation features are selected, users again can specify a buffer distance around the selected feature for searching land use features.

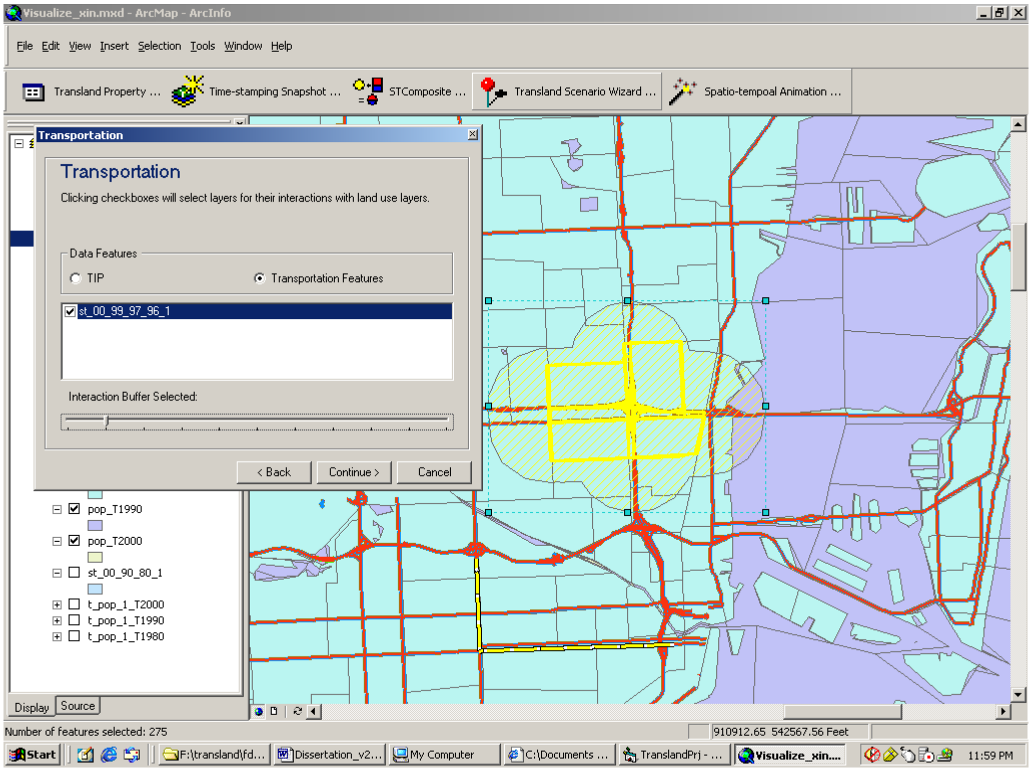
The next step offers a dialog window for users to select the specific attribute data fields associated with various snapshot times that will be included in the analysis and in the final summary report (see Figure 4.16e). In other words, this step and the previous step together control the locations and the attributes of land use features that will be examined with respect to the fixed transportation component selected in an earlier step.



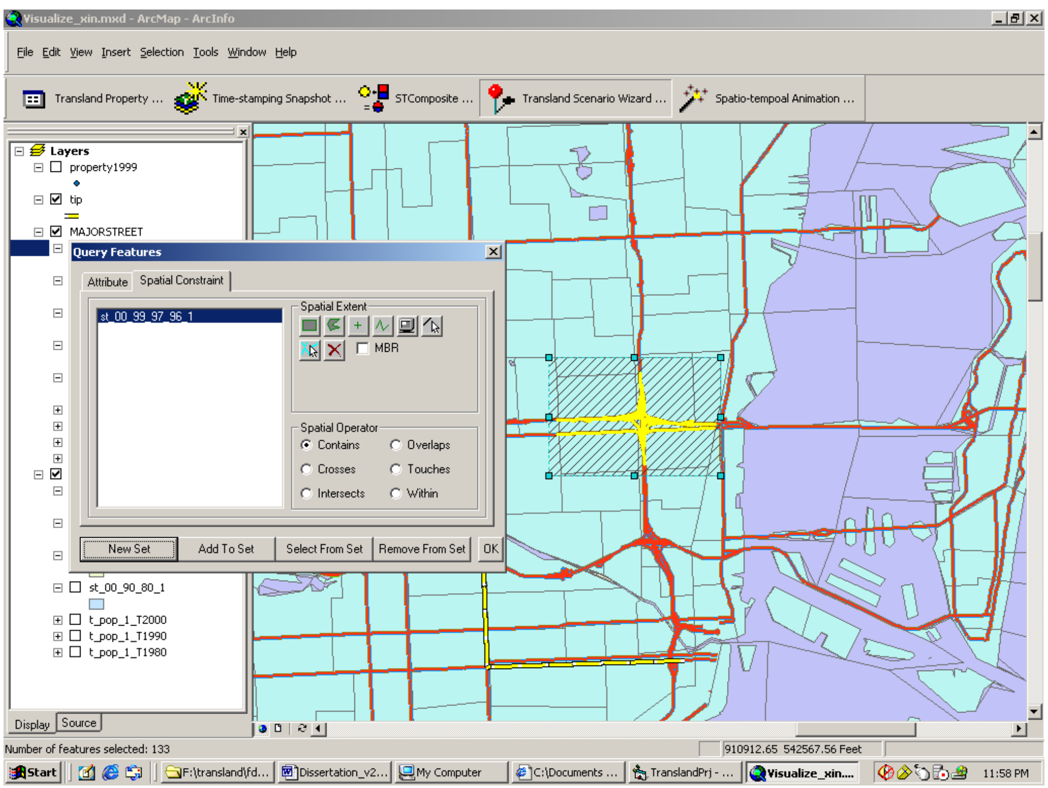
(a) Choose a scenario.



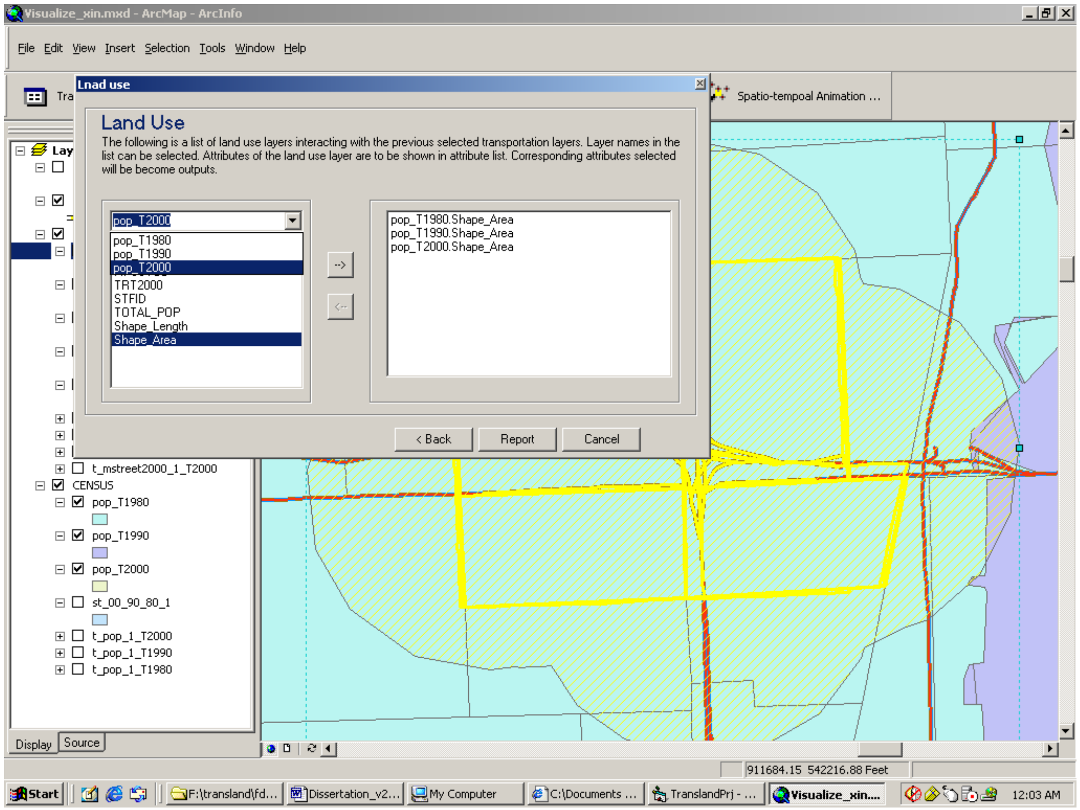
(b) Select TIPs or transportation features.



(c) Specify buffer distance around selected TIPs.



(d) Specify attribute query and spatial constraint of selected transportation features.



(e) Specify land use attribute data fields associated with various snapshot times to be included in the analysis and final report.

Figure 4.16 Key Dialog Windows for Exploratory Analysis of Scenario 1

With transportation features selected as the fixed component and land use features specified as the controlled component, the custom VB programs developed for the Transland Scenario Wizard will display the land use features located within the user-specified search distance as animated maps in an Internet Explorer window. In addition, a summary report that includes user-selected land use attribute data fields for different snapshot times is listed with the animated map display (see Figure 4.17). The animated map display area is automatically refreshed to display the selected land use features (with a yellow highlight color as shown in Figure 4.17) based on the time sequence. For example, if the user selects land use layers of 1994, 1998 and 2000 for the analysis, the animated map display area will continuously show the three snapshot maps in the sequence of 1994, 1998 and 2002. Users therefore can easily visualize the spatial change patterns of land use from 1994, through 1998, to 2000. In addition, the summary report shows when land use characteristics took place. For example, Figure 4.17 shows that the land use features with OID of 418 and 421 remained as parks (i.e., land use code = 517) from 1994 to 1998, while the features with OID of 443 and 455 changed their land use types between 1994 and 1998. Finally, the output file automatically reports some basic statistics (i.e., minimum, maximum, mean, and standard deviation) of those numerical data fields included in an analysis. Table 4.8 shows an example of the basic summary statistics associated with the selected land use features.

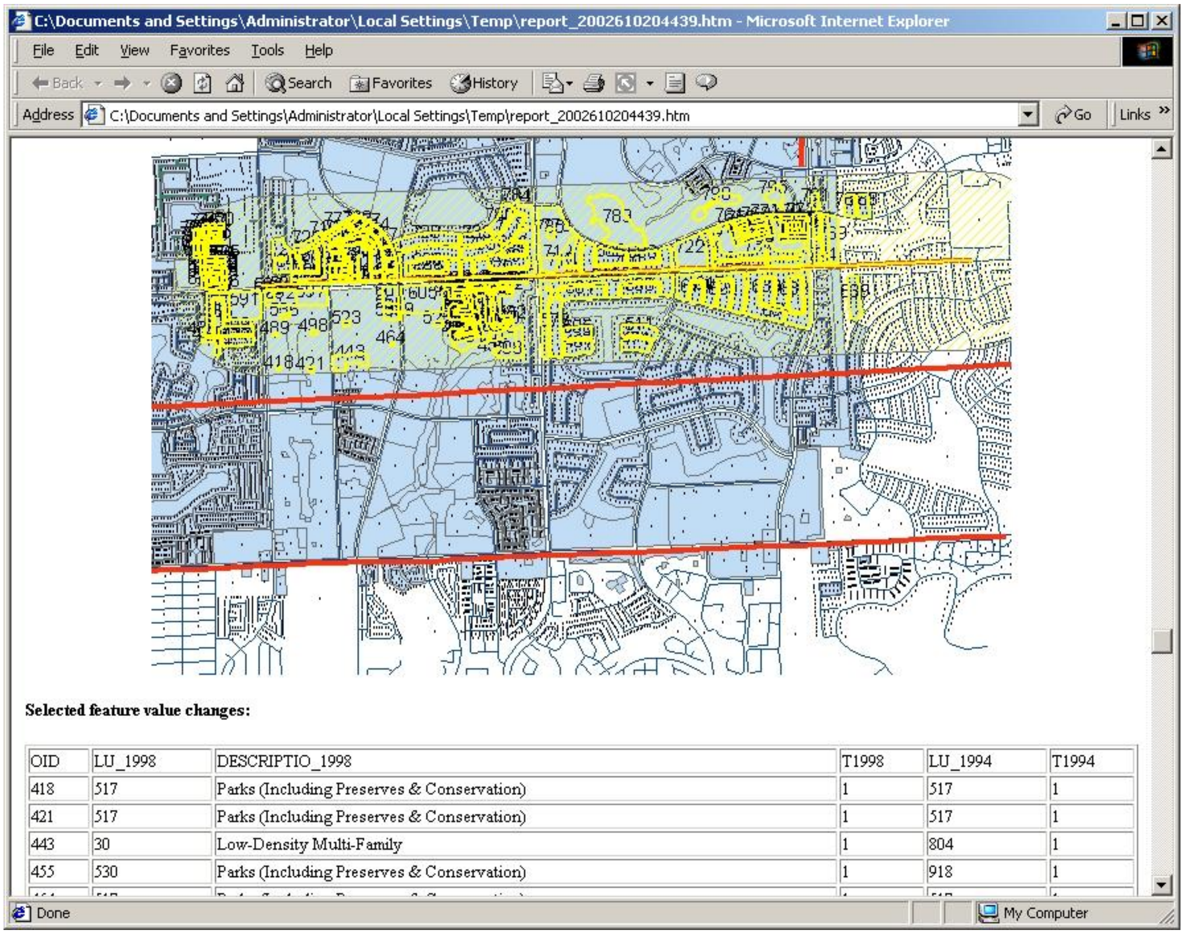


Figure 4.17 An Example of Animated Maps and Summary Report Displayed in an Internet Explorer Window

Table 4.8 An Example of Summary Statistics Included in the Transland Report File

Field Name	Minimum	Maximum	Mean	Standard Deviation
Shape_Area (selected features)	13409416.442228	22174761.0782468	17923143.2021348	4635208.85224721
Shape_Area (all features)	-2371.65704665466	26428935781.0803	228674520.370887	1891363333.58952

All of the output animated maps and tables are saved in a Hypertext Markup Language (HTML) file that is generated by custom VB programs developed for this project. The HTML file automatically displays in an Internet Explorer window at the end of each scenario analysis. Users can close the window after viewing it and open it at a later time to review the analysis results or to compare it with the results derived from other scenario analyses. This shows the power of exploratory analysis, which allows users to interactively examine land use and transportation interactions with different spatial, temporal, and attribute parameters in each run of the scenario analysis. The analysis results then can be compared to see if different change patterns are revealed at varying spatial and temporal scales.

The above discussion provides detailed information of the steps involved in the analysis for scenario 1. Scenarios 2 through 6 implement similar analysis procedures as scenario 1. The main difference among them is the components that are used to fix, control, and measure interactions between land use and transportation. Table 4.9 and Figure 4.18 illustrate the analysis steps and collaboration diagram for scenario 2, respectively. This scenario uses time as the control component; therefore, step 4 will display a dialog window for users to specify time instants or time intervals (see Figure 4.12) to be included in the analysis.

Analysis steps and collaboration diagram for scenario 3 are given in Table 4.10 and Figure 4.19, respectively. Under this scenario, users fix the time component and control transportation features in order to measure changes of land use patterns. As a result, the temporal query dialog window will appear first for users to specify the time component that will be used to search for relevant transportation and land use features. For example, if the user selects the year of 1996 only, the scenario 3 will limit its analysis to transportation and land use interaction in 1996. Users also can choose different time intervals for land use and transportation features (see Figure 4.12). If a user chooses 1992 to 1995 for land use features and 1995 to 2000 for transportation features, this enables an analysis of time lag effect of land use changes on transportation systems. In other words, we can examine the transportation system changes (e.g., traffic counts) during the years (i.e., 1995-2000) after land use changes had taken place (i.e., 1992-1995).

The spatiotemporal interaction framework presented in Table 3.1 indicates that scenarios 1, 2, and 3 are mirror cases of scenarios 5, 6, and 4, respectively. For instance, the only difference between scenario 1 and scenario 5 is the reversed roles of transportation and land use as the fixed component and the controlled component. Tables 4.11 – 4.13 and Figures 4.20 – 4.22 show the analysis steps and the collaboration diagrams of scenarios 4, 5, and 6, respectively. They represent different ways for users to explore land use and transportation interactions over time. According to the specific scenario chosen by a user, the custom VB programs automatically determine the sequence of dialog windows that will be presented to the user. The user then interactively specifies spatial and temporal scales, along with selected attribute data, to explore land use and transportation interactions. The analysis procedure can be repeated using different spatial, temporal, and attribute specifications under the same scenario or between different scenarios. At the end of each scenario analysis, a report of both animated maps and data tables is presented in HTML format to help users visualize the change patterns and review the summary data. These HTML files are saved in the default temp directory and are available for users to display at any time until users decide to delete them from the hard disk.

Table 4.9 Analysis Steps of Transland Scenario 2

Scenario 2 example: Where were the areas that experienced greater than 20% land value increases between 1995 and 2000 within a half-mile zone of I-595?	
Steps	Transland Classes Related to Each Step
1. Choose scenario 2	A list of all scenarios
2. Choose TIPs or transportation features	A list of TIP semantic class and a list of all transportation STComposite feature classes included in ArcMap
3. Specify spatial extent and/or attribute query	Selected TIP semantic classes or selected features from transportation STComposite feature classes
4. Specify the time instants or time intervals	TimeObject and Time2Space object classes
5. Generate final display and summary report	Selected features from land use snapshot and STComposite feature classes

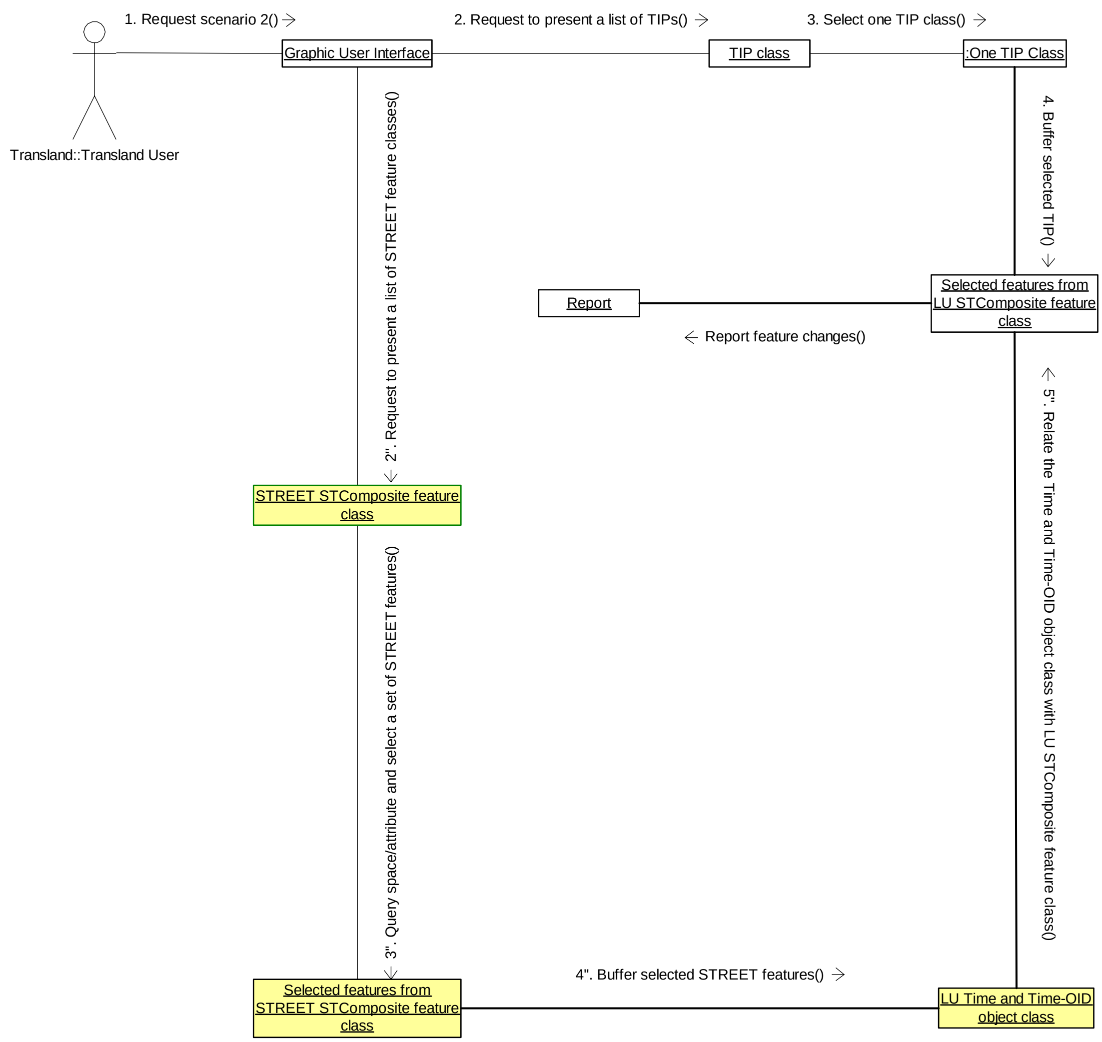


Figure 4.18 UML Collaboration Diagram of Scenario 2

Table 4.10 Analysis Steps of Transland Scenario 3

Scenario 3 example: Where were the vacant land parcels within a 1-mile zone of ongoing transportation projects in 1996?	
Steps	Transland Classes Related to Each Step
Choose scenario 3	A list of all scenarios
1. Specify time instants or time intervals	TimeObject and Time2Space object classes
2. Specify TIPs or transportation features	A list of TIP semantic class and a list of all transportation STComposite feature classes included in ArcMap
3. Specify spatial extent and/or attribute query	Selected TIP semantic classes or selected features from transportation STComposite feature classes within the temporal constraint specified in step 2
4. Specify land use data fields to be included in the analysis	Selected land use STComposite and snapshot feature classes within the constraints specified in step 4
5. Generate final display and summary report	Selected features from land use snapshot and STComposite feature classes

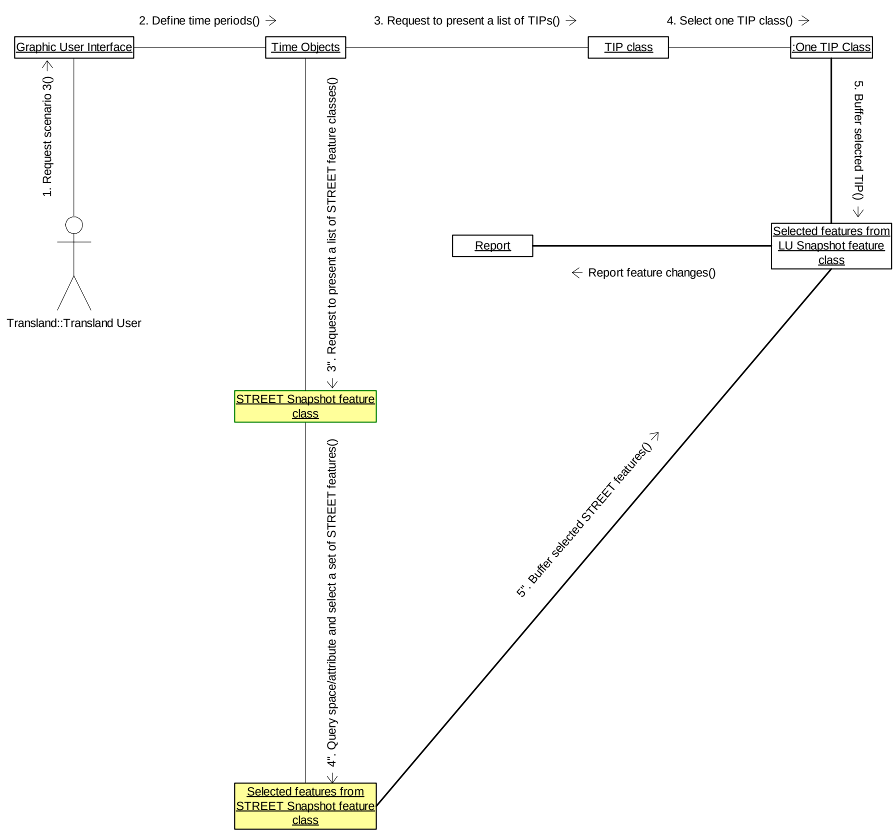


Figure 4.19 UML Collaboration Diagram of Scenario 3

Table 4.11 Analysis Steps of Transland Scenario 4

Scenario 4 example: Which street segments that are adjacent to commercial land use in the study area had a traffic volume/capacity ratio greater than 1 in 1998?	
Steps	Transland Classes Related to Each Step
1. Choose scenario 4	A list of all scenarios
2. Specify time instants or time intervals	TimeObject and Time2Space object classes
3. Specify land development semantic objects or land use features	A list of land development semantic class and a list of all land use STComposite feature classes included in ArcMap
4. Specify spatial extent and/or 5. Attribute query	Selected land development semantic classes or selected features from land use STComposite feature classes within the temporal constraint specified in step 2
6. Specify transportation data fields to be included in the analysis	Selected transportation STComposite and snapshot feature classes within the constraints specified in step 4
7. Generate final display and summary report	Selected features from transportation snapshot and STComposite feature classes

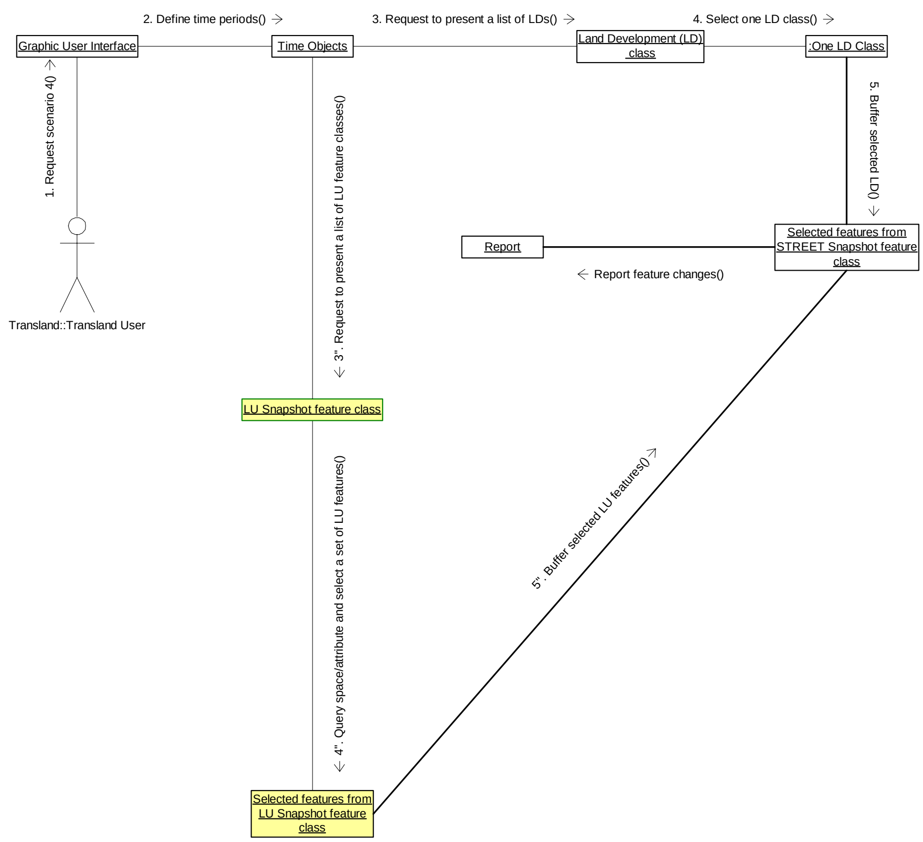


Figure 4.20 UML Collaboration Diagram of Scenario 4

Table 4.12 Analysis Steps of Transland Scenario 5

Scenario 5 example: When did traffic volumes on major streets increase by more than 30% in those traffic analysis zones reaching an employment density of 10,000 per square mile?	
Steps	Transland Classes Related to Each Step
1. Choose scenario 5	A list of all scenarios
2. Choose land development projects or land use features	A list of land development semantic class and a list of all land use STComposite feature classes included in ArcMap
3. Specify spatial extent and/or attribute query	Selected land development semantic classes or selected features from land use STComposite feature classes
4. Specify land use data fields to be included in the analysis	Transportation STComposite and snapshot feature classes within the constraints specified in step 3
5. Generate final display and summary report	Transportation snapshot, STComposite, TimeObject and Time2Space classes

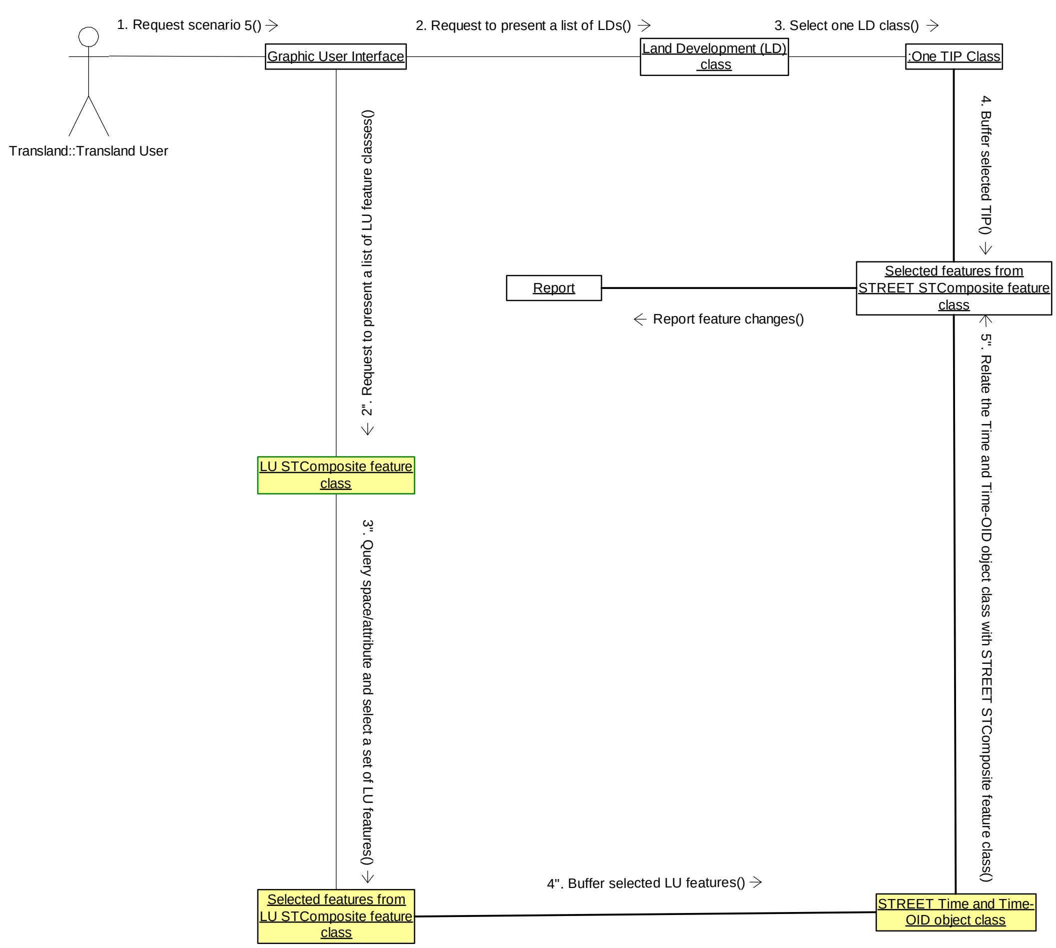


Figure 4.21 UML Collaboration Diagram of Scenario 5

Table 4.13 Analysis Steps of Transland Scenario 6

Scenario 6 example: What were the traffic volumes on major streets in each census tract with an annual population growth rate over 5% from 1990 to 2000?	
Steps	Transland Classes Related to Each Step
1. Choose scenario 6	A list of all scenarios
2. Choose land development projects or land use features	A list of land development semantic class and a list of all land use STComposite feature classes included in ArcMap
3. Specify spatial extent and/or attribute query	Selected land development semantic classes or selected features from land use STComposite feature classes
4. Specify the time instants or time intervals	TimeObject and Time2Space object classes
5. Generate final display and summary report	Selected features from transportation snapshot and STComposite feature classes

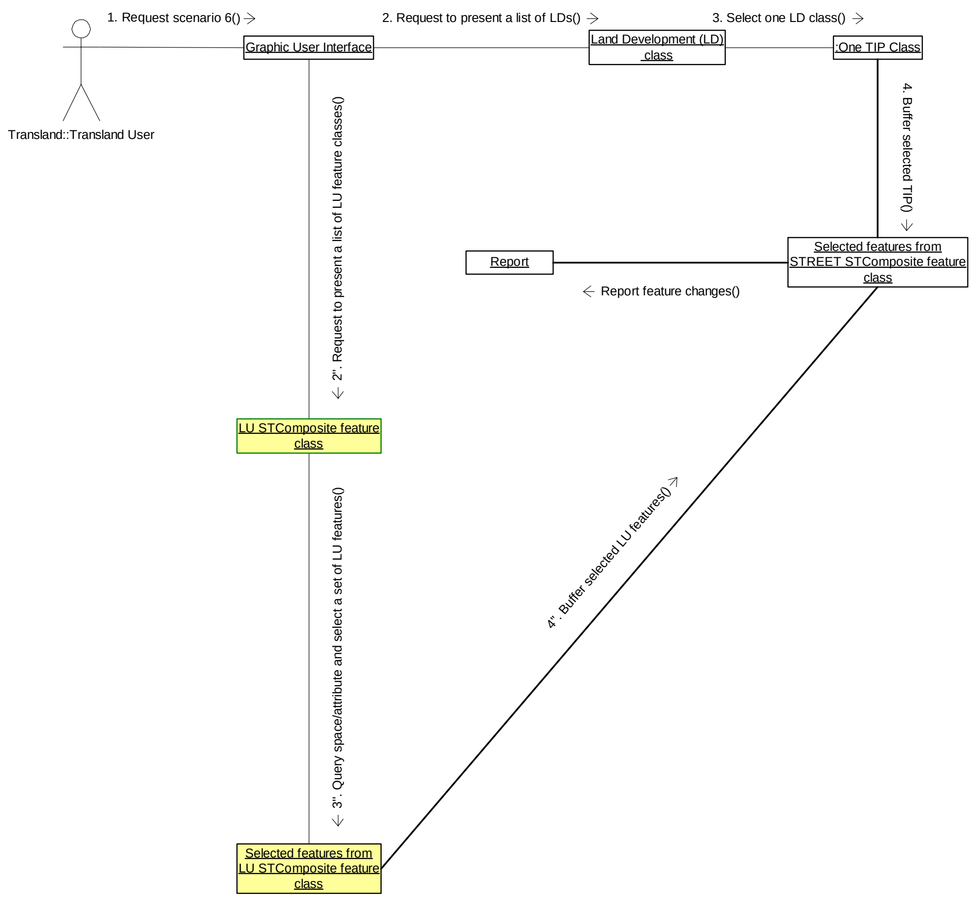


Figure 4.22 UML Collaboration Diagram of Scenario 6

4.2.3 Summary of Visual Basic Programs for Transland Scenario Wizard Implementation

This project develops a set of custom Visual Basic programs to implement the Transland scenario wizard. They include various VB modules and forms to manage, query, analyze, and display spatiotemporal interactions between the various temporal GIS databases created for the project. Table 4.14 presents a summary of the VB modules and forms developed for the Transland scenario wizard.

Table 4.14 VB Modules and Forms for Transland Scenario Wizard Implementation

Class Module Name	Class Module Description
DataPrepare.cls	Class module that includes subroutines for frmProperty.frm
ApplicationStatus.cls	Basic module that dimensions all global variables
IApplicationStatus.cls	Module that dimensions global variables for scenarios
IStatusMessage.cls	Module that includes subroutines for frmQuery.frm

Module Name	Module Description
basSettings.bas	Basic module that dimensions all global variables
dataScenario.bas	Module that dimensions global variables for scenarios
dataProperty.bas	Module that includes subroutines for frmProperty.frm
dataQuery.bas	Module that includes subroutines for frmQuery.frm
dataRelate.bas	Module that contains functions to establish relationships among feature classes and object classes.
dbUtil.bas	Module that contains utility functions for managing geodatabases
modMenu.bas	Module that assists for building right-click context menus
stringParseUtil.bas	Module that contains utility functions for string parsing.

Form Name	Form Description
frmScenario.frm	Simple form for scenario selections
frmTime.frm	Form that confines scenario timeframes
frmQuery.frm	Form that defines attribute and spatial constraints for queries
frmBefore.frm	Form that constructs semantics
frmAfter.frm	Form that collects attributes for final reports
frmProperty.frm	Form that dimensions reporting options
frmResources.frm	Form defining an icon representing dataPrepare.dll when it is added to ArcMap GUI as a button

Figure 4.23 is an activity diagram of the Transland scenario wizard implementation. It reflects the common VB forms and modules shared among the six scenarios. For example, scenarios 1, 2, 5, and 6 all use the same VB form to fix a component. The only difference between scenarios 1 and 2 and scenarios 5 and 6 is the former pair displays the transportation databases and the latter pair displays the land use databases for users to choose the fixed features. Scenarios 2 and 6 then access the VB form of temporal query to control the time component. On the other hand, scenarios 3 and 4 first access the VB form of temporal query for users to fix the time component. This figure shows that many VB forms and modules are shared among the six scenarios. However, they are accessed in difference sequences, and consequently will display and process different temporal GIS databases under each scenarios.

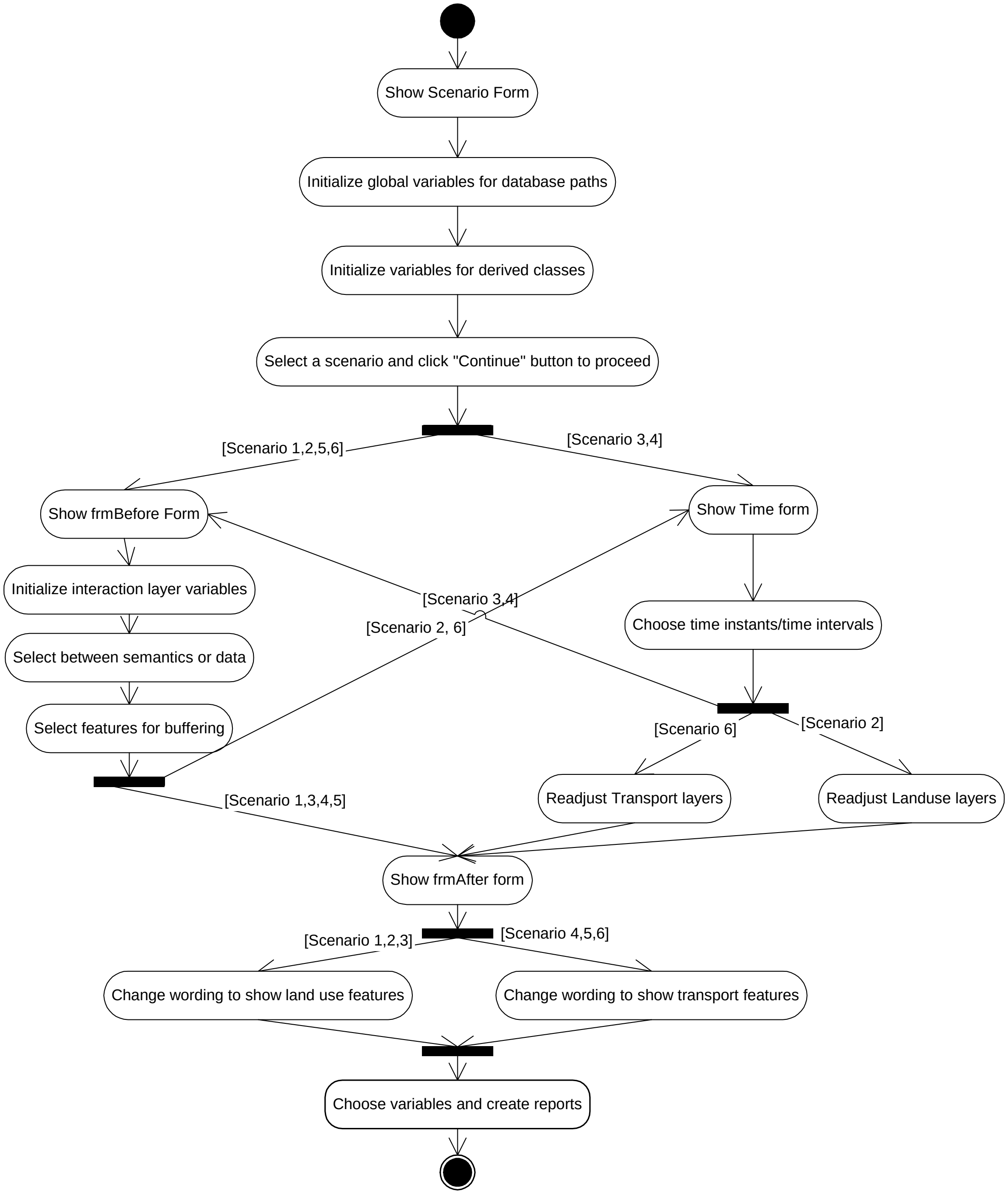


Figure 4.23 Activity Diagram of the Transland Scenario Wizard Implementation

Figure 4.24 illustrates how the above VB modules and forms reference each other as indicated by dotted arrow lines. Several notes about this figure are provided below.

1) frmScenario.frm, dataScenario.bas, stringParseUtil.bas, basSetting.bas, and dbUtil.bas are used to set up scenarios.

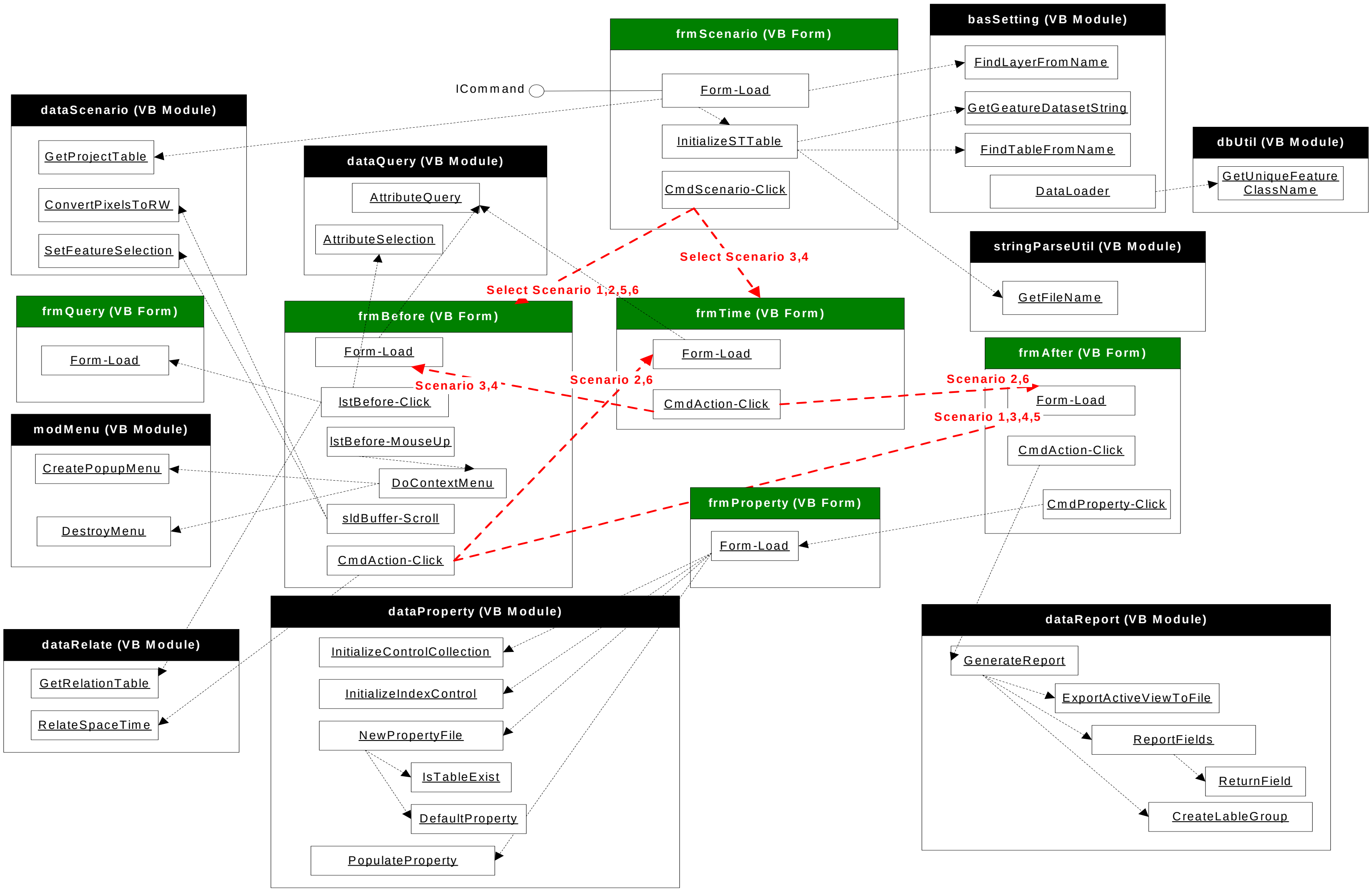


Figure 4.24 References among VB Forms and Modules in Transland Scenario Wizard Implementation

2) frmBefore.frm, frmQuery.frm, dataRelate.bas, modMenu.bas, and dataQuery.bas are used to select semantic objects or feature classes via the space-attribute query window.
3) frmAfter.frm, frmReport.bas, frmProperty.frm, and dataProperty.bas are used to determine the content and format of final report.
4) frmScenario.frm and frmAfter act as the start and the end forms.
5) frmTime.frm and frmBefore.frm are accessed in different orders depending on the type of scenario. Scenario 1, 2, 5, and 6 use frmBefore.frm first and then frmTime.frm. They differ from scenario 3 and 4, which access frmTime.frm first. Scenario 3 and 4 are snapshot scenarios with fixed time component. They only explore transportation and land use interactions at one specific time instant (e.g. 1996) that is fixed at the very beginning of scenario analysis. On the other hand, scenarios 1, 2, 5, and 6 allow users to evaluate land use and transportation interaction over different time periods.

If users only want to visualize temporal changes of various GIS databases without running through the Transland scenario wizard, this project also develops an animated visualization tool, written in Java script, to display land use and transportation layers by time. Users simply click the Spatiotemporal Animation button icon in the custom toolbar (see Figure 4.13). The system then displays a dialog window for users to specify the time interval, the map layers, and the display method for the animated visualization (see Figure 4.25). An example of this animated visualization is shown in Figure 4.26. Analysts can use this animated visualization tool to adjust the display speed, to pause the animation, and to replay the animated displays.

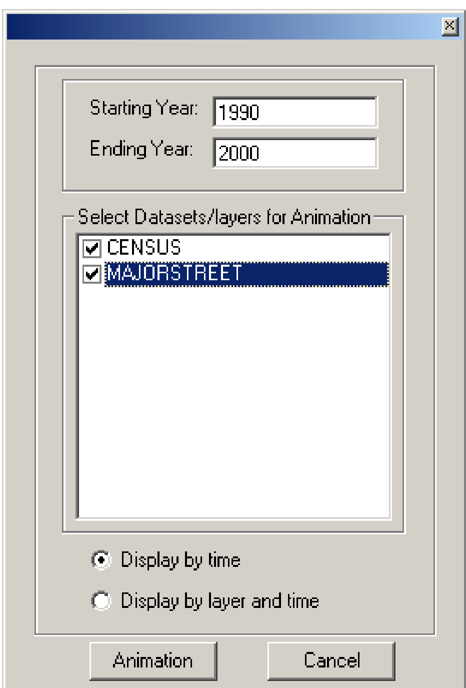


Figure 4.25 Dialog Window of Spatiotemporal Animation Tool

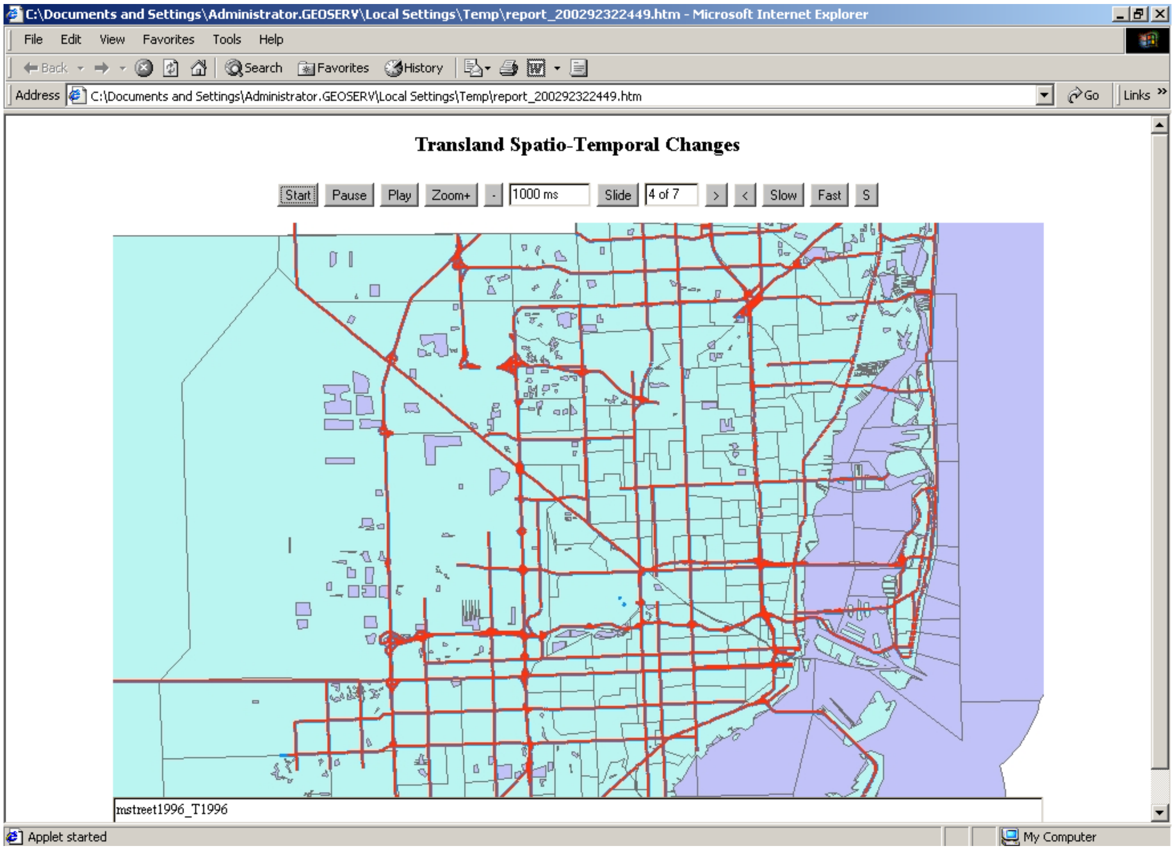


Figure 4.26 An Example of Spatiotemporal Animated Visualization in Transland

5. TIME SERIES ANALYSIS

The methodology utilized in this research is based on time series analysis. A time series is defined as an ordered sequence of values of a variable observed at equispaced time intervals. For instance, the following two time series may be observations of two events, X_t and Y_t, at $t = 1, 2, \ldots, n$:

$$x_1, x_2, \ldots, x_n$$
$$y_1, y_2, \ldots, y_n$$

Time series analysis is designed to describe the dynamic consequences of time series by developing models and to forecast the future of the system based on historical trends. Basic assumption required to model time series in many time series techniques is that the time series is stationary. If mean value of a time series is same over time, that time series is said to be stationary. Also a stationary process can be defined as a series without trend, constant variance over time, a constant autocorrelation structure over time, and no seasonality. If the time series is non-stationary, it can be transformed to a stationary process applying one of following techniques: differencing, curve fitting, typically with a straight line, and transformation. A stationary time series can be expressed in terms of the mean term and past and present error or innovation vectors. This form of the process is called the moving average (MA) representation. The error should be a white noise, which is a random vector with zero mean by assumption. A lag can be defined as a fixed time displacement. If there are observations, $y_1, y_2, \ldots, y_n$, over time, the lag between y_2 and y_7 is 5 (= 7 – 2).

Two types of time-series analysis approaches were utilized: univariate models and multivariate models. The main difference between univariate and multivariate models is that in a univariate model there is one endogenous variable whereas a multivariate model has multiple endogenous variables. An endogenous variable is one that is internal to the modeled system, the value of which is the result of solving the system. In contrast, an exogenous variable is one that is external to the system being modeled, and the value of an exogenous variable must be specified externally and input into the system.

For univariate models, the key aspect is the estimation of a discrete transfer function. There are two basic multivariate models: vector autoregressive model and vector autoregression with exogenous variables. Univariate models and discrete transfer functions will be described in Section 5.1, and multivariate models will be discussed in Section 5.2.

5.1 Univariate Methodology

In this section, the discrete transfer function, which is a univariate function, is discussed. With discrete transfer function, we can deal with pairs of observations (X_t, Y_t) at equispaced intervals of time of an input X, an exogenous variable, and an output Y, an endogenous variable, from some dynamic system. A description of the discrete transfer function is followed by parameter estimation based on the Least Square method. Finally diagnostic tests are presented to check model validity and adequacy.

5.1.1 The Discrete Transfer Function

In many cases, both X_t and Y_t are often observed at discrete times, therefore X_t and Y_t may be discrete series rather than a continuous process. The relationship between X_t and Y_t may be represented in a *linear filter* form as shown below (Box and Jenkins 1976):

$$\begin{aligned} Y_t &= v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \Lambda \\ &= (v_0 + v_1 B + v_2 B^2 + \Lambda\,) X_t \\ &= v(B) X_t \end{aligned} \qquad \text{[Eqn 5.1]}$$

where B is the backward shift operator (lag operator), which is defined as $B^k x_t = x_{t-k}$, and the weights v_0, v_1, v_2, … are called the *impulse response coefficients* of the system. The operator $v(B)$, called the *transfer function* of the filter, may be represented as

$$v(B) = \sum_{j=0}^{\infty} v_j B^j \qquad \text{[Eqn 5.2]}$$

If there is an initial period of pure delay, b, before the response to a given input change begins to take effect, Eqn 5.1 could be rewritten as

$$Y_t = v(B) X_{t-b} \qquad \text{[Eqn 5.3]}$$

By substituting in Eqn 5.1 the values $Y_t = g$, $1 = X_t = X_{t-1} = X_{t-2} = \ldots$, the long-run or steady state gain of Y_t from a change in X_t will be

$$\sum_{j=0}^{\infty} v_j = g\,. \qquad \text{[Eqn 5.4]}$$

Thus, for a stable system the sum of the impulse response weights converges and is equal to the long-run or steady state gain of the system.

Let x_t denote the incremental changes of X_t and y_t denote the incremental changes of Y_t, i.e.,

$$x_t = X_t - X_{t-1} = \nabla X_t \qquad \text{[Eqn 5.5]}$$

$$y_t = Y_t - Y_{t-1} = \nabla Y_t \qquad \text{[Eqn 5.6]}$$

where ∇ is the first-difference operator.

Eqn 5.3 may also be written as

$$y_t = v(B) x_{t-b} \qquad \text{[Eqn 5.7]}$$

A general model for describing discrete dynamic systems is the linear difference equation

$$(1 + \xi_1 \nabla + \Lambda + \xi_r \nabla^r) Y_t = g(1 + \eta_1 \nabla + \Lambda + \eta_s \nabla^s) X_{t-b}, \qquad \text{[Eqn 5.8]}$$

where ξ_i (i = 1, 2, …, r) and η_j (j = 1, 2, …, s) are model parameters. Eqn 5.8 is referred to as a transfer function model of order (r, s), in which r and s represent the numbers of past values of Y and X, respectively. Using the backward shift operator, $B = 1 - \nabla$, the above equation may be written as

$$(1 - \delta_1 B - \Lambda - \delta_r B^r) Y_t = (\omega_0 - \omega_1 B - \Lambda - \omega_s B^s) X_{t-b}$$

or

$$\delta(B) Y_t = \omega(B) X_{t-b} \qquad \text{[Eqn 5.9]}$$

Equivalently, suppose $\Omega(B) = \omega(B) B^b$, the model may be written as

$$\delta(B) Y_t = \Omega(B) X_t \qquad \text{[Eqn 5.10]}$$

Comparing Eqn 5.10 with Eqn 5.1, we obtain the transfer function of this model as

$$v(B) = \delta^{-1}(B) \Omega(B) \qquad \text{[Eqn 5.11]}$$

5.1.2 Estimation (LS)

The univariate model is assumed to be

$$y_i = \beta_0 x_{i,t} + \beta_1 x_{i,t-1} + \Lambda + \beta_k x_{i,t-k} + e_i \qquad \text{[Eqn 5.12]}$$

where the y_i (i = 1, 2, …, n) are observations, x_t, x_{t-1}, …, x_{t-k} are the lagged independent variables, the β_i are unknown parameters to be estimated, and the e_i are uncorrelated errors having zero means and the same variance σ^2. Eqn 5.12 can be written in compact form

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{e} \qquad \text{[Eqn 5.13]}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \mathrm{M} \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x_{1,t} & x_{1,t-1} & \Lambda & x_{1,t-k} \\ x_{2,t} & x_{2,t-1} & \Lambda & x_{2,t-k} \\ \mathrm{M} & \mathrm{M} & & \mathrm{M} \\ x_{n,t} & x_{n,t-1} & \Lambda & x_{n,t-k} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \mathrm{M} \\ \beta_k \end{bmatrix}, \text{ and } \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \mathrm{M} \\ e_n \end{bmatrix}.$$

It is assumed that the rank of $\boldsymbol{X}$ is $k + 1$. The estimates $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k)$ of the parameters β are obtained by minimizing the sum of squares.

$$S(\boldsymbol{\beta}) = e^T e = (y - \mathbf{X\beta})^T (\mathbf{y} - \mathbf{X\beta}) \qquad \text{[Eqn 5.14]}$$

Hence,

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \qquad \text{[Eqn 5.15]}$$

5.1.3 Diagnostic Tests

There are a number of tests for checking model validity and adequacy. In this subsection, three model diagnostic tests including Q statistics, ARCH test, and Ramsey's RESET test are introduced. Q statistics provide the criterion value to check for whiteness of residuals while ARCH test examines the presence of heteroskedasticity. Ramsey's RESET test checks for the model specification.

5.1.3.1 Q Statistics

As discussed in the section 4.1.2, the e_i in Eqn 5.12 are assumed to be random and uncorrelated white noise. Thus it is necessary to check whether the e_i from the estimated model is close enough to be white noise. The Q-statistic at lag L is a test statistic that determines if there is autocorrelation up to order L, and is given as

$$Q = N(N+2)\sum_{k=1}^{L}\frac{r_k^2}{N-k} \qquad \text{[Eqn 5.16]}$$

where r_k is the k-th autocorrelation, L is the number of lags, and N is the number of observations. The null hypothesis to be tested is that there is no autocorrelation up to order L. Under the null hypothesis, Q-statistic is asymptotically distributed as a χ^2-distribution with degrees of freedom equal to the number of autocorrelations.

The Q-statistic is often used as a test of whether the time series has white noise. There remains the practical problem of choosing the order of a lag to use for the test. If a lag is too small, the test may not detect serial correlation at high-order lags. However, if a lag is too large, the test may have low power since the significant correlation at one lag may be diluted by insignificant correlations at other lags.

5.1.3.2 ARCH Test

Let y_t denote an autoregressive process of order p (this process will be discussed in more detail in the subsection 5.2.1), AR(p), y_t may be expressed as

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \Lambda + \phi_p y_{t-p} + w_t \qquad \text{[Eqn 5.17]}$$

where c is a constant, $\phi_1, \phi_2, \mathrm{K}, \phi_p$ are parameter estimates of the process, and w_t is a white noise satisfying

$$E(w_t) = 0$$

$$E(w_t w_\tau^T) = \begin{cases} \sigma^2, & t = \tau \\ 0, & t \neq \tau \end{cases}$$

However, the conditional variance of w_t may change over time while the unconditional variance of w_t is the constant σ^2.

Engle (1982) derived the *Autoregressive Conditional Heteroskedasticity* (ARCH) test based on the Lagrange multiplier principle. This test checks whether the residuals w_t from a regression model exhibits time-varying heteroskedasticity. When the residuals w_t for different time have different variances but are uncorrelated with each other, then the residuals w_t are said to exhibit heteroskedasticity, in which case it is inappropriate to estimate model parameters using the least squares method. Instead, the model needs to be estimated using the GARCH method. Suppose $\hat{w}_t^2$ be regressed on a constant and m of its own lagged values then it is represented as

$$\hat{w}_t^2 = \zeta + \alpha_1 \hat{w}_{t-1}^2 + \alpha_2 \hat{w}_{t-2}^2 + \Lambda + \alpha_m \hat{w}_{t-m}^2 + e_t \qquad \text{[Eqn 5.18]}$$

where $t = 1, 2, \ldots, T$ and e_t is a new white noise process satisfying $E(e_t) = 0$ and $E(e_t e_\tau^T) = \lambda^2$ for $t = \tau$, otherwise 0. The linear projection of the squared error of a forecast of y_t on the previous m squared forecast errors is given by

$$\hat{E}(w_t^2 | w_{t-1}^2, w_{t-2}^2, \ldots) = \zeta + \alpha_1 w_{t-1}^2 + \alpha_2 w_{t-2}^2 + \Lambda + \alpha_m w_{t-m}^2 \qquad \text{[Eqn 5.19]}$$

A white noise process w_t satisfying Eqn 5.18 is described as an *autoregressive conditional heteroskedastic* process of order m, denoted w_t~ARCH(m). The sample size T times the centered R^2_u from the regression of Eqn 5.17 converges in distribution to a χ^2 variable with m degrees of freedom under the null hypothesis that w_t is assumed to be independent and identically distributed (i.i.d.) $N(0, \sigma^2)$ (Hamilton 1994).

5.1.3.3 Ramsey's RESET Test

Ramsey (1969) developed a test called RESET (Regression Specification Errors Test) to test the hypothesis that no relevant explanatory variables have been omitted from the regression equation. Suppose the true model is

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon} \qquad \text{[Eqn 5.20]}$$

where $\boldsymbol{X}_1$ is an ($N \times K$) matrix of nonstochastic and observable process and $\mathbf{X}_2$ is unobservable and $\boldsymbol{\varepsilon}$ is white noise following all basic assumptions. The hypothesis to be tested is

H_0: $\boldsymbol{\beta}_2 = 0$,
H_A: $\boldsymbol{\beta}_2 \neq 0$.

Since $\mathbf{X}_2$ is unobservable, $\mathbf{X}_2\boldsymbol{\beta}_2$ has to be approximated by $\mathbf{Z}\theta$, where $\mathbf{Z}$ is a ($N \times K$) set of observable, nonstochastic test variables and θ is a $(P \times 1)$ vector of corresponding coefficients. The method of least squares is then applied to

$$\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{Z}\theta + \mathbf{u} \qquad \text{[Eqn 5.21]}$$

resulting in

$$\vec{\theta} = (\mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{y} \qquad \text{[Eqn 5.22]}$$

where $\mathbf{M}_1 = \mathbf{I} - \mathbf{X}_1(\mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1)^{-1}\mathbf{X}^{\mathrm{T}}_1$. Therefore,

$$E(\vec{\theta}) = (\mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 \qquad \text{[Eqn 5.23]}$$

Thus, under H_0,

$$E(\vec{\theta}) = 0$$

and, under H_A,

$$E(\vec{\theta}) \neq 0 \quad (\text{unless } \mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 = 0).$$

This means that a *test of significance of* $\vec{\theta}$ would provide an enough evidence to reject H_0 as long as $\mathbf{Z}^{\mathrm{T}}\mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 \neq 0$.

5.1.4 Efficient Estimation (GARCH)

Suppose that the white noise w_t in Eqn 5.17 is

$$w_t = \sqrt{h_t} \cdot v_t \qquad \text{[Eqn 5.24]}$$

where v_t is an i.i.d. sequence with zero mean and unit variance and where h_t evolves according to

$$h_t = \zeta + \alpha_1 w_{t-1}^2 + \alpha_2 w_{t-2}^2 + \Lambda + \alpha_m w_{t-m}^2 \qquad \text{[Eqn 5.25]}$$

The generalized process for which the conditional variance depends on an infinite number of lags of w^2_{t-j} with lag operator B is represented as

$$h_t = \zeta + \pi(B) w_t^2 \qquad \text{[Eqn 5.26]}$$

where

$$\pi(B) = \sum_{j=1}^{\infty} \pi_j B^j \qquad \text{[Eqn 5.27]}$$

If $\pi(B)$ is parameterized as the ratio of two finite-order polynomials, then

$$\pi(B) = \frac{\alpha(B)}{1-\delta(B)} = \frac{\alpha_1 B^1 + \alpha_2 B^2 + \Lambda + \alpha_m B^m}{1 - \delta_1 B^1 - \delta_2 B^2 - \Lambda - \delta_r B^r} \qquad \text{[Eqn 5.28]}$$

where we assume that the roots of $1 - \delta(z) = 0$ are outside the unit circle. By multiplying $1 - \delta(B)$, Eqn 5.26 becomes

$$[1 - \delta(B)]h_t = [1 - \delta(B)]\zeta + \alpha(B)w^2{}_t$$

or

$$\begin{aligned} h_t &= \kappa + \delta_1 h_{t-1} + \delta_2 h_{t-2} + \Lambda + \delta_r h_{t-r} \\ &\quad + \alpha_1 w_{t-1}^2 + \alpha_2 w_{t-2}^2 + \Lambda + \alpha_m w_{t-m}^2 \end{aligned} \qquad \text{[Eqn 5.29]}$$

for $\kappa = [1 - \delta_1 - \delta_2 - \ldots - \delta_r]\zeta$. Eqn 5.29 is the *generalized autoregressive conditional heteroskedasticity* model, denoted as w_t~GARCH(r, m).

5.2 Multivariate Methodology

In this section, two multivariate models are described: vector autoregressive model (VAR) and vector autoregression with an exogenous variable (VARX). VAR identifies the interactions between endogenous variables, while VARX allows one to examine the impact from policy instruments or control variables (exogenous variables) to variables (endogenous variables) in the system. An exogenous variable is defined outside of the system.

After the discussion of two types of multivariate models, restricted system is described. A large estimation uncertainty for the model coefficients leads to imprecision in the estimation of cumulative impacts, which is one of the interests for this research. This problem may be cured by imposing constraints on the coefficients, resulting in a restricted system.

5.2.1 Vector AutoRegressive (VAR) Model

In this sub-section, VAR model is described in terms of model structure and estimation. Five criteria to select an appropriate order of model are presented. Four diagnostic tests including AR root test, Correlogram, Lagrange Multiplier test, and Portmanteau test are first described. Granger Causality test, which provides test statistics to check whether there is a causal relationship between two variables, is then introduced. Finally impulse response is discussed to examine the response of one variable to an impulse in another variable in a system that contains a number of other variables as well.

5.2.1.1 Model

In forecasting, it is plausible to predict the future value of the variable of interest based at least partially on data collected in the past since the growth rate in a certain period time has a significant influence to the next period time. In other words, the past value of the variable of interest helps to forecast the value of variable in period t. Let y_t denote the value of the variable at time t and w_t is a random variable. Then a dynamic equation relating these variables is given below:

$$y_t = \phi y_{t-1} + w_t \qquad \text{[Eqn 5.30]}$$

This equation is *a linear first-order difference equation* because only first lag of variable (y_{t-1}) appears in the equation. By generalizing the dynamic system described in Eqn 5.12 to model the value of y at time t with p of its own lags and the current value of the input variable w_t, a *linear* p^{th}*-order difference equation* is obtained as shown below:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \Lambda + \phi_p y_{t-p} + w_t \qquad \text{[Eqn 5.31]}$$

By substituting the input variable w_t in Eqn 5.30 with $w_t = c + \varepsilon_t$, we have *a first-order autoregression* denoted $AR(1)$ and represented as the following difference equation:

$$Y_t = c + \phi Y_{t-1} + \varepsilon_t \qquad \text{[Eqn 5.32]}$$

where ε_t is a *white noise process* satisfying conditions: $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma^2$, and $E(\varepsilon_t\varepsilon_\tau) = 0$ for $t \neq \tau$. A white noise ε_t is assumed to be normally independently identically distributed with zero mean and unit variance, $\varepsilon_t \sim$ i.i.d. $N(0, 1)$. By successively substituting the preceding value of the variable in Eqn 5.32, Y_t could be expressed as below.

$$\begin{aligned}
Y_t &= c + \phi Y_{t-1} + \varepsilon_t \\
&= (c + \varepsilon_t) + \phi(c + \phi Y_{t-2} + \varepsilon_{t-1}) \\
&= (c + \varepsilon_t) + \phi(c + \varepsilon_{t-1}) + \phi^2 Y_{t-2} \\
&\mathrm{M} \qquad \text{[Eqn 5.33]} \\
&= (c + \varepsilon_t) + \phi(c + \varepsilon_{t-1}) + \phi^2(c + \varepsilon_{t-2}) + \Lambda \\
&= (c + \phi c + \phi^2 c + \Lambda) + (\varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \Lambda) \\
&= [c/(1-\phi)] + \varepsilon_t + \phi\varepsilon_{t-1} + \phi^2\varepsilon_{t-2} + \Lambda
\end{aligned}$$

The result of taking expectations of Eqn 5.33 is

$$E(Y_t) = [c/(1 - \phi)] + 0 + 0 + \ldots,$$

so that the mean of $AR(1)$ process is $\mu = c/(1 - \varphi)$.

A p^{th}-*order autoregression*, denoted AR(p), satisfies

$$\mathbf{Y}_t = \mathbf{c} + \phi_1 \mathbf{Y}_{t-1} + \phi_2 \mathbf{Y}_{t-2} + \ldots + \phi_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t. \qquad \text{[Eqn 5.34]}$$

However, the value of the variable of interest is not only related to its past values but also depends on past values of other variables. Moreover, it is often of interest to learn about the dynamic interrelationships among a set of variables. Multivariate autoregressive model or vector autoregressive (VAR) model investigating the dynamic interrelationships is discussed in this section. The VAR model is merely a multiple time series generalization of the AR model. A p^{th}-*order vector autoregression*, denoted VAR(p), is a vector generalization of Eqn 5.34.

$$y_t = c + \boldsymbol{\Phi}_1 y_{t-1} + \boldsymbol{\Phi}_2 y_{t-2} + \ldots + \boldsymbol{\Phi}_p y_{t-p} + u_t \qquad \text{[Eqn 5.35]}$$

Here $\mathbf{y}_t = (y_{1t}, \ldots, y_{Nt})^T$ is a N-dimensional multiple time series, $\mathbf{c} = (c_1, \ldots, c_N)^T$ denotes an $(N \times 1)$ vector of constants and $\boldsymbol{\Phi}_i$ is an $(N \times N)$ matrix of autoregressive coefficients for $i = 1, 2, \ldots, p$. The $(N \times 1)$ vector $\mathbf{u}_t$ is a vector generalization of white noise satisfying $E(\mathbf{u}_t) = 0$ and $E(\mathbf{u}_t\mathbf{u}_t^T) = \Omega$ for $t = \tau$, otherwise $\mathbf{0}$.

5.2.1.2 Estimation

Let's define the following vectors and matrices:

$$\begin{aligned}
Y &= \begin{bmatrix} y_1, & \mathrm{K}, & y_N \end{bmatrix} && (N \times T), \\
B &= \begin{bmatrix} v, & \Phi_1, & \mathrm{K}, & \Phi_p \end{bmatrix} && (N \times (Np + 1)), \\
Z_t &= \begin{bmatrix} 1 \\ y_t \\ \mathrm{M} \\ y_{t-p+1} \end{bmatrix} && ((Np + 1) \times 1), \\
\mathbf{Z} &= \begin{bmatrix} Z_0, & \mathrm{K}, & Z_{T-1} \end{bmatrix} && ((Np + 1) \times T), \qquad \text{[Eqn 5.36]} \\
U &= \begin{bmatrix} u_1, & \mathrm{K}, & u_T \end{bmatrix} && (N \times T), \\
\mathbf{y} &= \mathrm{vec}(Y) && (NT \times 1), \\
\boldsymbol{\beta} &= \mathrm{vec}(B) && ((N^2 p + N) \times 1), \\
\mathbf{b} &= \mathrm{vec}(B') && ((N^2 p + N) \times 1), \\
\mathbf{u} &= \mathrm{vec}(U) && (NT \times 1)
\end{aligned}$$

where vec is the column stacking operator and T is a sample size or time series length. Using the above notations, the VAR(p) model in Eqn 5.35 can be written in compact form

$$Y = BZ + U \qquad \text{[Eqn 5.37]}$$

or

$$\begin{aligned}\mathrm{vec}(Y) &= \mathrm{vec}(BZ) + \mathrm{vec}(U)\\ &= (Z' \otimes I_N)\mathrm{vec}(B) + \mathrm{vec}(U)\end{aligned}$$

or

$$\mathbf{y} = (Z^{\mathrm{T}} \otimes I_N)\boldsymbol{\beta} + \mathbf{u} \qquad \text{[Eqn 5.38]}$$

where $\otimes$ is the Kronecker product. Note that the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{u}$ is

$$\boldsymbol{\Sigma}_{\mathbf{u}} = I_T \otimes \boldsymbol{\Sigma}_{\mathrm{u}}. \qquad \text{[Eqn 5.39]}$$

Thus, multivariate *LS* estimation of $\boldsymbol{\beta}$ means to choose the estimator that minimizes

$$\begin{aligned}S(\boldsymbol{\beta}) &= \boldsymbol{u}^{\mathrm{T}}(I_T \otimes \boldsymbol{\Sigma}_{\mathrm{u}})^{-1}\mathbf{u} = \mathbf{u}^{\mathrm{T}}(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{u}\\ &= (\mathbf{y} - (Z' \otimes I_N)\boldsymbol{\beta})^{\mathrm{T}}(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})(\mathbf{y} - (Z' \otimes I_N)\boldsymbol{\beta})\\ &= \mathrm{vec}(Y - BZ)^{\mathrm{T}}(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\,\mathrm{vec}(Y - BZ)\\ &= \mathrm{tr}[(Y - BZ)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}}(Y - BZ)]. \qquad \text{[Eqn 5.40]}\end{aligned}$$

Note that tr(M) or tr M is the trace of M. In order to find the minimum of this function we note that

$$\begin{aligned}S(\boldsymbol{\beta}) &= \mathbf{y}'(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y} + \boldsymbol{\beta}'(Z \otimes I_N)(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})(Z' \otimes I_N)\boldsymbol{\beta} - 2\boldsymbol{\beta}'(Z \otimes I_N)(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y}\\ &= \mathbf{y}'(I_T \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y} + \boldsymbol{\beta}'(ZZ' \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\boldsymbol{\beta} - 2\boldsymbol{\beta}'(Z \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y}\end{aligned}$$

Hence,

$$\frac{\partial S(\beta)}{\partial \beta} = 2(ZZ' \otimes \boldsymbol{\Sigma}_u^{-1})\beta - 2(Z \otimes \boldsymbol{\Sigma}_u^{-1})y.$$

Equating to zero gives the *normal equations*

$$(ZZ' \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\hat{\beta} = (Z \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y} \qquad \text{[Eqn 5.41]}$$

and consequently, the LS estimator is

$$\begin{aligned}\hat{\beta} &= ((ZZ')^{-1} \otimes \boldsymbol{\Sigma}_{\mathrm{u}})(Z \otimes \boldsymbol{\Sigma}^{-1}{}_{\mathrm{u}})\mathbf{y}\\ &= ((ZZ')^{-1} \otimes I_N)\mathbf{y}\end{aligned}$$

Note that this estimator is also the Maximum Likelihood (ML) estimator, $\hat{\delta}$.

5.2.1.3 Order Selection

The approximate mean squared error (MSE) matrix of the one-step predictor will increase with the order p (Lütkepohl 1993). Thus, applying unnecessarily large p will result in imprecise of estimates of the corresponding VAR(p) model. Consequently the imprecise parameter estimates will reduce the estimation precision of the impulse responses. Therefore it is necessary to have procedures or criteria for choosing an adequate VAR order.

Obviously there is not just one correct VAR order for the process $\mathbf{y}_t$ in Eqn 5.35. Lütkepohl (1993) defined VAR order, so-called the order of the process, as a unique number p if $\Phi_p \neq 0$ and $\Phi_i = 0$ for $i > p$ so that p is the smallest possible order.

There are five criteria to determine the order of the process discussed in this subsection. They are driven by different reasoning and objectives. First, the principle of the likelihood ratio (LR) testing is based on comparing the maxima of the log-likelihood function over the unrestricted and restricted parameter space. The likelihood ratio statistics is

$$\lambda_{LR} = 2[\ln l(\bar{\beta}) - \ln l(\tilde{\beta})] \qquad \text{[Eqn 5.42]}$$

where $\bar{\beta}$ is the unrestricted maximum likelihood estimator and $\tilde{\beta}$ is the restricted maximum likelihood estimator.

To determine the correct VAR order, i.e., the number of observations of the endogenous variable Y to be included in the model, it is necessary to set up sequence of tests checking if coefficient matrix $\mathbf{\Phi}$ is zero. If coefficient matrix $\mathbf{\Phi}_M$ at the VAR order M is zero (H_0: $\mathbf{\Phi}_M = 0$), the VAR order M is not significant. Thus the next VAR order $M-1$ is tested, and this test continues until the VAR order becomes 1. For instance, assuming M is an upper bound for the VAR order, the following sequence of null and alternative hypotheses may be tested using LR tests:

$$\begin{array}{ll}
H^1{}_0: \mathbf{\Phi}_M = 0 & \text{against } H^1{}_1: \mathbf{\Phi}_M \neq 0 \\
H^2{}_0: \mathbf{\Phi}_{M-1} = 0 & \text{against } H^2{}_1: \mathbf{\Phi}_{M-1} \neq 0 \mid \mathbf{\Phi}_M = 0 \\
\vdots & \\
H^i{}_0: \mathbf{\Phi}_{M-i+1} = 0 & \text{against } H^i{}_1: \mathbf{\Phi}_{M-i+1} \neq 0 \mid \mathbf{\Phi}_M = \ldots = \mathbf{\Phi}_{M-i+2} = 0 \\
\vdots & \\
H^M{}_0: \mathbf{\Phi}_1 = 0 & \text{against } H^M{}_1: \mathbf{\Phi}_1 \neq 0 \mid \mathbf{\Phi}_M = \ldots = \mathbf{\Phi}_2 = 0
\end{array} \qquad \text{[Eqn 5.43]}$$

The likelihood ratio statistic for testing the i-th null hypothesis is

$$\lambda_{LR}(i) = T(\ln|\tilde{\mathbf{\Sigma}}_u(M-i)| - \ln|\tilde{\mathbf{\Sigma}}_u(M-i+1)|) \qquad \text{[Eqn 5.44]}$$

where $\tilde{\mathbf{\Sigma}}_u(m)$ denotes the maximum likelihood estimator of $\mathbf{\Sigma}_u$ when a VAR(m) model is fitted to a time series of length T. For testing, since we have N^2 restrictions, we can use critical values from a $\chi^2(N^2)$-distribution. Alternatively one may use $\lambda_{LR}(i)/N^2$ in conjunction with the $F(N^2, T - N(M - i + 1) - 1)$-distribution.

It should be noted that the significance levels of the individual tests must be distinguished from the Type I error of the entire procedure because rejection of $H^i{}_0$ means that $H^{i+1}{}_0, \ldots, H^M{}_0$ are automatically rejected, too.

If we are interested in obtaining a good model for prediction, it makes sense to choose the order such that forecast mean squared error (MSE) is minimized. Hence we will discuss criteria based on the forecasting objective. One of such criteria is the *final prediction error* (FPE) criterion:

$$
\begin{aligned}
FPE(m) &= \det\left[\frac{T+Nm+1}{T}\frac{T}{T-Nm-1}\tilde{\Sigma}_u(m)\right] \\
&= \left[\frac{T+Nm+1}{T-Nm-1}\right]^N \det(\tilde{\Sigma}_u(m))
\end{aligned}
\qquad \text{[Eqn 5.45]}
$$

Based on the FPE criterion the estimate $\hat{p}(FPE)$ of p is chosen such that

$$FPE[\hat{p}(FPE)] = \min\{FPE(m) | m = 0, 1, \ldots, M\}.$$

That is, VAR models of orders m = 0, 1, …, M are estimated and the corresponding $FPE(m)$ values are computed. The order minimizing the FPE values is then chosen as the estimate for p.

Another criterion is called *Akaike's Information Criterion* (AIC). For a VAR(m) process the criterion is defined as

$$
\begin{aligned}
AIC(m) &= \ln|\tilde{\Sigma}_u(m)| + \frac{2}{T}(\textit{number of freely estimated parameters}) \\
&= \ln|\tilde{\Sigma}_u(m)| + \frac{2mN^2}{T}
\end{aligned}
\qquad \text{[Eqn 5.46]}
$$

The estimate $\tilde{p}(AIC)$ for p is chosen so that this criterion is minimized.

Two other criteria have been used popularly in recent studies. The first one is *Hannan-Quinn criterion,* denoted as HQ:

$$
\begin{aligned}
HQ(m) &= \ln|\tilde{\Sigma}_u(m)| + \frac{2\ln\ln T}{T}(\textit{number of freely estimated parameters}) \\
&= \ln|\tilde{\Sigma}_u(m)| + \frac{2\ln\ln T}{T} mN^2
\end{aligned}
\qquad \text{[Eqn 5.47]}
$$

The estimate $\hat{p}(HQ)$ is the order that minimizes HQ(m) for m = 0, 1, …, M.

SC is the other criterion devised by Schwarz (1978) using Bayesian arguments and is defined as

$$SC(m) = \ln|\widetilde{\Sigma}_u(m)| + \frac{\ln T}{T}(number\ of\ \ freely\ estimated\ parameters)$$
$$= \ln|\widetilde{\Sigma}_u(m)| + \frac{\ln T}{T} mN^2 \qquad \text{[Eqn 5.48]}$$

Again the order estimate $\hat{p}(SC)$ is chosen to minimize the value of the criterion.

5.2.1.4 Diagnostic Tests

Diagnosis tests described here include AR Roots, Correlogram, LM test, and Portmanteau test. These tests are for the purposes of checking model adequacy. AR roots checks the stability of model. Correlogram detects autocorrelation of a series. LM test examines serial correlation of the residuals. Portmanteau test checks for the whiteness of the residuals.

AR roots

The estimated VAR(p) process is *stable* (stationary) if its *reverse characteristic polynomial*, $\det(I_{Np} - \boldsymbol{\Phi} z) = \det(I_N - \Phi_1 z - \ldots - \Phi_p z^p)$, has no roots in and on the complex unit circle (Lütkepohl 1991). In other words, y_t is stable if

$$\det(I_N - \Phi_1 z - \ldots - \Phi_p z^p) \neq 0 \qquad \text{for } |z| \leq 1. \qquad \text{[Eqn 5.49]}$$

This condition is called the *stability condition*. There will be $k \times p$ roots, where k is the number of endogenous variables and p is the largest lag. If the VAR(p) process is not stable, certain results (such as impulse response standard errors) are not valid. Non-stationary time series may be transformed to a stationary series by differencing once or more times.

Correlogram

Correlogram is a graph showing autocorrelation of a series with plus or minus two times the asymptotic standard errors of the lagged correlations. To check the white noise assumption for the residuals of a VAR(p) process in Eqn 5.35, the estimated autocorrelation matrices of the u_t, denoted by R_i, can be computed as

$$R_i = D^{-1} C_i D^{-1} \qquad i = 0, 1, \ldots, h, \qquad \text{[Eqn 5.50]}$$

where D is a $(N \times N)$ diagonal matrix with the standard deviations of the components of u_t on the diagonal and C_i is the autocovariance matrices of u_t. The autocovariance matrices of u_t are estimated as

$$C_i = \frac{1}{T}\sum_{t=i+1}^{T} u_t u_{t-i}' = \frac{1}{T}\sum_{t=i+1}^{T}(u_t - \overline{u})(u_{t-i} - \overline{u})' \qquad \text{[Eqn 5.51]}$$

where $\bar{u} = \frac{1}{T}\sum_{t=1}^{T} u_t$.

The diagonal elements of D are the square roots of the diagonal elements of C_0. Denoting the covariance between $u_{m,i}$ and $u_{n,i}$ by $c_{\mathrm{mn,i}}$, the diagonal elements $c_{11,0}$, …, $c_{NN,0}$ of C_0 are the variances of u_{1t}, …, u_{Nt}. Thus,

$$D^{-1} = \begin{bmatrix} 1\big/\sqrt{c_{11}(0)} & & 0 \\ & \mathrm{O} & \\ 0 & & 1\big/\sqrt{c_{NN}(0)} \end{bmatrix} \qquad \text{[Eqn 5.52]}$$

and the correlation between $u_{m,\mathrm{i}}$ and $u_{n,\mathrm{i}}$ is

$$r_{mn,i} = \frac{c_{mn,i}}{\sqrt{c_{mm,0}}\sqrt{c_{nn,0}}}, \qquad \text{[Eqn 5.53]}$$

which is the mn^{th} element of R_i.

Lagrange Multiplier Test

The Lagrange Multiplier (LM) test provides the test statistics for residual serial correlation up to the specified order. The hypothesis for the test is shown as

$$H_0\text{: } \varphi(\boldsymbol{\beta}) = 0 \text{ against } \qquad H_1\text{: } \varphi(\boldsymbol{\beta}) \neq 0, \qquad \text{[Eqn 5.54]}$$

where $\boldsymbol{\beta}$ is $(N^2p + N)$-dimensional parameter vector and $\varphi(\cdot)$ is a twice continuously differentiable function with values in the K-dimensional space. In other words, $\varphi(\boldsymbol{\beta})$ is an $(K \times 1)$ vector and it is assumed that the matrix $\partial\varphi/\partial\boldsymbol{\beta}^{\mathrm{T}}$ of first partial derivatives has rank K at the true parameter vector.

The LM statistic for testing the hypothesis in Eqn 5.54 is

$$\lambda_{LM} = s(\tilde{\beta}_r)'\tilde{I}_a(\tilde{\beta}_r, \tilde{\Sigma}_u^r)^{-1} s(\tilde{\beta}_r) / T \qquad \text{[Eqn 5.55]}$$

where

$$s(\tilde{\beta}_r) = \left.\frac{\partial \ln l_0}{\partial \beta}\right|_{\tilde{\beta}_r} = \sum_{t=1}^{T}\left[\left.\frac{\partial u_t(\bar{y},\beta)'}{\partial \beta}\right|_{\tilde{\beta}_r}\right](\tilde{\Sigma}_u^r)^{-1}\tilde{u}_t(\bar{y},\tilde{\beta}_r) \qquad \text{[Eqn 5.56]}$$

is the score vector evaluated at the restricted estimator $\tilde{\beta}_r$ and

$$\tilde{I}_a(\tilde{\beta}_r, \tilde{\Sigma}_u^r) = \frac{1}{T}\sum_{t=1}^{T}\left[\left.\frac{\partial u_t(\bar{y},\beta)'}{\partial \beta}\right|_{\tilde{\beta}_r}\right](\tilde{\Sigma}_u^r)^{-1}\left[\left.\frac{\partial u_t(\bar{y},\beta)}{\partial \beta'}\right|_{\tilde{\beta}_r}\right] \qquad \text{[Eqn 5.57]}$$

is an estimator of the asymptotic information matrix based on the restricted estimator $\tilde{\beta}_r$. Here

$$\tilde{\Sigma}_u^r = \frac{1}{T}\sum_{t=1}^{T}\tilde{u}_t(\bar{y},\tilde{\beta}_r)\tilde{u}_t(\bar{y},\tilde{\beta}_r)'. \qquad \text{[Eqn 5.58]}$$

Under the null hypothesis of no serial correlation of order h, the statistic λ_{LM} is asymptotically distributed as a $\chi^2(h)$. Using the LM test it is particularly easy to test larger VAR orders against a maintained model because we only need an estimator of the coefficients of the VAR(p) process when testing a given VAR(p) specification against a VAR($p + s$) model.

Portmanteau tests

Portmanteau test is a popular tool for checking whiteness of the residuals and is used to check the overall significance of the residual autocorrelations up to lag h. The hypothesis to test is

$$H_0: \mathbf{R}_h = (R_1, \ldots, R_h) = 0 \quad \text{against} \quad H_1: \mathbf{R}_h \neq 0. \qquad \text{[Eqn 5.59]}$$

The test statistic is

$$\bar{P}_h = T^2 \sum_{i=1}^{h}(T-i)^{-1} tr(\hat{C}_i' \hat{C}_0^{-1} \hat{C}_i \hat{C}_0^{-1}) \qquad \text{[Eqn 5.60]}$$

where tr(M) is trace of M and

$$\hat{C}_i = \frac{1}{T}\sum_{t=i+1}^{T}\hat{u}_t \hat{u}_{t-i}' \qquad \text{[Eqn 5.61]}$$

The test statistic has an approximate asymptotic χ^2-distribution.

5.2.1.5 Granger Causality Test

Granger (1969) proposed testable definitions of causality based on the predictability of a stochastic series. For instance, if only past values of a stationary time series Y_t help in the prediction of a stationary time series X_t, then Y_t is said to cause X_t. If X_t and Y_t are two stationary time series with zero means, the general causal model with instantaneous causality is

$$X_t + b_0 Y_t = \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \varepsilon_t \qquad \text{[Eqn 5.62]}$$

$$Y_t + c_0 X_t = \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_t \qquad \text{[Eqn 5.63]}$$

where ε_t, η_t are taken to be two uncorrelated white-noise series, i.e., $E[\varepsilon_t \varepsilon_s] = E[\eta_t \eta_s] = 0$, $s \neq t$, and $E[\varepsilon_t \varepsilon_s] = 0$ for all t, s. In practice, it is assumed that m is finite and shorter than the given time series. If $b_0 = c_0 = 0$, then the given model will be the simple causal model without considering instantaneous causal relations.

To further understand Granger's idea, suppose that $Y_t(h|\Omega_t)$ is the optimal h-step predictor of the process Y_t with minimum MSE at origin t, based on the information in Ω_t containing all the relevant information in the universe available up to and including period t. Let $\Sigma_z(h|\Omega_t)$ denote the corresponding forecast MSE. The process X_t is said to *cause Y_t in Granger's sense* if

$$\Sigma_z(h|\Omega_t) < \Sigma_z(h|\Omega_t \setminus \{X_s | s \leq t\}) \text{ for at least one } h = 1, 2, \ldots \qquad \text{[Eqn 5.64]}$$

where $\Omega_t \setminus \{X_s | \ s \leq t\}$ is the set containing all the relevant information in the universe except for the information in the past and present of the X_t process.

5.2.1.6 Impulse Response

Granger-causality tests the causal relationship between two variables. In practice we often need to know the response of one variable to an impulse in another variable in a system that contains a number of other variables as well. This type of causality may be traced to find out the effect of an innovation (or a sudden change or a shock) in one of the variable on other variables.

Any VAR(p) process as described in Eqn 5.35 may be written in the VAR(1) form (Lütkepohl 1993). The corresponding VAR(1) process is

$$Y_t = v + \Phi Y_{t-1} + U_t$$

where

$$\mathbf{Y}_t = \underset{(Np \times 1)}{\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}}, \quad \mathbf{v} = \underset{(Np \times 1)}{\begin{bmatrix} v \\ 0 \\ \vdots \\ 0 \end{bmatrix}}, \quad \mathbf{\Phi} = \underset{(Np \times Np)}{\begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\ I_N & 0 & \cdots & 0 & 0 \\ 0 & I_N & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_N & 0 \end{bmatrix}}, \text{ and } \mathbf{U}_t = \underset{(Np \times 1)}{\begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}}.$$

If the stability condition, $\det(I_N - \Phi_1 z) \neq 0$ for $|z| \leq 1$, is satisfied, $\mathbf{Y}_t$ may be represented as

$$\mathbf{Y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \boldsymbol{\Phi}^i \mathbf{U}_{t-i} \quad \text{[Eqn 5.65]}$$

where $\mathbf{Y}_t$ is expressed in terms of past and present error or innovation vectors $\mathbf{U}_t$ and the mean term $\boldsymbol{\mu}$. This representation is known as the *moving average* (MA) *representation*. Multiplying Eqn 5.65 by the matrix $\mathbf{J} = [I_N\ 0\ \ldots\ 0]$, we have an MA representation of y_t, which is shown as

$$y_t = \mathbf{J}\mathbf{Y}_t = \mathbf{J}\boldsymbol{\mu} + \sum_{i=0}^{\infty} \mathbf{J}\boldsymbol{\Phi}^i \mathbf{J}'\mathbf{J}\mathbf{U}_{t-i} = \mu + \sum_{i=0}^{\infty} \Pi_i u_{t-i} \quad \text{[Eqn 5.66]}$$

Under the assumptions that $\mathbf{v} = 0$, $y_t = 0$ for $t < 0$, $u_t = 0$ for $t > 0$, and $y_0 = u_0$ is a N-dimensional unit vector with a one at the k^{th} coordinate and zeroes elsewhere, the impulse responses are the elements of the upper left-hand $(N \times N)$ block of $\boldsymbol{\Phi}^i$ in Eqn 5.65. Also as represented in Eqn 5.67 the matrix $\boldsymbol{\Phi}_i$ is the i-th coefficient matrix Π_i of the MA representation above. Therefore, $\Phi_{jk,i}$, or the jk-th element of Π_i, represents the response of the j^{th} variable of the system to a unit shock of variable k at i periods ago.

5.2.2 Vector AutoRegression with an eXogenous Variable

Since transportation improvement projects are determined outside the system of land use, it should be treated as an exogenous variable while land use variables are endogenous variables. Vector autoregression with an exogenous variable, denoted as VARX(p, s), allows one to investiage impacts from policy instruments or control variables (exogenous variables).

In this subsection, the structure of VARX models and model estimation are described. The VARX order p for endogenous variables may be decided based on five order selection criteria: the likelihood ratio (LR), final prediction error (FPE), Akaike information criterion (AIC), Schwarz information criterion (SC), and Hannan-Quinn information criterion (HQ). These criteria have been described earlier in Section 5.2.1.3. The VARX order s for an exogenous variable can be determined via trial-and-error. Since diagnostic tests for VARX models are the same as those for VAR models, they are not described in this sub-section again. Multiplier analysis to investiage impacts from an exogenous variable to an endogenous variable is presented.

5.2.2.1 VARX Model

In Section 5.2.1, the joint generation process of variables of interest has been discussed. In practice the generation process may be affected by other variables from outside of the system of interest, or exogenous variables. Reinsel (1997) claimed that if a variable X_t causes Y_t but Y_t does not cause X_t, X_t is referred to as an exogenous variable. By reformulating Eqn 5.35 with an endogenous variable, $\mathbf{y}_t = (y_{1t}, \ldots, y_{kt})^T$, and an exogenous variable, $\mathbf{x}_t = (x_{1t}, \ldots, x_{rt})^T$, VARX($p$, s) may be written as

$$y_t = \delta + \sum_{i=1}^{p} \Phi_i y_{t-i} + \sum_{i=0}^{s} \Theta_i x_{t-i} + \varepsilon_t$$

or $$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_t = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} + \sum_{i=1}^{p} \begin{bmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1N} \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N1} & \Phi_{N2} & \cdots & \Phi_{NN} \end{bmatrix}_i \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}_{t-i} + \cdots + \sum_{j=0}^{s} \begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \vdots \\ \Theta_N \end{bmatrix}_j [x]_{t-j} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}_t \quad \text{[Eqn 5.67]}$$

where δ is a $(N \times 1)$ constant vector and $\mathbf{\Phi}_i$ and $\mathbf{\Theta}_i$ are $(N \times N)$ matrices of coefficients for $i = 1, 2, \ldots, p$ and $i = 1, 2, \ldots, s$, respectively.

Order selection is done in the same way as discussed in the previous section and model estimation is also accomplished by LS estimation. Thus diagnostics tests discussed in the previous section are applicable to the VARX model.

5.2.2.2 Multipliers Analysis

In the context of transportation and urban planning, it is often of interest to investigate the marginal impact of changes in the exogenous variables. Suppose that the exogenous variables are instruments or control variables for transportation system, transportation planners or policy-makers will desire to know the consequences of changes in these control variables. For example, transportation planners or policy-makers may want to know the effects of a change in toll price on travel behavior. In this case the endogenous variable is travel behavior and the exogenous variable is toll price.

Using the lag operator notation, the process may be written as

$$\mathbf{\Phi}(L)y_t = \mathbf{\Theta}(L)x_t + \varepsilon_t, \quad \text{[Eqn 5.68]}$$

where $\delta = 0$, $\mathbf{\Phi}(L) = I_N - \mathbf{\Phi}_1 L - \ldots - \mathbf{\Phi}_p L^p$ and $\mathbf{\Theta}(L) = \mathbf{\Theta}_0 + \mathbf{\Theta}_1 L + \ldots + \mathbf{\Theta}_s L^s$.

It is useful to solve the system Eqn 5.68 for the endogenous variables in order to see the effect of a change in an exogenous variable on the endogenous variables. Assuming that the inverse of $\mathbf{\Phi}(L)$ exists and multiplying both side of Eqn 5.49 with $\mathbf{\Phi}(L)^{-1}$, the system becomes

$$y_t = \mathbf{D}(L)x_t + \mathbf{\Phi}(L)^{-1}\varepsilon_t, \quad \text{[Eqn 5.69]}$$

where $\mathbf{D}(L) = \mathbf{\Phi}(L)^{-1}\mathbf{\Theta}(L)$. The operator $\mathbf{D}(L)$ represents the *transfer function* that transfers the observable inputs into the outputs of the system (Lütkepohl 1993). This transfer function may be represented by its coefficient matrices $D_i = (d_{kj,i})$ as

$$\mathbf{D}(L) = \sum_{i=0}^{\infty} D_i L^i \quad \text{[Eqn 5.70]}$$

The coefficient matrices D_i contain the effects of changes in the exogenous variables on the endogenous variables. A unit change in the j^{th} exogenous variable in period t induces a change of $d_{kj,i}$ units in the k^{th} endogenous variable in period $t + i$ under the condition that everything else remains constant. The elements of the D_i matrices are called *dynamic multipliers*. The accumulated effects contained in $\sum_{i=0}^{n} D_i$ are the n^{th} *interim multipliers* and the elements of $\sum_{i=0}^{\infty} D_i$ are *long-run effects* or *total multipliers* (Lütkepohl 1993).

5.2.3 Restricted System

In Section 5.2.1, the estimation of the parameters of N-dimensional stationary, stable VAR(p) process has been discussed. However, in application, some of the coefficients estimates may be found to be not significantly different from zero throughout the estimation procedure. These zero coefficients may be interpreted in two ways. First, zero coefficients are found in the estimation results if some variables do not have causal relationship with other variables. Second, insignificant coefficient estimates are found if the information in the data is inadequate to provide sufficiently precise estimates with confidence intervals that do not contain zero (Lütkepohl 1993).

Since the number of parameters is often quite substantial relative to the available sample size or time series length, a large estimation uncertainty for the VAR coefficients leads to poor forecasts and imprecision in the estimates of the impulse responses and forecast error variance components. Various cures for this practical problem have been studied and putting constraints on the coefficients has been found to be a solution (Lütkepohl 1993).

Write Eqn 5.35 in compact form

$$Y = BZ + U \qquad \text{[Eqn 5.71]}$$

where

$$Y = \left[y_1, \ \ldots, \ y_T\right], \quad Z = \left[Z_0, \ \ldots, \ Z_{T-1}\right], \ \text{with} \ Z_t = \begin{bmatrix} 1 \\ y_t \\ \vdots \\ y_{t-p+1} \end{bmatrix}, \quad B = \left[v, \ A_1, \ \ldots, \ A_p\right],$$

$$U = \left[u_1, \ \ldots, \ u_T\right].$$

Suppose that linear constraints for B are given in the form

$$\boldsymbol{\beta} = \text{vec}(B) = \mathbf{R}\boldsymbol{\gamma} + r, \qquad \text{[Eqn 5.72]}$$

where $\boldsymbol{\beta} = \text{vec}(B)$ is a $N(Np + 1) \times 1$ vector, $\mathbf{R}$ is a known $N(Np + 1) \times M$ matrix of rank M, $\boldsymbol{\gamma}$ is an unrestricted $M \times 1$ vector of unknown parameters, and $\mathbf{r}$ is a $N(Np + 1)$-dimensional vector of known constants. This equation expresses all the linear restrictions and permits imposing the constraints by a simple reparameterization of the original model. Vectorizing Eqn 5.72 and replacing $\boldsymbol{\beta}$ by $\mathbf{R}\boldsymbol{\gamma} + \mathbf{r}$ gives

$$\mathbf{y} = \mathrm{vec}(Y) = (Z' \otimes I_N)\,\mathrm{vec}(B) + \mathrm{vec}(U) = (Z' \otimes I_N)(\mathbf{R}\boldsymbol{\gamma} + \mathbf{r}) + \mathbf{u}$$

or

$$\mathbf{z} = (Z' \otimes I_N)R\gamma + \mathbf{u} \qquad \text{[Eqn 5. 73]}$$

where $\mathbf{z} = \mathbf{y} - (Z' \otimes I_N)\,\mathbf{r}$ and $\mathbf{u} = \mathrm{vec}(U)$.

5.2.3.1 Efficient Estimation (GLS and EGLS)

Denoting by $\boldsymbol{\Sigma}_u$ the covariance matrix of $\mathbf{u}$, the vector $\hat{\gamma}$ minimizing

$$S(\gamma) = \mathbf{u}'(I_T \otimes \boldsymbol{\Sigma}_u^{-1})\mathbf{u} \qquad \text{[Eqn 5.74]}$$

with respect to $\boldsymbol{\gamma}$ is easily seen to be

$$\hat{\gamma} = [\mathbf{R}'(ZZ' \otimes \boldsymbol{\Sigma}_u^{-1})\mathbf{R}]^{-1}\mathbf{R}'(Z \otimes \boldsymbol{\Sigma}_u^{-1})\mathbf{z} \qquad \text{[Eqn 5.75]}$$

This estimator $\hat{\gamma}$ is commonly called a *generalized LS* (*GLS*) estimator because it minimizes the generalized sum of squared errors $S(\gamma)$ rather than the sum of squared errors $\mathbf{u}^T\mathbf{u}$. The GLS estimator is in general asymptotically more efficient than the LS estimator.

Unfortunately the estimator $\hat{\gamma}$ is of limited value in practice because its computation requires knowledge of $\boldsymbol{\Sigma}_u$. Since this matrix is usually unknown, it has to be replaced by an estimator. Using any consistent estimator $\bar{\boldsymbol{\Sigma}}_u^{-1}$ instead of $\boldsymbol{\Sigma}_u$ in Eqn 5.76 an *EGLS* (*estimated GLS*) estimator may be obtained

$$\hat{\hat{\gamma}} = [R'(ZZ' \otimes \bar{\boldsymbol{\Sigma}}_u^{-1})R]^{-1}R'(Z \otimes \bar{\boldsymbol{\Sigma}}_u^{-1})\mathbf{z} \qquad \text{[Eqn 5.76]}$$

We refer readers to Lütkepohl (1993) for the properties of GLS and EGLS estimators.

6. CASE STUDIES OF TRANSPORTATION AND LAND USE INTERACTIONS

6.1 Background

As shown in Figure 6.1, roadway improvements increase "accessibility" and remove "concurrency requirements" through additional capacity. Accessibility refers to the ease of reaching destinations with opportunities for employment, social and recreational activities, and other economic and personal activities. Concurrency requirements are constraints imposed by state or local agencies' laws and statues. In Florida, concurrency is mandated by the Florida State Statues (s. 163.3180), which requires public facilities and services to accommodate new development with the designated level of service. If a new development does not satisfy concurrency requirements, developers are required to remedy any level of service deficiencies caused by the new development. When roadway capacity is exhausted, new developments are discouraged because of the concurrency requirements due to the extra costs for impact mitigation. As accessibility improves and concurrency requirement is removed, the land becomes more attractive to developers due to lowered development costs, thereby encouraging more activities to locate in these places (Forkenbrock 2001). When new developments occur, they introduce additional demand on the transportation facilities. If the increased traffic reaches the roadway capacity again, travel demand and land development will be constrained. When the development growth is constrained, an "equilibrium point" is reached.

In this cycle of transportation improvement-new development, there are lag effects between these actions as depicted in Figure 6.1. Lag1 represents a time span after the expansion of roadways is completed and before developers apply building permits for new developments. Lag2 may be institutional lags for the building department to review applications and issue permits. Lag3 is simply the construction period for new developments. Lag2 and Lag3 may be estimated from the building permit data since each building permit record has application date, issue date, and certificate of occupancy (CO) date. CO is issued on the completion of building construction. In this study, it was assumed that travel demand responded instantaneously upon the completion of developments, and Lag1 was considered in the modeling procedure.

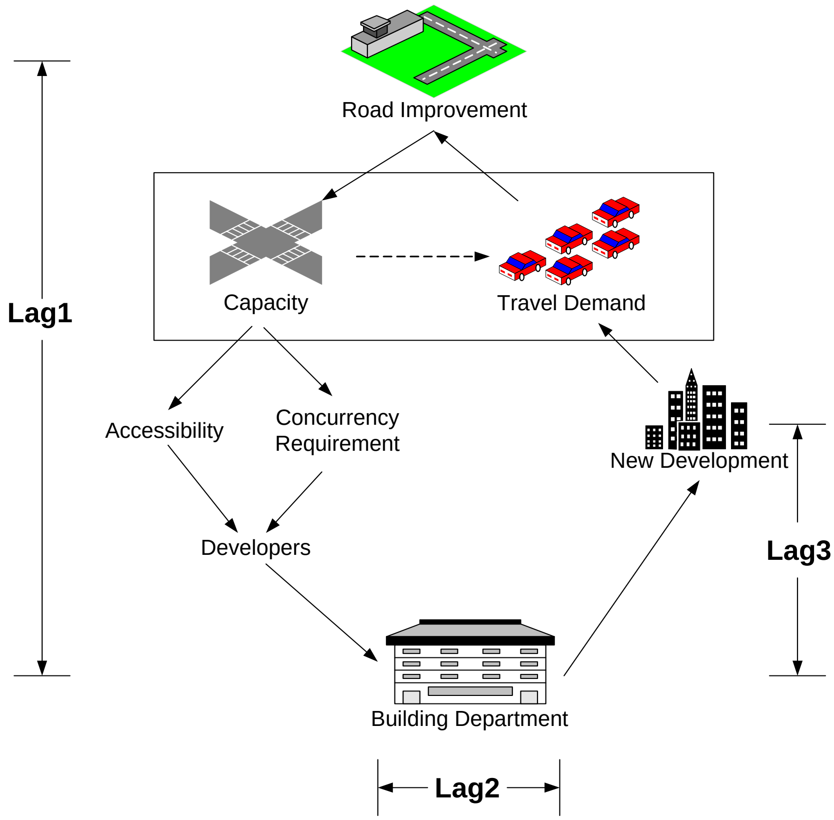


Figure 6.1 Interactions between Land Use Changes and Transportation Improvements

6.2 Study Data and Study Area Selection

A main challenge in studying the interactions or feedback mechanism between land use and transportation investments is to collect historical data on both land use and transportation. A significant amount of data and numerous databases are required because the interactions between transportation and land use may occur and last for a long period of time. The length of such periods during which interactions occur may be influenced by many factors such as availability of land, economic conditions, population growth, employment growth, and so on. Thus, effort has been made to collect land use, transportation, socioeconomic, and other data related to either land use or transportation. This effort has been a significant one because historical data were often either difficult to find or were not in a readily usable digital format due to the fact that computer use did not become wide spread until the last decade. For this research, Miami-Dade County was chosen as the study area, mostly for the reason of its excellent availability of digital, and in many cases GIS, data when compared to most of the Florida counties.

This study involved the use of historical building permit data and transportation improvement project data. The building permit database was obtained from Miami-Dade County and included building permit records from 1987 to 2001. The building permit records contained three fields: permit application date, issue date, and certificate of occupancy (CO) date. The original database had more than 1.4 million records including permits of all types issued by the county Building Department. Only records on new building constructions were extracted from the original database, resulting in 61,132 records for the unincorporated areas in Miami-Dade

County. All records were geocoded by their street addresses.

Transportation improvement project information came from the Transportation Improvement Programs (TIPs), prepared every year by the Miami-Dade County Metropolitan Planning Organization. The TIP specifies proposed transportation facility improvements to be implemented within five years in Miami-Dade County, which included capacity improvements, safety improvements, resurfacing, lighting, and so on. A total of twenty-three hard copies of TIP reports covering from 1978 to 2000 were obtained from the Miami-Dade County MPO. The temporal GIS database for these TIP reports was created by geocoding all projects listed in the reports based on the 2000 street network. A total of 883 projects were geocoded.

Three corridors were selected from the Miami-Dade County for this study based on two criteria: (1) isolation of impacts from other areas and (2) transportation improvements in the corridor (occurrence of growth in population, employment, and development). The study areas need to be relatively isolated in order to minimize impacts from adjacent street improvements or impacts from land development in other areas. The southwest of the Miami Dade County, bounded by Florida Turnpike and the Everglades National Park, is relatively isolated from other impacts. Secondly, corridors selected must have had transportation improvements. Based on these two criteria, three study corridors were selected: Tamiami Trail from SW 112th Avenue to SW 152nd Avenue, Bird Drive (or SW 42nd Street) from SW 117th Avenue to SW 157th Avenue, where the street ended, and North Kendall Drive (or SW 88th Street) between the Florida Turnpike and Krome Avenue (or SW 177th Avenue). Locations of the study corridors are depicted in Figure 6.2.

Tamiami Trail (also US 41 and SW 8th ST) is an east-west principal arterial in the southwest of Miami-Dade County. The section between SW 112th Avenue and SW 177th Avenue has a canal running on the north side of the street. Thus, most commercial developments have been located on the south side of the study corridor between the Florida Turnpike and SW 137th Avenue. The Miccosukee Indian Reservation begins at the intersection with SW 177th Avenue, where a large hotel/casino has been built. Dominant land use type to the north of the study corridor is residential. The study area around the Tamiami Trail corridor was defined by the buffer created along the study corridor. The size of buffer was varied along the corridor to capture the area impacted by the study corridor improvement. For instance, the buffer size of the north side of the study corridor between the Florida Turnpike and SW 127th Avenue was 0.8 mile, which was the halfway distance between the study corridor and the next parallel arterial, NW 12th Street; the buffer size of the south side of the study corridor between the Florida Turnpike and SW 137th Avenue was 0.6 mile, which was the half distance between the study corridor and the next parallel arterial, SW 26th Street; the buffer size for the rest of the corridor was one mile. In Figure 6.3(a), the buffer area of the Tamiami Trail corridor is illustrated, with selected building permits shown as yellow dots.

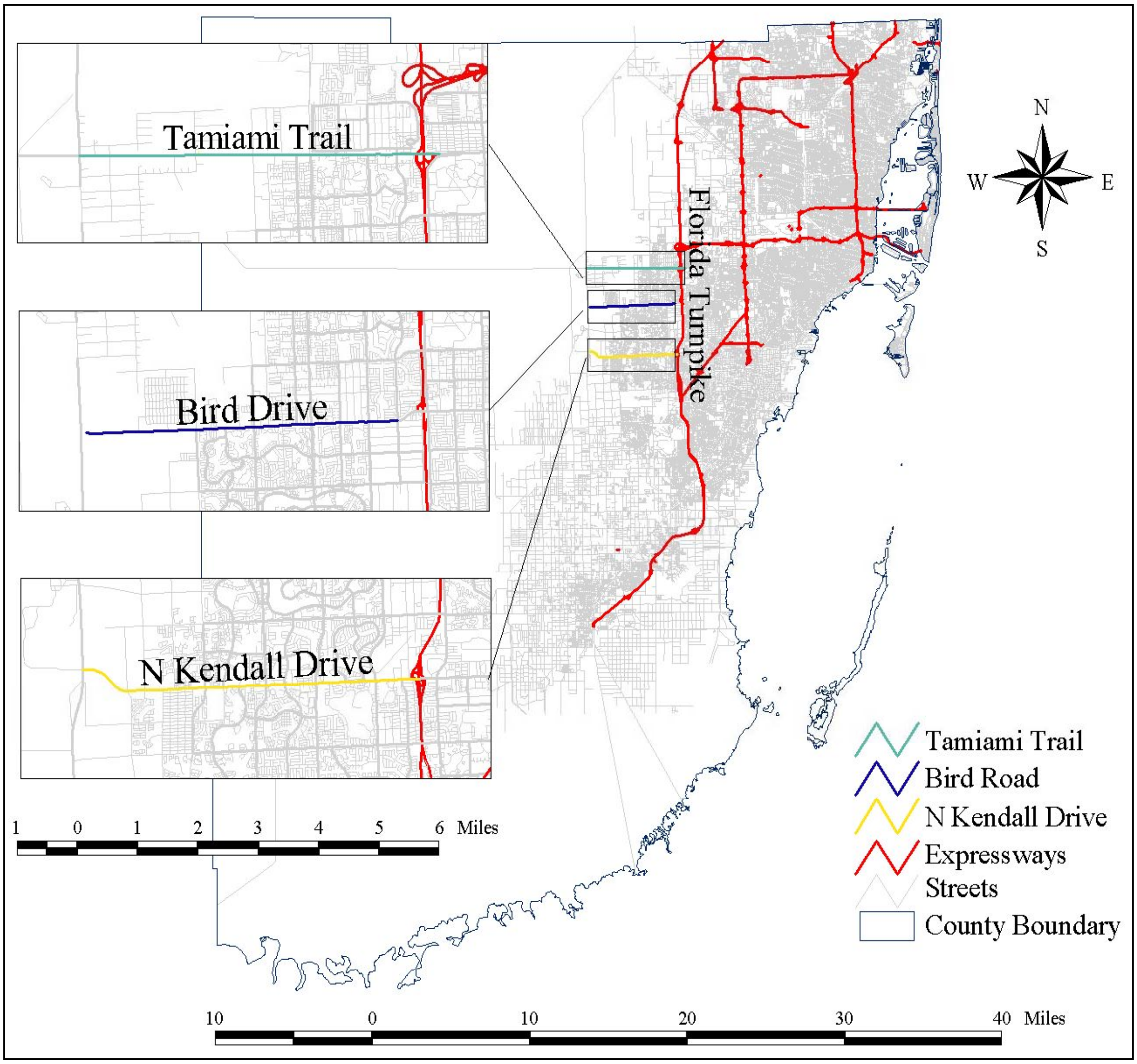
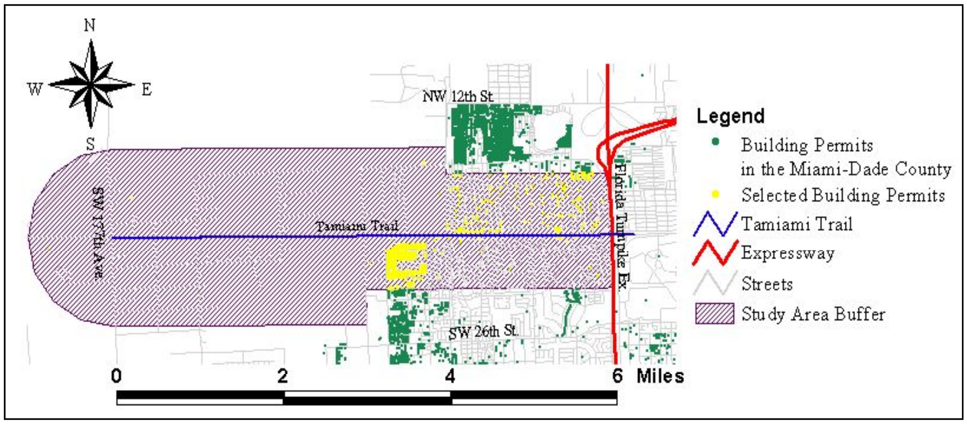


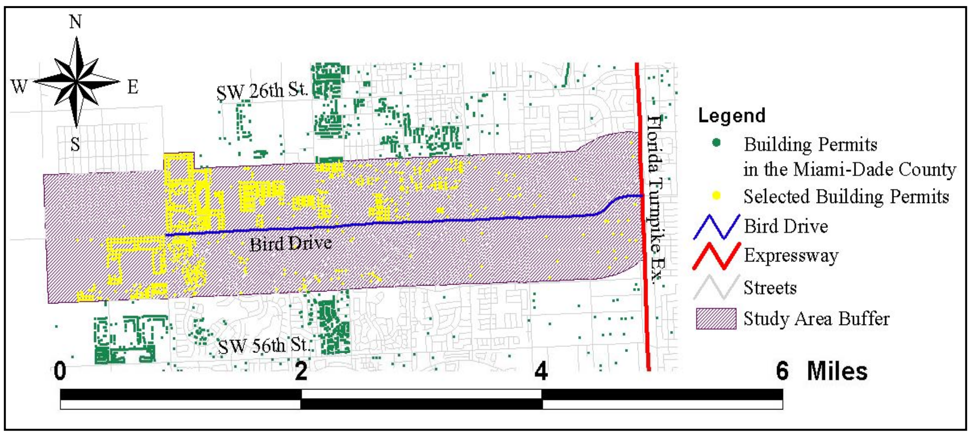
Figure 6.2 Location of Selected Study Corridors in the Miami-Dade County

Bird Drive is an east-west minor arterial in the southwest of Miami-Dade County. The section of interest was from SW 117th Avenue close to the Florida Turnpike, west to the end of the road, which becomes unpaved at SW 157th Avenue. The study area was defined as a ½-mile buffer area around Bird Drive bounded by the Florida Turnpike to the east and by SW 162nd Avenue to its west. Figure 6.3(b) shows the selected building permits within the ½-mile buffer around Bird Drive.

The North Kendall Drive corridor is also located in the southwest part of the county and is an east-west principal arterial with two other parallel arterials, Sunset Drive (SW 72nd Street) and Killian Drive (SW 104th Street), on either side at about one-mile spacing. The study area was defined as a ½-mile buffer area around North Kendall Drive bounded by the Florida Turnpike to the east and by SW 177th Avenue or Krome Avenue, which is the urban boundary, to its west. Beyond Krome Avenue lies the Everglades. Figure 6.3(c) shows the selected building permits within the ½-mile buffer around North Kendall Drive.



(a) Tamiami Trail



(b) Bird Drive

(c) North Kendall Drive

Figure 6.3 Spatial Distributions of Building Permits in the Study Corridors

In the original building permit data, each building permit record had a detailed proposed land use type. Land uses were simplified to five types: residential, commercial, warehouse, industrial, and other. Commercial land use included restaurants, retail stores, and gas stations. Industrial land use included industrial and office buildings. The "other" category included mostly government use and those that do not belong to other categories.

Figures 6.4 through 6.6 depict the temporal and spatial distribution of building permit applications using the simplified land use categories as snap shots between 1987 and 2001 for the Tamiami Trail and North Kendall Drive corridors and between 1988 and 2001 for the Bird Drive corridor. Colors indicate simplified land use categories. The yellow dots and red triangles represent residential and commercial land use, respectively. Annual building square feet of permit applications are summarized in Table 6.1 through 6.3 to provide a better understanding of the amount of development.

Between 1987 and 1991, new developments applied were located on the north side of Tamiami Trail from the Florida Turnpike to SW 137th Avenue and on the south side of Tamiami Trail between SW 137th Avenue and SW 147th Avenue. In 1991, residential developments started to increase and there was commercial development in 1992. The intensity of residential developments remained relatively unchanged from 1992 on until 2001 except in 1996 when permits for large residential developments were applied. Residential developments were followed by commercial developments, and most commercial developments were found on the south side of Tamiami Trail, as shown in the figure. There were industrial developments as well as developments of other land use categories.

Compared with other study areas, the study area around Bird Drive had steady commercial developments over the time period studied. From 1991, residential developments began to increase on the north side of Bird Drive and the east of SW 147th Avenue, and the south side of Bird Drive and the east of SW 157th Avenue. In 1993, residential developments continued to increase and expanded to the west of SW 147th Avenue. The block north of Bird Drive between SW 147th Avenue and SW 157th Avenue began to be developed for residential use in 1993 and large residential development occurred in that block in 1995. In 1996, residential developments decreased by half from the previous year, but started to grow again with a trend. Most of residential developments after 1996 took place in that particular block between SW 147th Avenue and SW 157th Avenue. In late 1995 and early 1996, large commercial developments occurred.

In the North Kendall Drive corridor, new developments scattered around SW 137th Avenue and SW 147th Avenue between 1987 and 1992. Since 1993, the area between SW 157th Avenue and SW 167th Avenue has been intensively developed and the dominant developments were single-family housing, with 91.4% of the total building permit in square-feet from 1987 to 2001. There were significant commercial and residential developments in the study area in 1994. Between 1994 and 2001 the intensity of residential development continued to decrease while an above the average number of permits for commercial developments was applied, with 1998 as an exception.

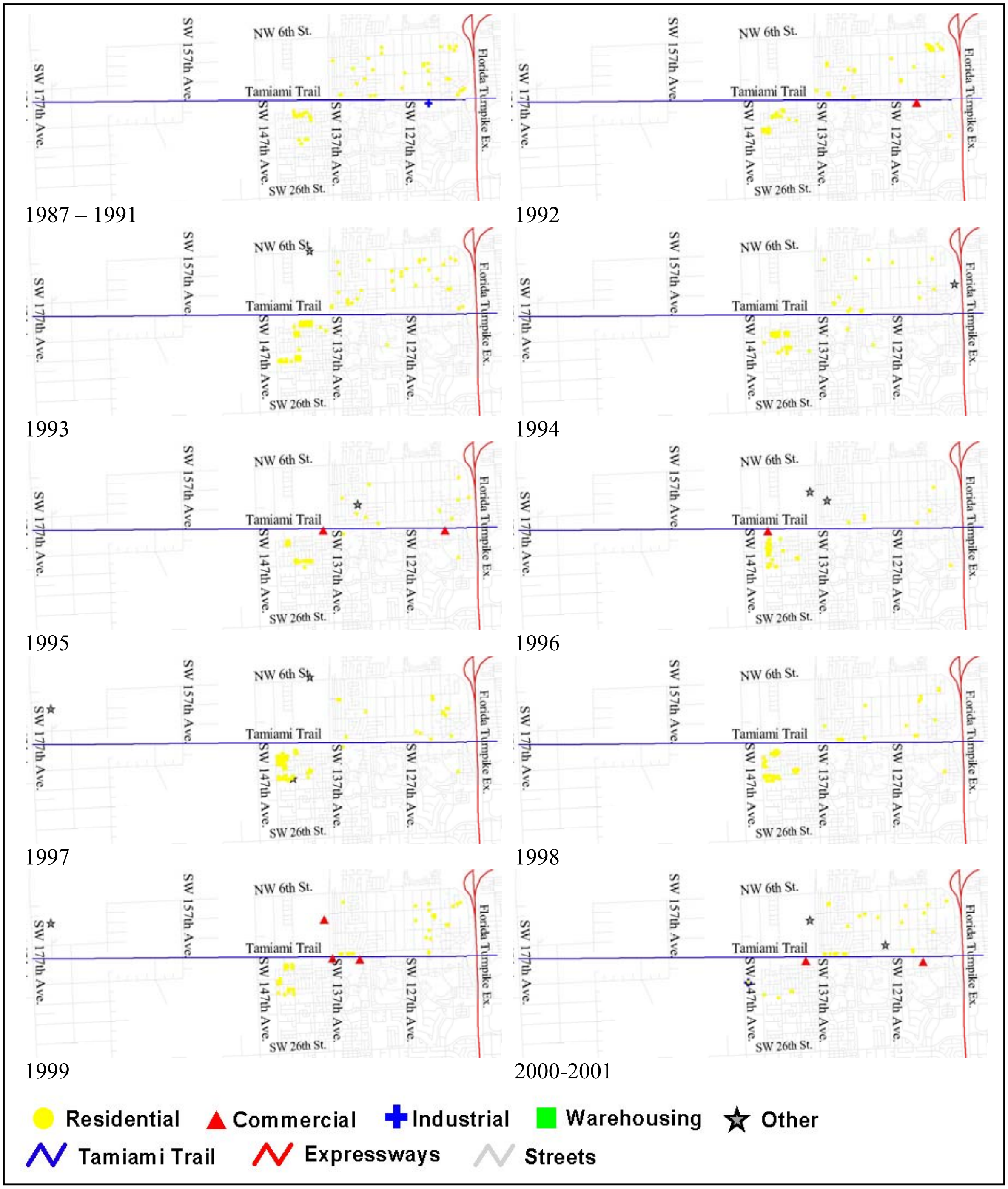


Figure 6.4 Temporal and Spatial Distribution of New Building Permit Applications in the Tamiami Trail Corridor

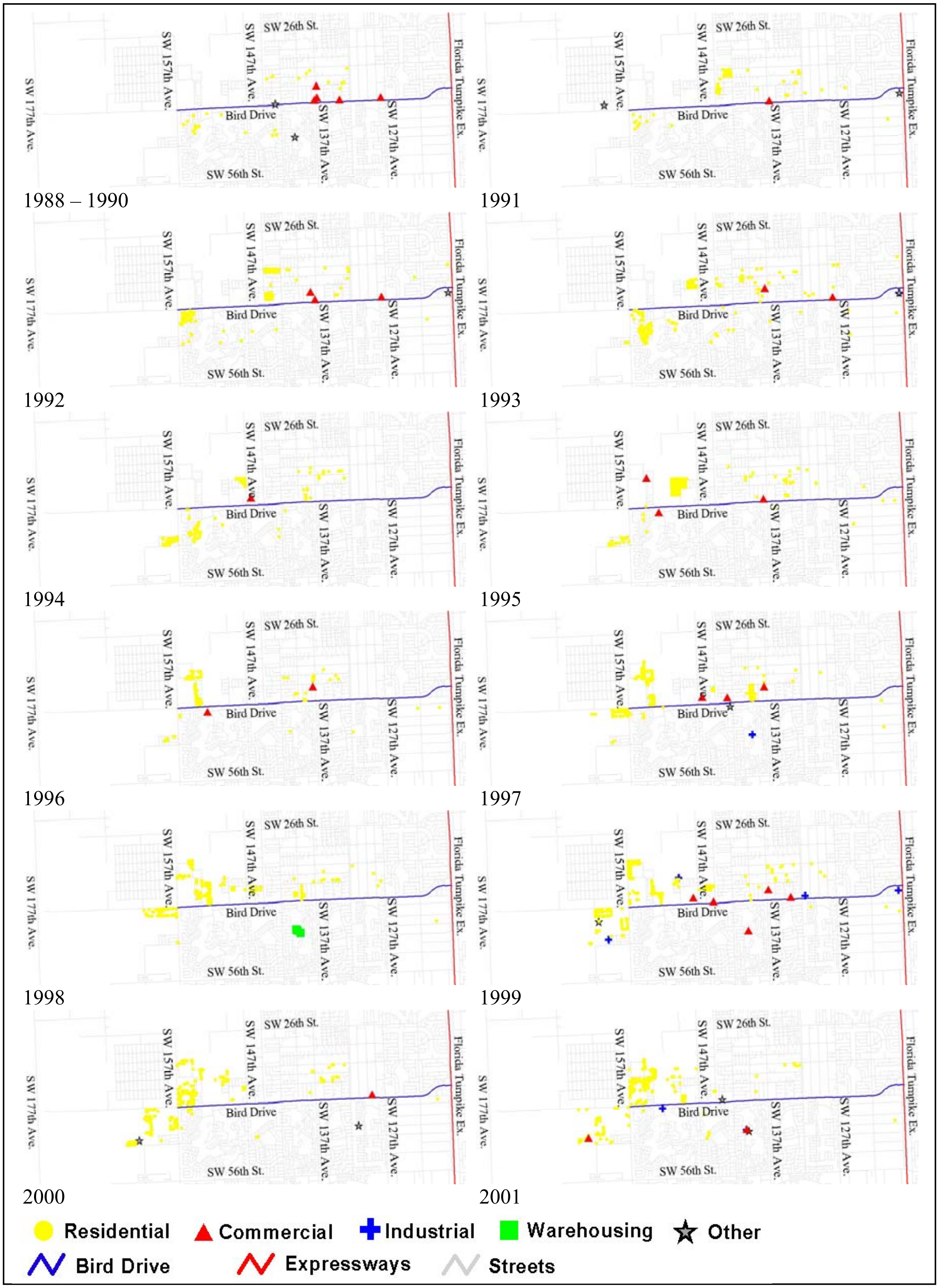


Figure 6.5 Temporal and Spatial Distribution of New Building Permit Applications in the Bird Drive Corridor

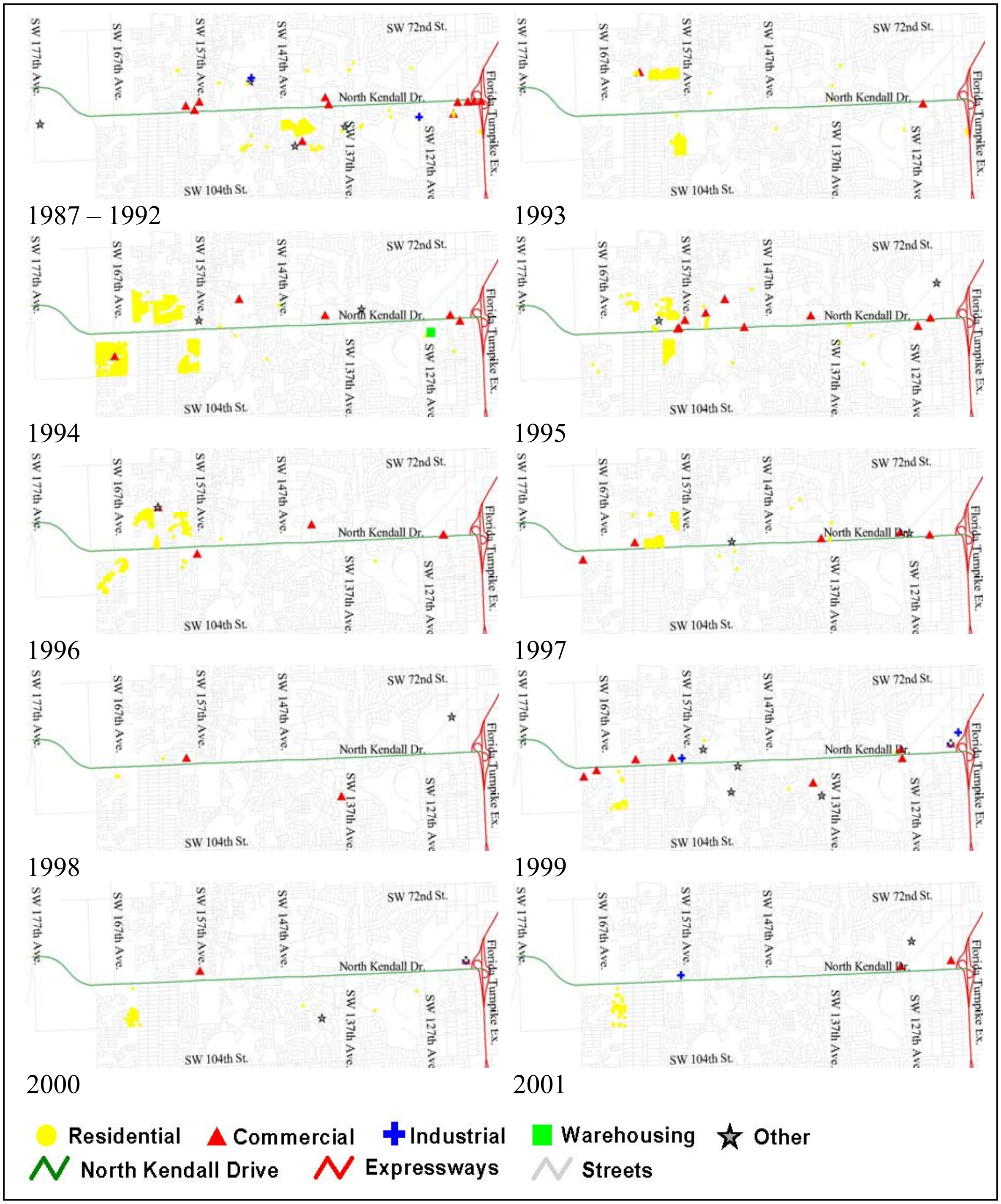


Figure 6.6 Temporal and Spatial Distribution of New Building Permit Applications in the North Kendall Drive Corridor

Table 6.1 Annual Building Square-Feet from Permit Application in the Tamiami Trail Corridor

YEAR	Residential (sq-ft)	Commercial (sq-ft)	Industrial (sq-ft)	Other (sq-ft)
1987	12,604	0	0	0
1988	11,959	0	0	0
1989	3,244	0	0	0
1990	40,542	0	7,157	0
1991	100,099	0	0	0
1992	179,406	12,057	0	0
1993	202,768	0	0	540
1994	163,273	0	0	6,324
1995	167,741	5,547	0	484
1996	1,089,998	2,200	0	425
1997	225,616	0	0	1,231
1998	215,446	0	0	0
1999	146,748	18,381	0	144
2000	35,231	100	0	450
2001	50,989	1	1027	0
Total	2,645,664	38,286	8,184	9,598
Percentage	97.92	1.42	0.30	0.36

Table 6.2 Annual Building Square-Feet from Permit Application in the Bird Drive Corridor

YEAR	Residential (sq-ft)	Commercial (sq-ft)	Industrial (sq-ft)	Warehousing (sq-ft)	Other (sq-ft)
1985	2,930	0	0	0	0
1987	36,081	0	0	0	0
1988	7,957	32,528	0	0	6,260
1989	28,311	4,500	0	0	98,674
1990	71,765	1,917	0	0	0
1991	209,064	485	0	0	560
1992	400,113	18,164	0	0	0
1993	459,587	12,786	132,000	0	19,800
1994	280,855	2,845	0	0	0
1995	515,572	91,133	0	0	0
1996	242,492	50,750	0	0	0
1997	524,208	8,110	6,084	0	1
1998	608,454	0	6,084	10376	0
1999	657,356	35,939	45,714	0	450
2000	781,954	825	0	0	965
2001	1,001,522	19,848	5,888	0	1,800
Total	5,828,221	279,830	195,770	10376	128,510
Percentage	90.46	4.34	3.04	0.16	1.99

Table 6.3 Annual Building Square-Feet from Permit Application in North Kendall Drive Corridor

YEAR	Residential (sq-ft)	Commercial (sq-ft)	Industrial (sq-ft)	Warehousing (sq-ft)	Other (sq-ft)
1987	14,744	3100	0	0	0
1988	62,778	16,030	0	0	6,260
1989	153,289	7,000	38,827	0	0
1990	244,531	19,176	0	0	5,992
1991	23,584	127,895	0	0	600
1992	361,971	15,000	0	0	0
1993	460,610	0	0	0	0
1994	2,069,492	331,044	0	19,050	360
1995	747,406	110,855	0	0	12,549
1996	536,126	73,229	0	0	160
1997	417,741	189,758	0	0	10,816
1998	12,068	1	0	0	10
1999	174,313	106,139	72,532	0	3,909
2000	192,375	31,411	1	0	3,681
2001	100,365	102,387	756	0	2,825
Total	5,571,393	1,133,025	112,116	19,050	47,162
Percentage	80.95	16.46	1.63	0.28	0.69

6.3 Specification of Model Variables

Using the simplified land use classification, two land use variables, *commercial development* and *residential development*, were developed. Commercial and residential developments were represented in terms of the sum of building square footage of applied building permits in a time unit, defined as month. Building square footage for commercial and residential developments was aggregated as below.

Commercial Developments (*dcom*):

$$dcom_t = \sum_{i=1}^{N} (BLDG\_COM)_{i,t} \qquad \text{[Eqn 6.1]}$$

where N was the number of building permits within a study area in month t and *BLDG_COM* was the building square footage for commercial use measured in thousand square feet.

Residential Developments (*dres*):

$$dres_t = \sum_{i=1}^{N} (BLDG\_RES)_{i,t} \qquad \text{[Eqn 6.2]}$$

where N was the number of building permits within a study area in month t and *BLDG_RES* was the building square footage for residential use measured in thousand square feet.

A third variable was the lane mile increase, an index to measure transportation improvements in terms of capacity. It was computed as the product of the number of lanes and length of the improved section. The first order difference of lane miles, $dlanemile_t$ (= $lanemile_t - lanemile_{t\text{-}1}$), was used as a variable in modeling.

From the geocoded TIP database, transportation improvement projects in the selected study corridors were retrieved and were verified with Florida Department of Transportation (FDOT) District VI and Miami-Dade County. There were one transportation improvement project for Tamiami Trail, three for Bird Drive, and two for North Kendall Drive between 1990 and 1995. They are summarized in Table 6.4 with the increases measured in lane miles.

Table 6.4 Summary of Transportation Improvement Projects for Three Study Corridors with Increase of Lane Miles

Corridor	Segment		Length (mile)	Opening Month	Lanes Added	Increase of Lane Miles
	From	To				
Tamiami Trail	SW 112th Ave	SW 127th Ave	1.52	04/1995	4	6.08
	SW 127th Ave	SW 137th Ave	1.00	–	0	–
	SW 137th Ave	SW 177th Ave	4.01	–	0	–
	Total		6.53			
Bird Drive	SW 117th Ave	SW 122nd Ave	0.56	02/1994	2	1.12
	SW 122nd Ave	SW 127th Ave	0.50	–	0	–
	SW 127th Ave	SW 142nd Ave	1.50	08/1991	2	3.00
	SW 142nd Ave	SW 147th Ave	0.50	03/1992	2	1.00
	SW 147th Ave	SW 157th Ave	1.00	–	0	–
	Total		4.06			
North Kendall Drive	Florida Turnpike	SW 132nd Ave	1.15	09/1993	4	4.60
	SW 132nd Ave	SW 152nd Ave	2.06	09/1993	2	4.12
	SW 152nd Ave	SW 177th Ave	2.46	–	0	–
	Total		5.67			

There were three transportation improvement projects in the Tamiami Trail corridor. One widened Tamiami Trail from four lanes to eight lanes between SW 112th Avenue and SW 127th Avenue. The other two projects added two more lanes to the existing four-lane road between SW 127th Avenue and SW 137th Avenue and between SW 137th Avenue and SW 152nd Avenue. The construction of the project for the section between SW 112th Avenue and SW 127th Avenue started in February 1991 and was completed in March 1995. The other two projects were not considered in this study since their constructions were either finished in 2002 or were to be completed in 2003, which were beyond the study horizon. The lane-miles were calculated based only for the first project and summarized in Table 6.4. The total lane-miles before and after the improvement project were 11.96 and 18.04, respectively.

There were four transportation improvement projects on Bird Drive in the study area. Three of them were considered in this study since the fourth project was completed in 1985, before the period of time considered in this study. The first project widened Bird Drive from two lanes to four lanes between SW 127th Avenue and SW 142nd Avenue. This project was completed in

August, 1991. The second project added two more lanes to the existing two-lane facility from SW 142nd Avenue to SW 147th Avenue, which was completed in March 1992. The third project also added two more lanes to the existing two-lane facility from SW 117th Avenue to SW 122nd Avenue. This project was completed in February 1994. The lane-miles were calculated for each project as shown in Table 6.4.

Two transportation improvement projects took place between 1990 and 1993 on North Kendall Drive: one widening it from four lanes to eight lanes between the Florida Turnpike and SW 132nd Avenue, and the other adding two more lanes to the existing four-lane road between SW 132nd Avenue and SW 152nd Avenue. The construction of both projects started on October 29, 1990 and was completed on August 31, 1993. The lane-miles are given in Table 6.4 with the total lane-miles before and after the improvement projects being 22.68 and 31.40, respectively.

Figures 6.7, 6.9, and 6.11 provide the computed lane miles for the study corridors along with the monthly variation of the applied building permits in million square feet for commercial and residential land use. In Figures 6.8, 6.10, and 6.12, the aggregated semi-annual variation of the applied building permits in million square feet for commercial and residential land use and the computed lane miles was depicted to show the trend of development growth more clearly. In Figures 6.9 and 6.10, the dashed line indicates a trend in the data, which needs to be treated in the time series analysis. The descriptive statistics of selected variables to investigate the interactions between land use and transportation for this study area are summarized in Table 6.5.

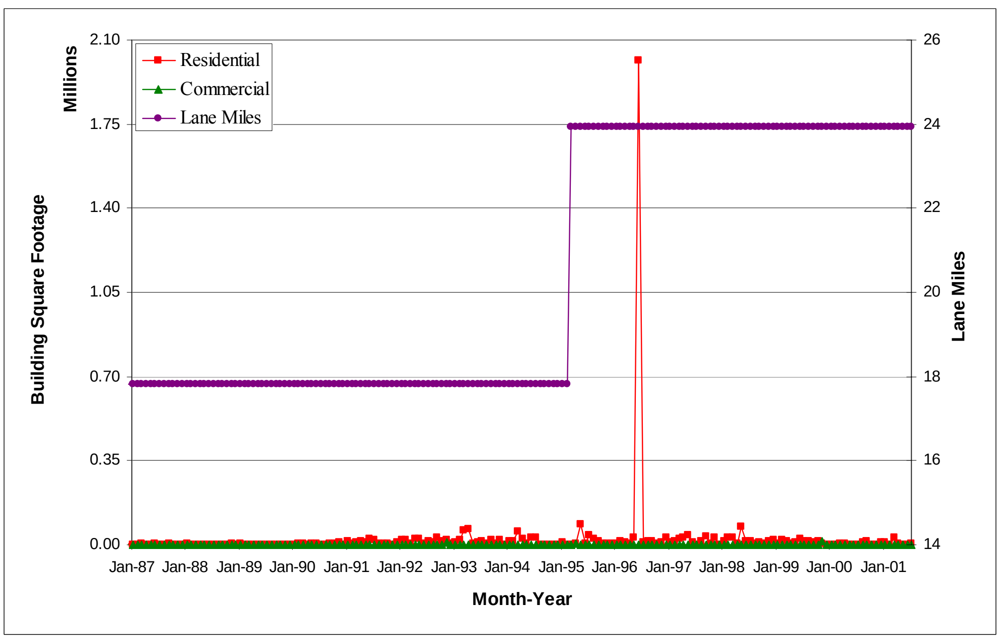


Figure 6.7 Monthly Building Square Footage for Residential and Commercial Use with Lane Miles in the Tamiami Trail Corridor

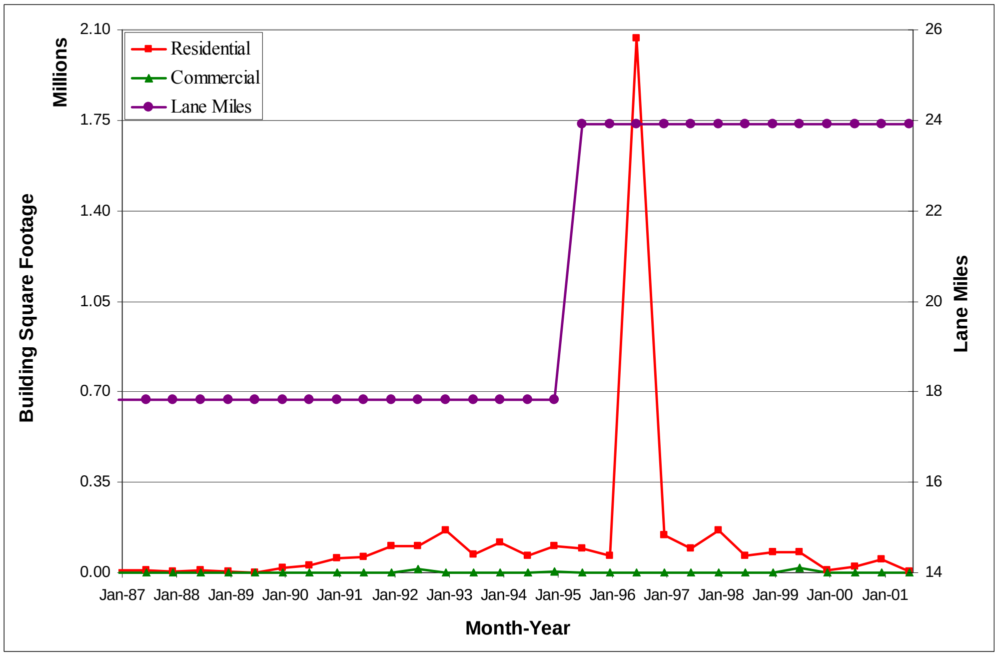


Figure 6.8 Semi-Annual Building Square Footage for Residential and Commercial Use with Lane Miles in the Tamiami Trail Corridor

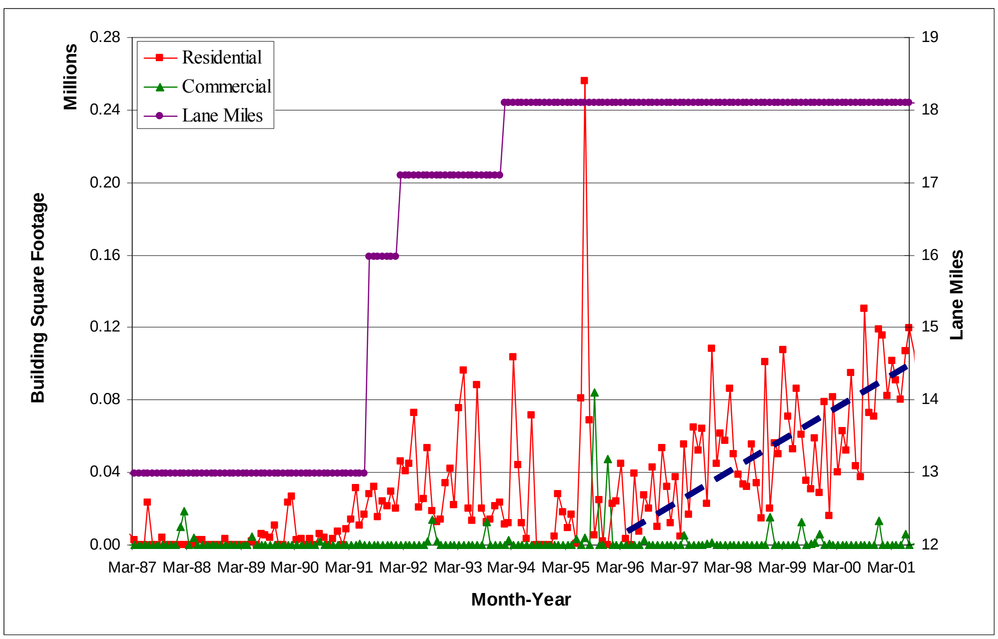


Figure 6.9 Monthly Building Square Footage for Residential and Commercial Use with Lane Miles in the Bird Drive Corridor

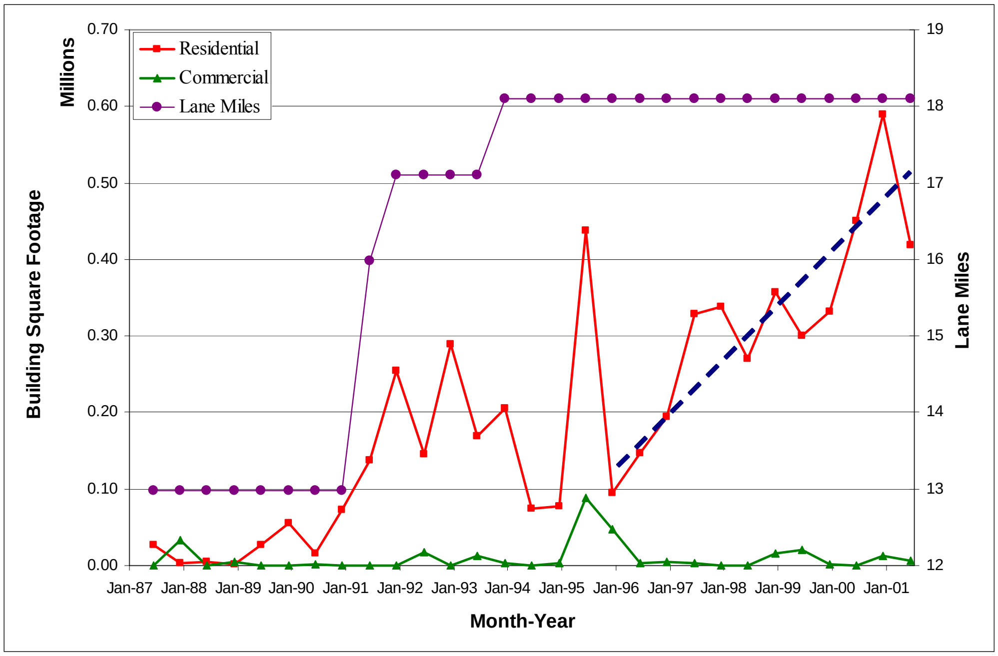


Figure 6.10 Semi-Annual Building Square Footage for Residential and Commercial Use with Lane Miles in the Bird Drive Corridor

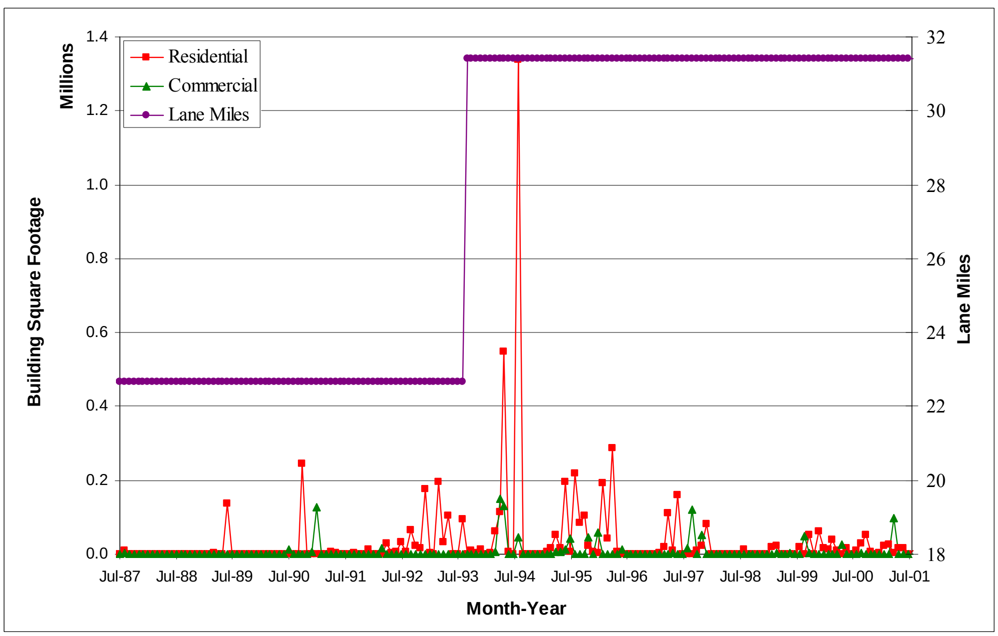


Figure 6.11 Monthly Building Square Footage for Residential and Commercial Use with Lane Miles in the North Kendall Drive Corridor

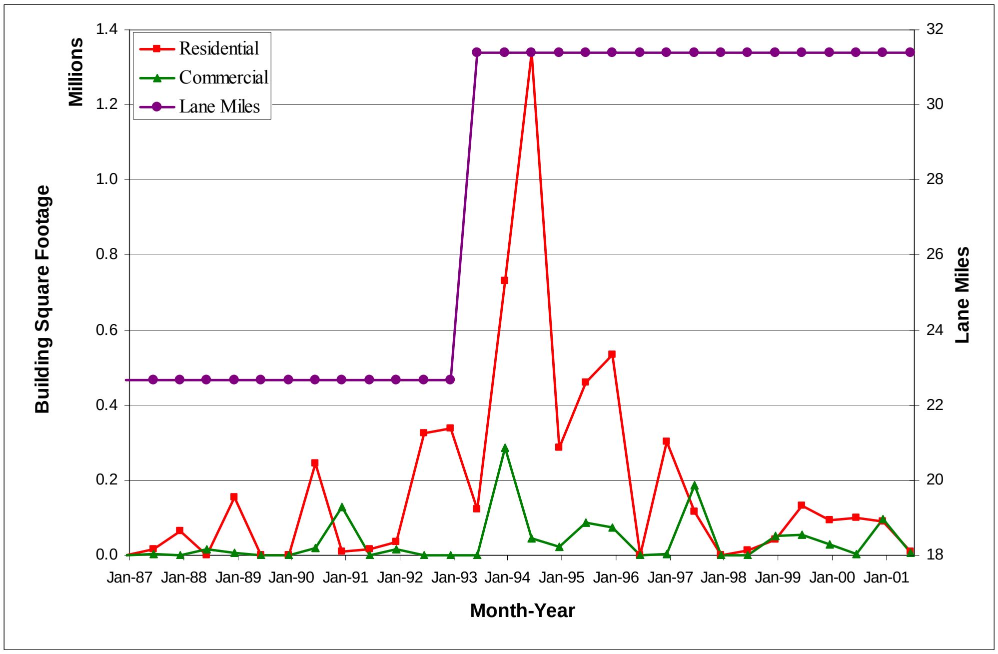


Figure 6.12 Semi-Annual Building Square Footage for Residential and Commercial Use with Lane Miles in the North Kendall Drive Corridor

Table 6.5 Descriptive Statistics of Variables of Interest for the Study Corridors

Corridor	Statistics	Lane Mile	Residential	Commercial
Tamiami Trail	Number of Observations	176	176	176
	Mean	20.50	21.98	0.22
	Median	17.83	5.90	0.00
	Maximum	23.92	2,012.41	16.18
	Minimum	17.83	0.00	0.00
	Standard Deviation	3.03	151.52	1.56
Bird Drive	Number of Observations	177	177	177
	Mean	16.35	32.95	1.58
	Median	18.10	21.20	0.00
	Maximum	18.10	256.02	83.94
	Minimum	12.98	0.00	0.00
	Standard Deviation	2.27	36.98	7.71
North Kendall Drive	Number of Observations	169	169	169
	Mean	27.66	32.28	6.26
	Median	31.40	0.00	0.00
	Maximum	31.40	1337.10	150.53
	Minimum	22.80	0.00	0.00
	Std. Dev.	4.28	119.42	23.09

Three model structures were applied in the case studies: the univariate model, the VAR model, and the VARX model. These models were to describe different relationships between the transportation and land use variables. The structures of these models, their estimations, and the questions they answer are described in the follwing sections.

6.4 Univariate Model Estimations

The univariate model includes one endogenous variable and one exogenous variable, which were the total building square feet from the permit applications in a study area and the transportation improvement index, respectively. It was designed to answer the following questions:

(1) Is there a causal relationship between transportation improvement and land developments?
(2) How quickly did land development respond to improvement in accessibility? How far behind did land development lag transportation improvements?
(3) How strongly did land use respond to transportation improvements?

In this section, the model estimations are first presented, followed by descriptions of the results from the diagnostic tests and

6.4.1 Dynamic Transfer Function

To investigate the temporal interaction between transportation improvement programs and land use changes, we considered the transportation improvement index, *dlanemile* (= $lanemile_t - lanemile_{t\text{-}1}$), as the exogenous series and the total building square footage, *dtotal*, as the endogenous response series as described in Chapter 5.

Parameters of the dynamic systems were estimated by the least square method. They are summarized in Table 6.6 through Table 6.8. The models for the study areas were given below:

Tamiami Trail

$$\begin{aligned} dtotal(t) = {} & 11.4328 \times dlanemile(t\text{-}2) + 327.5704 \times dlanemile(t\text{-}15) \\ & + 0.254607 \times AR(1) + 0.130499 \times MA(4) + 0.260083 \times MA(12) + 11.36979 \end{aligned}$$

Bird Drive

$$\begin{aligned} dtotal(t) = {} & 0.196534 \times dtotal(t\text{-}1) + 0.173770 \times dtotal(t\text{-}15) + 0.280335 \times dtotal(t\text{-}25) \\ & + 38.56848 \times dlanemile(t\text{-}23) + 10.89416 + 0.137867 \times \text{dummy@trend} \end{aligned}$$

North Kendall Drive

$$\begin{aligned} dtotal(t) = {} & 0.082 \times dtotal(t\text{-}10) + 0.085 \times dtotal(t\text{-}12) + 29.958 \times dlanemile(t\text{-}7) \\ & + 77.195 \times dlanemile(t\text{-}8) + 158.94 \times dlanemile(t\text{-}11) + 0.256 \times MA(2) + \\ & 0.242 \times MA(3) + 21.39 \end{aligned}$$

where *AR* stood for autoregression and *MA* denoted moving average. There was a trend in the development in the Bird Drive corridor as shown in Figure 6.9 as a dashed line, which was accounted for by introducing a dummy variable, dummy@trend.

For the Tamiami Trail model, coefficients corresponding to *dlanemile* with two and 15 lags were found to be significant. Coefficient for *dlanemile* with two lags, *dlanemile*(*t*-2), was 11.4328 and that for *dlanemile* with 15 lags, *dlanemile*(*t*-15), was 327.5704. This implied that there were large numbers of developments for which permits were applied 15 months later after the construction of adding four lanes to the existing facilities was completed. R-square value was found to be 0.994406. The sum of coefficients of *dlanemile*(*t*-2) and *dlanemile*(*t*-15) was 339.0032, which meant that cumulated impact from unit increase in lane miles would be 339,003 square feet of new buildings in this study area. In other words, if there was one lane added to a one-mile section, it would bring building permit applications totaling 339,003.2 square feet. All coefficients have positive sign as expected. The estimated parameters of the dynamic system are summarized in Table 6.6, and Figure 6.13 shows the plotted *actual dtotal* and *fitted dtotal* with residuals. Performance of the model seemed to be very good. However, diagnostic tests on the estimated model were required.

Table 6.6 Parameter Estimation by Least Squares for the Tamiami Trail Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.36979	1.759385	6.462367	0.0000
DLANEMILE(-2)	11.43280	1.837568	6.221700	0.0000
DLANEMILE(-15)	327.5704	1.866886	175.4635	0.0000
AR(1)	0.254607	0.078757	3.232819	0.0015
MA(4)	0.130499	0.077767	1.678081	0.0954
MA(12)	0.260083	0.080470	3.232058	0.0015
R-squared	0.994406	Mean dependent var		24.47809
Adjusted R-squared	0.994224	S.D. dependent var		158.7931
S.E. of regression	12.06830	Akaike info criterion		7.855820
Sum squared resid	22429.15	Schwarz criterion		7.971139
Log likelihood	-622.4656	F-statistic		5474.717
Durbin-Watson stat	2.029542	Prob(F-statistic)		0.000000

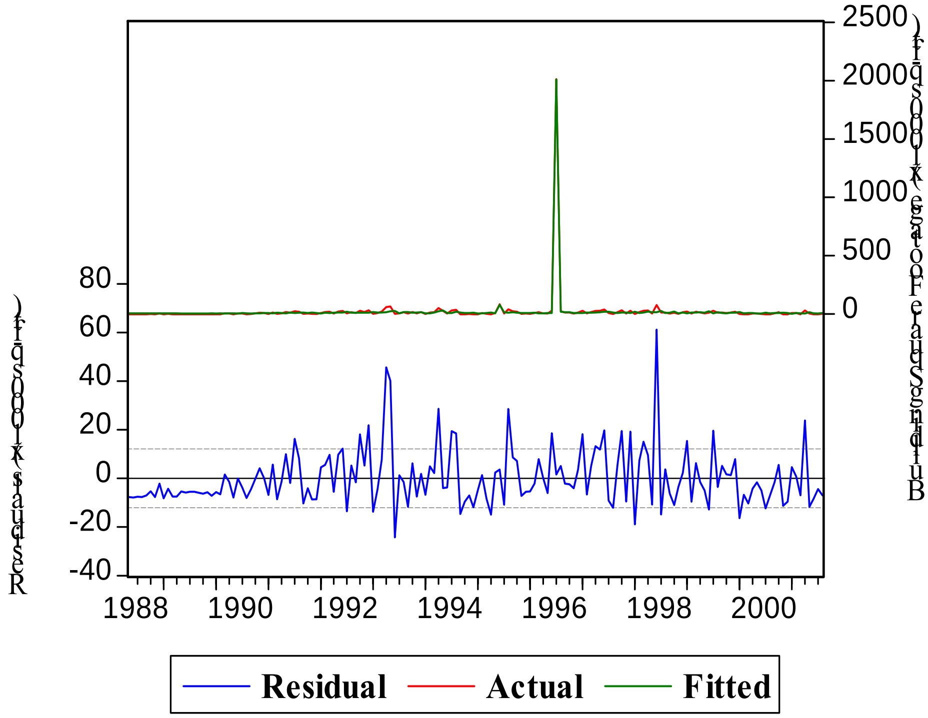


Figure 6.13 Plotted Actual *dtotal* and Fitted *dtotal* with Residuals for the Tamiami Trail Model

For the Bird Drive model, variables found to be significant included *dtotal*(*t*-1), *dtotal*(*t*-15), *dtotal*(*t*-25), *dlanemile*(*t*-23), a constant, and dummy variable. Past land use developments induced impacts with one, 15, and 25 lags while roadway improvement caused land use developments with a 23-month lag. The extent of impact from roadway improvement was much larger than the one from past land developments, which could be found from their coefficients. A dummy variable was added into the system to reflect the trend of increase from June 1997. R-squared value was 0.41. The coefficient of *dlanemile*(*t*-23) was 38.56848 and this may be interpreted as the extent of impacts on land use development from roadway improvement. For instance, a unit increase in lane miles would result in 38,685 square feet of buildings within 23 month. The estimated parameters of the dynamic system are summarized in Table 6.7, and Figure 6.14 shows the *actual dtotal* in red and *fitted dtotal* in green with residual plot at the bottom. Notice that there is a spike in Figure 6.14 between 1995 and 1996, which indicated that land development was not well captured in the model for that short period. The residuals as shown in Figure 6.14 are reasonable otherwise,.

Table 6.7 Parameter Estimation by Least Squares for the Bird Drive Model

Dependent Variable: DTOTAL				
Method: Least Squares				
Sample(adjusted): 1989:04 2001:11				
Included observations: 152 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
DTOTAL(-1)	0.196534	0.073306	2.681019	0.0082
DTOTAL(-15)	0.173770	0.073408	2.367189	0.0192
DTOTAL(-25)	0.280335	0.077372	3.623206	0.0004
DLANEMILE(-23)	38.56848	9.390324	4.107258	0.0001
C	10.89416	4.313974	2.525318	0.0126
DUMMY1*@TREND	0.137867	0.043530	3.167190	0.0019
R-squared	0.410730	Mean dependent var		41.88149
Adjusted R-squared	0.390550	S.D. dependent var		39.87206
S.E. of regression	31.12703	Akaike info criterion		9.752704
Sum squared resid	141458.3	Schwarz criterion		9.872068
Log likelihood	-735.2055	F-statistic		20.35287
Durbin-Watson stat	1.943767	Prob(F-statistic)		0.000000

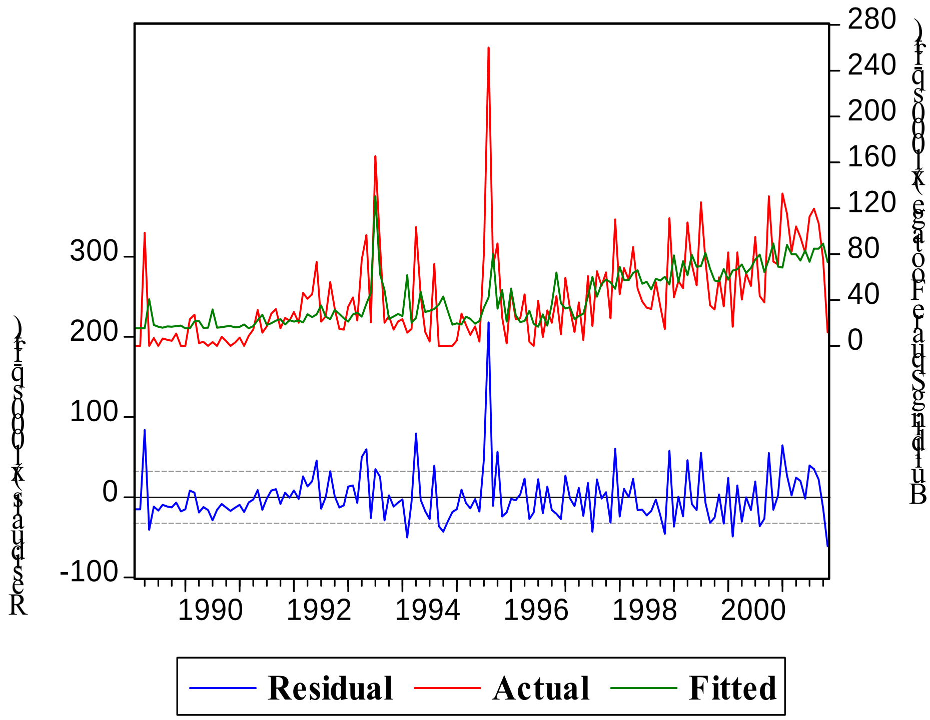


Figure 6.14 Plotted Actual *dtotal* and Fitted *dtotal* with Residuals for the Bird Drive Model

For the North Kendall Drive model, coefficients corresponding to *dlanemile* with seven, eight, and 11 lags were larger than the other coefficients. *Dlanemile* with 11 lags, *dlanemile*(t-11), had the largest coefficient, 158.94, which means there was a large amount of development for which permits were applied after the construction widening North Kendall Drive was completed. Additionally, *dlanemile*(*t*-7), *dlanemile*(*t*-8), and *dlanemile*(*t*-11) were also found to be significant in the model. The endogenous variable was well explained (about 80%) by the exogenous series and past values of endogenous variable, as can be seen from the R-square value. The sum of coefficients of *dlanemile*(*t*-7), *dlanemile*(*t*-8), and *dlanemile*(*t*-11), which was 266.0958, reflected the cumulated impacts (long-run effect) from roadway improvements. This meant that cumulated impact from a unit increase in lanemile would be 266,096 square feet of buildings. In other words, if there is one lane added to one mile section, it will bring building permit applications for 266,096 square feet. All coefficients have positive sign as we expected. Parameters of dynamic system estimated by the least square method are summarized in Table 6.8. Figure 6.15 shows the plotted actual *dtotal* and fitted *dtotal* with residuals. Performance of model seems to be very good. However, it was required to perform diagnostic tests on the estimated model above. Table 6.6 compared the results from the three univariate models.

Table 6.8 Parameter Estimation by Least Squares for the North Kendall Drive Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	21.39101	6.052073	3.534493	0.0005
DTOTAL(-10)	0.082333	0.031044	2.652128	0.0089
DTOTAL(-12)	0.085794	0.030822	2.783500	0.0061
DLANEMILE(-7)	29.95837	5.200177	5.761029	0.0000
DLANEMILE(-8)	77.19573	5.379442	14.35014	0.0000
DLANEMILE(-11)	158.9417	5.376841	29.56043	0.0000
MA(2)	0.256639	0.077386	3.316335	0.0011
MA(3)	0.242332	0.078432	3.089698	0.0024
R-squared	0.875891	Mean dependent var		43.32563
Adjusted R-squared	0.870061	S.D. dependent var		132.2363
S.E. of regression	47.66732	Akaike info criterion		10.61598
Sum squared resid.	338553.9	Schwarz criterion		10.77171
Log likelihood	-825.3545	F-statistic		150.2230
Durbin-Watson stat	1.889375	Prob(F-statistic)		0.000000

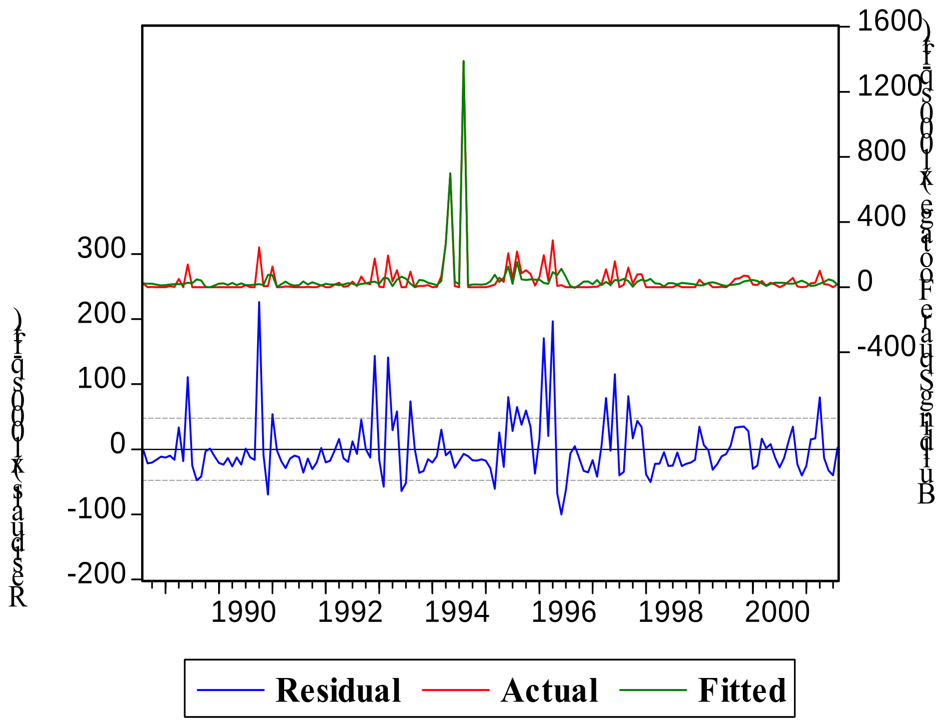


Figure 6.15 Plotted Actual *dtota*l and Fitted *dtotal* with Residuals for the North Kendall Drive Model

Table 6.9 Summary of Univariate Models

Model Statistics	Model		
	Tamiami Trail	Bird Road	North Kendall Drive
R^2	0.9	0.41	0.8
Lags significant (months) for lane-miles	2, 15	23	7, 8, 11
Lags significant (months) for total development	-	1, 15, 25	10, 12
Cumulative impacts (sq-feet)	339,003	110,397	319,439

6.4.2 Diagnostic Tests

As discussed in Chapter 5, three diagnostic tests were conducted and they included Q-statistics, ARCH test, and Ramsey RESET test. Q-statistics tests if there is serial correlation in the univariate models estimated. For the Tamiami Trail model, as shown in Figure 6.16, there appeared to be no serial correlation in the model. During the modeling process, there was serial correlation found in the model only with *dlanemile*(*t*-2), *dlanemile*(*t*-15), and the constant. However, this problem was cured by adding the *AR* and *MA* terms. Ramsey's RESET test was also conducted to examine whether a linear specification was appropriate. As presented in Table 6.10, the p-value was found to be 0.142 therefore the null hypothesis that no relevant explanatory variables had been omitted from the regression equation cannot be rejected. However, the result from the ARCH test showed that the estimated model had conditional heteroscedasticity as shown also in Table 6.10, which could be also found in the diagram in Figure 6.16. Residuals before 16 lags were mostly positive and after 16 lags were negative. Since GARCH can provide efficient estimation in the presence of conditional heteroscedasticity, it was required to re-estimate the model by GARCH.

Figure 6.17 shows autocorrelation with Q-statistics for the estimated Bird Drive model. The null hypothesis was that there was no autocorrelation up to order k in the estimated model. Order k, which could be any number, was 36. The p-values, which is found in the last column in Figure 6.17, were too large to reject the null hypothesis. There was, therefore, no autocorrelation up to order 36 according to the Q-statistics. Ramsey's RESET test was conducted to verify that no relevant explanatory variables had been omitted from the estimated model. Test results, which are summarized in Table 6.10, showed that the p-value was too large to reject the null hypothesis, i.e., the estimated model did not omit any relevant explanatory variables. Finally, ARCH test was conducted to check whether there was conditional heteroscedasticity in the estimated model. The ARCH test result is shown in Table 6.10. It was found that there was no conditional heteroscedasticity in the estimated model according to the test statistics.

For the North Kendall Drive model, as shown in Figure 6.18, there appeared to be no serial correlation in the model. The Ramsey's RESET test was also conducted to examine whether a linear specification was appropriate, which was confirmed by the test results, as shown Table 6.10. However, the results from the ARCH test indicated that there was conditional heteroscedasticity in the model (see Table 6.10). Therefore, the model needed to be re-estimated using GARCH.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.017	-0.017	0.0488	
		2	0.041	0.041	0.3307	
		3	0.025	0.027	0.4354	
		4	0.027	0.026	0.5575	0.455
		5	-0.004	-0.005	0.5596	0.756
		6	0.164	0.161	5.0707	0.167
		7	0.022	0.027	5.1491	0.272
		8	-0.027	-0.040	5.2694	0.384
		9	0.062	0.052	5.9188	0.432
		10	0.057	0.055	6.4863	0.484
		11	0.077	0.079	7.5186	0.482
		12	0.017	-0.011	7.5721	0.578
		13	0.099	0.084	9.2840	0.505
		14	0.070	0.085	10.153	0.517
		15	0.148	0.135	14.059	0.297
		16	0.128	0.118	16.996	0.199
		17	-0.012	-0.038	17.022	0.255
		18	-0.025	-0.040	17.132	0.311
		19	-0.047	-0.086	17.541	0.351
		20	-0.016	-0.060	17.585	0.415
		21	0.051	0.007	18.080	0.450
		22	0.059	0.014	18.727	0.475
		23	-0.096	-0.105	20.488	0.428
		24	0.024	-0.007	20.594	0.484
		25	-0.095	-0.123	22.321	0.441
		26	-0.003	-0.049	22.322	0.501
		27	-0.016	-0.064	22.372	0.557
		28	-0.010	-0.061	22.392	0.613
		29	-0.033	-0.040	22.610	0.655
		30	-0.013	-0.050	22.644	0.704
		31	-0.117	-0.141	25.392	0.606
		32	-0.009	-0.012	25.407	0.657
		33	-0.068	-0.040	26.357	0.657
		34	0.177	0.264	32.792	0.379
		35	-0.090	-0.030	34.489	0.350
		36	-0.066	-0.043	35.393	0.356

Figure 6.16 Autocorrelation Plot with Q-statistics for the Tamiami Trail Model

Table 6.10 Summary of Diagonostic Test Results

Model	Ramsey	ARCH
Tamiami Trail	0.142099	0.024741
Bird Drive	0.142813	0.707147
North Kendall Drive	0.274012	0.010206

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.009	-0.009	0.0116	0.914
		2	0.044	0.044	0.3112	0.856
		3	0.055	0.056	0.7821	0.854
		4	0.075	0.075	1.6773	0.795
		5	-0.067	-0.070	2.3846	0.794
		6	0.044	0.034	2.7005	0.845
		7	-0.021	-0.023	2.7694	0.905
		8	0.036	0.036	2.9853	0.935
		9	-0.113	-0.107	5.0653	0.829
		10	-0.177	-0.194	10.233	0.420
		11	0.093	0.108	11.672	0.389
		12	0.052	0.080	12.133	0.435
		13	-0.123	-0.096	14.689	0.327
		14	-0.048	-0.069	15.084	0.372
		15	0.007	-0.017	15.092	0.445
		16	0.056	0.107	15.632	0.479
		17	-0.102	-0.087	17.421	0.426
		18	0.015	-0.016	17.461	0.492
		19	-0.059	-0.107	18.067	0.518
		20	-0.043	-0.056	18.388	0.562
		21	-0.112	-0.021	20.621	0.482
		22	-0.002	-0.024	20.622	0.544
		23	-0.075	-0.122	21.650	0.541
		24	0.058	0.054	22.260	0.564
		25	-0.127	-0.061	25.214	0.450
		26	-0.080	-0.094	26.418	0.440
		27	0.094	0.061	28.081	0.407
		28	0.016	0.013	28.128	0.458
		29	0.060	0.089	28.815	0.475
		30	-0.026	-0.100	28.948	0.520
		31	-0.027	-0.077	29.086	0.565
		32	-0.021	-0.034	29.176	0.610
		33	-0.010	-0.025	29.196	0.657
		34	-0.016	0.004	29.249	0.700
		35	-0.008	-0.105	29.261	0.741
		36	0.015	-0.015	29.305	0.778

Figure 6.17 Autocorrelation Plot with Q-statistics for the Bird Drive Model

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.055	0.055	0.4894	
		2	-0.006	-0.009	0.4959	
		3	0.035	0.036	0.6988	0.403
		4	-0.092	-0.096	2.0733	0.355
		5	0.109	0.122	4.0111	0.260
		6	0.050	0.032	4.4188	0.352
		7	-0.039	-0.035	4.6733	0.457
		8	0.029	0.017	4.8144	0.568
		9	-0.144	-0.133	8.2998	0.307
		10	-0.143	-0.132	11.791	0.161
		11	-0.140	-0.155	15.145	0.087
		12	-0.078	-0.051	16.182	0.095
		13	-0.138	-0.171	19.501	0.053
		14	0.019	0.039	19.561	0.076
		15	-0.074	-0.074	20.534	0.083
		16	0.050	0.100	20.973	0.102
		17	-0.021	-0.041	21.051	0.135
		18	-0.052	-0.011	21.533	0.159
		19	0.058	0.017	22.136	0.180
		20	0.062	0.024	22.837	0.197
		21	0.003	-0.060	22.839	0.244
		22	-0.017	-0.122	22.890	0.294
		23	0.041	0.021	23.197	0.334
		24	-0.062	-0.166	23.915	0.352
		25	0.016	0.008	23.966	0.406
		26	0.123	0.075	26.866	0.311
		27	-0.029	0.005	27.026	0.355
		28	-0.085	-0.144	28.419	0.338
		29	0.099	0.196	30.322	0.300
		30	0.092	0.119	32.004	0.274
		31	0.002	-0.046	32.005	0.320
		32	-0.088	-0.134	33.567	0.298
		33	-0.078	-0.062	34.785	0.292
		34	0.047	0.014	35.230	0.318
		35	0.162	0.103	40.576	0.171
		36	-0.055	-0.068	41.199	0.185

Figure 6.18 Autocorrelation Plot with Q-statistics for the North Kendall Drive Model

6.4.3 Efficient Estimation by GARCH

Because of the presence of conditional heteroscedasticity in the Tamiami Trail and North Kendall Drive models, these models were re-estimated using GARCH. Table 6.11 presents a summary of the GARCH estimation of the Tamiami Trail model. There were slight changes in the model estimation in terms of model performance measures such as R-square, standard errors of regression, and sum of squared residuals. Coefficients had also changed. For instance, coefficients of *dlanemile*(*t*-2) and *dlanemile*(*t*-15) were changed from 11.4328 and 327.5704 to 9.846803 and 327.8634, respectively. However, they still remained significant variables.

Table 6.12 presents a summary of the GARCH estimation of the North Kendall Drive model. There were slight differences in the models estimation by LS and GARCH in terms of model performance measures. Coefficients of *dlanemile*(-7), *dlanemile*(-8), and *dlanemile*(-11) changed from 29.95837, 77.19573, and 158.9417 to 30.38312, 75.58657, and 159.1483 respectively, but still remained as significant variables.

Table 6.11 Parameter Estimation by GARCH for the Tamiami Trail Model

	Coefficient	Std. Error	z-Statistic	Prob.
C	2.168639	1.190409	1.821760	0.0685
DLANEMILE(-2)	9.846803	1.101416	8.940133	0.0000
DLANEMILE(-15)	327.8634	3.890717	84.26810	0.0000
AR(1)	0.430777	0.104361	4.127743	0.0000
MA(4)	0.096855	0.077138	1.255598	0.2093
MA(12)	0.441659	0.059067	7.477308	0.0000
	Variance Equation			
C	0.413077	0.388232	1.063996	0.2873
ARCH(1)	0.227074	0.076715	2.959964	0.0031
GARCH(1)	0.830839	0.054223	15.32249	0.0000
R-squared	0.993396	Mean dependent var		24.47809
Adjusted R-squared	0.993046	S.D. dependent var		158.7931
S.E. of regression	13.24202	Akaike info criterion		7.587729
Sum squared resid	26478.03	Schwarz criterion		7.760708
Log likelihood	-598.0184	F-statistic		2839.121
Durbin-Watson stat	2.277254	Prob(F-statistic)		0.000000

Table 6.12 Parameter Estimation by GARCH for the North Kendall Drive Model

	Coefficient	Std. Error	z-Statistic	Prob.
C	20.37742	8.009264	2.544232	0.0110
DTOTAL(-10)	0.070770	0.017071	4.145583	0.0000
DTOTAL(-12)	0.082993	0.024828	3.342728	0.0008
DLANEMILE(-7)	30.38312	12.98859	2.339217	0.0193
DLANEMILE(-8)	75.58657	22.12256	3.416719	0.0006
DLANEMILE(-11)	159.1483	18.53391	8.586867	0.0000
MA(2)	0.267779	0.078087	3.429250	0.0006
MA(3)	0.340186	0.050106	6.789383	0.0000
	Variance Equation			
C	1181.819	317.7913	3.718854	0.0002
ARCH(1)	0.097781	0.063074	1.550239	0.1211
GARCH(1)	0.918528	0.353941	2.595146	0.0095
GARCH(2)	-0.566883	0.297382	-1.906243	0.0566
R-squared	0.873992	Mean dependent var		43.32563
Adjusted R-squared	0.864432	S.D. dependent var		132.2363
S.E. of regression	48.68873	Akaike info criterion		10.56334
Sum squared resid	343735.9	Schwarz criterion		10.79694
Log likelihood	-817.2226	F-statistic		91.42889
Durbin-Watson stat	1.881292	Prob(F-statistic)		0.000000

6.4.4 Summary of the Univariate Models

Table 6.13 summarizes the statistics of the three univariate models. According to the univariate models, there appeared to be a causal relationship between transportation improvement and land developments and transportation improvements appeared to have encouraged land developments. The land development lagged behind transportation improvements between two to 23 months, and there was a wide range of the lags depending on the corridors. The response in land use to transportation was also stronger in the Tamiami Trail and North Kendall Drive corridors than in the Bird Drive corridor as indicated by the cumulative impacts shown in Table 6.13. For instance, the models predicted over 300,000 square feet of new building permits for every one-mile lane added to the transportation system. Although the Bird Drive and North Kendall Drive models indicated that previous developments also affected the current developments (row 7 of Table 6.13), their effects were much smaller than those from transportation improvements as the the coefficients of the variables representing previous land developments were hundred or thousand times smaller than those of variables representing transportation improvements.

Table 6.13 Summary of the Final Univariate Models

Model Statistics	Model		
	Tamiami Trail	Bird Road	North Kendall Drive
Estimation Method	GARCH	LS	GARCH
R^2	0.9	0.41	0.8
SER	13.24202	31.12703	48.68873
SSR	26478.03	141458.3	343735.9
Lags significant (months) for lane-miles	2, 15	23	7, 8, 11
Lags significant (months) for total development	-	1, 15, 25	10, 12
Cumulative impacts of lane miles (sq-feet)	339,003	110,397	319,439

6.5 VARX Model

Recall from Chapter 5 that a multivariate VARX model includes one or more endogenous and one or more exogenous variables. In this study, the endogenous variables were the total residential and commercial building square feet from the permit applications in a study area, and the exogenous variable was the transportation improvement index. The VARX models in this study were designed to answer the following questions:

(1) Was there a causal relationship between transportation improvements and land developments in terms of commercial development and residential development?

(2) What type of land use development was stimulated by transportation improvements? If both residential and commercial developments are affected, which one was affected to a large degree?

(3) Did residential developments stimulate commercial development? How strong

was the influence from residential developments on commercial developments? Conversely, do commercial developments cause residential developments and how strong is this causal relationship if it does exist?

(4) How quickly did residential and commercial developments respond to improvement in accessibility and how far did residential or commercial development lag behind transportation improvement projects?

(5) What were the cumulative impacts of transportation improvements on land development?

The VARX model development involved order selection, parameter estimation, diagnostic tests, and restricted system estimation. An unnecessarily large order will result in imprecise estimates of the corresponding VARX(*p, s*) model. It is therefore necessary to choose an adequate VAR order.

The least square method provided parameter estimates of the VARX(*p*, *s*) model. Any violations of the basic assumptions were checked using diagnostic tests. Finally, a restricted system was estimated by applying zero constraints on the coefficients of insignificant variables. These processes and results are described in this section.

6.5.1 Model Structure

The endogenous variables represented land use changes and included the applied building permits for residential and commercial uses. The exogenous variable was the lane mile index. Transportation projects such as adding capacity to the existing road are the results of transportation planning process instead of being freely driven by pure market forces. Although transportation projects are motivated by demand generated from land development, the transportation system does not respond to the development instaneously. Therefore, a road improvement index may be considered as an instrument variable of transportation policy while applied building permits is part of the land use system as they are driven by the land use market. As mentioned in Chapter 5, X_t is referred to as an exogenous variable where variable X_t causes Y_t and Y_t does not cause X_t (Reinsel 1997). To verify that *dlanemile* was an exogenous variable, granger causality test was performed. The test result showed that *dlanemile* caused *dcom* and *dres* but *dres* and *dcom* did not cause *dlanemile* at 95% confidence level.

Let $\mathbf{y}_t$ denote a *N*-dimensional vector of endogenous variables in period *t* and $\mathbf{x}_t$ be an exogenous variable. A VARX(*p*, *s*) model, where *p* is the order for the endogenous variables and *s* is the order for the exogenous variable, has the form

$$\mathbf{y}_t = \mathbf{c} + \sum_{i=1}^{p} \mathbf{\Phi}_i \mathbf{y}_{t-i} + \sum_{j=0}^{s} \mathbf{\Theta}_j \mathbf{x}_{t-j} + \mathbf{u}_t \qquad \text{[Eqn 6.3]}$$

where

$\mathbf{c}$ = a *N*-dimensional constant vector,

$\mathbf{\Phi}_i$ = a $(N \times N)$ coefficient matrix for $i = 1, 2, \ldots, p$,

$\mathbf{\Theta}_i$ = a $(N \times 1)$ coefficient matrix for $j = 1, 2, \ldots, s$, and

$\mathbf{u}_t$ = a *N*-dimensional error vector.

It is assumed that $\mathbf{u}_t$ is a white noise vector, which is an independently identically distributed random vector with zero mean. In the VARX model, coefficient matrices of $\mathbf{\Phi}_i$ and $\mathbf{\Theta}_i$ and constant vector, **c**, can be estimated using the least square (LS) method.

6.5.2 Order Selection

Five order selection criteria were used in this study to select appropriate order *p*. They included the likelihood ratio (LR) [Eqn 5.44], final prediction error (FPE) [Eqn 5.45], Akaike information criterion (AIC) [Eqn 5.46], Schwarz information criterion (SC) [Eqn 5.47], and Hannan-Quinn information criterion (HQ) [Eqn 5.48].

For the Tamiami Trail corridor, the likelihood ratio suggested 24 as the VARX order while FPE, AIC, and HQ statistics found the order to be one. SC selected the zero lag order for the model, which did not seem to be plausible. Therefore, one- and 24-month lag orders were chosen to develop the VARX model. For the Bird Drive corridor, the three-month lag order was suggested by FPE, SC, and HQ and the six-month lag order by LR and AIC. In the case of North Kendall Drive, only LR gave plausible suggestion of the 24 lag order.

The significant orders suggested by different criteria and the final choice of orders for different models are summarized in Table 6.14.

Table 6.14 Order p Selection by Five Criteria (Significance Levels = 0.05)

Criterion	Tamiami Trail	Bird Drive	North Kendall Drive
LR	24	6	23
FPE	1	3	0
AIC	1	6	0
SC	0	3	0
HQ	1	3	0
Selection	1, 24	3, 6	23

The selection of order *s* was done via a trial-and-error process when the univariate models were developed.

6.5.3 Parameter Estimation

The suggested VARX orders based on order selection criteria were one and 24 for the Tamiami Trail corridor. Thus two models, VARX(1, 16) and VARX(24,16), were estimated. Based on *R*-square values and results from the diagnostic tests, VARX(1, 16) was a better model. Table 6.15 shows the coefficient estimates for all variables with corresponding standard errors and *t*-statistics. The bold numbers indicate significant variables. The model was estimated using the least squared method. All lagged endogenous variables and the exogenous variable for commercial development were insignificant at the 95% confidence level. The coefficient estimate of one-month lagged residential development was 0.344524 and its corresponding *t*-statistics was 4.36310. Critical value of *t*-statistics with 13 degrees of freedom was 2.160. Five lagged exogenous variables representing lane mile changes were significant. The lengths of lagged effects were two, three, four, 14, and 16 months. The coefficient of 15-month lagged lane-mile changes was estimated as 327.3019 and its corresponding *t*-statistics was 159.351. Considering endogenous variables measured in thousand square feet, this may be interpreted as that building permits involving 327,301.9 building square feet for residential use were applied 15 months after one lane-mile was added to Tamiami Trail. Coefficient estimates of three and 16-month lagged lane-mile changes had a negative sign. However, the sum of coefficients of lagged exogenous variables was positive, which meant that the overall impacts from roadway improvement were positive, as expected. The estimated model for commercial development had a low *R*-square value (0.044645), but the *R*-square value for the residential development model was high as 0.994596.

Table 6.15 Parameter Estimates of VARX(1,16) for the Tamiami Trail Corridor

	DCOM			DRES		
	Coeff.	Std. Error	t-stat.	Coeff.	Std. Error	t-stat.
DCOM(-1)	-0.016461	(0.08415)	[-0.19562]	-0.584289	(0.61569)	[-0.94900]
DRES(-1)	0.005001	(0.01079)	[0.46340]	**0.344524**	**(0.07896)**	**[4.36310]**
DLANEMILE(-1)	0.522681	(0.27961)	[1.86935]	-0.233209	(2.04584)	[-0.11399]
DLANEMILE(-2)	-0.022122	(0.28242)	[-0.07833]	**12.47008**	**(2.06642)**	**[6.03464]**
DLANEMILE(-3)	-0.094735	(0.30697)	[-0.30862]	**-5.080748**	**(2.24604)**	**[-2.26209]**
DLANEMILE(-4)	-0.030471	(0.27913)	[-0.10916]	**5.251321**	**(2.04237)**	**[2.57119]**
DLANEMILE(-5)	-0.060002	(0.28396)	[-0.21130]	0.434476	(2.07772)	[0.20911]
DLANEMILE(-6)	0.314806	(0.27990)	[1.12473]	0.327818	(2.04795)	[0.16007]
DLANEMILE(-7)	-0.034818	(0.28049)	[-0.12413]	-1.483078	(2.05232)	[-0.72263]
DLANEMILE(-8)	-0.028813	(0.27929)	[-0.10316]	-0.649495	(2.04352)	[-0.31783]
DLANEMILE(-9)	-0.029920	(0.27918)	[-0.10717]	-0.772216	(2.04272)	[-0.37803]
DLANEMILE(-10)	0.331205	(0.27920)	[1.18626]	-0.359076	(2.04287)	[-0.17577]
DLANEMILE(-11)	-0.025733	(0.28039)	[-0.09178]	1.190849	(2.05156)	[0.58046]
DLANEMILE(-12)	-0.039055	(0.27905)	[-0.13996]	-0.763731	(2.04176)	[-0.37406]
DLANEMILE(-13)	-0.032878	(0.27898)	[-0.11785]	-0.913950	(2.04129)	[-0.44773]
DLANEMILE(-14)	-0.029998	(0.27917)	[-0.10745]	3.280978	(2.04266)	[1.60622]
DLANEMILE(-15)	-0.049985	(0.28072)	[-0.17806]	**327.3019**	**(2.05397)**	**[159.351]**
DLANEMILE(-16)	-1.677301	(3.55423)	[-0.47192]	**-113.3580**	**(26.0058)**	**[-4.35896]**
C	0.160738	(0.18494)	[0.86914]	**7.262640**	**(1.35317)**	**[5.36712]**

Similar to the Tamiami Trail model, two orders, three and six, were considered in modeling the dynamic system of the Bird Drive corridor. A trend was found in residential development variable as shown in Figure 6.9. Therefore, four models, for each combination of the orders and with or without trend, were estimated. The VARX(6,18) with trend had the best performance in terms of *R*-square value, Akaike Information Criteria (AIC), and Schwarz Criteria (SC). Table 6.16 summarizes the parameter estimates for the VARX(6,18) model. Significant variables are shown in bold. Three- and six-month lagged commercial development and two-month lagged residential development were found to be significant for commercial development in current time period (or a given time *t*). Their coefficients were estimated as 0.444189, -0.198105, and 0.138026, respectively. None of the lagged exogenous variables were significant to commercial development, which may be interpreted as that the impact from roadway improvement on commercial development was negligible in the Bird Drive corridor. Coefficient estimates of one-month lagged residential development and 18-month lagged roadway improvement were significant to residential development and their values were 0.230819 and 29.95745, respectively. R-square values for both models for commercial and residential development were 0.470811 and 0.461018, respectively.

Table 6.16 Parameter Estimates of VARX(6,18) for the Bird Drive Corridor

	DCOM			DRES		
	Coeff.	Std. Error	t-stat.	Coeff.	Std. Error	t-stat.
DCOM(-1)	-0.150352	(0.08658)	[-1.73658]	-0.057260	(0.40990)	[-0.13969]
DCOM(-2)	0.115921	(0.08859)	[1.30850]	-0.351369	(0.41942)	[-0.83774]
DCOM(-3)	**0.444189**	**(0.08887)**	**[4.99846]**	0.114018	(0.42072)	[0.27101]
DCOM(-4)	0.110873	(0.08545)	[1.29750]	-0.376771	(0.40456)	[-0.93132]
DCOM(-5)	-0.057239	(0.07394)	[-0.77413]	0.008512	(0.35006)	[0.02432]
DCOM(-6)	**-0.198105**	**(0.07126)**	**[-2.78020]**	0.503029	(0.33735)	[1.49112]
DRES(-1)	-0.010838	(0.01826)	[-0.59369]	**0.230819**	**(0.08643)**	**[2.67069]**
DRES(-2)	**0.138026**	**(0.01859)**	**[7.42347]**	0.003825	(0.08803)	[0.04346]
DRES(-3)	0.031092	(0.02167)	[1.43498]	0.084005	(0.10258)	[0.81891]
DRES(-4)	-0.038492	(0.02213)	[-1.73969]	0.137115	(0.10475)	[1.30894]
DRES(-5)	0.007128	(0.02256)	[0.31599]	-0.181890	(0.10679)	[-1.70318]
DRES(-6)	-0.024221	(0.02177)	[-1.11248]	0.145725	(0.10308)	[1.41377]
DLANEMILE(-1)	-0.087079	(2.06298)	[-0.04221]	3.801177	(9.76691)	[0.38919]
DLANEMILE(-2)	-0.894280	(2.06106)	[-0.43389]	8.440157	(9.75784)	[0.86496]
DLANEMILE(-3)	-1.167536	(2.06556)	[-0.56524]	4.634416	(9.77911)	[0.47391]
DLANEMILE(-4)	-1.718949	(2.06435)	[-0.83268]	0.671910	(9.77339)	[0.06875]
DLANEMILE(-5)	-1.551922	(2.05745)	[-0.75429]	1.771397	(9.74075)	[0.18185]
DLANEMILE(-6)	-0.360059	(2.06115)	[-0.17469]	3.627241	(9.75827)	[0.37171]
DLANEMILE(-7)	-0.114230	(2.06081)	[-0.05543]	3.605424	(9.75665)	[0.36953]
DLANEMILE(-8)	-0.912825	(2.15577)	[-0.42343]	2.614394	(10.2062)	[0.25616]
DLANEMILE(-9)	-0.877175	(2.14562)	[-0.40882]	1.907910	(10.1582)	[0.18782]
DLANEMILE(-10)	-0.881765	(2.14055)	[-0.41193]	7.242150	(10.1342)	[0.71463]
DLANEMILE(-11)	-0.935288	(2.14395)	[-0.43624]	-9.450253	(10.1503)	[-0.93103]
DLANEMILE(-12)	-2.001432	(2.06486)	[-0.96928]	1.592158	(9.77580)	[0.16287]
DLANEMILE(-13)	-0.175960	(2.06748)	[-0.08511]	11.06242	(9.78823)	[1.13018]
DLANEMILE(-14)	3.444544	(2.07370)	[1.66106]	1.215661	(9.81765)	[0.12382]
DLANEMILE(-15)	-1.377126	(2.10431)	[-0.65443]	-1.533715	(9.96260)	[-0.15395]
DLANEMILE(-16)	-1.755109	(2.10725)	[-0.83289]	-5.073126	(9.97650)	[-0.50851]
DLANEMILE(-17)	-1.655757	(2.09489)	[-0.79038]	13.51607	(9.91800)	[1.36278]
DLANEMILE(-18)	0.139900	(2.10298)	[0.06652]	**29.95745**	**(9.95629)**	**[3.00890]**
C	-0.277902	(0.95016)	[-0.29248]	**9.287373**	**(4.49840)**	**[2.06460]**
TREND DUMMY	**-0.037569**	**(0.01334)**	**[-2.81563]**	**0.194148**	**(0.06317)**	**[3.07342]**

For the North Kendall Drive corridor, the VARX(23,21) model was estimated. The estimation results are summarized in Table 6.17. The bold numbers indicate significant variables. All coefficient estimates of lagged commercial development were insignificant at the 95% confidence level. Three-month lagged residential development had a coefficient of 0.193724 with corresponding t-statistics of 4.17134. The coefficient estimates of seven-, eight-, and 14-month lagged roadway improvement were 17.19923, 14.38736, and -31.51322, respectively, and their corresponding t-statistics were 6.86341, 4.51184, and -3.94422, respectively. Commercial development in the current time period was affected by residential development with a three-month lag and roadway improvement with seven-, eight-, and 14-month lags. Cumulative impact from roadway improvement was almost zero. For residential development there were nine significant variables including one commercial development with a 23-month lag, four residential developments with 12-, 14-, 18-, and 20-month lags, and four roadway improvements with eight-, 11-, 17-, and 21-month legs. They all had positive signs. Cumulative impacts from roadway improvement calculated by summarizing the coefficient estimates of lagged exogenous variables were 33.840008 and 235.071577 for commercial development and residential development, respectively. It means that there were an increase of 33,840.008 and 235,071.577 building square feet in commercial and residential development permit applications within 21 months after the North Kendall Drive corridor was widened. R-square values for commercial development and residential development were 0.673213 and 0.944553, respectively.

Table 6.17 Parameter Estimates of VARX(23,21) for the North Kendall Drive Corridor

	DCOM			DRES		
	Coeff.	Std. Error	t-stat.	Coeff.	Std. Error	t-stat.
DCOM(-1)	-0.004306	(0.11107)	[-0.03877]	-0.283565	(0.23612)	[-1.20093]
DCOM(-2)	0.059267	(0.10753)	[0.55116]	-0.160642	(0.22861)	[-0.70270]
DCOM(-3)	-0.021335	(0.10733)	[-0.19877]	0.261966	(0.22819)	[1.14804]
DCOM(-4)	-0.071832	(0.10897)	[-0.65918]	-0.128735	(0.23167)	[-0.55568]
DCOM(-5)	0.013433	(0.12643)	[0.10625]	-0.278959	(0.26878)	[-1.03789]
DCOM(-6)	0.011024	(0.12755)	[0.08643]	-0.168252	(0.27117)	[-0.62047]
DCOM(-7)	0.017535	(0.12839)	[0.13657]	0.246640	(0.27296)	[0.90358]
DCOM(-8)	-0.107666	(0.11175)	[-0.96346]	-0.218216	(0.23757)	[-0.91852]
DCOM(-9)	-0.081886	(0.11306)	[-0.72427]	-0.344051	(0.24036)	[-1.43138]
DCOM(-10)	-0.046686	(0.11324)	[-0.41228]	-0.170729	(0.24074)	[-0.70918]
DCOM(-11)	0.107004	(0.11275)	[0.94903]	-0.311029	(0.23970)	[-1.29756]
DCOM(-12)	-0.065356	(0.11433)	[-0.57166]	0.231927	(0.24306)	[0.95421]
DCOM(-13)	-0.055179	(0.11260)	[-0.49004]	-0.174053	(0.23939)	[-0.72708]
DCOM(-14)	-0.068679	(0.11285)	[-0.60860]	-0.101237	(0.23991)	[-0.42198]
DCOM(-15)	-0.040221	(0.11002)	[-0.36558]	-0.411635	(0.23390)	[-1.75990]
DCOM(-16)	-0.094123	(0.11032)	[-0.85320]	0.024557	(0.23453)	[0.10471]
DCOM(-17)	-0.115417	(0.10984)	[-1.05082]	-0.182482	(0.23351)	[-0.78148]
DCOM(-18)	-0.046800	(0.10900)	[-0.42937]	0.115623	(0.23172)	[0.49897]
DCOM(-19)	0.094926	(0.10736)	[0.88417]	-0.172796	(0.22825)	[-0.75706]
DCOM(-20)	-0.055338	(0.10811)	[-0.51185]	0.107789	(0.22985)	[0.46896]
DCOM(-21)	-0.151687	(0.10424)	[-1.45522]	-0.200755	(0.22160)	[-0.90593]
DCOM(-22)	0.108738	(0.10241)	[1.06184]	-0.072972	(0.21771)	[-0.33518]
DCOM(-23)	-0.137130	(0.10242)	[-1.33895]	**0.664667**	**(0.21773)**	**[3.05268]**
DRES(-1)	-0.063255	(0.04565)	[-1.38580]	0.046521	(0.09704)	[0.47940]
DRES(-2)	-0.001629	(0.04655)	[-0.03498]	0.000367	(0.09897)	[0.00371]
DRES(-3)	**0.193724**	**(0.04644)**	**[4.17134]**	0.195315	(0.09873)	[1.97821]
DRES(-4)	-0.032346	(0.04688)	[-0.69001]	-0.112814	(0.09966)	[-1.13197]
DRES(-5)	-0.011648	(0.04753)	[-0.24509]	0.135337	(0.10104)	[1.33944]
DRES(-6)	-0.002672	(0.04753)	[-0.05622]	-0.182651	(0.10104)	[-1.80773]
DRES(-7)	0.025027	(0.04916)	[0.50914]	0.067896	(0.10450)	[0.64970]
DRES(-8)	-0.045249	(0.05221)	[-0.86675]	0.124724	(0.11099)	[1.12377]

	DCOM			DRES		
	Coeff.	Std. Error	t-stat.	Coeff.	Std. Error	t-stat.
DRES(-9)	-0.065689	(0.05183)	[-1.26750]	-0.013956	(0.11018)	[-0.12667]
DRES(-10)	-0.029829	(0.05668)	[-0.52626]	-0.234737	(0.12050)	[-1.94801]
DRES(-11)	0.031652	(0.02343)	[1.35117]	0.010095	(0.04980)	[0.20270]
DRES(-12)	0.012269	(0.02316)	[0.52969]	**0.217136**	**(0.04924)**	**[4.40946]**
DRES(-13)	0.003227	(0.02461)	[0.13116]	0.007566	(0.05231)	[0.14463]
DRES(-14)	0.001668	(0.02438)	[0.06843]	**0.116314**	**(0.05182)**	**[2.24436]**
DRES(-15)	-0.007416	(0.02492)	[-0.29756]	-0.096993	(0.05298)	[-1.83070]
DRES(-16)	-0.015897	(0.02388)	[-0.66573]	0.068570	(0.05077)	[1.35072]
DRES(-17)	0.047192	(0.02365)	[1.99568]	-0.090639	(0.05027)	[-1.80296]
DRES(-18)	0.014553	(0.02358)	[0.61722]	**0.166847**	**(0.05013)**	**[3.32839]**
DRES(-19)	0.027668	(0.02351)	[1.17697]	-0.050637	(0.04998)	[-1.01319]
DRES(-20)	0.033192	(0.02179)	[1.52305]	**0.164367**	**(0.04633)**	**[3.54764]**
DRES(-21)	-0.001118	(0.01914)	[-0.05844]	-0.024599	(0.04069)	[-0.60454]
DRES(-22)	0.000170	(0.01873)	[0.00907]	0.017947	(0.03983)	[0.45063]
DRES(-23)	-0.025533	(0.01820)	[-1.40294]	-0.055123	(0.03869)	[-1.42464]
DLANEMILE(-1)	-0.877321	(2.76256)	[-0.31757]	0.677451	(5.87311)	[0.11535]
DLANEMILE(-2)	-2.525417	(2.61789)	[-0.96468]	-6.086672	(5.56555)	[-1.09363]
DLANEMILE(-3)	0.698166	(2.62132)	[0.26634]	-4.278199	(5.57283)	[-0.76769]
DLANEMILE(-4)	0.467212	(2.64609)	[0.17657]	-0.939691	(5.62549)	[-0.16704]
DLANEMILE(-5)	-0.896141	(2.49784)	[-0.35877]	-1.774276	(5.31033)	[-0.33412]
DLANEMILE(-6)	-0.846388	(2.53502)	[-0.33388]	2.410933	(5.38936)	[0.44735]
DLANEMILE(-7)	**17.19923**	**(2.50593)**	**[6.86341]**	7.660197	(5.32752)	[1.43785]
DLANEMILE(-8)	**14.38736**	**(3.18880)**	**[4.51184]**	**61.22733**	**(6.77927)**	**[9.03155]**
DLANEMILE(-9)	1.029581	(4.53065)	[0.22725]	2.302162	(9.63200)	[0.23901]
DLANEMILE(-10)	-4.012478	(4.42321)	[-0.90714]	-8.077370	(9.40358)	[-0.85897]
DLANEMILE(-11)	-7.228457	(4.51069)	[-1.60252]	**137.8058**	**(9.58957)**	**[14.3704]**
DLANEMILE(-12)	11.79751	(7.84310)	[1.50419]	2.162763	(16.6742)	[0.12971]
DLANEMILE(-13)	-1.848534	(8.04609)	[-0.22974]	-0.640190	(17.1057)	[-0.03743]
DLANEMILE(-14)	**-31.51322**	**(7.98972)**	**[-3.94422]**	-29.06354	(16.9859)	[-1.71104]
DLANEMILE(-15)	5.496984	(8.20564)	[0.66990]	10.23850	(17.4449)	[0.58690]
DLANEMILE(-16)	7.060149	(8.12707)	[0.86872]	-19.46240	(17.2779)	[-1.12644]
DLANEMILE(-17)	6.364461	(8.09743)	[0.78599]	**39.70256**	**(17.2149)**	**[2.30630]**
DLANEMILE(-18)	-4.669815	(8.56201)	[-0.54541]	9.827119	(18.2025)	[0.53988]
DLANEMILE(-19)	4.832108	(8.58392)	[0.56293]	-18.12854	(18.2491)	[-0.99339]
DLANEMILE(-20)	11.79060	(8.39723)	[1.40411]	-11.18305	(17.8522)	[-0.62642]
DLANEMILE(-21)	7.134418	(8.96126)	[0.79614]	**60.69069**	**(19.0513)**	**[3.18564]**
C	6.905126	(3.97430)	[1.73744]	16.81899	(8.44923)	[1.99059]

6.5.4 Diagnostic Tests

Three model diagnostic tests including AR root test, Lagrange Multiplier test, and Portmanteau test were conducted to check model validity and adequacy. AR root test was to check if the estimated VARX models for study corridors were stable. If the stability condition, $\det(I_N - \Phi_1 z - \ldots - \Phi_p z^p) \neq 0$ for $|z| \leq 1$, is satisfied, then the estimated model is said to be stable. All three models for study corridors satisfied this condition and therefore determined stable.

The Portmanteau test was conducted to check whiteness of the residuals of the estimated model and the overall significance of the residual autocorrelations. The null hypothesis was that there were no residual autocorrelations up to the specified lag. The test is valid only for lags larger than the lag order of endogenous variables for VARX models. The tests were performed up to lags of four, 12, and 48 for the Tamiami Trail, Bird Drive, and North Kendall Drive models, respectively. The *p*-values of tested lags for both the Tamiami Trail and the Bird Drive were too large to reject the null hypothesis. Some autocorrelation was found up to 38 lags in the estimated VARX model residuals for the North Kendall Drive model, but no more autocorrelation were found for lags higher than 38.

Under the null hypothesis that there was no serial correlation at the specified lag. LM test was conducted to check whether the estimated model was appropriate. The p-values corresponding to LM test statistics were 0.4712, 0.6409, and 0.4613 at lag one, six, 23 for the Tamiami Trail, Bird Drive, and North Kendall Drive models, respectively. These values were too large to reject the null hypothesis (at the 95% confidence level or higher), which meant that there was no serial correlation at the lag order of each model estimated for the study corridors.

6.5.5 Restricted System

Results from the parameter estimation described earlier showed some of the coefficient estimates were not significantly different from zero. These zero coefficients may be interpreted in two ways. First, variables corresponding to zero coefficients did not have causal relationship with other variables. Second, the information in the data was inadequate to provide sufficiently precise estimates with confidence intervals that did not contain zero. In the latter case imposing zero constraints on the insignificant coefficients has been found as a solution (Waddell 2001). Insignificant variables were removed by placing zero constraints on the coefficients and the restricted systems with only significant variables were estimated by the least square method. Since none of the coefficients were significant to the endogenous variable of commercial development for the Tamiami Trail corridor, only the dynamic system of residential development was modeled with significant variables. The restricted systems for each study corridor are shown below.

Tamiami Trail

$$\begin{aligned} dres(t) = {} & 1.379023 \times dcom(t\text{-}4) - 1.327276 \times dcom(t\text{-}6) + 2.321187 \times dcom(t\text{-}16) \\ & - 2.324449 \times dcom(t\text{-}21) + 0.471442 \times dres(t\text{-}1) + 0.011933 \times dres(t\text{-}15) \\ & + 0.010363 \times dres(t\text{-}17) + 0.034389 \times dres(t\text{-}23) - 0.014000 \times dres(t\text{-}24) \\ & + 12.38801 \times dlanemile(t\text{-}2) - 6.585278 \times dlanemile(t\text{-}3) \\ & + 5.502976 \times dlanemile(t\text{-}4) + 327.0074 \times dlanemile(t\text{-}15) \\ & - 154.2895 \times dlanemile(t\text{-}16) + 4.638388 \end{aligned} \qquad \text{[Eqn 6.1]}$$

Bird Road Corridor

$$\begin{aligned} dres(t) = {} & 0.234462 \times dres(t\text{-}1) + 22.79471 \times dlanemile(t\text{-}18) + 15.76523 \\ & + 0.224561 \times TrendDummy \end{aligned} \qquad \text{[Eqn 6.2]}$$

$$dcom(t) = 0.388299 \times dcom(t\text{-}3) - 0.173641 \times dcom(t\text{-}6) + 0.059454 \times dres(t\text{-}2) \qquad \text{[Eqn 6.3]}$$

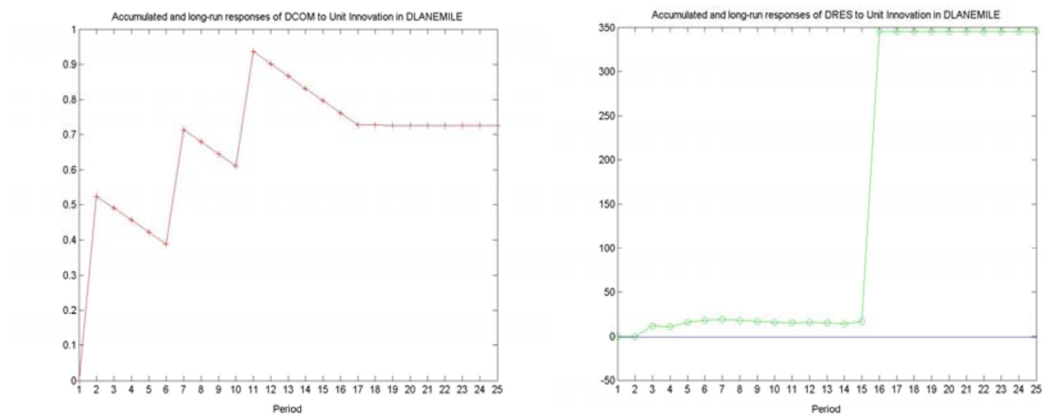
North Kendall Drive Corridor

$$dres(t) = 0.096\times dres(t\text{-}10) + 0.126\times dres(t\text{-}12) + 0.098\times dres(t\text{-}18) + 0.099\times dres(t\text{-}20) + 0.648\times dcom(t\text{-}23) + 61.223\times dlanemile(t\text{-}8) + 151.979\times dlanemile(t\text{-}11) \quad \text{[Eqn 6.4]}$$

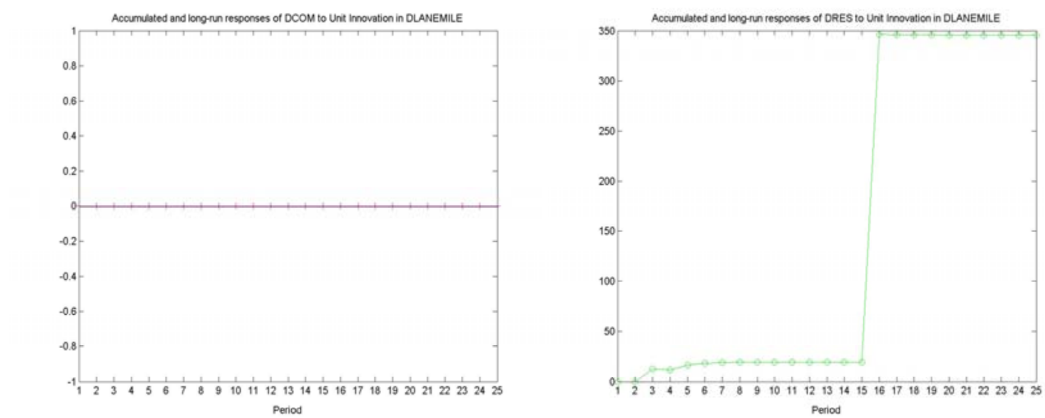
$$dcom(t) = 0.043\times dres(t\text{-}17) + 16.997\times dlanemile(t\text{-}7) + 13.804\times dlanemile(t\text{-}8) + 3.695 \quad \text{[Eqn 6.5]}$$

6.5.6 Multiplier Analysis

Based on the estimations of VARX models and corresponding restricted systems for the study corridors, it was of interest to know the effect of a shock or innovation in an exogenous variable, i.e., lane miles, on endogenous variables, i.e., commercial and residential developments. As discussed in Chapter 5, it is possible to trace the response of endogenous variables by estimating dynamic multipliers, D_i. Dynamic multipliers can be computed as the sum of the lags of exogenous variable ($\mathbf{\Theta}(L) = \mathbf{\Theta}_0 + \mathbf{\Theta}_1 L + \ldots + \mathbf{\Theta}_s L^s$) divided by one minus the sum of all lags of the endogenous variables, i.e., $\mathbf{\Phi}(L) = I_N - \mathbf{\Phi}_1 L - \ldots - \mathbf{\Phi}_p L^p$. Accumulated dynamic multipliers represent cumulative impacts of an exogenous variable on endogenous variables, while one dynamic multiplier is a response of endogenous variable at a particular time. Figure 6.19 through 6.18 depict the accumulated dynamic multipliers up to 25 months for the study corridors. In the figures, the charts on the left represent commercial developments and those on the right represent the residential developments. Dynamic multipliers depicted in Figures 6.19(a), 6.20(a), and 6.21(a) were estimated based on VARX models while those in Figures 6.19(b), 6.20(b), and 6.21(b) were estimated based on the restricted system. Accumulated dynamic multipliers for 25 months suggested that the cumulative impact for the residential development was 345,679.5 square feet and that there was no impact on the commercial development in the Tamiami Trail study corridor. The cumulative impact was 38,945.8 and 7,632.8 square feet for residential and commercial developments, respectively, in the Bird Drive study corridor. It was 349,765.2 and 60,654.9 square feet for residential and commercial developments, respectively, in the North Kendall Drive study corridor. The commercial development model for the Bird Drive did not have any significant exogenous variables, but cumulative impact to unit increase of roadway improvement was found. This might be the indirect impact caused by roadway improvement through residential development. For instance, roadway improvement affected residential development and, in turn, residential development caused commercial development.

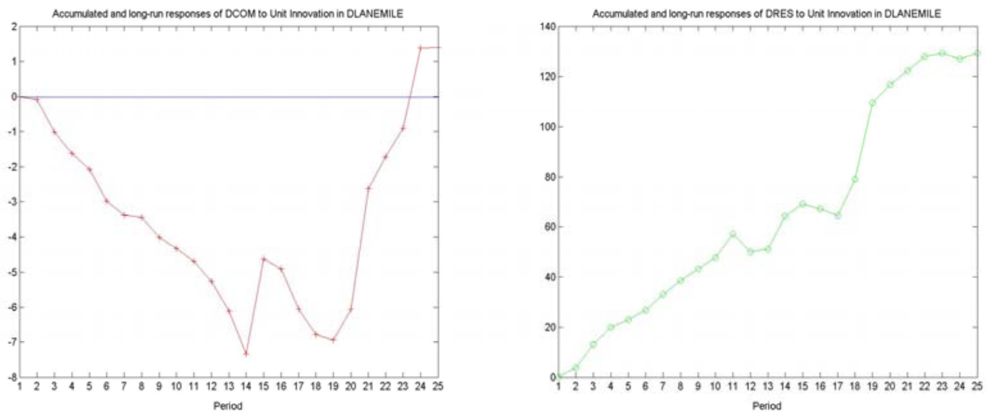


(a) VARX model

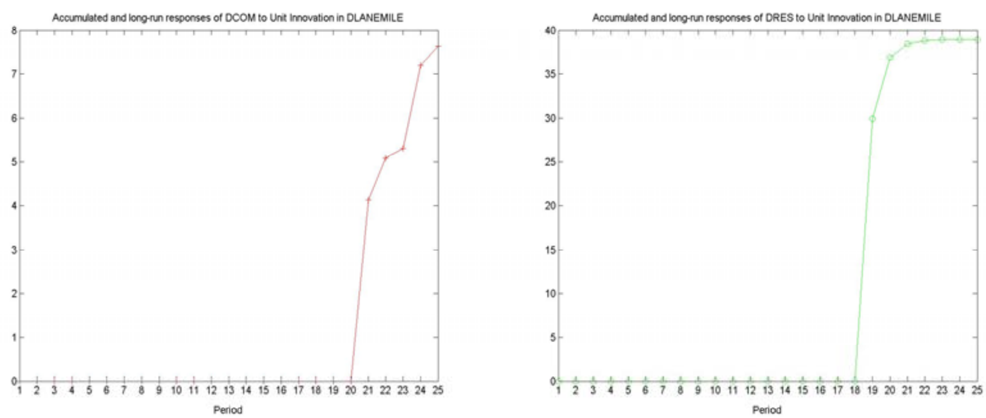


(b) Restricted System

Figure 6.19 Accumulated Response to Unit Innovation in an Exogenous Variable in the Tamiami Trail Model

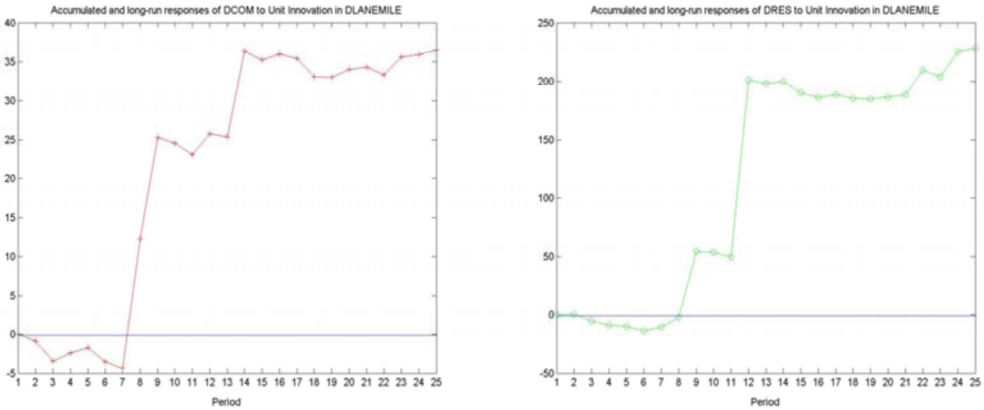


(a) VARX model

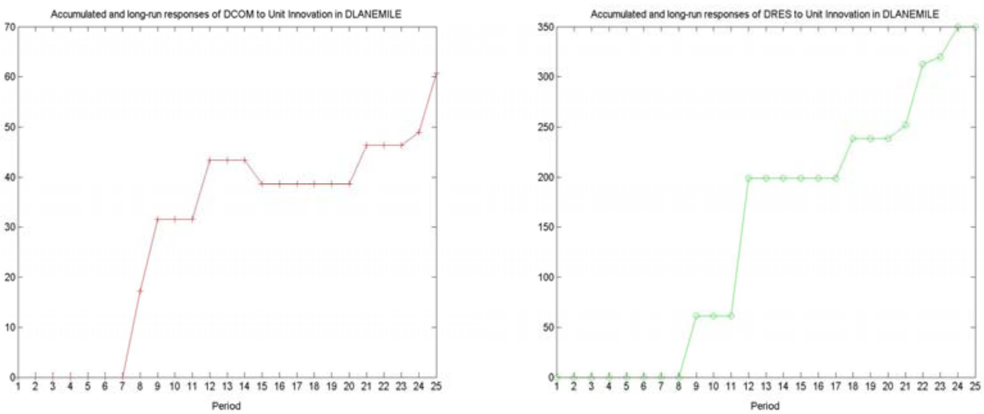


(b) Restricted System

Figure 6.20 Accumulated Response to Unit Innovation in an Exogenous Variable in the Bird Drive Model



(a) VARX model



(b) Restricted System

Figure 6.21 Accumulated Response to Unit Innovation in an Exogenous Variable in the North Kendall Drive Model

6.5.7 Summary of VARX Models

Table 6.19 summarizes the important statistics of the three VARX models, one for each study area. The estimated VARX models indicated that transportation improvements did have impacts on residential development with different lag effects in all study corridors. No evidence was found that roadway improvements caused commercial development in Tamiami Trail and Bird Drive corridors. However, as discussed in Section 6.5.6, there might be indirect impact of roadway improvement on commercial development around Bird Drive as a result of the residential developments. Results for the North Kendall Drive model showed that roadway improvements caused both commercial and residential developments.

Table 6.18 Summary of VARX Models

Model Statistics	Model					
	Tamiami Trail		Bird Drive		North Kendall Drive	
	dcom	*dres*	*dcom*	*dres*	*dcom*	*dres*
Estimation Method	LS		LS		LS	
R^2	0.04	0.99	0.47	0.46	0.67	0.94
SER	1.69	12.40	6.47	30.65	19.27	40.96
SSR	404.82	21,672.41	5322.27	119,294.9	28,954.69	130,867.3
Lags significant (months) for lane-mile		2, 3, 4, 15, 16		18	7, 8, 14	8, 11, 17, 21
Lags significant for commercial development (months)			3, 6			23
Lags significant for residential development (months)		1	2	1	3	12, 14, 18, 20
Cumulative impacts from land-mile (sq-feet)	0	345,679.5	7,632.8	38,945.8	60,654.9	349,765.2

Again, previous developments appeared to have some effects over the current developments, but the effects were much smaller compared to the effects of transportation improvements as indicated by the magnitudes of the coefficients representing land developments and transportation improvements.

Both commercial and residential developments in all case studies were stimulated by roadway improvements directly or indirectly. The only exception was the commercial developments in the Tamiami Trail corridor, which did not appear to be affected by the transportation improvement. In the North Kendall Drive corridor, both types of developments were caused by transportation improvements, and the cumulative impacts of transportation improvements were 60,654.9 and 349,765.2 for commercial and residential developments, respectively. Thus residential development was more intensely affected by transportation improvements. This was also found true in other corridors in this study.

Building permit applications began to increase noticeably as soon as two months and as late as 21 months after transportation improvements. For instance, residential building permit applications responded to transportation improvement within two months in the Tamiami Trail corridor, and there was still impact 16 months later. In the Bird Drive case, the impact of transportation improvement on residential developments became evident in 18 months through building permit applcations. In the North Kendall Drive case, it took eight months for residential development to start to manifest in the form of building permit applications, and the impact was still felt in the 21st month. Commercial permit applications responded with seven, eight and 14 month lags. It seemed that in the North Kendall Drive Corridor, commercial developments were more severely constrained than residential developments were.

Instantaneous impacts from transportation improvements may be found from coefficient estimation. In the residential development model for the Tamiami Trail corridor, the coefficients of the significant exogenous variables, 12.38801, –6.585278, 5.502976, 327.0074, and -154.2895, may be interpreted as instantaneous impact while the long-run impact may be estimated as 345.6795. For the Bird Drive corridor, instantaneous impact on residential development was estimated as 22.79471 and long-run impact was estimated as 38.9458 whereas there was no instantaneous impact and an indirect impact of 7.6328 on commercial development. In the North Kendall Drive case, the instantaneous impacts on the residential development were estimated to be 61.223 and 151.979, and instantaneous impacts on the commercial development were 16.997 and 13.804. The long-run impact was measured as 349.7652 and 60.6549 for residential and commercial developments, respectively.

To compare the results of the VARX models with those from the univariate models discussed in Section 6.4, it is helpful to consider the distribution of developments based on land use type as shown in Table 6.1 through 6.3. The dominant land use type of the new developments around study corridors was found to be of residential use. In the Tamiami Trail corridor, residential development was 97.92% of the total new development. In the Bird Drive corridor and the North Kendall Drive corridor residential developments accounted for 90.46% and 80.95% of the total new development. Percentages of other land use type excluding residential and commercial were 0.66, 5.19, and 2.59 for Tamiami Trail, Bird Drive, and North Kendall Drive corridors, respectively.

In terms of significant lags for lane-miles, the VARX models found a larger number of lags than the univariate models except for Bird Drive, for which there was only one significant lag for lane-miles in both the residential and commercial models. The VARX model found more lags because univariate models employed total development as an endogenous variable, which was a variable aggregating all types of land use, while VARX model treated both commercial and residential developments as endogenous variables. It was, therefore, possible to track interactions between land use and transportation more closely. However, the univariate models for the Bird Drive corridor had more lags for the land use variables than the VARX models. In the VARX modeling process, when more lags were included, the models had serious autocorrelation, which violated the assumption of variable randomness.

According to the VARX model estimations, the largest cumulative impact of roadway improvements on land development was found in the North Kendall Drive corridor among the study corridors, estimated as 410,420.1 (= 60,654.9 + 349,765.2) square feet, while the univariate model estimated it as 319,439 square feet. The difference between the two model estimation results could be explained by considering the largest lag included in the model. The largest lag for lane-miles was 21 in the VARX model and 11 in the univariate model, and the largest lag for land development was 23 in the VARX model and 12 in the univariate model. Therefore, cumulative impact estimated by VARX was larger then the univariate model estimation. The largest lag included in the univariate model for the Tamiami Trail corridor was 15 and the one in the VARX model was 16. Thus cumulative impacts estimated by VARX model and univariate model were close to each other and they were 345,679.5 square feet and 339,003 square feet, respectively. For the Bird Drive corridor, the largest lag for lane-miles was 18 in the VARX model and 23 in the univariate model, and the largest lag for land development was six in the VARX model and 25 in the univariate model. It was found that there was a significant difference in the largest lag for land development between the univariate model and the VARX model. The difference influenced the cumulative impact estimation. The VARX model for the Bird Drive corridor estimated the cumulative impact as 46,578.6 (=7,632.8 + 38,945.8) square-feet while the univariate model estimated the cumulative impact as 110,397 square-feet. Controlling those endogenous variables with 15 and 25 month lags, the cumulative impact estimations became close to each other.

6.6 Examination of Traffic Conditions, Development Timing, and Density

As discussed in Section 6.1, development is influenced by transportation supply or accessibility, but is also determined by other factors. One important factor is the congestion level. Congestion not only reduces transportation accessibility thus the attractiveness of the land, but also prevents further development from taking place through the concurrency requirements. It has been shown by the univariate models and the VARX models that the effects of transportation investments varied in different study corridors. To try to understand why such variations exist, the traffic conditions in two corridors, Tamiami Trail and North Kendall Drive corridors, both state roads with historical traffic data (AADT) available, were examined.

A good measure of accessibility or congestion may be levels of service (LOS), which not only consider the number of lanes but also the roadway geometry, intersection geometry, and traffic control at intersections. Historical LOS data were, however, unavailable. Therefore, historical AADT data for state highway system from the Florida Traffic Information (FTI) CD-ROM series were used. To estimate the degree of saturation on different sections of the two corridors, volume over capcity (v/c) ratios were roughly approximated. The peak hour volume was assumed to be the product of AADT and the corresponding K-factor. The capacity of corridors was assumed to be 1,800 vehicles per hour per lane and was adjusted by the green to signal cycle length ratio, assumed as 0.6. The v/c ratios were calculated as the hourly traffic volume divided by the capacity of the road.

The annual estimated v/c ratios for the North Kendall corridor is presented in Figure 6.22 with the completion date of the two transportation improvement projects indicated by the red line. Three count stations at SW 127th Avenue, SW 137th Avenue, and SW 147th Avenue were in the improved sections. Thus the v/c ratios for these locations were calculated based on improved capacity after the construction. It may be seen that the v/c ratios at SW 127th Avenue and SW 137th Avenue were higher than one before the construction, indicating that these sections had serious congestions. The fluctuation of the v/c ratios between 1985 and 1993 might be the result of partial traffic diversion although there were no data to confirm this suspicion. However, it seemed that the decrease in traffic volume before the completion of roadway improvement projects might have been due to the construction. The traffic volume at SW 137th Avenue surpassed that at SW 127th Avenue in 1992, possibly due in part to the concentration of commercial development at the location. In 2000, the v/c ratio at SW 127th Avenue again was approaching 1.0, which might have hindered further development of the area. This, however, will need to be verified.

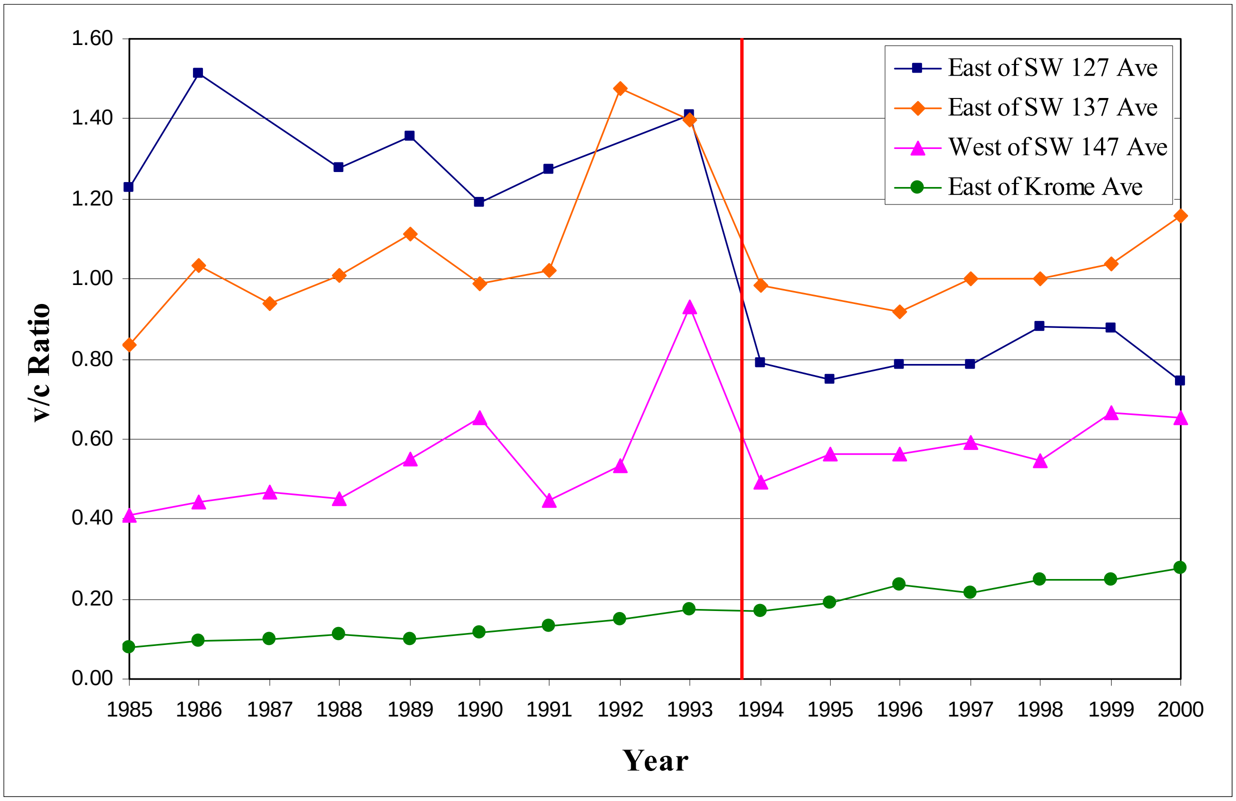


Figure 6.22 V/C Ratios in the North Kendall Drive Corridor

The estimated annual v/c ratios for Tamiami Trail are plotted in Figure 6.23. The section between SW 112th Avenue and SW 127th Avenue was improved and the v/c ratio at SW 122nd Avenue was the only one affected by the road improvement. Thus the v/c ratio at SW 122nd Avenue was computed with the improved capacity beginning in 1995. There was a decrease in the v/c ratio from 1989 to 1994, but after the improvement, the v/c ratio started to grow again. The v/c ratios at SW 137th Avenue increased between 1993 and 1995, reflecting increase of residential developments in the area between SW 137th Avenue and SW 147th Avenue in 1993 as shown in Figure 6.3(a). The difference of the v/c ratio between SW 137th Avenue and SW 139th Avenue was likely due to the travel originated from the area south of Tamiami Trail, which was loaded onto Tamiami Trail through SW 137th Avenue.

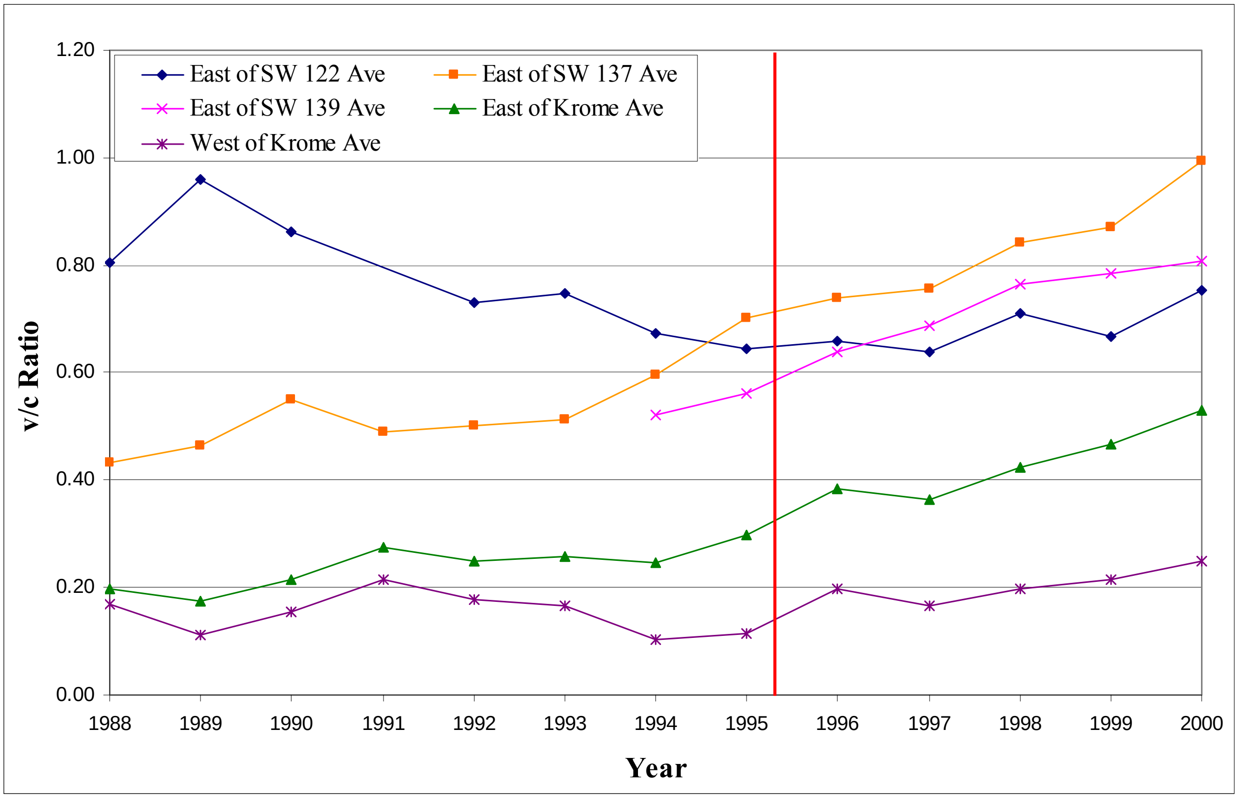


Figure 6.23 V/C Ratios in the Tamiami Trail Corridor

The univariate and VARX models both indicated that there were correlations between transportation projects and land development. To examine the timing of the transportation projects and land development, the time series of the cumulative building floor areas in square feet based on building permit records in the unincorporated Miami-Dade County area and the same in the three study corridors are plotted in Figures 6.24 through 6.27. The purpose is to determine if it is reasonable to suggest that the developments were likely the results of transportation improvements and not a reflection of overall growth trend.

From Figure 6.24, it may be see that there was a continuous, upward development trend, mostly residential, in the unincorporated Miami-Dade County, slow at the beginning but picking up after 1991. A spike also occurred mid 1994, after which the growth rate remained rather steady. In comparison, the development trend in the Tamiami Trail corridor exhibited a similar pattern but the spike occurred in mid 1996 (see Figure 6.25). The transportation improvement project was completed in 1995, which was before the Tamiami Corridor development spike and after the one for the unincorporated Miami-Dade County. In the Bird Drive corridor, the transportation improvements projects took place between 1991 and 1994, with accelerated increases in development ocuuring after the first stage of the transportation improvement project and continued after the the completion of the transportation improvement. The situation in North Kendall Drive was similar in that the growth rate sharply increased only after transportation improvement.

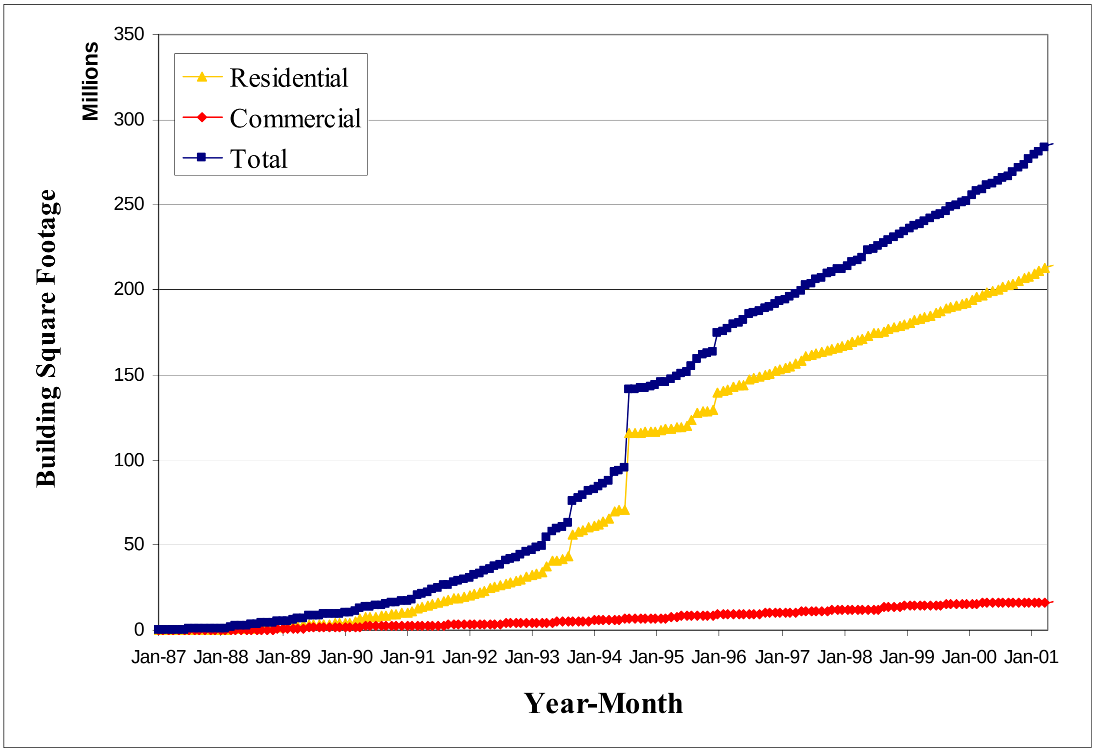


Figure 6.24 Cumulative Floor Space (Sq-Feet) based on Building Permit Applications in Unincorporated Miami-Dade County

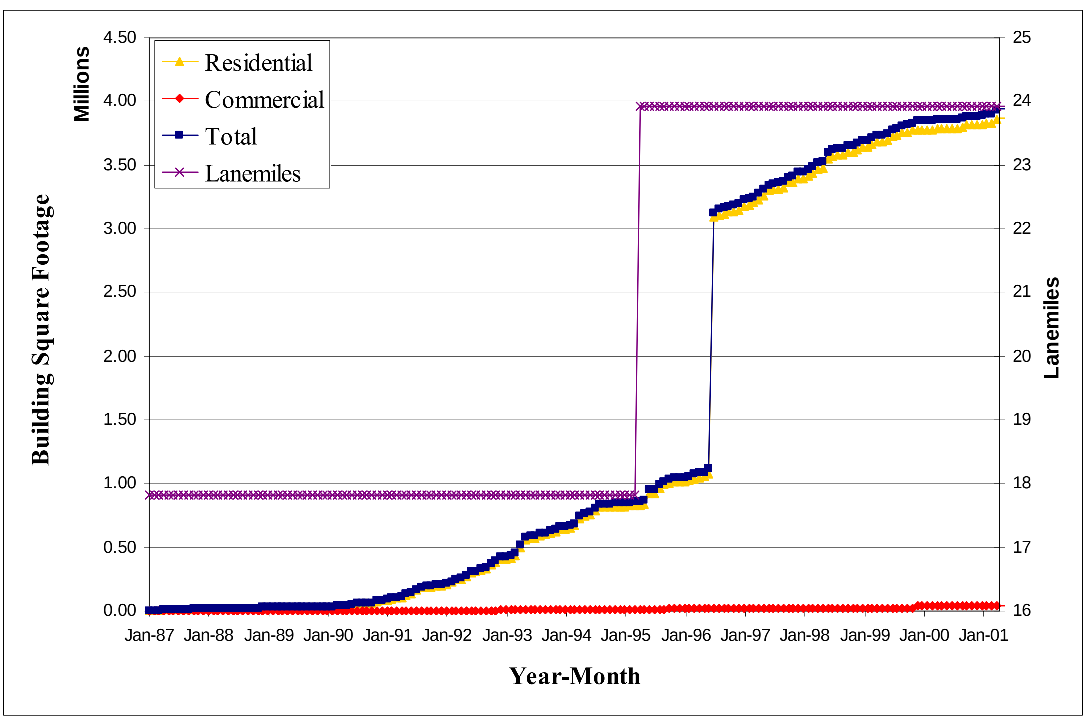


Figure 6.25 Cumulative Floor Space (Sq-Feet) based on Building Permit Applications in the Tamiami Trail Corridor

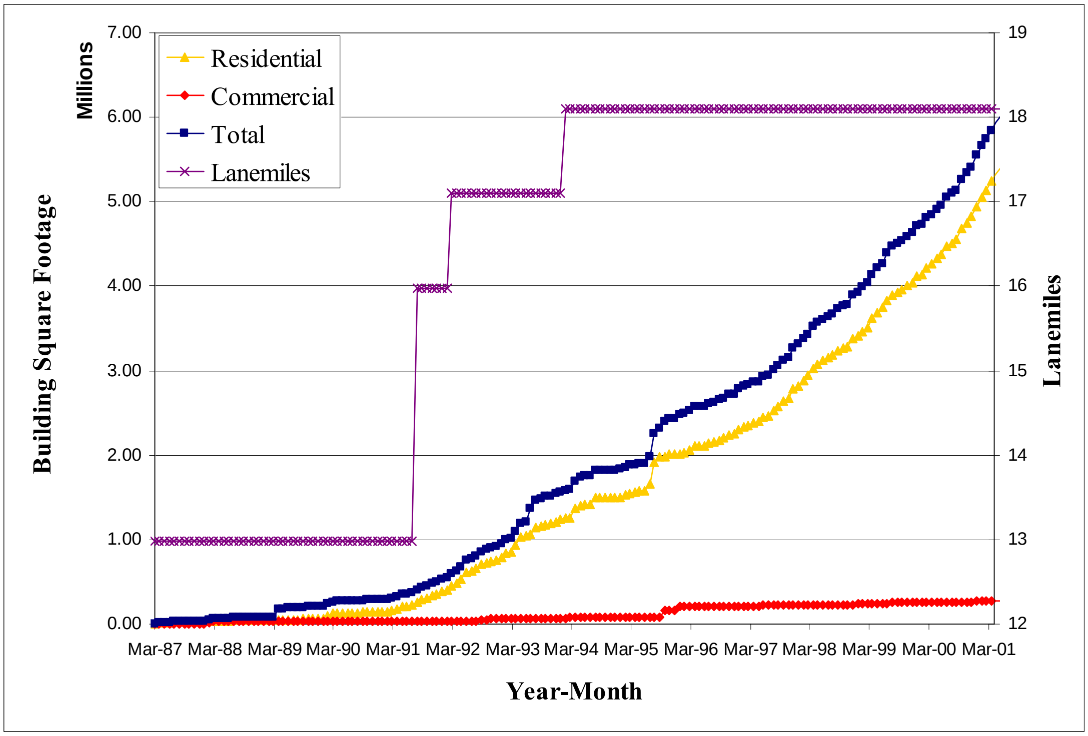


Figure 6.26 Cumulative Floor Space (Sq-Feet) based on Building Permit Applications in the Bird Drive Corridor

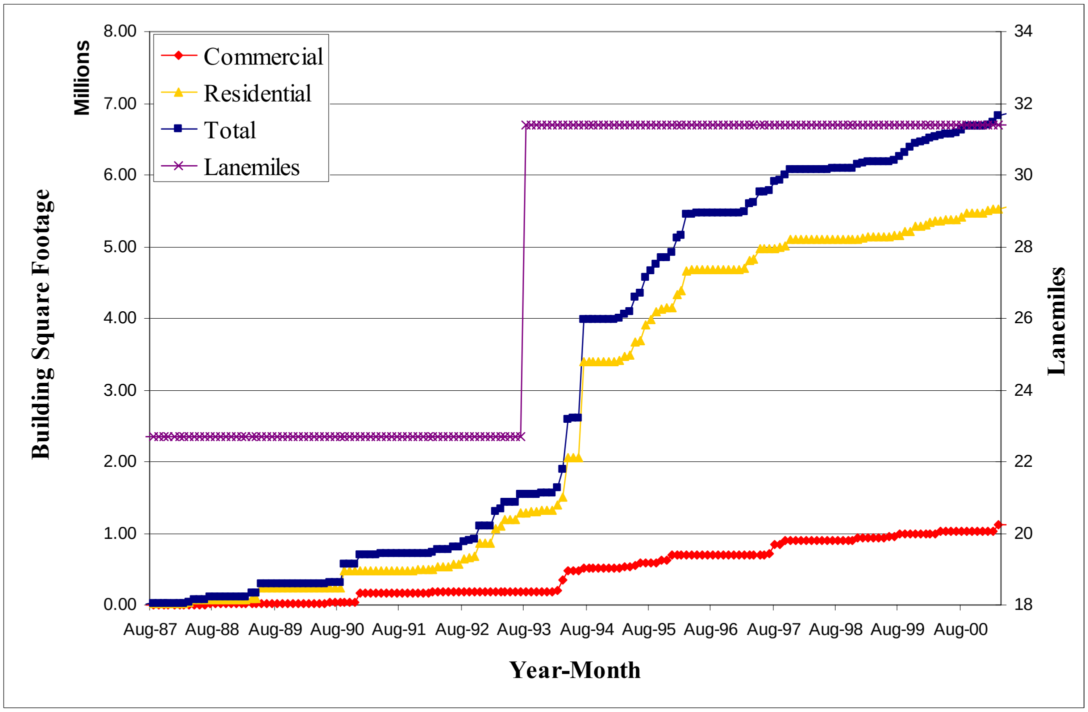


Figure 6.27 Cumulative Floor Space (Sq-Feet) based on Building Permit Applications in the North Kendall Drive Corridor

In Figure 6.28 the development densities are plotted for the period between 1987 and 2000 for the unincorporated Miami-Dade County and the three study corridors. To calculate the densities, the sum of building floor areas from the property tax database was calculated for each year, which was then divided by the total areas for the same year excluding roadways. It may be seen that the density growth rate in unincorporated Miami-Dade County remained steady between 1987 and 2000, while the densities as well as the density growth rates in the Bird Drive and North Kendall Drive corridors were both higher than those of the unincorporated Miami-Dade County. Tamiami Trail corridor area had the lowest development density and also the lowest rate of density increase. The difference may not be adequately explained by transportation, as other market factors also play important roles.

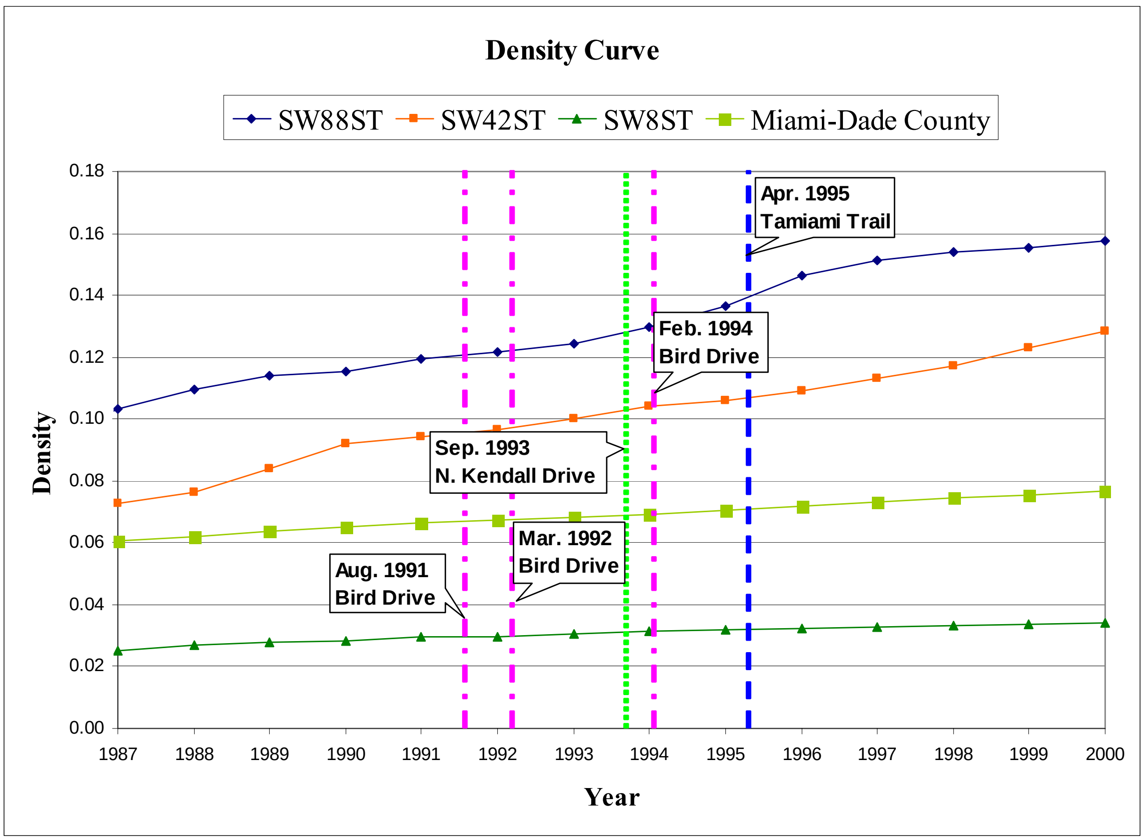


Figure 6.28 Density of Development in the Study Areas and Unincorporated Miami-Dade County

7. CONCLUSIONS

Land use and transportation interaction has been a research topic for several decades. Many theories and models have been suggested to study this well-known, yet extremely complex, process. Using a confirmatory analysis approach based on prior theories and models certainly has helped us gain some insights into this complex process of land use and transportation interaction. However, empirical studies have suggested that land use and transportation interaction patterns can be highly variable in different geographic areas. Each geographic area tends to have its own unique characteristics that may lead to a different pattern of land use and transportation interactions. The patterns are also likely to vary as we examine them at different spatial and temporal scales. As geographic processes often exhibit properties of both spatial dependency and spatial heterogeneity, the challenge is to identify the spatiotemporal patterns underlying these complex geographic variations.

With recent advancements of GIS technology and research progress on temporal GIS, we are now equipped with better tools to tackle complex geographic processes. This project successfully implemented a temporal GIS, coupled with an exploratory analysis approach, that allow systematic and interactive ways of analyzing land use and transportation interaction among various data sets and at user-selected spatial and temporal scales. Temporal GIS databases implemented in this project makes it feasible for the analysis of spatiotemporal interaction patterns in a more efficient and effective way than the conventional snapshot GIS approach. Extending Sinton's measurement framework into a spatiotemporal interaction framework also provides a systematic means of exploring land use and transportation interactions.

In addition to develop temporal GIS tools to support the study of land use and transportation interactions and to support land use and transportation modeling, historical time series data on land use and transportation have been analyzed for selected corridors in the Miami-Dade County. For corridor level analysis, the geographic scale of land use data in temporal GIS databases should be property parcel to accurately reflect the development pattern and the temporal scale should be month.

Both univariate models and multivariate models were developed. Since the dominant type of land developments was residential, data for residential developments had variations whereas many data points for commercial developments were found to be 0. Therefore, residential models were well expressed in terms of roadway improvements and had shown more significant impacts from roadway improvements than commercial models had.

The magnitude of impacts was quantified in terms of building square feet from building permit applications per unit lane mile increase. Cumulated impacts on land use in the Tamiami Trail and North Kendall Drive corridors were larger than that in the Bird Drive corridor. The v/c ratios estimated from historical traffic data in the Tamiami Trail and North Kendall Drive corridors indicated that several sections had low levels of service before the improvements and traffic on those sections increased after the improvements, likely the results of both new developments and traffic diversion and other travel behavior changes.

The lag effects between transportation improvements and land-use development were examined.

By comparing the results of model estimations, it was found that in two principal arterial corridors, Tamiami Trail and North Kendall Drive, the responses from land-use development to roadway improvements were faster than that in the Bird Drive corridor. A plausible explanation may be that the traffic congestions in these two corridors were more severe and were constraining the development because of the Florida concurrency requirements. Consequently, increased accessibility might have resulted in faster response in land use. Impacts of roadway improvement on land use may continue to be felt under the conditions that population and employment growth continues in the urban area, that developable land is still available, and that traffic in the corridor does not reach saturation level, among other market factors.

Estimated lag effect for Lag1, which is the time span between the completion of a roadway expansion project and increased building permit applications by developers for new developments, ranged from a few months to one and one half years, which was shorter than lags found in Cervero's study (2003) (two to three years). However, it should be pointed out that land development may respond much earlier to the transportation improvement projects. The developers may have begun assembling land when the transportation projects were planned. Such activities, however, are more difficult to detect and track.

The averages of Lag2 and Lag3, which are institutional lags for the building department to review applications and issue permits and the construction period for new developments, respectively, were found to be four months and 10 months with standard deviations of four and nine months, respectively. Consequently, it took two to four years for travel demand to respond to road investments in the growing areas in Miami-Dade County.

The historical data available to this study were still limited in the sense that there was only one significant improvement in the transportation system in each of the study corridor. A longer modeling period that includes more than one cycle of transportation improvement-land use development will allow a better understanding of the transportation-land use interactions and improvement of the accuracy of the models. Additionally, the lack of traffic data also prevented the effects of congestion on land development to be adequately considered. Preservation of historical data remains a serious challenge.

8. RECOMMENDATIONS

Given the state-of-the-art nature of this research topic, there are many follow-up research opportunities to further improve the results from this project. The spatiotemporal interaction framework used in this study offers a general structure to lead users in a systematic manner to explore land use and transportation interactions. This framework, however, may be too cumbersome to follow for transportation planners with limited GIS background. Additional custom application tools are needed to make it a more user-friendly system. Integration of existing land use and transportation models into the system will be another possible improvement area. The rich set of temporal GIS databases can be used to generate input data for the existing land use and transportation models. Results from these models can be compared with the interaction patterns identified from spatiotemporal exploratory analyses to validate the models. The temporal GIS design in this system also has room for other improvements. For example, temporal GIS databases can quickly grow into large files. Therefore, innovative approaches to database design and spatiotemporal analysis procedures are needed to achieve a reasonable performance level. Additionally, remote sensing data, especially with the availability of high-resolution images, may be useful for analyzing land use and transportation interactions. However, incorporation of remote sensing data will require an extension of the temporal exploratory GIS design developed in this project to facilitate the simultaneous analysis of both vector and raster GIS data. These improvements will help make spatiotemporal exploratory data analysis a useful approach to studying land use and transportation interaction.

As most data sets available to this project were at an annual or a coarser temporal scale, the temporal resolution level chosen for the implementation of this project was year. It should be noted that the data model and the computer programs developed to time-stamp snapshot GIS databases and to create space-time composite databases are capable of handling temporal data at any resolution levels (i.e., year, month, day, hour, minute, or second). Various data sets may be collected and stored at different temporal scales. However, care must be taken if various databases are of different temporal resolutions. For example, if a user performs an analysis of data from annual land use databases and hourly traffic counts databases, the results could be misleading since the traffic counts data may vary by hours, days, or months while the land use data may remain constant in any given year. In this case, the traffic count data should be aggregated to AADT before a spatiotemporal analysis is performed. When data sets become available at finer temporal resolution levels, the temporal GIS user interfaces developed in this project can be modified to store and manage different temporal scales (e.g., monthly, daily, or hourly). For example, the temporal scales shown in the dialog windows, developed as Visual Basic forms, may be changed from years to months, days, or hours. The computer programs will automatically retrieve and compare the temporal data to ensure consistent temporal scales before proceeding with spatiotemporal analyses.

A critical aspect of this and similar studies is the preservation of historical land use and transportation data, the latter including both transportation improvement project data and traffic and LOS data. In addition to being useful to empirical studies, disaggregate microsimulation models of land use and transportation also need a significant amount of historical data. With the advances in information technology including temporal GIS and ever-improving information systems at public agencies, more data will beocome available in the future. However, it is extremely important for different agencies and even different divisions within the same organization to share data and develop an integrated and shared information system in place of today's often separate, disconnected, and in some cases, incompatible or even conflicting databases. Establishment of enterprise databases including GIS databases will involve coordination among different agencies as well as modification of the business processes.

In allocating growth in the long-range planning process, it will be necessary to consider the growth trend in the past, the "need" for commercial development in an area, the congestion level, and the availability of developable land. For a highly developed area with considerable congestion, transportation improvements may result in new land developments within a shorter period of time. In other words, the initial lags may be shorter. For an area with a significant amount of developable land and low congestion level, the lags may be longer. However, the results from this study cannot be generalized because of the limited number of case studies. Therefore, the study pointed to the need to consider the lag effect in developing land use data.

This study focused on finding empirical evidence to support land use and transportation modeling. The full market mechanism was not modeled but is necessary for a land use model to capture in order to take into consideration of other factors such as developable land, land price, housing price, quality of schools, accessibility, etc. While results from this study supports more detailed analysis of land use and transportation interactions, more powerful models that are based on economics theories are needed to explain the mechanism of their interactions.

REFRENCES

Abelson, P.W. and D.A. Hensher (2001). *Induced travel and user benefits: Clarifying Definitions and Measurement for Urban Road Infrastructure*, Working paper No. ITS-WP-01-12, Institute of Transport Studies, University of Sydney, Australia.

Adams, J.S. (1970). Residential Structure of Mid-Western Cities. *Annals of the Association of American Geographers*, Vol. 60, pp. 37-62.

Adams, T.M., N.A. Koncz, and A.P. Vonderohe (2001). *Guidelines for the Implementation of Multimodal Transportation Location Referencing System*. Transportation Research Board, National Research Council, Washington D.C.

Alonso, W. (1964). *Location and Land Use*. Harvard University Press, Cambridge, MA.

Anas, A. (1982). *Residential Location Models and Urban Transportation: Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models*. Academic Press, New York.

Anas, A. and L.S. Duann (1986). Dynamic Forecasting of travel Demand, Residence Location, and Land Development: Policy Simulations with the Chicago Area Transportation/Land Use Analysis System. In Hutchinson, B.G. and Batty, M. (Eds.), *Advances in Urban Systems Modelling*. North-Holland, Amsterdam, the Netherlands, pp. 299-322.

Barr, L.C. (2000). *Testing for Significance of Induced Highway Travel Demand in Metropolitan Areas*. The 79th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Bland, B.H. (1979). *Land Use Patterns and the Viability of Bus Services: A Study Using the Transport and Road Research Laboratory LUTE Model*. Presented at the Planning and Transportation Research and Computation Summer Annual Meeting, Warwick University, England.

Boarnet, M.G. and A.F. Haughwout (2001). *Do Highways Matter? - Evidence and Policy Implications of Highways' Influence on Metropolitan Development*, Discussion Paper, the Brookings Institution Center on Urban and Metrolitan Policy.

Booch, G. (1993). *Object-Oriented Analysis and Design with Applications*, 2nd edition. Benjamin Cummings, Redwood City, CA.

Box, G.E.P. and G.M. Jenkins (1976). *Time Series Analysis: forecasting and control*, Holden-Day Inc., San Francisco, CA.

Boyce, D.E. (1980). A Framework for Constructing Network Equilibrium Model of Urban Location. *Transportation Science*, Vol. 14, pp. 77-96.

Boyce, D.E. (1990). Network Equilibrium Models of Urban Location and Travel Choice: New Research Agenda. In Chatterji, M. and Kuene, R. (Eds.), *New Frontiers in Regional Science,* Vol. 1. Macmillan Publishers Limited, London, UK, pp. 238-256.

Buffington, J.L., L.M. Crane, B. Clifton, and J. Speed (1992). Methodology for Estimating Economic Impacts of Highway Improvements: Two Case Studies in Texas. *Transportation Research Record 1359*, Transportation Research Board, National Research Council, Washington, D.C., pp. 156-165.

Buffington, J.L., M.K. Chui, and J.L. Memmott (1985). Effects of Freeway Stage Construction on Nearby Land Uses and Vehicle User Costs. *Transportation Research Record 1046*, Transportation Research Board, National Research Council, Washington, D.C., pp. 62-69.

Butler, J.A. and Dueker, K.J. (2001). Implementing the Enterprise GIS in Transportation Database Design. Manuscript submitted to *Journal of the Urban and Regional Information Systems Association*, available at www.urisa.org.

Cascetta, E., A. Nuzzolo, and P. Coppola (2000). A System of Mathematical Models for the Simulation of Land-Use and Transport Interaction. *Proceedings of the Sixth International Conference on Urban Transport and the Environment for the 21st Century*, pp. 167-177.

Cervero, R. and S. Seskin (1995). An Evaluation of the Relationships between Transit and Urban Form. *TCRP Research Results Digest*, No. 7, Transportation Research Board, National Research Council, Washington, D.C.

Cervero, R. (2003). Road Expansion, Urban Growth and Induced Travel: a Path Analysis. *Journal of the American Planning Association*, Vol.69, No. 2, pp.145-161.

Chen, P. (1991). *The Entity Relationship Approach to Logical Database Design*. Wellesley, QED Information Sciences, MA.

Cohen, H.S. (1995). *Review of Empirical Studies of Induced Traffic, Special Report 245: Expanding Metropolitan Highways: Implications for Air Quality and Energy Use*, Transportation Research Board, National Research Council, Washington, D.C., pp. 295-309.

Cosby, P.J. and J.L. Buffington (1978a). *Land Use Impact of Widening State Highway 30 in an Undeveloped Area in College Station, Texas*, Research Report 225-3, Texas Transportation Institute, Texas A&M University, College Station, TX.

Cosby, P.J. and J.L. Buffington (1978b). *Land Use Impact of Widening East 29th Street in a Developed Residential Area in Bryan, Texas*, Research Report 225-5, Texas Transportation Institute, Texas A&M University, College Station, TX.

Cosby, P.J. and J.L. Buffington (1979). *Land Use Impact of Improving Collins in a Developed Area of Arlington, Texas*, Research Report 225-13, Texas Transportation Institute, Texas A&M University, College Station, TX.

DeCorla-Souza, P. and H. Cohen (1998). *Accounting for Induced Travel in Evaluation of Metropolitan Highway Expansion*. The 77th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Dowling, R.G. and S.B. Colman (1998). Effects of Increased Highway Capacity: Results of a Household Travel Survey. *Highway Capacity Expansion and Induced Travel – Evidence and Implications*, *Transportation Research Circular*, No. 481, Transportation Research Board, National Research Council, Washington, D.C., pp. 12-32.

Downs, A. (1962). The Law of Peak-Hour Expressway Congestion. *Traffic Quarterly*, Vol. 16, pp. 393-409.

Dueker, K.J. and J.A. Butler (1998). GIS-T Enterprise Data Model with Suggested Implementation Choices. *Journal of the Urban and Regional Information Systems Association*, Vol. 10, No. 1, pp. 12-36.

Dunphy, R.T. (1998). Widening the Roads: Data Gaps and Philosophical Problems, in *Highway Capacity Expansion and Induced Travel: Evidence and Implications*, Transportation Research Board, National Research Council, Washington, D.C., pp. 16-20.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, Vol. 50, pp. 987-1007.

Federal Geographic Data Committee (FGDC) (2001). *NSDI Framework Transportation Identification Standard—Public Review Draft*. Bureau of Transportation Statistics, United States Department of Transportation, Washington, D.C.

Flink, J.J. (1975). *The Car Culture*. MIT Press, Cambridge, Massachusetts.

Forkenbrock, D.J. (2001). *Transportation Investments and Urban Form*. Presented at the 81st Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Fujii, S. and R. Kitamura (2000). Evaluation of Trip-Inducing Effects of New Freeways Using a Structural Equations Model System of Commuters' Time Use and Travel. *Transportation Research B*, Vol. 34, pp. 339-354.

Fulton, L.M., D.J. Meszler, R.B. Noland, and J.V. Thomas (2000). A Statistical Analysis of Induced Travel Effects in the U.S. Mid-Atlantic Region. The 79th Annual Meeting of the Transportation Research Board, Washington, D.C.

Giuliano, G. (1995). Land Use Impacts of Transportation Investments: Highway and Transit. In Hansen, S. (Ed.), *The Geography of Urban Transportation*. Guilford Press, New York, pp. 305-341.

Goodchild, M.F. (2000). Spatial Analysis: Methods and Problems in Land Use Management. In Hill, M.J. and Aspinall, R.J. (Eds.), *Spatial Information for Land Use Management*. Gordon and Breach Science Publishers, Amsterdam, the Netherlands, pp. 39-51.

Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, Vol. 37, No. 3, pp. 424-438.

Hamilton, D.J. (1994). *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Hansen, M. (1998). The Traffic Inducement Effect: Its Meaning and Measurement. *Transportation Research Circular*, Vol. 481, pp. 7-15, Transportation Research Board, National Research Council, Washington, D.C.

Hansen, M. and Y. Huang (1997). Road Supply and Traffic in California Urban Areas. *Transportation Research A,* Vol. 31, pp. 205-218.

Hansen, M., D. Gillen, A. Dobbins, Y. Huang, and M. Puvathingal (1993). *The Air Quality Impacts of Urban Highway Capacity Expansion: Traffic Generation and Land Use Change*, Institute of Transportation Studies, University of California, Berkeley.

Hansen, W.G. (1959). How Accessibility Shapes Land Use. *Journal of the American Institute of Planners,* 25, pp. 73-76.

Headicar, P. (1996). The Local Development Effects of Major New Roads: M40 Case Study. *Transportation*, Vol. 23, pp. 55-69.

Heanue, K. (1998). Highway Capacity and Induced Travel: Issues, Evidence and Implications. *Transportation Research Circular*, No. 481, Transportation Research Board, National Research Council, Washington, D.C., pp. 33-45.

Herndon, C. W. Jr. and J. L. Buffington (1979). *Land Use Impact of Widening West 43rd Avenue in a Developed Residential Area of Houston, Texas*, Research Report 225-7, Texas Transportation Institute, Texas A&M University, College Station, Texas.

Hirschman, I. and M. Henderson (1990). Methodology for Assessing Local Land Use Impacts of Highways. *Transportation Research Record 1274*, Transportation Research Board Meetings, National Research Council, Washington, D.C., pp. 35-40.

Huang, W. (1994). *The Effects of Transportation Infrastructure on Nearby Property Values: A Review of Literature*. Working Paper 620, Institute of Urban and Regional Development, University of California at Berkeley.

HUD (2000). *The State of the Cities, Annual Report*. The US Department of Housing and Urban Development, Washington, D.C.

Hunt, J.D. (1997). *A Description of the MEPLAN Framework for Land Use and Transport Interaction Modeling*. The 77th Annual Transportation Research Board Meetings, Washington, D.C.

Jacobson, I. (1992). *Object-Oriented Software Engineering*. Reading, Addison-Wesley, MA.

Johnston, J. and J. Dinardo (1997). *Econometric Methods*, 4th edition, McGraw-Hill, New York, NY.

Kain, J.F. (1970). Postwar Changes in Land Use in the American City. In *Toward a National Urban Policy*, D.P. Moynihan (ed.), Basic Book, NY.

Kawamura, K. (2001). *Empirical Examination of the Relationship between Firm Location and Transportation Facilities*. The 80th Annual Meeting of the Transportation Research Board, Washington, D.C.

Kiefer, M. and S.R. Mehndiratta (1998). *If We Build It, Will They Really Keep Coming? - A Critical Analysis of the Induced Demand Hypothesis*. 1998 TRB Annual Meeting Preprint Paper #980937, Transportation Research Board, National Research Council, Washington, D.C.

Kim, T.J. (1983). A Combined Land Use-Transportation Model When Zonal Travel Demand is Endogenously Determined. *Transportation Research B,* Vol. 17B, pp. 449-462.

Kim, T.J. (1989*). Integrated Urban System Modeling: Theory and Applications*. Kluwer Academic Publishers, Boston, Massachusetts.

Kim, T.J., J.H. Rho, and S. Suh (1989). *Integrated Urban Systems Modeling: Theory and Applications*. Kluwer Academic Publishers, Dordrecht, Netherlands.

Knight, R. and L. Trygg (1977). Evidence of Land Use Impacts of Rapid Transit Systems. *Transportation*, Vol. 6, pp. 231-247.

Koncz, N. and T.M. Adams (2001). A Data Model for Multidimensional Transportation Locational Referencing Systems. Manuscript submitted to *Journal of the Urban and Regional Information Systems Association*, available at www.urisa.org.

Ladner, R., K. Shaw, and M. Abdelguerfi (eds.) (2002). *Mining Spatio-Temporal Information Systems*. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Langran, G. (1992). *Time in Geographic Information Systems*. Taylor & Francis, London, United Kingdom.

Langran, G. and N. Chrisman, (1988). A Framework for Temporal Geographic Information. *Cartographica,* 25, pp. 1-14.

LANL (1999). *Version 1.0 Documentation, Volume 2 – Software, Part 1 – Modules*, Los Alamos National Laboratory, LA-UR-99-2574.

Litman, T. (1999). *Generated Traffic – Implications for Transport Planning*. Technical Report, Victoria Transport Policy Institute, Victoria, BC, Canada.

Lomax, T. and D. Schrank (1996). *Trends in Urban Congestion: 1982 – 1993*. Technical Report, Texas Transportation Institute, College Station, TX.

Lütkepohl, H. (1993). Introduction to Multiple Time Series Analysis, Second Edition, Springer-Verlag.

MacDonald, A. (1999). *Building a Geodatabase*. ESRI Press, Redlands, CA.

Mackett, R.L. (1990a). *MASTER Model* (*Micro-Analytical Simulation of Transport, Employment and Residence*). Contract Report 237, Transport and Road Research Laboratory, Transport Research Laboratory, Crowthorne, Berkshire, England.

Mackett, R.L. (1990b). Comparative Analysis of Modeling Land-Use Transport Interaction at the Micro and Macro Levels. *Environment and Planning*, Vol. 22 A, pp 459-475.

Marshall, N.L. (2000). *Evidence of Induced Demand in the Texas Transportation Institute's Urban Roadway Congestion Study Data Set*. The 79th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Middleton, W. (1967). *The Time of the Trolley*. Kalmbach Publishing, Milwaukee, WI.

Miller, H.J. and J. Han (eds.) (2001). *Geographic Data Mining and Knowledge Discovery*. Taylor and Francis, London.

Moynihan, D.P. (1960). New Roads and Urban Chaos. *The Reporter* (April 14), pp. 13-20.

Muller, P.O. (1986). Transportation and Urban Form: Stages in the Spatial Evolution of the American Metropolis. *The Geography of Urban Transportation*, pp. 24-48.

Noland, R. (2001). "Relationships between Highway Capacity and Induced Vehicle Travel," *Transportation Research A,* Vol. 35, pp. 47-72. www.epa.gov/tp/trb-rn.pdf.

Noland, R. and W.A. Cowart (2000). Analysis of Metropolitan Highway Capacity and the Growth in Vehicle Miles of Travel. *Transportation*, Vol. 27, No. 4, pp. 263-290.

Openshaw, S. (1994). Two Exploratory Space-Time-Attribute Pattern Analysers Relevant to GIS. In Fotheringham, S. and Rogerson, P. (Eds.), *Spatial Analysis and GIS*. Taylor and Francis, London, pp. 83-104.

Palmquist, R. (1982). Impact of Highway Improvements on Property Values in Washington State. *Transportation Research Record 887*, Transportation Research Board, National Research Council, Washington, D.C., pp. 22-29.

Payne-Maxie Consultants (1980). *The Land Use and Urban Development Impacts of Beltways*. Final Report No. DOT-OS-90079, U.S. Department of Transportation and Department of Housing and Urban Development, Washington, D.C.

Peuquet, D.J. (1994). It's about time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Annals of the Association of American Geographers*, 84(3), pp. 441-461.

Peuquet, D.J. and N. Duan (1995). An Event-Based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data. *International Journal of Geographical Information Systems*, Vol. 9, pp. 7-24.

Prastacos, P. (1986). An Integrated Land-Use-Transportation Model for the San Francisco Region: 1. Design and Mathematical Structure. *Environment and Planning*, Vol. 18 A, pp. 307-322.

Putman, S.H. (1983). *Integrated Urban Models*. Pion Limited, London. England.

Ramsey, J.B. (1969). Tests for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistical Society*, Series B, Vol. 31, pp. 350-371.

Reinsel, G.C. (1997). *Elements of Multivariate Time Series Analysis*, Second Edition, New York: Springer-Verlag.

Richter, C. (1999). *Designing Flexible Object-Oriented Systems with UML*. Macmillan Technical Publishing, Indianapolis, IN.

Ryan, S. (1999). Property Values and Transportation Facilities: Finding the Transportation-Land Use Connection. *Journal of Planning Literature*, Vol. 13, No. 4, pp. 412-427.

SACTRA (1994). *Truck Roads and the Generation of Traffic*. Standard Advisory Committee on Truck Road Assessment, Department of Transport (UK), HMSO, London.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, Vol. 6, pp. 461-464.

Sinton, D.F. (1978). The Inherent Structure of Information as a Constraint to Analysis: Mapped Thematic Data as a Case Study. In Dutton, G. (Ed.), *Harvard Papers on Geographic Information Systems*, Vol. 6, Addison-Wesley, Reading, MA, pp. 1-17.

Smerk, G. (1967). The Streetcar: Shaper of American Cities. *Traffic Quarterly*, October, pp. 134-146.

Smith, W. (1984). Mass Transit for High-Rise, High-Density Living. *Journal of Transportation Engineering*, ASCE, Vol. 110, No. 6, pp. 521-535.

Southworth, F. (1995). *A Technical Review of Urban Land Use-Transportation Models as Tools for Evaluating Vehicle Travel Reduction Strategies*. Technical Report ORNL-6881, U.S. Department of Energy, Washington, D.C.

Strathman, J.G., K.J. Dueker, T. Sanchez, J. Zhang, and A.E. Riis (2000). *Analysis of Induced Travel in the 1995 NPTS*. Center for Urban Studies, Portland State University, Portland, OR.

Swenson, D., L. Eathington, and D. Otto (1998). Economic Growth, Property Valuation Change, and Transportation Investments. *Crossroads 2000*, pp. 227-231.

TRB (1995). *Special Report 245: Expanding Metropolitan Highways: Implications for Air Quality and Energy Use*. Committee for Study of Impacts of Highway Capacity Improvements on Air Quality and Energy Consumption, TRB, National Research Council, Washington, D.C.

Tukey, J.W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

UrbanSim (2001). *UrbanSim Description*, http://www.urbansim.org/.

Vadali, S.R. and C. Sohn (2001). *Tracking Changes in Spatially Segregated Location Premiums Using GIS: An Alternative Evaluation Method of Assessing Residential Land Use Impacts of Transportation Projects*. The 80th Annual Meeting of the Transportation Research Board, National Research Council, Washington, D.C.

Visio, Corp. (2000). *Microsoft Visio 2000 Online Help Document*.

Voith, R. (1992). City and Suburban Growth: Substitutes or Complements? *Business Review*, Federal Reserve Bank of Philadelphia, Philadelphia, Pennsylvania.

Vonderohe, A.P., C.L. Chou, F. Sun, and T.M. Adams (1997). A Generic Data Model for Linear Referencing Systems. *National Cooperative Highway Research Results Digest 218*, Transportation Research Board, National Academy Press, Washington, D.C.

Waddell, P. (2001). UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. *Journal of the American Planning Associatio*n, forthcoming.

Wegener, M. (1986). Transport Network Equilibrium and Regional Deconcentration. *Environment and Planning*, Vol. 18 A, pp. 437-456.

Wegener, M. (1995). Current and Future Land Use Models. In Travel Model Improvement Program Land Use Modeling Conference Proceedings, pp. 13-40, Travel Model Improvement Program, U.S. Department of Transportation, Washington, D.C.

Woodhull, J. (1992). How Alternative Forms of Development Can Reduce Traffic Congestion. In *Sustainable Cities: Concepts and Strategies for Eco-City Development.* Eco-Home Media.

Worboys, M.F., (1998). A Generic Model for Spatio-Bitemporal Geographic Information. In M.J. Egenhofer and R.G. Golledge (Eds.), *Spatial and Temporal Reasoning in Geographic Information Systems*. Oxford University Press, NY, pp. 25-39.

Yuan, M. (1996). Modeling Semantics, Temporal, and Spatial Information in Geographic Information Systems. In Craglia, M. and Couclelis, H. (Eds.), *Geographic Information Research: Bridging the Atlantic*. Taylor and Francis, London, pp. 334-347.

Yuan, M. (1999). Use of a Three-Domain Representation to Enhance GIS Support for Complex Spatiotemporal Queries. *Transactions in GIS*, Vol. 3(2), pp. 137-159.

Zeiler, M. (1999). *Modeling Our World: The ESRI's Guide to Geodatabase Design*. ESRI Press, Redlands, CA.

Zeiler, M. (ed.) (2001). *Exploring ArcObjects*. ESRI Press, Redlands, CA.

APPENDIX A. REFERENCE LIST OF TEMPORAL DATABASE AND SPATIOTEMPORAL GIS

Abraham, T.R. and J.F. Roddick (1999). Survey of Spatio-Temporal Databases. *GeoInformatica*, Vol. 3, pp. 61-99.

Agouris, P. and A. Stefanidis (1999). Integrated Spatial Databases: Digital Images and GIS (International Workshop ISD '99, Portland, ME). *Lecture Notes in Computer Science*, Vol. 1737, Springer.

Alesheikh, A.A. (1998). *Modeling and Managing Uncertainty in Object-Based Geospatial Information Systems*, Civil Engineering/Computer Science, University of Calgary, Canada.

Allen, K.M.S. (1988). Trade Networks and European Contact: A Case Study Using Geographic Information Systems. *Proceedings of the 3rd International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 367-385.

Allen, K.M.S., S.W. Green, and E.B.W. Zubrow (Eds.) (1990). *Interpreting Space: GIS and Archaeology*. Taylor & Francis.

Al-Taha, K. and A. Frank (1993). What a Temporal GIS can do for Cadastral Systems. *Proceedings of the First Sharjah Conference on Geographic Information Systems and Applications GISA 93*, Sharjah Municipality, pp. 13.1-13.17.

Al-Taha, K.K. (1992). Temporal Reasoning in Cadastral Systems, Ph.D. Dissertation, Department of Surveying Engineering, University of Maine, ME.

Al-Taha, K.K. and R. Barrera (1990). Temporal Data and GIS: An Overview. *Proceedings of GIS/LIS '90*, Bethesda, MD, pp. 244-254.

Al-Taha, K.K. and A. Frank (1991). Temporal GIS Keeps Data Current. *1991-1992 International*

GIS Sourcebook, H. D. Parker (Ed.), GIS World Inc., pp. 384-388.

Al-Taha, K.K., R.T. Snodgrass, and M.D. Soo (1993). Bibliography on Spatiotemporal Databases. *SIGMOD*, Vol. 22, pp. 59-67.

Al-Taha, K.K., R.T. Snodgrass, and M.D. Soo (1994). Bibliography on Spatiotemporal Databases. *International Journal of Geographical Information Systems*, Vol. 8, pp. 95-103.

Aref, W.G. and H. Samet (1991). Extending a DBMS with Spatial Operations. *Advances in Spatial Databases Lecture Notes in Computer Science 525*, O. Gunther and H.J. Schek (Eds.), Springer-Verlag, pp. 299-318.

Ariav, G. (1984). Preserving the Time Dimension in Information Systems, Ph.D. Dissertation, Computer Science Department, University of Pennsylvania, PA.

Armenakis, C. (1992). Estimation and Organization of Spatio-Temporal Data. *Proceedings of the Canadian Conference on GIS '92*, Ottawa, Canada, The Canadian Institute of Geomatics, pp. 900-911.

Armenakis, C. (1993). Map Animation and Hypermedia: Tools for Understanding Changes in Spatio-temporal Data. *Proceedings of the Canadian Conference on GIS '93*, Ottawa, Canada, The Canadian Institute of Geomatics, pp. 859-868.

Armenakis, C. and E.M. Sierkierska (1991). Issues on the Visualization of Time-Dependent Geographical Information. *Proceedings of the Canadian Conference on GIS '91*, Ottawa, Canada, The Canadian Institute of Geomatics, pp. 584-595.

Armstrong, M. (1988). Temporality in Spatial Databases. *Proceedings of GIS/LIS '88*, Bethesda, MD, ASPRS/ACSM, pp. 880-889.

Bagg, J. and N. Ryan (1997). Modeling Historical Change in Southern Corsica: Temporal GIS Development Using an Extensible Database System. *Innovations in GIS 4*, Z. Kemp (Ed.), Taylor & Francis, pp. 42-55.

Baker, B. and P. Gersmehl (1991). Temporal Trends in Soil Productivity Evaluation. *Professional Geographer*, pp. 304-318.

Barrera, R. and K.K. Al-Taha (1990). *Models in Temporal Knowledge Representation and Temporal DBMS*, Technical Report, NCGIA, University of Maine, ME.

Barrera, R., A. Frank, and K. Al-Taha (1990). Temporal Relations in GIS: A Workshop. *Temporal GIS Workshop*.

Barrera, R., A. Frank, and K.K. Al-Taha (1991). Temporal Relations in Geographic Information Systems: A Workshop at the University of Maine. *SIGMOD*, Vol. 20, pp. 85-91.

Basoglu, U. and J. Morrison (1978). The Efficient Hierarchical Data Structure for the U.S. Historical Boundary File. *Harvard Papers on GIS*, G. Dutton (Ed.), Addison-Wesley, pp. 1-26.

Batty, M. and Y. Xie (1994a). Modeling Inside GIS: Part 1. Model Structures, Exploratory Spatial Data Analysis and Aggregation. *International Journal of Geographical Information Science*, Vol. 8, pp. 291-307.

Batty, M. and Y. Xie (1994b). Modeling Inside GIS: Part 2. Selecting and Calibrating Urban Models Using ARC-INFO. *International Journal of Geographical Information Systems*, VOl. 8, pp. 451-470.

Bell, S.B.M., R.A. Chadwick, A.P.R. Coopert, D.C. Mason, M. O'Connell, and P.A.V. Young (1990). Handling Four Dimensional Geo-Coded Data. *Proceedings of the 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission of GIS, pp. 918-927.

Beller, A. (1991). Spatial/Temporal Events in a GIS. *Proceedings of GIS/LIS '91*, Bethesda, MD, ASPRS/ACSM, pp. 766-775.

Beller, A., T. Giblin, L.K. Litz, and D. Schimel (1991). A Temporal GIS Prototype for Global Change Research. *Proceedings of GIS/LIS '91*, Bethesda, MD, ASPRS/ACSM, pp. 752-765.

Berry, B.J.L. (1964). Approaches to Regional Analysis: A Synthesis. *Annals of the Association of American Geographers*, Vol. 54, pp. 2-11.

Berry, M.W. and R.O. Flamm (1994). The Land-Use Change Analysis System (LUCAS) for Evaluating Landscape Management Decision. Technical Report, Computer Science Department, The University of Tennessee.

Bertino, E., B.C. OOI, R. Sacks-Davis, K.-L. Tan, J. Zobel, B. Shidlovsky, and B. Catania (1997). *Indexing Techniques for Advanced Database Systems*. Kluwer Academic Publishers.

Bimbo, A.D., E. Vicario, and D. Zingoni (1995). Symbolic Description and Visual Querying of Image Sequences Using Spatio-Temporal Logic. *TKDE*, Vol. 7, pp. 609-622.

Blaser, A. and M. Egenhofer (2000). A Visual Tool for Querying Geographic Databases. *Advanced Visual Interface -- AVI 2000*, L. Tarantino (Ed.)

Böhlen, M.H., C.S. Jensen, and M.O. Scholl (1999). Spatio-temporal Database Management (Proceedings of International Workshop STDBM'99, Edinburgh, Scotland). *Lecture notes in computer science;* Vol. 1678, Springer.

Bonfatti, F. and P.D. Monari (1994). Spatio-temporal Modeling of Complex Geographical Structure. *Computer Support for Environment Impact Assessment (Proceedings of the IFIP TC5/WG5.11 Working Conference on Computer Support for Environment Assessment, CSEIA 93)*, Como, Italy, Elsevier.

Brodsky, A. (1997). Constraint Databases: Promising Technology or Just Intellectual Exercise?

Constraints, Vol. 2, pp. 35-44.

Brodsky, A.S., V.E. Segal; J. Chen, and P.A. Exarkhopoulo (1997). The CCUBE Constraint Object-Oriented Database System. *Constraints*, Vol. 2, pp. 245-277.

Burrough, B.J.L., W.V. Deursen, and G. Heuvelink (1988). Linking Spatial Process Models and GIS: A Marriage of Convenience or a Blossoming Partnership. *Proceedings of GIS/LIS '88*, Bethesda, MD, ASPRS/ACSM, pp. 598-607.

Calkins, H.W. (1984). Space-time Data Display Techniques. *Proceedings of International Symposium on Spatial Data Handling*, Zurich, Switzerland, Geographisches Institit, Abteiling Kartographie/EDV, pp. 324-331.

Charlton, M.E., S. Openshaw, and C. Wymer (1990). Some Experiments with an Adaptive Data Structure in the Analysis of Space-Time Data. *Proceedings of the 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 1030-1039.

Chen, J. and J. Jie (2000). An Event-Based Approach to Spatio-Temporal Data Modeling in Land Subdivision Systems. *GeoInformatica*, Vol. 4, pp. 387-402.

Cheng, T.S. and S.K. Gadia (1992). A Seamless Object-Oriented Model for Spatio-Temporal Databases.

Cheng, T.S. and S.K. Gadia (1994). A Pattern Matching Language for Spatio-Temporal Databases. *Proceedings of the Conference on Information and Knowledge Management*. Gaithersburg, MD.

Cheng, T.S., S.K. Gadia, and S. Nair (1992). Relational and Object-Oriented Parametric Databases. TR-92-42, Computer Science Dept., Iowa State University.

Choi, A. and W.S. Luk (1992). Using an Object-oriented Database System to Construct a Spatial Database Kernel for GIS Applications. *Computer Systems and Engineering*, pp. 100-121.

Chrisman, N. (1983). The Role of Quality Information in the Long-Term Functioning of a Geographic Information System. *Proceedings of the Sixth International Symposium on Automated Cartography*, Ottawa, Ontario, Canada, The Steering Committee for the Sixth International Symposium on Automated Cartography, pp. 303-321.

Chrisman, N. (1984). The Role of Quality Information in the Long-Term Functioning of a Geographic Information System. *Cartographica*, Vol. 21, pp. 79-87.

Chrisman, N.R. (1993). Beyond Spatio-temporal Data Models: A Model of GIS as a Technology Embedded in Historical Context. *Proceedings AUTO-CARTO*, Vol. 11, pp. 23-32.

Parent, C. (1998). Modeling Spatial Data in the MADS Conceptual Model. *Proc. of the International Symposium on Spatial Data Handling*, Vancouver, Canada, pp. 11-15.

Claramunt, C. and M. Theriault (1995). Managing Time in GIS: An Event-Oriented Approach. *Recent Advances in Temporal Databases*, Springer-Verlag, pp. 142-161.

Claramunt, C. and M. Theriault (1996). Toward Semantics for Modelling Spatio-temporal Process within GIS. *Advances in GIS Research II: Proceedings of the Seventh International Symposium of Spatial Data Handling*, M.-J. Kraak, M. Molenaar, and E. M. Fendel (Eds.), Taylor & Francis, pp. 47-63.

Clifford, J. (1982). A Model for Historical Databases. *Proceedings of Workshop on Logical Bases for Data Bases*. Toulouse, France, ONERA-CERT (DER-Informatique).

Clifford, J. (1983). Formal Semantics for Time in Database. *ACM Transactions on Database Systems*, Vol. 8, pp. 214-254.

Clifford, J. and A.U. Tansel (1983). Formal Semantics for Time in Databases. *ACM Transactions on Database Systems*, Vol. 8, pp. 214-254.

Clifford, J. and A. Tansel (1985). On an Algebra for Historical Relational Databases: Two Views. *SIGMOD RECORD*, S. Navathe (Ed.), pp. 247-265.

Cu, W.W., I.T. Ieong, and R.K. Taira (1994). A Semantic Modeling Approach for Image Retrieval by Content. *VLDB Journal*, Vol. 3, pp. 445-478.

Davis, B. and R. Williams (1989). The Five Dimensions of GIS. *Proceedings of ACSM/ASPRS*, Bethesda, MD, ACSM/ASPRS, pp. 382-388.

Dikau, R. (1990). Geomorphic Landform Modeling Based on Hierarchy Theory. *Proceedings of Second International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 230-239.

Dobson, M.W. (1986). Spatial Decision Support Systems for the Early Warning of Disaster Driven Social Emergencies. *Proceedings of Second International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 349-352.

Dragicevic, S. and M. Danielle (2000). A Fuzzy Set Approach for Modeling Time in GIS. *International Journal of Geographical Information Science*, Vol. 14, pp. 225-245.

Duke-Sylvester, S.M. and L.J. Gross (1999). *Integrating Spatial Data Into and Agent-based Modeling System: Ideas and Lessons from the Development of the Across Tropic Level System Simulation (ATLSS)*. Technical Report, The Institute for Environmental Modeling (TIEM), The University of Tennessee, Knoxville, TN.

Dutta, S. (1991). Topological Constraints: A Representational Framework for Approximate Spatial and Temporal Reasoning. *Advances in Spatial Databases--Second Symposium SSD '91*, NY, Spinger-Verlag, pp. 161-180.

Easterfield, M., R. Newell, and D. Theriault (1990). Management of Time and Space. SmallWorld Systems Ltd.

Easterfield, M., R. Newell, and D. Theriault (1990). Version Management in GIS-Applications and Techniques, Cambridge, England, Smallworld Systems Ltd.

Easterfield, M., R.G. Newell, and D. Theriault (1991). Modeling Spatial and Temporal Information. *Proceedings of EGIS '91*, Brussels, EGIS Foundation, pp. 294-304.

Edsall, R. and D. Peuquet (1996). Graphical Query Techniques for Temporal GIS. *Proceeding of ACSM/ASPRS Annual Conference,* Seattle, WA, pp. 182-189.

Edwards, G. (1992). The Integration of Remotely Sensed Data Analysis into GIS: Time and Uncertainty Management Needs. *Proceedings of The Canadian Conference on GIS*, Ottawa, Canada, The Canadian Institute of Geomatics, pp. 432-440.

Edwards, G., P. Gagnon, and Y. Benard (1993). Spatio-temporal Topology and Causal Mechanisms in Time-integrated GIS: from Conceptual Model to Implementation Strategies. *Proceedings of The Canadian Conference on GIS '93*, Ottawa, Canada, The Canadian Institute of Geomatics, pp. 842-851.

Effenberg, W.W. (1992). Time in Spatial Information Systems. *First Regional Conference on GIS Research in Vistoria and Tasmania*. Ballarat, Victoria.

Egenhofer, M. (1991). Extending SQL for Graphical Display. *Cartography and Geographic Information Systems*, Vol. 18, pp. 230-245.

Egenhofer, M. (1992). Why Not SQL! *International Journal of Geographical Information Systems*, Vol. 6, pp. 71-85.

Egenhofer, M. (1993). What is Special about Spatial? - Database Requirements for Vehicle Navigation in Geographic Space. *Proceedings of SIGMOD '93 International Conference on Management of Data*, Washington, D.C., ACM Press, pp. 398-402.

Egenhofer, M. (1994). Spatial SQL: A Query-based Presentation Language. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 6, pp. 86-95.

Egenhofer, M. and K. Al-Taha (1992). Reasoning about Gradual Changes of Topological Relationships. A. U. Frank, I. Campari, and U. Formentini (Eds.), *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pp. 169-219.

Egenhofer, M., J. Glasgow, O. Gunther, J.R. Herring, and D.J. Peuquet (1999). Progress in computational methods for representing geographical concepts. *International Journal of Geographical Information Science*, Vol. 13, pp. 775-796.

Erwig, M. (1999). Spatio-temporal Data Types: An Approach to Modeling and Querying Moving Objects in Database. *GeoInformatica*, Vol. 3.

Erwig, M., M. Schneider, and R.H. Guting (Eds.) (1998). *Temporal Objects for Spatio-Temporal Data Models and a Comparison of Their Representations*. *Advances in Database Technologies (ER '98 Workshop on Spatio-Temporal Data Management), LNCS 1552 (1999)*, Springer Verlag, pp. 454-465.

Faria, G., C.B. Mederios, and M.A. Nascimento (1998). *An Extensible Framework for Spatio-temporal Database Applications*. Technical Report, TIMECENTER, Denmark.
Feuchtwanger, M. (1993). Towards A Geographic Semantic Database Model, *Geography/Information Science*.

Fisher, P.F. (1992). Real-Time Randomization for the Visualization of Uncertain Spatial Information. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 491-494.

Flewelling, D., A. Frank, and M. Egenhofer (1992). Constructing Geological Cross Sections with a Chronology of Geological Events. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 544-553.

Frank, A.U. (1998). Different Types of "Times" in GIS. *Spatial and Temporal Reasoning in Geographic Information Systems*, M. J. Egenhofer and R. G. Golledge (Eds.), Oxford University Press, pp. 40-62.

Frank, A.U., I. Campari, and U.F. (Eds.) (1992). *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. Vol. 639, Springer-Verlag.

Frazier, G.F. (1992). Temporal Quandtree. *Proceedings of Information and Knowledge Management CLKM-92*, pp. 658-659.

Frazier, G.F. (1993a). Formal Models for Managing Spatiotemporal Data (Database Management, Temporal Database, Ph.D. Dissertation, Computer Science, University of Kansas.

Frazier, G.F. (1993b). An Incremental Algorithm for Building Temporal Quadtrees. *Proceedings of ACSM 1993 Computer Science Conference*, Denver, CO, ACM Press, pp. 446-452.

Gabbay, D. and P. McBrien (1991). Temporal Logic and Historical Databases. *Proceedings of the 17th International Conference on Very Large Data Bases*, San Mateo, CA, Morgan Kaufmann Publishers, Inc., pp. 423-430.

Gadia, S. (1986). Toward a Multihomogeneous Model for a Temporal Database. *Proceedings of the International Conference on Data Engineering*, pp. 390-397.

Gadia, S. (1988). A Homogeneous Relational Model and Query Languages for Temporal Databases. *ACM Transactions on Database Systems*, Vol. 13, pp. 418-448.

Gadia, S. and C. Yueng (1985). A Query Language for a Homogenous Temporal Database. *Proceedings of the International Conference on Principles of Database Systems*, pp. 51-56.

Gadia, S. and S.S. Nair (1993). Temporal Databases: a Prelude to Parametric Data. *Temporal Databases: Theory, Design, and Implementation*, A. Tansel, J. Clifford, S. Gadia, A. Jajodia, N. Sarda, and R. Snodgrass (Eds.), Benjamin/Cummings Publishing Company, pp. 28-66.

Gadia, S., V. Chopra, and U.S. Tim (1993). An SQL-like Seamless Query Language for Spatio-temporal Data. *Proceedings of the International Workshop on an Infrastructure for Temporal Databases*, R. T. Snodgrass (Ed.).

Gadia, S. (1993). Seamless Integration of Spatial, Temporal, Belief and Ordinary Data. *SIGMOD Record*, Vol. 22, pp. 15-20.

Gadia, S. and V. Chopra (1992). A Relational Model and SQL-Like Language for Seamless Query of Temporal Data. TR-92-05, Computer Science Department, Iowa State University.

Gadia, S. and V. Chopra (1993). A Relational Model and SQL-like Query Language for Spatial Databases. *Advanced Database Systems, Lecture Notes in Computer Science 759*, N.R. Adam and B.K. Bhargava (Eds.), Springer-Verlag.

Gadia, S.K., V. Chopra, and U.S. Tim (1992). Seamless SQL-Like Query of Spatiotemporal Data and a Case Study in Agricultural Environment Management. TR-92-21, Computer Science Department, Iowa State University.

Gadia, S.K., V. Chopra, S. Nair, and U.S. Tim (1992). An SQL-Like Seamless Query Language and Optimization of Spatio-Temporal Data. TR-92-40, Computer Science Department, Iowa State University.

Gaede, V. (1996). Constraint Databases and Applications: CP '96 Workshop on Constraints and Databases, Cambridge, MA, *Lecture notes in computer science 1191*, Springer-Verlag.

Galetto, R. and F. Viola (1994). *Time as the Fourth Dimension of the Topographic Databases*. *Mapping and Geographic Information Systems*, International Society for Photogrammetry and Remote Sensing.

Garret, M. and G.A. Jeffress (1993). Managing Oil Spills: Texas GIS Incorporates Near-Real Time Environment Data. *Geographical Information Systems*, Vol. 3, pp. 28-35.

Gillespie, R. and W.A. Davis (1981). Tree Data Structures for Graphics and Image Processing. *Proceedings of the 7th Conference of the Canadian Man-Computer Communications Society (CMCCS '81 / ACCHO '81)*, pp. 155-162.

Gold, C.M. (1992). An Object-based Dynamic Spatial Model, and its Application in the Development of a User-friendly Digitizing System. *Proceedings of the 5th International*

Symposium on Spatial Data Handling, Columbia, South Carolina, IGU Commission on GIS, pp. 495-504.

Gold, C.M. (1993). An Outline of an Event-Driven Spatial Data Structure for Managing Time-Varying Maps. *Canadian Conference on GIS*, pp. 880-888.

Golledge, R.G. (Ed.) (1978). *Learning about the Urban Environments*. *Timing Space and Spacing Time*, Edward Arnold.

Goralewalla, I.A. and M.T. Ozsu (1993). Temporal Extensions to a Union Behavioral Object Model. *Proceedings of the 10th International Conference on the ER Approach*, pp. 110-121.

Grossman, W.D. and J. Schaller (1990). Connecting Dynamic Feedback Mode with Geographic Information Systems. *Proceedings of the 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 501-511.

Grossman, W.D. and S. Eberhardt (1992). Geographical Information Systems and Dynamic Modeling: Potentials of a New Approach. *Annals of Regional Science*, Vol. 26, pp. 53-66.

Guesgen, H.W. (1989). *Spatial Reasoning Based on Allen's Temporal Logic*. TR-89-049, International Computer Science Institute, Berkeley, CA.

Guptill, S. C. (1992). Multiple Representations of Geographic Entities through Space and Time. *Proceedings of the 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 859-868.

Güting, R.H., D. Papadias, and F.H. Lochovsky (1999). Advances in Spatial Databases, *Proceedings of the 6th International Symposium, SSD '99: Hong Kong, China, July 20-23, 1999,* Springer-Verlag.

Hachem, N.I., M. Gennert, and M. Ward (1992). A DBMS Architecture for Global Change Research. *Proceedings of ISY Conf. on Earth and Space Science*, Pasadena, CA.

Hachem, N.I., K. Qiu, M. Gennert, and M. Ward (1993). Managing Derived Data in the Gaea Scientific DBMS. Very Large Data Bases, Vol. 12.

Hadzilacos, T. and N. Tryfona (1996). Logical Data Modeling for Geographical Applications. *International Journal of Geographical Information Systems*, Vol. 10, pp. 179-203.

Hallin, P. (1991). New Paths for Time-geography. *Geographiska Annaler*, Vol. 73B, pp. 199-207.

Hamre, T. (1994). An Object-oriented Conceptual Model for Measured and Derived Data Varying in 3D Space and Time. *Advances in GIS Research*.

Hamre, T. (1995). Development of Semantic Spatio-temporal Data Models for Integration of

Remote Sensing and in situ Data In a Marine Information System, University of Bergen, Norway.

Hao, Q. (1995). A Hyper-Relational Framework for Unified Geographic Modeling (Temporal Structure), Computer Science/Geography/Geodesy, Southern Illinois University at Carbondale.

Hasting, H., R. Pekelney, R. Monticciolo, D. Vun-Kannon, and D. Del-Monte (1982). Time Scales, Persistence and Patchiness. *BioSystems*, Vol. 15, pp. 281-289.

Hazelton, N.W.J. (1991). *Integrating Time, Dynamic Modeling and Geographical Information Systems: Development of Four-Dimensional GIS*. Technical Report, Department of Surveying and Land Information, the University of Melbourne, Australia.

Hazelton, N.W.J. (1991a). Extending GIS to Include Dynamic Modeling. *Proceedings of the Third Colloquium of the Spatial Information Research Centre*, Dunedin, New Zealand, University of Otago, pp. 73-82.

Hazelton, N.W.J. (1991b). Using GIS with Non-Point-Source and Distributed Models: The `How' and `Why' of an Information Opportunity. *Proceedings of the AURISA '91 Conference*, Washington, URISA, pp. 298-314.

Hazelton, N.W.J. (1991c). *Integrating Time, Dynamic Modeling and Geographical Information Systems: Development of Four-Dimensional GIS*, Department of Surveying and Land Information, the University of Melbourne, Australia.

Hazelton, N.W.J. (1992a). Beyond the 2-D Map: A New Metaphor for Multi-Temporal 4-D GIS. *Proceedings of GIS/LIS '92*, Bethesda, MD, ASPRS/ACSM, pp. 303-313.

Hazelton, N.W.J. (1992b). Developments in Spatio-Temporal GIS. *Geographical Information Sciences Research in Victoria and Tasmania Conference*.

Hazelton, N.W.J. (1992c). Time in GIS. *The Monash University GIS Seminar Series*.

Hazelton, N.W.J. (1998). Some Operational Requirements for a Multi-Temporal 4-D GIS. *Spatial and Temporal Reasoning in Geographic Information Systems*, M. J. Egenhofer and R. G. Golledge (Eds.), Oxford University Press, pp. 63-73.

Hazelton, N.W.J. and S. Pigot (1992). Basic Concepts in 3-D and 4-D Geographic Information Systems and Their Application in the Earth Sciences. *Proceedings of the Conference on the Geological Applications of Geographic Information Systems*, Perth, Australia, Australia Institute of Geoscientists, pp. 9-19.

Hazelton, N.W.J. and L.M. Bennett (1992). *Topological Relationships in Four-Dimensional Geographical Information Systems*. Technical Report, Department of Surveying and Land Information, the University of Melbourne, Australia.

Hazelton, N.W.J., F.J. Leahy, and I.P. Williamson (1990). On the Design of Temporally Referenced 3-D Geographical Information Systems: Development of Four-Dimensional GIS. *Proceedings of GIS/LIS '90*, Bethesda, MD, ACSM, pp. 357-372.

Hazelton, N.W.J., F.J. Leahy, and I.P. Williamson (1990). Incorporating Dynamic Modeling into GIS. *Proceedings of the AURISA '90 / URPIS 18 Conference*, WA, URISA, pp. 77-84.

Hazelton, N.W.J., L. Bennett, and J. Masel (1992). Topological Structures for 4-Dimensional Geographic Information Systems. *Computers, Environment, and Urban Systems*, Vol. 16, pp. 227-237.

Hazelton, N.W.J., F.J. Leahy, and I.P. Williamson (1993). Integrating Dynamic Modelling and Geographical Information Systems. *URISA*.

Herbert, M.J. and D.B. Kidner (1992). Spatial and Temporal Visualization of Three-Dimensional Surfaces for Environment Management. *Proceedings of Auto-Carto II*, Minneapolis.

Herring, J.R. (1991). Using Spline Functions to Represent Distributed Attributes. *Proceedings of Auto-Carto 10: Technical Papers of the 1991 ACSM-ASPRS Annual Convention*, Bethesda, MD, ASPRS/ACSM, pp. 46-58.

Hintz, R. and H. Onsrud (1990). A Methodology for Upgrading Real Property Boundary Information in a GIS Using a Temporally Efficient Automated Survey Measurement Management System.

Holmberg, S.C. (1994). DML: a Map Language for Dynamics. *Proceedings of EGIS 94. The 5th European Conference and Exhibition on Geographical Information Systems*, Paris, France.

Hornsby, K. and M.J. Egenhofer (2000). Identity-based Change: A Foundation for Spatio-temporal Knowledge. *International Journal of Geographical Information Science*, Vol. 14, pp. 207-224.

Hornsby, K.E. (1999). Identity-Based Reasoning About Spatio-Temporal Change, Computer Science, University Of Maine.

Hough, R.R. (1992). Time and Space: an Economic Model. *Theories and Models of Spatio-Temporal Reasoning in Geographic Space*, A. Frank, I. Campari, and U. Formentini (Eds.), Springer-Verlag, pp. 78-96.

Hunter, G. (1988). Non-Current Data and Geographical Information Systems. A Case for Data Retention. *International Journal of Geographical Information Systems*, Vol. 2, pp. 281-286.

Hunter, G. and I.P. Williamson (1990). The Development of a Historical Digital Cadastral Database. *International Journal of Geographical Information Systems*, Vol. 4, 169-179.

Hunter, G.J. (1986). Non-Current Data and Land Information Systems, Department of Surveying and Land Information, the University of Melbourne, Australia.

Hunter, G.J. (1986). Selection Criteria for Archiving Non-Current Data in Land Information Systems. *CISM Journal ACSGC*, Vol. 46, pp. 255-264.

Ibrahim, M., J. Küng, and N. Revell (2000). *Database and Expert Systems Applications (Proceedings of 11th international conference, DEXA 2000, London, UK)*. *Lecture notes in computer science 1873*, Springer-Verlag.

Itami, R.M. (1988). Cellular Automatons as a Framework for Dynamic Simulations in Geographic Information Services. *Proceedings of GIS/LIS '88*, pp. 590-597.
Itami, R.M. and J.D. Clark (1992). Spatial Dynamic Simulation Using Discrete Time and Discrete Event Theory in Cell Based GIS Systems. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 702-712.

Jensen, C.S., R. Snodgrass, and M.D. Soo (1992). *Extending Normal Forms to Temporal Relations*. TR 92-17, Dept. of Computer Science, University of Arizona.

Jensen, C.S., R. Snodgrass, and M.D. Soo (1994). *Extending Existing Dependency Theory to Temporal Databases*. University of Aalborg Institute for Electronic Systems Dept. of Mathematics and Computer Science.

Jones, S., P. Mason, and R. Stamper (1979). LEGOL 2.0: a Relational Specification Language for Complex Rules. *Information Systems*, Vol. 4, pp. 293-305.

Kafer, W. and H. Schoning (1992). Realizing a Temporal Complex-Object Data Model. *SIGMOD Conference 1992*, pp. 266-275.

Kafer, W., N. Ritter, and H. Schoning (1990). Support for Temporal Data by Complex Objects. *Proceedings of the International Conference on Very Large Databases (VLDB)*, pp. 24-35.

Kambayashi, Y. (1999). Advances in Database Technologies: Proceedings of ER 98 Workshops on Data Warehousing and Data Mining, Mobile Data Access, and Collaborative Work Support and Spatio-Temporal Data Management, Singapore. *Lecture notes in computer science 1552*, Springer-Verlag.

Kampke, T. (1994). Storing and Retrieving Changes in a Sequence of Polygons. *International Journal of GIS*, Vol. 8, pp. 493-513.

Kang, S.-B. and Y.-C. Choy (1995). Object-Relational Data Modeling For GIS. *GIS/LIS '95 Annual Conference and Exposition*.

Kelmelis, J.A. (1991). *Time and Space in Geographic Information: Toward a Four-Dimensional Spatio-Temporal Data Model*, Ph.D. Dissertation, Department of Geography, Pennsylvania State University.

Kemp, Z. and A. Kowalczyk (1994). Incorporating the Temporal Dimension in a GIS.

Innovations in GIS, M.F. Worboys (Ed.), Taylor & Francis, pp. 89-103.

Koeppel, I.J. and S.D. Ahlmer (1993). Integrating the Dimensions of Time into AM/FM Systems. *Proceedings of AM/FM XVI International Annual Conference*, Aurora, CO.

Kolovson, C. (1990). Indexing Techniques for Multi-Dimensional Spatial Data and Historical Data in Database Management Systems, University of California, Berkley.

Kolovson, C. (1993). Indexing for Historical Databases. *Temporal Databases: Theory, Design, and Implementation*, A. Tansel, J. Clifford, S.K. Gadia, S. Jajodia, A. Segev, and R. Snodgrass, (Eds.), Benjamin/Cummings, pp. 418-432.

Kolovson, C. and M. Stonebraker (1991). Segment Indexes: Dynamic Indexing Techniques for Multi-Dimensional Interval Data. *Proceedings of ACM SIGMOD 1991*, Denver, CO, ACM Press, pp. 138-147.

Koomen, J.A.G.M. (1989). Reasoning About Recurrence (Knowledge Representation), Computer Science/Artificial Intelligence, The University of Rochester.

Koussoulakou, A. and M.-J. Kraak (1992). Spatio-Temporal Maps and Cartographic Communication. *The Cartographic Journal*, Vol. 29, pp. 101-108.

Kraak, M.-J. and A.M. Maeachren (1994). Visualization of the Temporal Component of Spatial Data. *Advances in GIS Research. Proceedings of the 6th Symposium. on Spatial Data Handling*, Taylor & Francis.

Kubo, S., S. Takamura, and S. Yoshino (1990). Multimedia GIS on PC. *Proceedings of the 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 363-370.

Langran, G. (1988a). Temporal GIS Design Tradeoffs. *Proceedings of GIS/LIS '88*, Bethesda, MD, ASPRS/ACSM, pp. 890-899.

Langran, G. (1988b). A Framework for Temporal Geographic Information Systems. *Cartographica*, Vol. 25, pp. 1-14.

Langran, G. (1989a). *Time in Geographic Information Systems*, Ph.D. Dissertation, University of Washington.

Langran, G. (1989b). Accessing Spatiotemporal Data in a Temporal GIS. *Proceedings of Auto-Carto 9*, Baltimore, MD, ASPRS/ACSM, pp. 191-198.

Langran, G. (1989c). A Review of Temporal Database Research and Its Use in GIS Applications. *International Journal of Geographical Information Systems*, Vol. 3, pp. 215-232.

Langran, G. (1990a). Tracing Temporal Information in an Automated Nautical Charting System.

Cartography and Geographic Information Systems, Vol. 17, pp. 291-299.

Langran, G. (1990b). Temporal GIS Design Tradeoffs. *Journal of the Urban and Regional Information Systems*, VOl. 2, pp. 16-25.

Langran, G. (1991). Issues in Implementing a Spatiotemporal System. *Proceedings of the International Cartographic Conference*, Bournemouth.

Langran, G. (1992a). States, Events, and Evidence: The Principal Entities of A Temporal GIS. *Proceedings of GIS/LIS '92*, ACSM/ASPRS, pp. 416-425.
Langran, G. (1992b). Preparing for a Temporal Forest GIS. *COMPILER*, Vol. 10, pp. 18-26.

Langran, G. (1992c). Forest Monitoring Using Temporal GIS Techniques. *Proceedings of GIS/LIS '92 Symposium*, Vancouver, Canada, Forestry, Canada, pp. 416-425.

Langran, G. (1992d). *Time in Geographic Information Systems*. Taylor & Francis.

Langran, G. (1993a). Issues of Implementing a Spatiotemporal System. *International Journal of Geographical Information Systems*, Vol. 7, pp. 305-314.

Langran, G. (1993b). Manipulation and Analysis of Temporal Geographic Information. *Proceedings of the Canadian Conference on GIS '93*, Ottawa, Canada, The Canadian Institute of Geometrics, pp. 869-879.

Langran, G. (1993c). One GIS, Many Realities. *Proceedings of GIS '93*, Vancouver, Canada, Ministry of Supply and Service, pp. 767-762.

Langran, G. and N. Chrisman (1988). A Framework for Temporal Geographic Information. *Cartographica*, Vol. 25, pp. 1-14.

Laurini, R. (1998). Spatial Multi-database Topological Continuity and Indexing: a Step Towards Seamless GIS Data Interoperability. *International Journal Geographical Information Science*, Vol. 12, pp. 373-402.

Lawrence, N. (1975). *Temporal Passage and Spatial Metaphor*. Springer-Verlag.

Lester, M. (1990). Tracking the Temporal Polygon: A Conceptual Model of Multidimensional Time of Geographic Information Systems. *Proceedings of the Temporal GIS Workshop*, Orono, ME, University of Maine.

Leung, Y., K.S. Leung, and J.Z. He (1999). A Generic Concept-based Object-Oriented Geographical Information System. *International Journal Geographical Information Science*, Vol. 13, pp. 475-498.

Li, C.-S. (1997). *Sequential processing for content-based retrieval of composite objects. Research report; RC 21067.*, IBM T.J. Watson Research Center.

Li, K.-J. (1992). *Contributions on Hypermedia Systems: Modeling and Indexing Spatiotemporal Objects (in French)*, Institut National des Sciences Appliquees (INSA).

Li, K.-J., N. Badji, and R. Laurini (1992). Modeling Moving Objects in Spatiotemporal Databases.

Li, K.-J., N. Badji, and R. Laurini (1992). Toward Spatiotemporal Databases: Modeling Moving Objects (in French). *Proceedings of the Eighth Conference on BDA, Advanced Databases, Tregastel*, Tregastel, France, INRIA, pp. 102-118.

Lin, H. and H.W. Calkins (1991). A Rationale for Spatiotemporal Intersection. *Proceedings of the Technical Report of the ACSM-ASPRS Annual Convention*, Bethesda, MD, ASPRES/ACSM, pp. 204-213.

Liu, Y. (1995). *Gdb: A Graphic Database Indexing and Retrieving by Visual Relations*, Computer Science, University Of Hawaii.

Lolonis, P. (1994). Knowledge-Based Spatial Decision Support Systems for Solving Service Provision Problems, Geography/Computer Science/Urban and Regional Planning, the University of Iowa.

Lorentzos, N. and R. Johnson (1988). Requirements Specification for a Temporal Extension to the Relational Model. *Database Engineering*, Vol. 11, pp. 204-211.

Lorentzos, N.A. (1988). A Formal Extension of the Relational Model for the Representation and Manipulation of Generic Intervals, Birbeck College, University of London.

Lorentzos, N.A. (1992). DBMS Support for Time and Totally Ordered Compound Data Types. *Information Systems*. Vol. 17.

Lorentzos, N.A. (1993a). The Interval Extended Relational Model and Its Application to Valid Time Databases. *Temporal Databases: Theory, Design, and Implementation*, A. Tansel, J. Clifford, S. Gadia, S. Jajodia, A. Segev, and R. Snodgrass (Eds.), Benjamin/Cummings, pp. 67-91.

Lorentzos, N.A. (1993b). *Axiomatic Generalization of the Relational Model to Support Valid Time Data*. Internal Report, Information Laboratory, Agricultural University of Athens.

Lorentzos, N.A. (1994). DBMS Support for Non-Metric Measurement Systems. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 5, pp. 67-91.

Lorentzos, N.A. and R.G. Johnson (1988). An Extension of the Relational Model to Support Generic Intervals. *Advances of Database Technology--EDBT '88*, J.W. Schmidt, S. Ceri, and M. Missikoff (Eds.), Springer-Verlag, pp. 528-542.

Lorentzos, N.A. and R.G. Johnson (1988). Extending Relational Algebra to Manipulate

Temporal Data. *Information Systems*, Vol. 13, pp. 289-296.

Lorentzos, N.A. and V. Kollias (1989). The Handling of Depth and Time Intervals in Soil-Information Systems. *Computers and Geosciences*, Vol. 15, pp. 396-401.

Lowell, K.E. (1991). Utilizing Discriminant Function Analysis with a Geographical Information System to Model Ecological Succession Spatially. *International Journal of Geographical Information Systems*, Vol. 5, pp. 175-192.

Lowell, K.E. (1994). Probabilistic Temporal GIS Modeling Involving More Than Two Map Classes. *International Journal of `Geographical Information Systems*, Vol. 8, pp. 73-93.

Lu, H. and B.C. Ooi (1993). Spatial Indexing: Past and Future. *IEEE Bulletin on Data Engineering*, Vol. 16, pp. 16-21.

Lu, H. and A. Zhou (2000). Web-age Information Management : First International Conference, WAIM 2000, Shanghai, China, June 2000: *Proceedings*. *Lecture Notes in Computer Science 1846*, Springer-Verlag.

Lum, V., P. Dadam, R. Erbe, J. Guenauer, P. Pistor, G. Walch, H. Walch, H. Werner, and J. Woodfill (1984). Designing DBMN Support for the Temporal Dimension. *ACM SIGMOD Record*, Vol. 14, pp. 115-130.

MacDougall, E.B. (1991). Dynamic Statistical Visualization of Geographic Information Systems. *Proceedings of GIS/LIS '91*, Bethesda, MD, ASPRS/ACSM, pp. 158-165.

MacEachren, A. (1993). Time as a Cartographic Variable. D. Unwin and H. Hearnshaw (Eds.), *Visualization in Geographic Information Systems*.

Mackaness, W. and B. Buttenfield (1991). Incorporating Time into Geographic Process: A Framework for Analyzing Process in GIS. *Proceedings of 15th Conference and 9th General Assembly of the International Cartographic Association*, Basel, Switzerland, International Cartographic Association, pp. 565-574.

Macor, F., V. Pouthas, and W. Friedman (1992). *Time-Action, and Cognition*. Kluwer Academic Publishers.

Maddux, R. (1990). *Some Algebras and Algorithms for Reasoning about Time and Space*. Technical Report, Department of Mathematics, Iowa State University.

Mark, D. (1992). Spatiotemporal Relations and GIS: A Geographer's View. *GIS From Space to Territory: Theories and Methods of Spatiotemporal Reasoning*.

Mason, D.C., M.A. O'Conaill, and S.B.M. Bell (1994). Handling Four-Dimensional Geo-referenced Data in Environmental GIS. *International Journal of Geographic Information System*, Vol. 8.

McDonald, J.R. and T.M. Pullan-Cuthbertson (1989). Time and Spatial Information Systems. *Proceedings of the Inaugural Colloquium of the Spatial Information Research Centre*, Otago, New Zealand, University of Otago, pp. 79-89.

Medak, D. (1999). Lifestyles - a Paradigm for the Description of Spatiotemporal Databases, Department of Geoinformation, University of Technology Vienna, pp. 1-169.

Miller, D. (1992). Analysis of Vegetation Succession within an Expert System. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 381-390.

Miller, H. (1991). Modelling Accessibility Using Space-Time Prism Concepts within Geographical Information Systems. *International Journal of Geographical Information Systems*, Vol. 5, pp. 287-301.

Miller, H. (1991). Modeling Accessibility Using Space-time Prism Concepts within GIS. Ph.D. Dissertation, Department of Geography, Ohio State University.

Monmonier, M. (1990). Strategies for the Visualization of Geographic Time-Series Data. *Cartographica*, Vol. 27, pp. 30-45.

Monmonier, M. (1990). Strategies for the Interactive Exploration of Geographic Correlation. *Proceedings of 4th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 512-521.

Monmonier, M. (1992). Time and Motion as Strategic Variables in the Analysis and Communication of Correlation. *Proceedings of 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 72-81.

Montgomery, K. (1991). Methodological and Spatio-temporal Contexts for Geomorphologic Knowledge: Analysis and Implications. *Canadian Geographer*, Vol. 35, pp. 345-352.

Moon, B. (1996). Data Placement For High Performance Spatio-temporal Databases (Clustering, Declustering, Hilbert Curve, Space Filling), Computer Science, University of Maryland College Park.

Moreira, J.M.M. (1997). Spatio-temporal Database Managements Optimizations, Ecole Nationale Superieure de Telecommunications de Paris (France) and Faculdade de Engenharia da Universidade do Porto (Portugal).

Muntz, R., E. Shek, and C. Zaniolo (1995). Using LDL++ for Spatio-temporal Reasoning in Atmospheric Science Databases. *Applications of Logic Databases*, R. R. (Ed.), Kluwer Academic.

Narciso, F.E. (1999). A Spatiotemporal Data Model for Incorporating Time in Geographic

Information Systems (Gen-Stgis), Computer Science/Remote Sensing, University Of South Florida.

Navathe, S. and R. Ahmed (1993). Temporal Extensions to the Relational Model and SQL. *Temporal Databases: Theory, Design, and Information*, J. C. A. Tansel, S. Gadia, S. Jajodia, A. Segev, and R. Snodgrass (Eds.), Benjamin/Cummings Publishing Company, pp. 92-109.

Navathe, S.B. and R. Ahmed (1988). TSQL: A Language Interface for History Databases. *Proceedings of the Conference on Temporal Aspects in Information Sciences*, pp. 113-128.

Navathe, S.B. and R. Ahmed (1989). A Temporal Relational Model and Query Language. *Information Sciences*, Vol. 49, pp. 147-175.

Needham, S. and B.E. Vieux (1989). A GIS for AgNPS Parameter Input and Mapping Output. *Proceedings of ASAE Winter Meeting, Paper Number 89-2673*.

Newell, R.G. and M. Easterfield (1990). Version Management--The Problem of the Long Transaction. *Mapping Awareness*, Vol. 4.

Newell, R.G., D. Theriault, and M. Easterfield (1990). Temporal GIS--Modeling the Evolution of Spatial Data in Time. *Functionality Conference*.

Newell, R.G., D. Theriault, and M. Easterfield (1992). Temporal GIS--Modeling the Evolution of Spatial Data in Time. *Computers and Geosciences*, Vol. 18, pp. 427-433.

Nittel, S., R.R. Muntz, and E. Mesrobian (1996). *GeoPOM : a Heterogeneous Geoscientific Object System*. Computer Science Dept. University of California.

Norrie, M.C. (1992). *A Collection Model for Data Management in Object-oriented Systems*, Ph.D. Dissertation, Department of Computer Science, University of Glasgow.

Norrie, M.C. (1993). An Extended Entity-Relationship Approach to Data Management in Object-Oriented Systems. *Proceedings of the 12th International Conference on the ER Approach*, pp. 390-401.

Onsrud, H.J. and R.J. Hintz (1989). Upgrading Boundary Information in a GIS Using an Automated Survey Measurement Management System. *Proceedings of ACSM Annual Convention*, Bethesda, MD, ASPRS/ACSM, pp. 275-284.

Ooi, B.C. (1988). Efficient Query Processing for Geographic Information Systems, Computer Science, Monash University.

Osborne, S. and M. Stoogenke (1989). Integration of a Temporal Element into a Natural Resource Decision Support System. *Proceedings of GIS/LIS '89*, Bethesda, MD, ASPRS/ACSM, pp. 221-227.

Papagno, G. (1992). Seeing Time Contribution to the Discussion on Space-Time Dynamic. *Four Contributions to GIS Theory and Application*, A.U. Frank, I. Campari, and U. Formentini (Eds.), Editrice Universitaria Litografia Felici, pp. 29-46.

Parent, C., S. Spaccapietra, and E. Zimanyi (1999). Spatio-Temporal Conceptual Models: Data Structure + Space + Time. *Proc. of the 7th ACM Symposium on Advances in Geographic Information Systems, GIS'99*, Kansas City, USA.

Parks, D. and N. Thrift (1980). *Times, Spaces, and Places*. John Wiley and Sons.

Paton, N.W., A.A.A. Fernandes, and T. Griffiths (2000). Spatio-temporal Databases: Contentions, Components and Consolidation.

Peuquet, D. (1993). A Framework for the Representation of Spatiotemporal Processes in Geographic Information Systems. *Proceedings of the International Workshop on an Infrastructure for Temporal Databases*, R. T. Snodgrass (Ed.).

Peuquet, D. and L. Qian (1995). An Integrated Database Design for Temporal GIS. *Advances in GIS Research II: Proceedings of the Seventh International Symposium on Spatial Data Handling*, Taylor & Francis, pp. 21-31.

Peuquet, D. J. (1992). *Toward the Representation and Analysis of Spatiotemporal Processes in Geographic Information Systems*. Technical Report.

Peuquet, D. J. (1994). It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information System. *Annals of the Association of American Geographers*, Vol. 84, pp. 441-461.

Peuquet, D.J. and E. Wentz (1994). An Approach for Time-based Analysis of Spatiotemporal Data. *Advance in GIS Research Proceedings of the 6th Int. Symposium. on Spatial Data Handling*, T.C. Waugh and R.G. Healy (Eds.), Taylor & Francis.

Peuquet, D.J. and N. Duan (1995). An Event-based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data. *International Journal of Geographical Information Systems*, Vol. 9, pp. 7-24.

Pigot, S. and B. Hazelton (1992). The Fundamentals of a Topological Model for a Four-Dimensional GIS. *Proceedings of the Fifth International Symposium on Spatial Data Handling*, Columbia, South Carolina, IGU Commission on GIS, pp. 580-591.

Preston, E.B. (1993). Remote Sensing GIS Technologies Support Sea Ice Motion Monitoring. *GIS World*, Vol. 6, pp. 36-39.

Price, S. (1989). Modeling the Temporal Element in Land Information Systems. *International Journal of Geographical Information Systems*, Vol. 3, pp. 233-243.

Qian, L. and D. Peuquet (1998). Design of a Visual Language for GIS. *Proceedings of 8th International Symposium on Spatial Data Handling*, Vancouver, Canada.

Qian, L., M. Wachowicz, D. Peuquet, and A. MacEachren (1997). Delineating Operations for Visualization and Analysis of Space-Time Data in GIS. *Proceedings of GIS/LIS '97*, Cincinnati, OH.

Qiu, K., N.I. Hachem, M.O. Ward, and M.A. Gennert (1992). Providing Temporal Support in Data Base Management Systems for Global Change Research. *Proceedings of the 6th International Working Conference on Scientific and Statistical Database Management*, Springer-Verlag, pp. 274-289.

Raafat, H., Q. Xiao, and D. Gauthier (1991). An Extended Relational Database for Remotely Sensed Image Data Management within GIS. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, pp. 651-655.

Raafat, H., Z. Yang, and D. Gauthier (1994). Relational Spatial Topologies for Historical Geographical Information. *International Journal of Geographical Information Systems*, Vol. 8, pp. 163-173.

Radic, U., D. Jenssen, and B. McInnes (1987). On the Surging Potential of Polar Ice Streams. Antarctic Surges-A Clear and Present Danger? Report NO. DOE/ER/60197-H1.

Raza, A. and W. Kainz (1998). Design and Implementation of a Temporal GIS for Monitoring the Urban Land Use Change. *Spatial Information Technology Towards 2000 and Beyond: The Proceedings of Geoinformatics '98 Conference*, pp. 417-427.

Rennison, E. and L. Strauss (1995). The Millennium Project: Constructing a Dynamic 3D Virtual Environment for Exploring Geographically, Temporally, and Categorically Organized Historical Information. *Proceedings of COSIT95*.

Renolen, A. (1995). Temporal Maps and Temporal Geographical Information Systems (Review of Research). ResearchIndex.

Ridd, M.R. (1991). Spatial and Temporal Scale Issues Related to the Integration of GIS and Remote Sensing. *Proceedings: Remote Sensing and Geographic Information Systems, ACSM-ASPRS*.

Ris, F. (1991). Toward a `Scientific GIS'. *Proceedings of GIS/LIS '91*, Bethesda, MD, ASPRS/ACSM, pp. 740-751.

Robertson, G.G., J.D. MacKinlay, and S.K. Card (1991). Cone Trees: Animated 3D Visualizations of Hierarchical Information. *Proceedings of the CHI '91 Conference*, NY, ACM Press, pp. 189-194.

Rodrigue, J.P. (1997). Parallel Modeling and Neural Networks: An Overview for

Transportation/Land Use Systems. *Transportation Record C*, Vol. 5, pp. 259-271.

Rojas-Vega, E. and Z. Kemp (1995). An Object-Oriented Data Model for Spatio-Temporal Data. *Proceedings of the Ninth Annual Symposium on Geographic Information Systems*, Vancouver, Canada.

Roman, G.C. (1990). Formal Specification of Geographic Data Processing Requirements. *TKDE*, Vol. 2, pp. 370-380.

Rose, E. (1993). A Temporal Object-oriented Algebra. *Proceedings ECOOP '93*, pp. 297-325.
Rose, E. and A. Segev (1991). A Temporal Object-Oriented Data Model with Temporal Constraints. *Proceedings of the 10th International Conference on the ER Approach*, pp. 205-230.

Roshannejad, A.A. (1996). The Management of Spatio-Temporal Data in A National Geographic Information System (Semantic Analysis, Object-Oriented), Computer Science/Remote Sensing, Universiteit Twente (The Netherlands).

Rossignol, J. and L.E. Wandsnider (1992). Space, Time, and Archaeological Landscapes.

Roussoppolous, N., C. Faloutsos, and T. Sellis (1988). An Efficient Pictorial Database System for PSQL. *IEEE Transactions on Software Engineering*, Vol. 14, pp. 639-650.

Rowe, L.A. and M.R. Stonebraker (1987). The POSTGRES Data Model. *Proceedings of the 13th LVDB Conference*, Brighton, pp. 83-96.

Sadeghi, R., W.B. Samson, and S.M. Deen (1988). HQI - A Historical Query Language. *Proceedings of the Sixth British National Conference on Database (BNCOD 6)*, Cardiff, Wales.

Salzberg, B. (1994). On Indexing Spatial and Temporal Data. Boston, MA.

Sandhu, J. (1990). Combining Exploratory Data Analysis and Scientific Visualization in the Study of Very Large Space-Time Data Sets, Department of Geography, Ohio State University.

Sarda, N. (1990). Extensions to SQL for Historical Databases. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 2, pp. 220-230.

Sarda, N. (1993). HSQL: A Historical Query Language. *Temporal Databases: Theory, Design, and Implementation*, A. Tansel, J. Clifford, S. Gadia, A. Jajodia, A. Segev, and R. Snodgrass, (Eds.), Benjamin/Cummings Publishing Company, pp. 110-138.

Schneider, R. and H.-P. Kriegel (1992). Indexing the Spatiotemporal Monitoring of a Polygonal Object. *Proceedings of the 5th International Symposium on Spatial Data Handling*, P. Bresnahan, E. Corwin, and D. Cowen (Eds.), IGU Commission on GIS, pp. 200-209.

Schoenauer, M. (2000). Parallel Problem Solving from Nature: PPSN VI, 6th International Conference, Paris, France, September 18-20, 2000, *Proceedings of Lecture Notes in Computer*

Science 1917, Springer-Verlag.

Scholl, M., M.H. Bohlen, and C.S. Jensen (Eds.) (1999). *Spatio-temporal Database Management: International Workshop {STDBM}'99, Edinburgh, Scotland, September 10-11, 1999: proceedings*. Lecture Notes in Computer Science ed. Vol. 1678. *Spatio-temporal database management: International Workshop {STDBM}'99, Edinburgh, Scotland, September 10-11, 1999: proceedings*, Springer-Verlag Inc.

Schumm, S. and R. Lichty (1965). Time, Space, and Causality in Geomorphology. *American Journal of Science*, Vol. 263, pp. 110-119.
Sellis, T. (1999). Research Issues in Spatio-temporal Database Systems. *SSD: Advances in Spatial Databases*, LNCS, Springer-Verlag.

Sharpley, A.N. and J.R. Williams. (Eds.) (1990). EPIC - Erosion / Productivity Impact Calculator. Vol. 1 (Model Documentation) and Vol. 2 (User Manual). Technical Bulletin NO. 1768.

Shaw, S-L. and D. Wang (2000). Handling Disaggregate Spatiotemporal Travel Data in GIS. *GeoInformatica*, Vol. 4, pp. 161-178.

Siebert, L.J. (1997). Creating a GIS Spatial History of Tokyo (Japan), Geography/Information Science/History, Asia, Australia and Oceania, University of Washington.

Sinton, D.E. (1978). The Inherent Structure of Information as a Constraint to Analysis: Mapped Thematic Data as a Case Study. *Harvard Papers on GIS*, Vol. 7.

Skjellaug, B. and A.-J. Berre (1997). Multi-dimensional Time Support for Spatial Data Models.

Slocum, T.A., J.C. Davis, and S.L. Eghert (1993). Developing Software for Exploring Temporal Spatial Data. *Proceedings of GIS/LIS 93*, Minneapolis, MN.

Smith, T. (1992). Towards a Logic-Based Language for Modeling and Database Support for Spatio-Temporal Domains. *Proceedings of the 5th International Conference on Spatial Data Handling*, P. Bresnahan, E. Corwin, and D. Cowen (Eds.), IGU Commission on GIS, pp. 592-601.

Smyth, S. (1992). A Representational Framework for Route Planning in Space and Time. *Proceedings of the 5th International Symposium on Spatial Data Handling*, P. Bresnahan, E. Corwin, and D. Cowen (Eds.), IGU Commission on GIS, pp. 692-701.

Snodgrass, R. (1995). *The TSQL2 Temporal Query Language*. Kluwer Academic Publishers.

Snodgrass, R.T. (1992). Temporal Databases. *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, A. U. Frank, I. Campari, and a. U. Formentini, Eds., Springer-Verlag, pp. 22-64.

Snodgrass, R.T. (2000). *Developing Time-Oriented Database Application in SQL*. Morgan Kaufmann Publishers.

Snodgrass, R.T. and I. Ahn (1985). A Taxonomy of Time in Databases. *Proceedings of the SIGMOD '85 Conference in Austin, Texas*, pp. 236-245.

STDBM'99, 1999: *International Workshop STDBM'99, Edinburgh, Scotland, September 1999. Proceedings of Lecture Notes in Computer Science 1678.*, Springer-Verlag.

Steiner, A. (1998). A Generalization Approach to Temporal Data Models and their Implementations, Swiss Federal Institute of Technology, Zurich.

Stickler, G.J., V.J. Shearn, and G.P. Browning (1992). *Handling Overlapping Parcels in Arc/Info*. Technical Report 4, ESRI, United Kingdom.

Stonebraker, M. (1987). The Design of POSTGRES Storage System. University of California, Berkeley, CA.

Stonebraker, M. and L.A. Rowe (1986). The Design of POSTGRES. University of California, Berkeley, CA.

Story, P.A. and M.F. Worboys (1995). An Object-Oriented Model of Time. Keele, Staffordshire, United Kingdom.

Su, S.Y.W. and H.M. Chen (1991). A Temporal Knowledge Representation Model OSAM*/T and its Query Language OQL/T. *Proceedings of the International Conference on Very Large Database (VLDB)*, pp. 431-442.

Szarmes, M.C. (1997). Modeling the Evolution of Spatio-Temporal Database Structures For GIS Applications, Computer Science, University of Calgary, Canada.

Tang, Q. (1993). A Dynamic Visualization Approach to the Exploration of Area-Based, Spatial-Temporal Data (Geographic Information Systems), Geography/Computer Science/Statistics/Remote Sensing, Ohio State University.

Theodoridis, Y., J.R.O. Silva, and M.A. Nascimento (1999). On the Generation of Spatio-temporal Datasets. TR-40, TimeCenter.

Thrift, N. (1977). *An Introduction to Time Geography*. *Concepts in Modern Geography*, Geo-Abstracts, Ltd.

Trepied, T. (1995). Dynamic Icons for a Visual Spatio-Temporal Query Language. *Proceedings of the Joint European Conference and Exhibition on Geographical Information (JEG-CT 95)*, The Hague.

Tryfona, N.J., Christian S. (1999). Conceptual Data Modeling for Spatiotemporal Applications.

GeoInformatica, Vol. 3.

Tsotras, V.J., C.S. Jensen, and R.T. Snodgrass (1997). A Notation for Spatiotemporal Queries. TR-10, TimeCenter.

Turner, H. (1988). Real-time Data Collection Quality Assurance Models in Three-dimensional Spatial Information Systems. *International Journal of Geographical Information Systems*, Vol. 2, pp. 295-306.

Valsecchi, P., C. Claramunt, and E. Peytchev (2000). OSIRIS: an Inter-Operable System for the Integration of Real-Time Traffic Data within GIS. *Sage Urban Studies Abstracts (Computers, Environment and Urban Systems (UK)245-257, July 1999)*, Vol. 28.

Varma, H., H. Boudreau, and W. Prime (1990). A Data Structure for Spatio-Temporal Databases. *International Hydrographic Review*, Vol. 67, pp. 71-92.

Vasiliev, I. (1990). Examples of the Treatment of Times as a Variable on Maps. *Proceedings of the Temporal GIS Workshop*, Orono, ME, University of Maine.

Vazirgiannis, M., Y. Theodoridis, and T. Sellis (1998). Spatio-temporal Composition and Indexing for Large Multimedia Applications. *ACM/Springer Multimedia Systems*, Vol. 6, pp. 284-298.

Vieux, B.E. (1991). Geographic Information Systems and Non-Point Source Water Quality and Quantity Modeling. *Hydrological Processes*, Vol. 5, pp. 101-113.

Vrana, R. (1989). Historical Data as an Explicit Component of Land Information Systems. *International Journal of Geographical Information Systems*, Vol. 3, pp. 33-49.

Vrana, R. (1997). Monitoring Urban Land Use Transition with Geographic Information Systems, Department of Geography, University of Washington.

Wachowicz, M. and R. G. Healey (1994). Towards Temporality in GIS. *Innovations in GIS*, M. Worboys (Ed.), Taylor and Francis, pp. 105-115.

Wakim, T. and F.B. Chedid (2000). On the Construction of Old Versions from a Spatio-temporal Database, pp. 1-13.

Wang, C. (1998). *Improvement of a Prototype Temporal Geographic Information System for Transportation Using ARC/INFO ODE*, Computer Science, Florida International University.

Wang, W. (1999). *Spatial Temporal Data Mining (Spatial Databases)*, Computer Science, University of California, Los Angeles, CA.

Weismiller, R.A., S.J. Kristof, D.K. Scholz, P.F. Anuta, and S.A. Momin (1977). Change Detection in Coastal Zone Environments. *Photogrammetric Engineering and Remote Sensing*, Vol. 43, pp. 1533-1539.

Wolf, A. (1989). The DASDBS GEO-Kernel: Concepts, Experiences, and the Second Step: Principles and Applications. *Design and Implementation of Large Spatial Databases, Proceedings of SSD '89*, Santa Barbara, CA, Springer-Verlag, pp. 115-125.

Worboys, M.F. (1990a). Reasoning About GIS Using Temporal and Dynamic Logics. *Proceedings of Temporal GIS Workshop*, Orono, ME, University of Maine.

Worboys, M.F. (1990b). Temporal Logic. *Temporal Relations Workshop*. Orono, ME, University of Maine.

Worboys, M.F. (1990c). The Role of Modal Logics in the Description of Knowledge in a Geographical Information System. *Cognitive and Linguistic Aspects of Geographic Space: an Introduction*, D. Mark and A. Frank, Eds., Kluwer Academic, pp. 403-414.

Worboys, M.F. (1992). A Model for Spatio-Temporal Information. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Columbia, SC, IGU Commission on GIS, pp. 602-611.

Worboys, M.F. (1992). Object-Oriented Models of Spatiotemporal Information. *Proceedings of GIS/LIS '92*, Bethesda, MD, ASPRS/ACSM, pp. 825-834.

Worboys, M.F. (1994). A Unified Model for Spatial and Temporal Information. *The Computer Journal*, Vol. 37, pp. 27-34.

Worboys, M.F. (1994). Object-Oriented Approaches to Geo-Referenced Information. *International Journal of Geographical Information Systems.*, Vol. 8, pp. 385-399.

Worboys, M.F., H.W. Hearnshaw, and D.J. Maguire (1990). Object-Oriented Data Modeling for Spatial Databases. *International Journal of Geographical Information Systems*, Vol. 4, pp. 369-383.

Worboys, M.F.S., P.A. (1995). A Design Support Environment for Spatio-temporal Database Applications. *Spatial Information Theory: A Theoretical Basis for GIS, International Conference COSIT '95, Sep. 21-23 1995, Semmering, Austria*.

Xiao, Q., H. Raafat, and D. Gauthier (1989). A Temporal/Spatial Database Structure for Remotely Sensed Image Data Management Within GIS. *Proceedings of GIS/LIS '89*, Bethesda, MD, ASPRS/ACSM, pp. 116-122.

Xu, X. (1990). *Extending Relational DBMS For Spatiotemporal Information*, Computer Science, Simon Fraser University, Canada.

Xu, X., J. Han, and W. Lu (1990). RT-Tree: An Improved R-Tree Index Structure for Spatiotemporal Databases. *Proceedings of the 4th International Symposium on Spatial Data*

Handling, Columbia, SC, IGU Commission on GIS, pp. 1040-1049.

Yao, E.C.M. (1994). *An Object-Oriented Data Model for Spatial Data*, Computer Science, University of Guelph, Canada.

Yates, P. and L.W. Crissman (1993). Design and Implementation of Historical Dimensions in the SIIASA/China GIS Project Databases. [Available online from http://www.asian.gu.edu.au/papers/historical.html.]

Yau, M.-M. and S.N. Srihari (1983). A Hierarchical Data Structure for Multidimensional Digital Images. *Communications of the Association of Computing Machinery*, Vol. 26, pp. 504-515.

Yearsley, C.M. and M.F. Worboys (1995). A Deductive Model of Complex Planar Spatio-Temporal Objects. *Innovations in Geographical Information Systems 2*, P. Fisher (Ed.), Taylor and Francis, pp. 43-51.

Yeh, T. (1992). Temporal Aspects of Geographical Databases. *Proceedings of EGIS '92*, Munich, EGIS Foundation, pp. 320-328.

Yeh, T. (1995). Modeling Highly Variable Spatio-temporal Data. *Australian Computer Science Communications*, Vol. 17, pp. 221-230.

Yeh, T.S. (1995). A Model for the Management of Highly Variable Spatio-temporal Data. *Proceedings of the Joint European Conference and Exhibition on Geographical Inferences. (JEG-CI 95)*, The Hague.

Yuan, M. (1996). Temporal GIS and Spatio-temporal Modeling. [Available online from http://ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papaers/yuan_may/may.htm.]

Zeiler, M. (1999). *Modeling Our World: The ESRI Guide to Geodatabase Design*. ESRI Press.

Zhang, Y. (1994). Supporting Scientific Data Analysis in a Visual Environment (Data Analysis), Computer Science, Worcester Polytechnic Institute.

Zhao, F. and J.O. Brien (1999). Considerations in the Design of a GIS-Based Tool for Visualizing Land Use-Transportation Interactions. *Geographic Information Systems for Transportation Symposium/Interoperability of Transportation GIS*, San Diego, CA.