Monitoring Land Use Activity Changes in the Twin Cities Metropolitan Region

Report # 7 in the series

Access to Destinations Study

Report # 2008-26

Sponsored by:

University of Minnesota
Center for Transportation Studies

Minnesota Department of Transportation

Hennepin

Metropolitan Council

Technical Report Documentation Page

1. Report No. MN/RC 2008-26	2.	3. Recipients Accession No.
4. Title and Subtitle Access to Destinations: Monitoring Land Use Activity Changes in the Twin Cities Metropolitan Region		5. Report Date July 2008 6.
7. Author(s) Michael Iacono, David Levinson, Ahmed El-Geneidy, Rania Wasfi and Shanjiang Zhu		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil Engineering University of Minnesota 500 Pillsbury Drive SE Minneapolis, Minnesota 55455		10. Project/Task/Work Unit No. 11. Contract (C) or Grant (G) No. (c) 89261 (wo) 25
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Research Services Section 395 John Ireland Boulevard St. Paul, Minnesota 55155		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes http://www.lrrb.org/PDF/200826.pdf Report #7 in the series: Access to Destinations Study		

16. Abstract (Limit: 200 words)

This study presents an effort to track and model land use change in the Twin Cities Metropolitan Region. To that end, we make use of a unique, high-resolution, cell-level set of land use data for the Twin Cities. The data represent 75 meter by 75 meter land use cells, observed at several points in time during the period from 1958 to 2005. These data are used to validate three different types of land use models, which then are used to forecast land use several decades into the future. The models applied in this study include Markov Chain models, Markov Chain- Cellular Automata (MC-CA) models, and an empirical model based on a logistic regression specification. The models are intended to have a simple, transparent structure that allows the user to identify sources of forecast error. Forecasts of land use are made both for the entire study area and also for a specific corridor along State Highway 610 in the northwestern suburbs of the Twin Cities. The study concludes with a brief discussion of the limitations of the models, and how they might meaningfully be expanded and applied.

17. Document Analysis/Descriptors Land use, Twin Cities (Minnesota), Mathematical models, Historical land use patterns,	Urban growth, Land use-forecasting, Markov Chain model	18. Availability Statement No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161	
19. Security Class (this report) Unclassified	20. Security Class (this page) Unclassified	21. No. of Pages 107	22. Price

Access to Destinations:

Monitoring Land Use Activity Changes in the Twin Cities Metropolitan Region

Report # 7 in the series

Final Report

Prepared by:

Michael Iacono
David Levinson
Ahmed El-Geneidy
Rania Wasfi
Shanjiang Zhu

Department of Civil Engineering
University of Minnesota

July 2008

Published by:

Minnesota Department of Transportation
Research Services Section
395 John Ireland Boulevard, MS 330
St. Paul, Minnesota 55155-1899

Contents

List of Tables

List of Figures

Executive Summary

As cities grow and change, transitions among types of land use are among the most visible outcomes. Vacant land at the urban fringe is converted to urban uses, such as housing, retail and office space. Abandoned central city parcels become candidates for redevelopment toward different uses and/or intensities. Major transportation corridors become magnets for clusters of activity. But how and why does land use in cities organize the way it does?

The answers are surprisingly complex, and recent efforts to try to model and forecast many of the elements of metropolitan regions, including land use, largely reflect this complexity. These models treat land use outcomes as a function of the interactions between transportation networks and urban land markets, with some even attempting to model flows within the urban economy. Decisions about the location and intensity of urban land use typically involve some measure of $accessibility$ as the moderating influence that ties together the various processes at work within metropolitan regions.

Efforts to model urban land use change at a greater level of detail have inevitably resulted in models that are resource intensive (in terms of data collection, operation and maintenance) and more difficult to interpret. In this study, we attempt to recast the process of modeling and forecasting land use change in simpler terms. We introduce some models that strip the process of land use change down to a few basic principles and then test the models using data for the Twin Cities from 1958 to 2005.

Three types of models are employed to attempt to reproduce patterns of land use change over the period from 1958 to 2005. The first type of model is called a *Markov Chain* model, and operates on the basic principle that land use at the next time period is solely a function of current land use. Transitions between land use types are embodied in a matrix of transition probabilities that are based on actual observed land use change between two time periods. If we assume that this set of transition probabilities remains constant over time, it can be used to forecast several periods into the future.

The second type of land use change model is an empirical model that attempts to identify the determinants of land use change over time. We begin with the concept of the Markov Chain model, using current land use as a predictor, then add several variables relating to neighboring land use, proximity to highways and regional accessibility. The functional form of the model is that of a *logistic regression* model, where the outcome (cell-level land use) is discrete and the model predicts the probability of transition to each outcome.

The third model is a *Cellular Automata-Markov Chain* (CA-MC) model. This model extends the framework of the Markov Chain model to include the effects of neighboring land use. Again, the concept of a transition matrix is used to define probabilities of change between land use types. The difference in this case is that initial states are defined not only by land use in a cell, but also by the two most common land use types in neighboring cells. This additional feature makes the number of possible types of transitions much larger, since each initial state is defined by three factors (land use in a central cell and two neighboring land uses).

A unique data set is created in this study to test the predictions of the various models. Building on a parcel-level data set furnished by the Metropolitan Council, the regional planning agency for the seven core counties of the Twin Cities metropolitan area, we create a new, cell-level data set that divides parcels into 75 meter by 75 meter square cells that are classified according to the predominant land use type in the cell. Following the original classification of the land use data, 10 land use classes are defined. The original parcel-level data were available going back as far as 1984, but were supplemented with additional land use data in the form of hard-copy maps of land use. These maps, once digitized, provide data for the years 1958, 1968 and 1978. In all, the data cover the period from 1958 to 2005 and were collected at 10 different points in time.

Our method of validating our models of land use change is to use the historical land use data to forecast land use for later years in the data set, a process commonly known as *backcasting*. Once validated, the models are used to forecast cell-level land use a few decades into the future. Our forecasts were made for two different study areas. One study area comprises most of the metropolitan region and corresponds to the 1958 boundaries of the region. The other is a much smaller section of the northwestern part of the region, corresponding to a corridor along State Highway 610, a recently-built, four-lane freeway. Since this corridor is still growing and contains much land that remains unimproved, it provides an opportunity to forecast land use change in response to a major, new link in the transportation network.

Results of the modeling efforts indicate none of the models presented are able to fully reproduce patterns of land use change over time. The Markov Chain model, because of its probabilistic nature, tends to produce more dispersed and mixed patterns of land use than actually occur. Adding the influence of neighboring land uses to form the CA-MC model reduces some of this error, but still leaves a significant number of cells incorrectly predicted. The logistic regression model does a better job of predicting the spatial clustering of different types of land use, particularly commercial and industrial land uses, but tends to overconcentrate some land uses (e.g. residential) and does permit as much mixing of land uses, as is observed in many of the older parts of the region.

While the models of land use change introduced here have somewhat limited predictive power, especially in the longer term, they do contain some desirable features. Their structure is simple and transparent, as are the assumptions that underlie them, which makes easier the task of tracing the source of land use change over time. This makes them ideal for certain sketch planning applications. Also, their simple structure allows them to be extended rather easily and to incorporate new features of urban growth processes to add complexity and realism to the models.

Chapter 1

Introduction

As cities grow and change, transitions among types of land use are among the most visible outcomes. Vacant land at the urban fringe is converted to urban uses, such as housing, retail and office space. Abandoned central city parcels become candidates for redevelopment toward different uses and/or intensities. Major transportation corridors become magnets for clusters of activity. But how and why does land use in cities organize the way it does?

The answers are surprisingly complex, and recent efforts to try to model and forecast many of the elements of metropolitan regions, including land use, largely reflect this complexity. These models treat land use outcomes as a function of the interactions between transportation networks and urban land markets, with some even attempting to model flows within the urban economy. Decisions about the location and intensity of urban land use typically involve some measure of $accessibility$ as the moderating influence that ties together the various processes at work within metropolitan regions.

Efforts to model urban land use change at a greater level of detail have inevitably resulted in models that are resource intensive (in terms of data collection, operation and maintenance) and more difficult to interpret. In this study, we attempt to recast the process of modeling and forecasting land use change in simpler terms. We introduce some models that strip the process of land use change down to a few basic principles and then test the models using data for the Twin Cities from 1958 to 2005.

The models introduced in this study rely heavily on one or more of the following principles. First, land use change is treated as a stochastic process, that is, there is some inherent randomness to the process of land use change. All models are abstractions and necessarily involve some degree of error that their predictions cannot account for. This is especially true in the case of land use models, where the process that the model attempts to replicate is exceedingly complex. Second, land use change is a slow process, and there is a great deal of inertia for existing land uses. This is especially true of developed areas, where an existing stock of buildings represents a formidable impediment to rapid change. Hence, over relatively short periods, parcels are likely to remain in the same land use. Third, land use at a particular location is strongly influenced by neighboring land uses. This influence should be especially strong in the case of residential land, where certain other land uses (e.g. industrial) may be viewed as incompatible. Thus, we would expect residential neighborhoods to remain residential over significant amounts of time. Fourth, land use is responsive to proximity to transportation networks and the accessibility they provide. Locations that become served by new transportation links or otherwise see their accessibility increase become likely candidates for land use transition. We would particularly expect to see this in the case of vacant land.

The remainder of this report is laid out in the following sections. The next section provides a brief introduction to many of the current and previous approaches to modeling land use change, with an emphasis on models that integrate land use with transportation networks. The following section introduces each of the three models that will be applied in this study and describes their relevant methodological aspects. The fourth section briefly describes the cell-level land use data set that is used in the modeling exercises. The fifth section describes the application and validation of the land use models using historical data for the Twin Cities. The sixth section describes some simulations of land use change using the validated models to predict land use change in future years, both for the region as a whole and for a specific study area along State Highway 610 in the northwestern part of the Twin Cities region. Lastly, the paper concludes by observing the strengths and weaknesses of the three models and commenting on their appropriateness for planning applications.

Chapter 2

Models of Land Use and Transportation Change

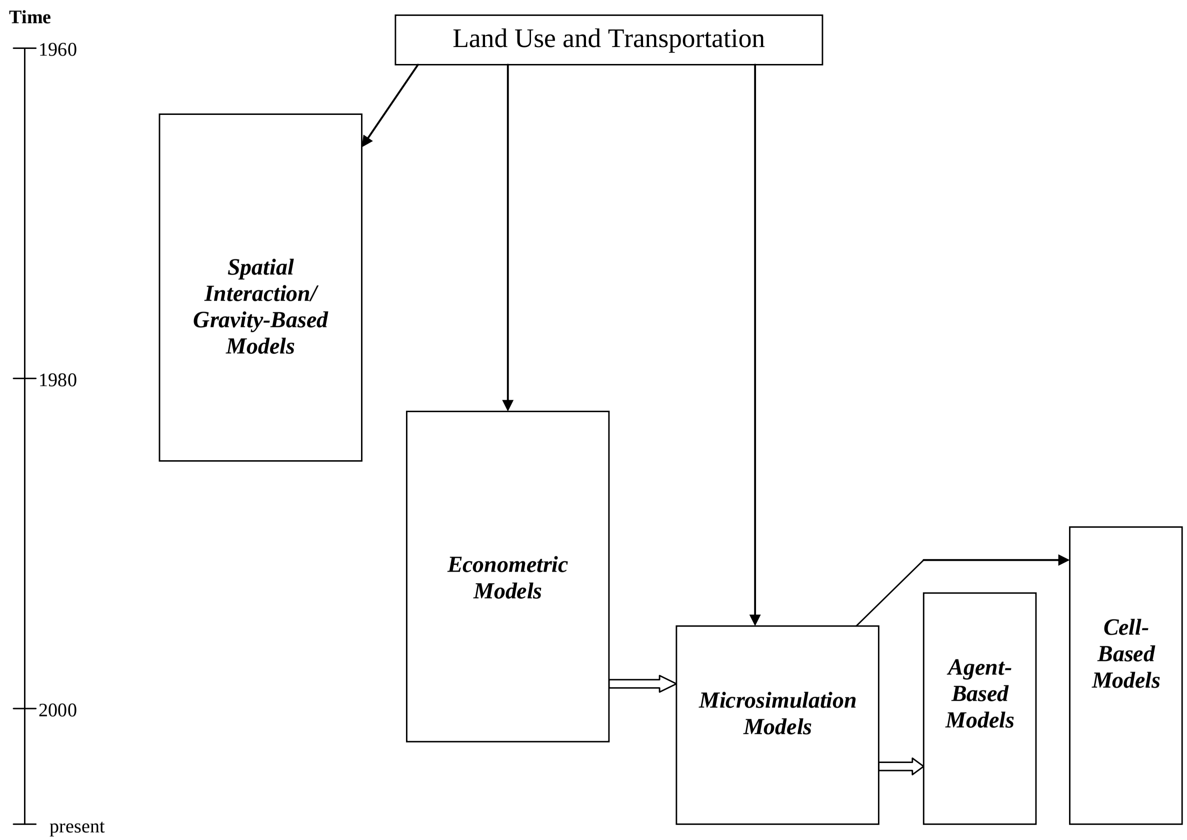
Currently, there are a number of operational models that have been developed to simulate the relationship between transportation networks and patterns of land use. Here we will give a brief overview of some of the modeling frameworks that help to distinguish different types of models. A more thorough treatment of this topic is provided in an appendix at the end of this report.

2.1 Chronology of Model Development

The history of simulation models of transportation and land use is dated back to the late 1950s (Batty, 1979). While models of regional travel demand had been established as far back as the early 1950s and some early experiments with transportation and land use models were carried out in the following years, it wasn't until the early 1960s that the first operational land use simulation model was built. The *Model of Metropolis* developed by Lowry (1964) is widely considered to be the first operational simulation model of urban land use. Lowry's model was the first of a generation of models based on theories of spatial interaction, including the gravity model that was popular in quantitative geography at the time. Models based on a spatial interaction framework continued to be developed through the early to mid-1980s, when they became largely replaced by models grounded in random utility theory and econometric methods.

Figure 2-1 describes this process and gives an approximate timeline for the adoption of various modeling frameworks within transportation and land use research. Several of the models that follow an econometric framework continue to be used today, although some, like the UrbanSim simulation system (Waddell, 2002*b*; Waddell et al., 2003) are being redeveloped within a microsimulation design. The broad class of transportation and land use models that could fall under the title of 'microsimulation' began to be developed in the early 1990s, in parallel with major improvements in computational power that allowed for their operation. These included prototype models of activity-based travel, cell-based models land use change and the introduction of multi-agent models for urban simulation. More recently, some researchers have begun to devote effort to developing comprehensive urban microsimulation models that fully reflect the dynamics of changes in the population and the urban environment within which they make choices.

Figure 1: Chronological development of land use and transportation models



38

Figure 2-1: Chronological development of land use and transportation models

2.2 Spatial Interaction Models

The earliest class of land use and transportation simulation models are a set of highly aggregate models based on principles of spatial interaction that were popular in the regional science and quantitative geography fields in the 1950s and 1960s. There were many different formulations of this type of model, though most revolved around variations of the gravity model, an adaptation from Newtonian physics. The derivation of the gravity model from principles of entropy maximization (Wilson, 1967, 1970) was a major accomplishment and formed the basis for many of the allocation mechanisms within spatial interaction models. A general form of the gravity model can be expressed as:

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) \tag{2.1}$$

where T_{ij} represents trips (or other measures of interaction) between two zones, O_i represents origins at zone i, D_j represents destinations to zone j, and A_i and B_j are balancing factors to ensure that total origins equal total destinations. The exponential term in the model is used to capture the effect of decreasing interaction as a function of travel cost, including travel time.

As mentioned previously, the first operational land use simulation model was the model developed by Lowry (1964) for the Pittsburgh region. This model has great importance, since many of the other land use and transportation models that follow a spatial interaction framework have similar structures. A detailed review of this model and its variations are provided in Horowitz (2004). Table 2-1 provides a list of some of the more well-known operational models based on a spatial interaction framework.

Table 2-1: Summary of Spatial Interaction / Gravity Models

Model	Reference	Distinguishing Features
Model of Metropolis	Lowry (1964); Garin (1966)	First recognized operational land use-transportation model; Garin provided matrix representation
TOMM	Crecine (1964)	Disaggregation of population; incorporation of inertia effects in activity allocation
PLUM	Goldner (1971)	Replaced standard gravity model with intervening opportunity model; use of county-specific dispersion parameters
ITLUP	Putman (1983)	First complete software package for integrated modeling; improved calibration techniques; improved network model with multiple modes; incorporation of congestion effects in activity allocation
LILT	Mackett (1983)	Use of accessibility function; car ownership submodel; land use model capable of handling demolition, changing occupancy and vacancy rates
IRPUD	Wegener (1982)	Contains seven separate submodels; microsimulation of land use; use of differing spatial scales for submodels; separates discretionary and non-discretionary travel

2.3 Econometric Approaches

One of the major shortcomings of the aggregate spatial interaction models was the absence or use of inappropriate theory to describe the behavior captured in the model. Developments in the use of random utility theory to describe choices among discrete alternatives, such as the choice of travel mode, provided the impetus for a new generation of models based on the study of disaggregate behavior. When it was shown that discrete choice models could be applied to problems such as residential location (Lerman, 1976; McFadden, 1978), researchers began to look for ways to model the interrelated choices individuals made in terms of location and travel behavior.

Land use and transportation models that follow econometric frameworks can be thought of as comprising two types of models: regional economic models and land market models. In these two types of simulation models the economic model and the land market model each form the core of a simulation system that includes the prediction of transportation flows. Both types tend to have improved representation of land markets that include endogenously-determined (determined within the model) prices and market clearing mechanisms. A summary of these models and their characteristics are provided in Table 2-2.

2.4 Disaggregate and Microsimulation Models

Since the late 1980s, advances in computing power and efficiency of data storage have allowed researchers to begin to build models that address many of the shortcomings associated with previous large-scale modeling efforts and represent important change processes in cities with the detail they require. Examples of these include activity-based models of travel behavior, multi-agent models of urban land use and transportation, and cell-based models of urban land use. The common conceptual underpinning of each of these models is that they attempt to represent processes of change from the bottom up, that is, they account for the behavior of individual agents in space and/or time, along with interactions between agents. The use of the term *microsimulation* can be applied to each of these types of models, though it requires some definition. As defined by Miller (2003), microsimulation relates to "a method or approach (rather than a model *per se*) for exercising a disaggregate model over time." All of the types of models identified above are what would be considered disaggregate models and all have a significant temporal element. Microsimulation methods are particularly effective for modeling systems that are dynamic and complex, which urban systems invariably are. A sample of some of these models is provided in Table 2-3.

Table 2-2: Summary of Econometric Models

Model	Reference	Distinguishing Features
CATLAS	Anas (1982)	Improved representation of economic agents and decision making; explicit treatment of housing markets; economic analysis capabilities
MEPLAN	Echenique et al. (1969); Echenique et al. (1990)	Incorporation of spatial input-output model with economic evaluation component; able to forecast commercial trip generation; travel treated as a derived demand
TRANUS	de la Barra (1989)	Development supply model simulates choices of developers; sophisticated travel model with combined mode-route choice
MUSSA	Martinez (1992)	Incorporation of bid-rent framework for land, floor space markets; detailed representation of transit network in travel model; high level of disaggregation of household types
METROSIM	Anas (1994)	Model extended to commercial real estate markets; addition of dynamic CH-PMM housing market model
NYMTC-LUM	Anas (1998)	Endogenous determination of housing prices, floor space rents, and wages; high level of spatial disaggregation suitable for transit and land use policy evaluation
DELTA	Simmonds (1999)	Microsimulation of demographic changes; treatment of quality in the market for space
PECAS	Hunt and Abraham (2005)	Regional econometric model with microsimulation of land development at the parcel level; ability to couple with an activity-based travel model and to apply at supra-regional level

Table 2-3: Summary of Microsimulation Models of Transportation and Land Use

Model	Reference	Distinguishing Features
ILUTE	Salvani and Miller (2005)	Comprehensive urban system microsimulation model; structured to accurately capture temporal elements of urban change; activity-travel model includes household member interactions; disequilibrium modeling framework
ILUMASS	Moeckel et al. (2003); Strauch et al. (2003)	Descendent of IRPUD model; incorporates microscopic dynamic simulation model of traffic flows and goods movement model; designed with environmental evaluation submodel;
Ramblas	Veldhuisen et al. (2000)	Entirely rule-based model framework; designed to simulate very large populations
UrbanSim	Waddell et al. (2003)	Land use model incorporating microsimulations of demographic processes and land development; parcel-level land use representation; high level of household type disaggregation; open-source software developed for general use

Models of transportation and land use change have evolved significantly since their early applications more than four decades ago. In the search to design models that capture the recursive relationship between transportation and land use, there has been a general trend toward the disaggregation of the representation of people and space. Newer models represent in greater detail the dynamics of the transportation-land use change process. Experiments with bottom-up approaches to modeling urban systems, especially those that recognize the interactions between agents, provide an alternative means for understanding their complexity. Yet, the ability to forecast these processes for policy applications remains an important goal. Most of the newer generation of microsimulation models are designed with the objective of making them more policy sensitive. Unfortunately, few of them have yet reached a point where they can be fairly evaluated on this criterion, and the older operational models still raise important questions about the utility of such complex tools.

Thus, at present we will turn our attention to simpler models that attempt to recreate processes of land use change at the microscopic level using a few basic principles. In the next section, the models to be applied in the present study will be introduced, along with their methodological underpinnings. The following sections will demonstrate the application of these models to land use in the Twin Cities metropolitan region.

Chapter 3

Methodology

The preceding review of models of land use and transportation change gives an indication of how land use is typically treated in operational models used for planning purposes. There has been a general trend toward the disaggregation of land use units from zones that comply with the types of units employed in travel demand analysis models to much-smaller units that more closely approximate neighborhood or even individual land parcels. It is this more disaggregate type of unit that we will employ in the present analysis. The data set used for this study will be discussed in more detail in the next section, but for now attention will be turned to the three types of land use models employed in this study. They are *Markov Chain* models, *Logistic Regression* models and *Markov Chain-Cellular Automata* models. The formulation of each model will be discussed, along with its method of application.

3.1 Markov Chain Model

The basic premise of the Markov chain model is that land use at some point in the future $(t+1)$ can be determined as a function of current land use (t), or mathematically,

$$X_{t+1} = f(X_t) \tag{3.1}$$

where X_{t+1} represents the land use at time $t+1$ and X_t represents land use at time t. The structure of the Markov chain model as applied to land use change involves a vector ($\boldsymbol{n}_t$) with dimension $mX1$ (where m represents the number of states, in this case land use classes) describing the distribution of land use among current states and an mXm matrix of transition probabilities (***P***) that governs the probability of transition between each pair of land uses, i and j. The model can then be written as a difference equation in matrix form (Baker, 1989)

$$\mathbf{n}_{t+1} = \mathbf{P}\mathbf{n}_t \tag{3.2}$$

where $\boldsymbol{n}_{t+1}$ is another $mX1$ column vector describing the distribution of land use at time $t+1$. Since the transitions are probabilities, it follows that:

$$\sum_{j=1}^{m} p_{ij} = 1 \quad i = 1, 2, ...m \tag{3.3}$$

meaning simply that the rows of the transition matrix must sum to 1. Maximum likelihood estimates of the transition probabilities can be obtained as (Anderson and Goodman, 1957):

$$\hat{p}_{ij} = n_{ij} \Bigg/ \sum_{j=1}^{J} n_{ij} \tag{3.4}$$

where p_{ij} is the probability of transition between i and j, J is the number of columns in the transition matrix, and n_{ij} denotes the number of transitions from i to j. These values can all be obtained empirically.

To test the validity of the Markov Chain model, a useful first step is to test the null hypothesis that land use at one point in time, $t + 1$, is statistically independent of land use at the preceding time period, t. This test can be conducted using standard contingency table techniques for cross-classified categorical data. The expected values for each cell indicating the number of transitions between i and j can be compared with the actual number of transitions to compute the test statistic, Pearson's chi-square, which is distributed χ^2 with $(M - 1)^2$ degrees of freedom, where M indicates the number of land use classes (in this case 10). Under the hypothesis of independence, the expected number of transitions in each cell of the transition matrix $(\hat{m}_{ij})$ can be calculated by:

$$\hat{m}_{ij} = n_{i+} n_{+j} \tag{3.5}$$

where n_{i+} denotes the marginal total of transitions for the i^{th} row of the transition matrix and n_{+j} denotes the marginal total for the j^{th} column of the transition matrix.

Using these expected values, the test statistic (K^2) then takes the form:

$$K^2 = \sum_{i=1}^{I} \sum_{j}^{J} \frac{(n_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}} \tag{3.6}$$

The test statistic is typically given the notation K^2 instead of X^2 to differentiate it from its distribution, which is chi-square. The null hypothesis of independence is almost universally rejected, indicating some level of dependency between successive land use states.

Stationarity is another important property of Markov chains, particularly as it applies to the transition probability matrix. This property is critical for applications in which a Markov chain model is to be used for forecasting. The transition probability matrix (***P***) is assumed to remain constant in successive periods, meaning that at any future period $t + k$, the matrix of cell transitions can be obtained by multiplying the vector of current land uses, $\boldsymbol{n}_t$ by the transition probability matrix ***P***, raised to the k^{th} power ($\boldsymbol{P}^k$) In most forecasting applications, the transition probability matrix is assumed to remain constant through successive time periods, and is seldom tested empirically. This study follows the work of Bourne (1971), who compared transition matrices for successive periods using simple correlations between cells of the matrix. By expressing the elements of one matrix ($\boldsymbol{P}_{t+1,t+2}$) as a function of another ($\boldsymbol{P}_{t,t+1}$), one can provide a rough check for stationarity by determining whether the correlation between matrix elements is significantly different from a value of one.

In order to use the Markov Chain model for prediction, an additional stochastic element is added. Since the transition probabilities represent estimates of the likelihood of conversion from one land use state at time t to one of 10 other states at time $t + 1$, a mechanism is added to introduce randomness to the model and its predictions of future states. Since each row of the transition probability matrix sums to one, predictions of future land use states are obtained by drawing a pseudo-random number between zero and one, rounded to four digits. If the number falls within the probability space allocated to a particular land use state according to the transition matrix, then that state is chosen for conversion. This process is repeated for each land use cell in the data set. Predicted land uses can then be compared to actual observed land uses to summarize the accuracy of the model's predictions.

3.2 Regression Modeling

In addition to the Markov chain technique, one could also formulate the problem of predicting land use change using a variant of the classical regression model. In this case, the objective would be to find a set of covariates that serve as reliable predictors of land use change, using the land use cells as units of analysis. Various factors relating to neighboring land uses, existing or previous land uses in a given cell, and position relative to transportation networks could be incorporated into the specification of the model.

The model used to predict land use change treats land uses as a set of discrete states, which conforms well to the notion of land use classes as qualitative variables. The logistic regression model is an adaptation of the linear regression model which allows the dependent variable to be specified as a discrete, rather than continuous outcome. The model predicts the probability of a given outcome, conditional upon the presence of a set of attributes. We can consider the various land uses as constituting J separate outcomes. The outcomes can then be related to the attributes ($\boldsymbol{x}_i$) by a linear predictor of the form:

$$\boldsymbol{y} = \boldsymbol{X\beta} \tag{3.7}$$

The linear predictor is analogous to the utility function commonly employed in econometric choice modeling. If the individual land use cells are each denoted with the subscript i, then for a specific land use type j, the probability that the observed outcome (y_i) will be equal to j is given by the expression:

$$P(y_i = j) = \frac{exp(\boldsymbol{X}_i\boldsymbol{\beta}_j)}{1 + \sum_j^J exp(\boldsymbol{X}_i\boldsymbol{\beta}_j)} \tag{3.8}$$

where $\boldsymbol{X}_i$ is a vector of explanatory variables (attributes) for land use cell i, $\boldsymbol{\beta}_j$ is a vector of unknown parameters to be estimated, typically by the method of maximum likelihood, and J denotes the set of all outcomes/land use types.

In estimating models of land use change, several land uses will be dropped from the analysis. In particular, parks, public land uses, airports, railways and cells covered by water will be excluded. These land uses are considered to either be fixed in nature or relatively unresponsive to land market forces that drive much of the change observed in the remaining land uses. This narrows the set of possible outcomes from ten land uses to five. Furthermore, in each model, one category of the dependent variable must be designated as a comparison category. This category is then omitted from the analysis. The parameter estimates are interpreted as *relative risk ratios*, which are the exponentiated beta coefficients. These ratios represent the change in the odds of being in the dependent variable category versus the comparison category associated with a one unit change in the independent variable.

Since we are interested in predicting the *change* in land use states over time, the models to be estimated will employ a dynamic specification. That is, we are interested in predicting land use in cell i at time t (L_{it}) using observations on several variables at a previous time period ($t-1$). The set of variables that are defined as covariates relate to a cell's previous land use and that of its neighboring land uses, the presence of a highway network in the cell and its neighbors, and the accessibility of the cell at time $t-1$. The variables are formally defined as follows:

- $L_{i,t}$ represents land use in cell i at time t
- $L_{i,t-1}$ represents land use in cell i at time $t-1$
- $L_{j,t-1}$ represents the number of neighboring cells in the adjacent *Moore neighborhood* in land use j at time $t-1$. This variable is defined for residential, commercial, industrial and vacant land uses and is denoted, respectively, as (L_R, L_C, L_I, and L_V).
- $A_{i,t-1}$ represents the regional accessibility to employment in cell i at time $t-1$. The measure is extracted from the larger transportation analysis zone within which cell i is located.
- $N_{i,t-1}$ is a dummy variable representing the presence/absence of land classified as 'highway' in cell i at time $t-1$, and serves as a measure of proximity to transportation networks which might be expected to influence the land use in i at time t. The 'highway' classification is applied to roads and adjacent highway-related land along state trunk highways and federal primary and secondary highways (Interstate and U.S. highway system) in the National Highway System.
- $N_{j,t-1}$ represents the number of neighboring cells containing land classified as 'highway' at time $t-1$.

The resulting model is then written in general form, relating land use at time t to the above variables:

$$L_{i,t} = f\left(L_{i,t-1}, \boldsymbol{L}_{j,t-1}, A_{i,t-1}, N_{i,t-1}, N_{j,t-1}\right) \tag{3.9}$$

This expression represents the probability of observing a particular land use in cell i at time t, given the set of covariates. The covariates represent the inputs to the linear predictor in (8), which are in turn used to predict the probabilities from (9).

Hypotheses can be stated regarding the expected effects of the variables introduced above. The variable $L_{i,t-1}$ is expected to increase the probability of observing the same land use in i at time t. The variables representing land use in adjacent cells, $L_{j,t-1}$, are expected to increase the probability of transition of land use in cell i to the predominant land use in the neighborhood (j). The accessibility to employment at time $t-1$ variable is expected to increase the likelihood of transition to residential land use at t. Finally, the two variables representing the presence of highway-related land in cell i and its neighborhood (j) are expected to increase the likelihood of transition to industrial or commercial uses, which might be expected to benefit more from proximity to highway networks.

3.3 Markov Chain-Cellular Automata (MC-CA) Modeling

The original Markov chain modeling approach had the desirable property that its structure is quite simple, yet retains the qualities of a stochastic representation of land use. In the logistic regression formulation of the land use change phenomenon, neighboring land uses were added to give the model additional explanatory power. If we combine this treatment of neighboring cells with the original structure of the Markov Chain model, it becomes possible to consider a model where the effects of neighboring cells can be incorporated directly into the transition probabilities specified in the Markov Chain model. The resulting model is referred to here as a Markov Chain-Cellular Automata model, since it combines the probabilistic elements of a Markov Chain (MC) with the neighborhood effects that are the hallmark of most models based on the cellular automata (CA) modeling framework.

The structure of the MCCA model is fundamentally similar to that of the original Markov Chain model, as are the assumptions that underlie its application. The information about neighboring land uses is contained in the definition of the cell states. Each cell state is comprised of a land use type, along with the two most predominant land uses in the neighboring cells (Moore neighborhood). "Highway" is defined as a separate land use, as in the Markov Chain model, since it is hypothesized to have an effect on other land uses, particularly as a neighbor. A cell state can then be defined by referring to the land use in a cell and its two most prevalent neighbors. For example, a state could define a cell as having Commercial land use with primarily Commercial and Industrial neighbors, denoted as C_{CI}.

Enumerating the possible outcomes from this model requires considering each of the combinations of current land use and neighboring land uses. With 10 land use types defined there are $10^2 = 100$ unique combinations of neighboring land uses for each initial land use in a cell, defining $10^3 = 1,000$ initial cell states. With 10 possible land uses as outcomes, there are a total of $10^4 = 10,000$ possible state transitions in the model, each of which has an observed probability of conversion (though many of these observed probabilities will be zero).

The model is run in a similar fashion to the Markov Chain model, with quasi-random numbers being drawn to simulate the process of cell transition during the period $t, t+1$, according to the probabilities specified for each state. The predictions of land use in each cell at $t+1$ can then be compared to actual observed outcomes to evaluate the model's accuracy. It is expected that the MCCA model should offer some improvement over the predictions of the basic Markov Chain model.

Chapter 4

Data

The land use data employed in this study build from a previous set of land use data used by Levinson and Chen (2005) in an earlier study of the Twin Cities. The expanded data set comprises a time series with observations for the years 1958, 1968, 1978, 1984, 1990, 1997, 2000 and 2005. Land use data for years prior to 1984 were manually digitized from paper copies of land use maps stored at the John R. Borchert Map Library at the University of Minnesota. Data for selected years from 1984 to 2005 were obtained from the Metropolitan Council, the Twin Cities' regional planning agency and designated metropolitan planning organization (MPO), which maintains a parcel-level land use inventory for the region that is updated every few years.

The parcel-level land use data was converted to a raster format and rectified to reduce geometric distortion. Some error remains due to the manual digitization process and the lower level of accuracy associated with earlier mapmaking processes. Differences in classification schemes for land use across years were addressed by adopting a common set of 10 generalized land use classes. These land use classes, along with their adopted abbreviations, include:

- Airports (AIRPOR)
- Commercial (COMM)
- Highway (HWY)
- Industrial (INDUST)
- Parks (PARKS)
- Public (PUBLIC)
- Railroads (RAILWA)
- Residential (RES)
- Vacant (VAC)
- Water (WATER)

The data set covers a large portion of the core seven counties of the Twin Cities region. Some portions of the region could not be covered due to a need to limit the analysis to the part of the region for which common land use data sets could be acquired for each year. The portions left out of the study area are comprised mostly of low-density residential and non-urban uses, which would likely be classified as vacant under the present scheme. The resulting study area covers approximately 3,426 square kilometers (1,322 square miles). The study area is partitioned into a grid of 75-meter by 75-meter cells, a spatial resolution much finer than the 188-meter square cells used in Levinson and Chen's study, leading to a roughly tenfold increase in the number of land use cells in the study area. This produces a data set containing over 610,000 cells. Each cell is assigned a land use class according to its predominant land use. Figure 4-1 presents a summary of trends among the land use classes from 1958 to 2005.

Virtually all land use classes have increased over this period, with the greatest increase in land use registered by the residential category. This growth has largely come at the expense of vacant (including agricultural) land, as the region has been able to accommodate growth over the years via outward expansion.

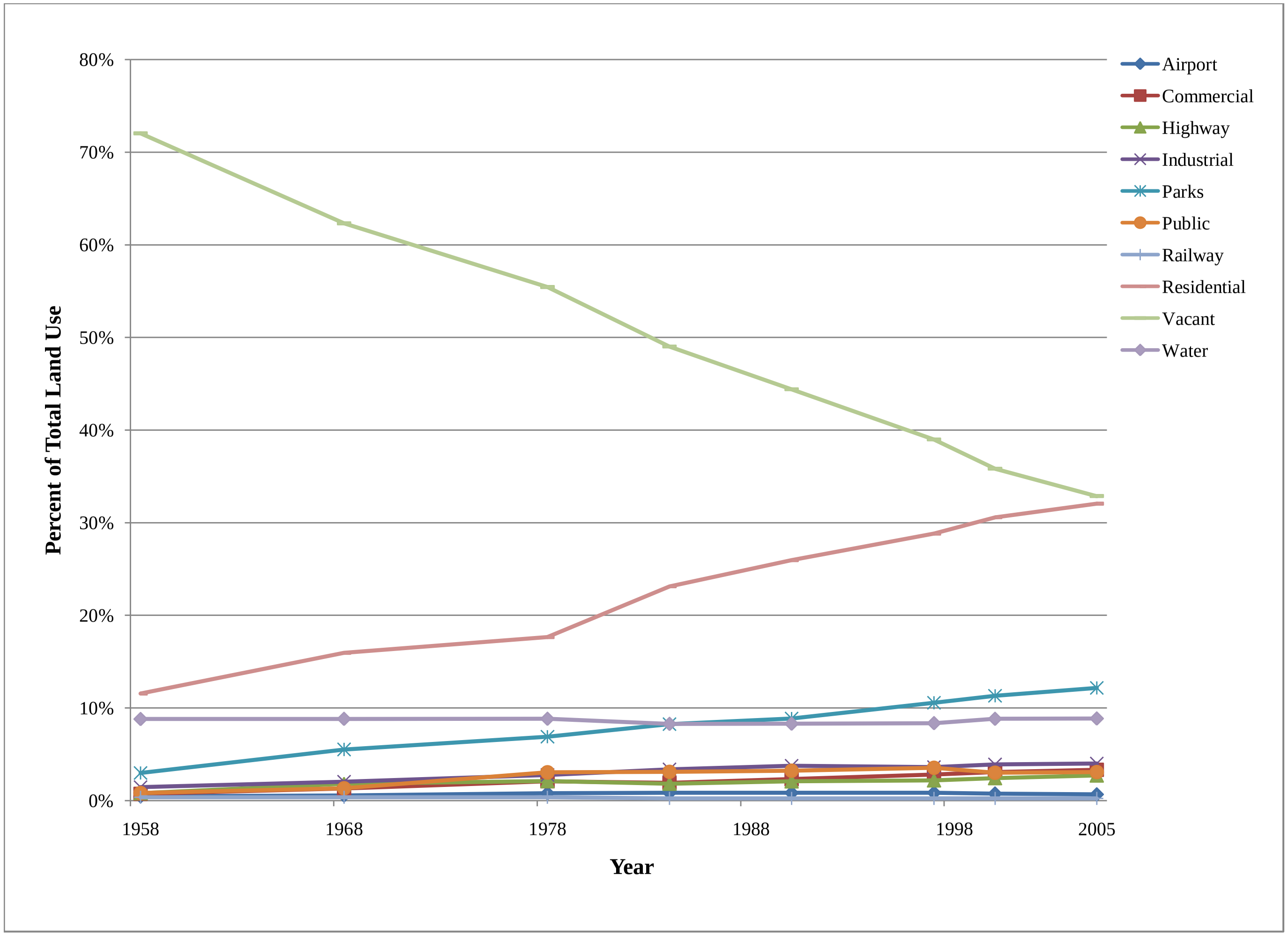


Figure 4-1: Trends in Twin Cities land use, 1958 to 2000

Chapter 5

Model Application and Validation

With a better knowledge of the model structures and data set in hand, we can now apply the models to the Twin Cities and validate their performance using historical data for the region. The first model to be applied will be Markov chain model, followed by the logistic regression and Markov chain-Cellular Automata models.

5.1 Markov Chain Model

The Markov chain model was applied to the Twin Cities land use data in order to test whether the land use dynamics observed could be described by a Markov process, and to provide forecasts of future land use patterns. Results are described in terms of the test for independence, examination of the stationarity of the transition probability matrices, accuracy of predictions made using historical data to backcast land use in successive years, and forecasts of land use decades into the future.

In describing the Markov chain model, it was noted that an elementary test for Markovian models involves the hypothesis of independence between successive time steps. Using the matrix of transitions for the period 1968 to 1978, the observed and expected cell frequencies can be applied to construct the test statistic (K^2). The matrices of observed and expected cell frequencies are shown in Tables 5-1 and 5-2.

The K^2 statistic can be compared to a χ^2 distribution with $(10\text{-}1)^2 = 81$ degrees of freedom. With a critical region of $\alpha = 0.05$, values of the test statistic less than approximately 100 would indicate that land uses in 1978 were independent of those in 1968.

With a computed K^2 of roughly 2.75 x 10^6, this is clearly not the case. Again, it should be noted that in the case of Markov chain models of land use, the hypothesis of independence is nearly always rejected. Historical dependence in land use is a strong force, as is indicated by the primacy of the diagonal elements of the observed transition matrix.

Table 5-1: Observed cell frequencies, 1968–1978

	AIRPOR	COMM	HWY	INDUST	PARKS	PUBLIC	RAILWA	RES	VAC	WATER	Row Totals
AIRPOR	2874	12	2	148	9	90		59	160		3354
COMM	9	3709	59	588	238	508	1	2257	635	2	8006
HWY		15	10989	14	3	13		27	57		11118
INDUST	4	825	138	7641	343	496	21	1114	1886	4	12472
PARKS	383	364	137	378	18440	1191	4	3096	9622	76	33691
PUBLIC	62	226	5	95	354	4984	1	1688	644	3	8062
RAILWA		3		9	3	4	2162	10	18		2209
RES	138	4597	261	1865	5579	5646	14	72155	7089	74	97418
VAC	1375	2976	1304	6198	17038	5676	35	27410	318548	174	380734
WATER	1	2	10	103	3		65	104	53636	53924	
Column Totals	4845	12728	12897	16946	42110	18611	2238	107881	338763 53969	610988	

Table 5-2: Expected cell frequencies, 1968–1978 from Markov Chain model

1968-1978	AIRPOR	COMM	HWY	INDUST	PARKS	PUBLIC	RAILWA	RES	VAC	WATER	Row Totals
AIRPOR	27	70	71	93	231	102	12	592	1860	296	3354
COMM	63	167	169	222	552	244	29	1414	4439	707	8006
HWY	88	232	235	308	766	339	41	1963	6164	982	11118
INDUST	99	260	263	346	860	380	46	2202	6915	1102	12472
PARKS	267	702	711	934	2322	1026	123	5949	18680	2976	33691
PUBLIC	64	168	170	224	556	246	30	1423	4470	712	8062
RAILWA	18	46	47	61	152	67	8	390	1225	195	2209
RES	773	2029	2056	2702	6714	2967	357	17201	54014	8605	97418
VAC	3019	7931	8037	10560	26241	11597	1395	67225	211098	33631	380734
WATER	428	1123	1138	1496	3717	1643	198	9521	29898	4763	53924
Column Totals	4845	12728	12897	16946	42110	18611	2238	107881	338763	53969	610988

Table 5-3: Summary of transition probability regressions from Markov Chain model

				95% C.I.	
Y	X	Adj. R^2	β	Lower	Upper
1968-78	1958-68	0.977	0.98	0.95	1.01
1978-90	1958-68	0.943	0.948	0.902	0.995
1990-2000	1958-68	0.962	1.029	0.988	1.07

Another way to examine the validity of the Markov chain framework is to test the stability or stationarity of the transition matrix. As described in an earlier section, one way to do so is to observe the correlation between the elements of matrices describing the transition probabilities. By regressing the matrix elements of a subsequent time period on a base period, it is possible to determine whether (and how far) the correlations between the two matrices deviate. The matrix of transition probabilities for the period from 1958 to 1968 will serve as a base period, since this is the earliest transition period for which data is available. Table 5-3 shows the results of three successive transition probability matrices being regressed on the original 1958 to 1968 matrix. The X and Y variables denote the response and predictor variables in the regression. The fit of the equation is summarized with the adjusted R^2 value.

The value of the slope coefficient (β) is indicated, along with the lower and upper bounds of a 95% confidence interval for the mean value. In two of the three cases the 95% confidence interval includes the value of one, and in the third case the upper bound falls just short of one. While these results do not provide entirely conclusive evidence on whether the transition matrix is stationary, they offer some confidence that dramatic changes in transition probabilities are not occurring over time. Moreover, even a lack of stationarity need not preclude the use of Markov models. As (Baker, 1989) has noted, stationarity can be assumed as a heuristic device for scenario generation using Markov chains.

It is possible to evaluate how well the Markov chain model predicts land use change by using the historical time series to produce "backcasts" of land use for previous points in time. For example, the 1958 to 1968 transition probability matrix can be used as a base to predict forward in roughly 10-year increments to the years 1978, 1990 and 2000. Due to the different sources of data and data-generating processes noted for the years before and after 1984, we can provide "control" forecasts for the newer data using the 1984 to 1990 transition probability matrix as a base year matrix. These forecasts are provided for the years 1997 and 2005. Again, the land use conversion process in the model is governed by a random number generation procedure that draws values that correspond to the transition probabilities in the matrix for each initial land use state. Forecasts covering more than 10 years use the predicted land use distribution from 10 years prior as inputs to the forecast (e.g. forecast land use for 1990 is used as an input, along with the 1958-1968 probability matrix, for a forecast to the year 2000). This links the forecasts forward through successive time steps and preserves the Markovian principle that future states are only influenced by the present state. Summaries of the accuracy of the forecasts are provided in Table 5-4.

As the results indicate, the accuracy of forecasts made using the 1958 to 1968 matrix of transition probabilities declines sharply over time. While all long-term forecasts can be expected to decline in accuracy the further they are asked to predict, there is a notable decline between the

Table 5-4: Forecast accuracy of Markov Chain model using historical time series data

Base year matrix	Forecast Year	% Correct
1958-1968	1978	70.0
1958-1968	1990	55.2
1958-1968	2000	47.8
1984-1990	1997	84.4
1984-1990	2005	78.5

forecast years 1978 and 1990. This period coincides with the use of different sources of land use data which may not be entirely consistent and which may introduce additional inaccuracy to the forecast. The monotonic decline in accuracy also indicates that errors in forecasts from previous periods are fed forward into subsequent predictions. On the other hand, the forecasts made using a more recent transition matrix (1984 to 1990) as an input show a higher degree of accuracy and a more moderate decline over the second time step. This may be a result of more consistent data as well as a shorter transition period (6 to 8 years).

Lastly, we are interested in using the Markov chain model to predict land use patterns several periods into the future. The most recent land use data are available for the years 1997, 2000 and 2005, indicating that the 1997 to 2005 period most closely matches the 10-year transition periods used throughout this study. Thus, a 1997 to 2005 transition probability matrix can be constructed and used for forecasting in 8-year increments. This matrix is reproduced below in Table 5-5.

The 1997 to 2005 matrix is used to forecast forward through three time steps, yielding land use forecasts for the years 2013, 2021 and 2029. These forecasts are shown below in Table 5-6, along with the land use distribution in 2005, the base year.

Table 5-5: Transition probability matrix for 1997 to 2005 from Markov Chain model

1997-2005	AIRPOR	COMM	HWY	INDUST	PARKS	PUBLIC	RAILWA	RES	VAC	WATER	Row Totals
AIRPOR	0.7388	0.001	0.0068	0.001	0.0325	0.0131	0	0.0055	0.1984	0.0029	1
COMM	0.0001	0.8187	0.0201	0.056	0.0045	0.0227	0.0002	0.0413	0.035	0.0015	1
HWY	0.0004	0.0107	0.9542	0.0054	0.0058	0.0031	0.0002	0.0094	0.0105	0.0001	1
INDUST	0.0004	0.071	0.0099	0.8371	0.0082	0.0086	0.001	0.0106	0.0517	0.0014	1
PARKS	0.0022	0.0036	0.0031	0.0025	0.9128	0.0062	0.0001	0.0116	0.0364	0.0214	1
PUBLIC	0.0001	0.0193	0.01	0.0384	0.0569	0.7364	0.0004	0.0223	0.1091	0.0071	1
RAILWA	0	0.0065	0.0142	0.0201	0.011	0.0032	0.9139	0.0168	0.013	0.0013	1
RES	0	0.0024	0.0024	0.0009	0.0041	0.0023	0.0001	0.9634	0.023	0.0013	1
VAC	0.0004	0.0141	0.0099	0.0156	0.0513	0.0057	0.0002	0.0988	0.792	0.012	1
WATER	0.0001	0.001	0.0003	0.0014	0.0136	0.0002	0	0.0055	0.0096	0.9684	1

Table 5-6: Land use forecasts for 2005 through 2029 from Markov Chain model

	Prediction					
Land use	2005	2013	2021	2029	Change (2005-29)	Change (%)
Residential	195,934	210,008	220,686	228,227	32,293	16.5%
Commercial	20,296	22,188	23,162	23,663	3,367	16.6%
Industrial	24,503	25,570	26,150	26,060	1,557	6.4%
Public	18,820	16,106	14,064	12,437	-6,383	-33.9%
Parks	74,251	86,485	97,245	106,497	32,246	43.4%
Vacant	200,837	167,865	141,224	120,643	-80,194	-39.9%
Highway	16,635	19,984	23,006	25,867	9,232	55.5%
Railway	1,505	1,578	1,642	1,702	197	13.1%
Airport	4,047	4,139	4,212	4,277	230	5.7%
Water	54,160	57,065	59,597	61,615	7,455	13.8%
Total	610,988	610,988	610,988	610,988		

Table 5-6 shows the land use distribution in each forecast year, along with the absolute and percentage changes through each time step. The land use forecasts for each period appear to be sensitive to abrupt, discontinuous changes that occur during the 1997 to 2005 period and are reflected in the transition matrix. The most notable effect is the prediction of a major decline in airport land. While there appears to have been a small decline from 1997 to 2005, this trend is projected out in each of the forecast periods, leading to a predicted decline of 44 percent from 2005 to 2029. This is unlikely in a growing metropolitan area that anticipates continued growth in air travel in the coming decades. The same can be said of the trend in land used for highways, which is projected by the model to grow by roughly 46 percent. It would be useful to attempt to decompose this predicted growth by class of highway. Interstate and state trunk highway networks are already in place and are not likely to experience sharp increases in the near future, yet county highway networks, which tend to be more robust, may see substantial growth in newly-developing parts of the region. The model also predicts a major increase in residential land use, mostly at the expense of vacant land. This largely reflects the effects of the real estate boom of the late 1990s and early 2000s in the Twin Cities. Due to this reliance on past trends, the model will probably overpredict the demand for residential land use in the 2005 to 2013 period. Once new data become available, this observation can be tested.

5.2 Logistic Regression Model

While the Markov chain model provides a simple, intuitive, but effective approach to land use transition, we would like to know a bit more about some of the factors that influence land use change. If some of these factors can be readily identified, then it will be possible to provide predictions of future land use states given some knowledge about past and present states. That is the purpose of the logistic regression modeling method.

As was discussed in section 3.2, the specification for the land use change regressions attempts to

Table 5-7: Predicted land use change, 1958–1968 from logistic regression model

Land Use	Actual 1968	Predicted 1968	Difference	Percent
Commercial	7,995	7,560	-435	-5.4%
Industrial	12,436	12,292	-144	-1.2%
Other	110,842	107,312	-3,530	-3.2%
Residential	97,326	97,094	-232	-0.2%
Vacant	378,337	382,678	4,341	1.1%

Table 5-8: Accuracy of logistic regression land use model predictions, 1968–2000

Prediction Year	Percent Correct
1968	95.4%
1978	80.8%
1990	78.8%
2000	83.3%

retain some of the information from the Markov chain model by including the previous land use in a cell as a predictor of current land use. The model expands on the basic Markov chain structure by introducing neighboring land use as an additional influence. Not only does this provide additional explanatory power, but it also helps to reduce some of the random spatial scattering of land use cells observed in the Markov chain model.

From the available land use data, logistic regression models were estimated for each period of (roughly) 10 years from 1958 to 2000. These estimated models were used to validate the basic structure of the land use change model, permitting further exploration of future land use change by forecasting cell-level land use with the parameters from the estimated 1990-2000 model.

The first model is fit to data for 1958 and used to predict land use in 1968. Table 5-7 indicates that the predictions for this period are quite good. For each of the five land use types identified (residential, commercial, industrial, vacant and other), the predictions are within about five percent of the actual cell counts. For four of the land use types, the predictions are slightly below the actual counts, while only vacant land was overpredicted. Table 5-8 indicates that across land use categories, the model's predictions were strikingly accurate, with over 95 percent of land use cells predicted correctly. Maps of the predicted land use cells, along with the actual land use in 1968 are provided in Figures 5-1 and 5-2.

Table 5-9: Estimated regression model of land use change, 1958–1968

Variable	Coefficient	S.E.	t	
Residential	Residential$_{t-1}$	1.98	0.04	50.49
	Commercial$_{t-1}$	2.00	0.13	15.75
	Industrial$_{t-1}$	2.55	0.11	24.18
	Vacant$_{t-1}$	2.10	0.04	54.05
	Residential$_{neighbor,t-1}$	1.73	0.01	162.13
	Commercial$_{neighbor,t-1}$	0.81	0.02	35.60
	Industrial$_{neighbor,t-1}$	0.76	0.02	34.30
	Vacant$_{neighbor,t-1}$	0.76	0.01	93.09
	Road$_{t-1}$	-2.16	0.07	-29.57
	Road$_{neighbor,t-1}$	0.17	0.02	10.94
	Constant	-8.47	0.06	-151.87
Commercial	Residential$_{t-1}$	2.66	0.13	19.76
	Commercial$_{t-1}$	2.20	0.16	13.51
	Industrial$_{t-1}$	3.07	0.19	16.03
	Vacant$_{t-1}$	2.77	0.13	20.96
	Residential$_{neighbor,t-1}$	0.88	0.02	39.61
	Commercial$_{neighbor,t-1}$	2.37	0.03	76.80
	Industrial$_{neighbor,t-1}$	0.88	0.04	22.75
	Vacant$_{neighbor,t-1}$	0.71	0.02	32.88
	Road$_{t-1}$	-1.91	0.14	-13.64
	Road$_{neighbor,t-1}$	0.25	0.03	8.22
	Constant	-10.91	0.17	-63.59
Industrial	Residential$_{t-1}$	3.05	0.12	26.05
	Commercial$_{t-1}$	3.11	0.20	15.53
	Industrial$_{t-1}$	3.68	0.12	30.62
	Vacant$_{t-1}$	3.14	0.10	31.49
	Residential$_{neighbor,t-1}$	0.74	0.02	34.12
	Commercial$_{neighbor,t-1}$	0.81	0.04	20.12
	Industrial$_{neighbor,t-1}$	2.23	0.03	85.41
	Vacant$_{neighbor,t-1}$	0.75	0.02	42.87
	Road $_{t-1}$	-1.90	0.13	-14.16
	Road$_{neighbor,t-1}$	0.26	0.03	9.02
	Constant	-10.88	0.14	-75.89
Vacant	Residential$_{t-1}$	2.85	0.05	56.08
	Commercial$_{t-1}$	2.65	0.16	16.65
	Industrial$_{t-1}$	3.88	0.10	40.15
	Vacant$_{t-1}$	4.19	0.03	137.92
	Residential$_{neighbor,t-1}$	0.77	0.01	90.51
	Commercial$_{neighbor,t-1}$	0.70	0.02	29.46
	Industrial$_{neighbor,t-1}$	0.83	0.02	44.36

Continued on next page

Variable	Coefficient	S.E.	t	
	$\text{Vacant}_{neighbor,t-1}$	1.37	0.01	219.46
	Road_{t-1}	-1.92	0.06	-32.87
	$\text{Road}_{neighbor,t-1}$	0.18	0.01	13.74
	Constant	-8.98	0.05	-198.71

N = 606,936; Likelihood ratio = 1,096,275.39; Log likelihood = -80,250.94; ρ^2 = 0.87

Table 5-9 provides the parameter estimates and model fit. The 'other' land use class is treated as the reference category in the model and sets of alternative-specific parameters are estimated for each of the remaining four land use classes. The model fit summary includes the values of the likelihood ratio, log-likelihood and ρ^2 statistics. The ρ^2 statistic is a goodness-of-fit measure analogous to the R^2 statistic in basic linear regression. Table 12 indicates that the value of this statistic is 0.87, meaning this model has fairly strong explanatory power.

Individual parameter estimates are also provided in Table 5-9 with model coefficients, standard errors and asymptotic t-statistics. Significance level indicators are omitted from the table since nearly all parameters are statistically significant at the $p < 0.01$ level. This and the rather high level of model goodness-of-fit are both at least in part a function of the extremely large sample size.

The effects of neighboring land uses on the probability of converting to residential land use in a subsequent time period are all positive relative to the 'other' category, though the effect of having residential as an initial state is more than twice as strong as any other state, indicating a resistance to change. Having roadway-related land uses in at least part of a cell has a markedly negative effect on the probability of transition to residential, though having roads in a neighboring cell appears to have the opposite effect.

Commercial and industrial land uses have somewhat similar characteristics in that each tends to be associated with the presence of roads in neighboring cells, though not in the immediate one. This is indicates that roadway access is an important component for both land uses. Also, commercial and industrial land uses tend to stay in areas where they are surrounded by similar land uses. Finally, vacant land tends to remain vacant when surrounded by like cells. Interestingly though, having industrial land in a cell at the previous time period also exerts a fairly strong influence on the probability of vacancy in the next (1968). Perhaps this was evidence of the intra-urban migration of manufacturing firms during the 1960s.

The second model estimated land uses in 1978 given observations on the 1968 variables. While not as accurate as the previous model, this model was able to accurately predict land use for over 80 percent of cells, as is indicated in Table 5-8. The resulting ρ^2 statistic was 0.44, lower than the 1958-68 model, but still notable for a cross-sectional data set.

Tables 5-10 and 5-11 provide the parameter estimates and the predicted and actual land use in 1978 by land use class. The predictions for 1978 again appear to underestimate each class of land use except for vacant and agricultural land. In particular, commercial and industrial land uses are sharply underpredicted, by 55 and 37 percent, respectively. The maps in Figures 5-3 and 5-4 show why this is so. The model seems to produce contiguous swaths of residential land, which tend to dominate local clusters of commercial and industrial activity. Commercial land uses along arterial streets and highways are masked by the expansion of adjacent residential neighborhoods. Major industrial areas in the central cities are also underpredicted, along with isolated industrial parks

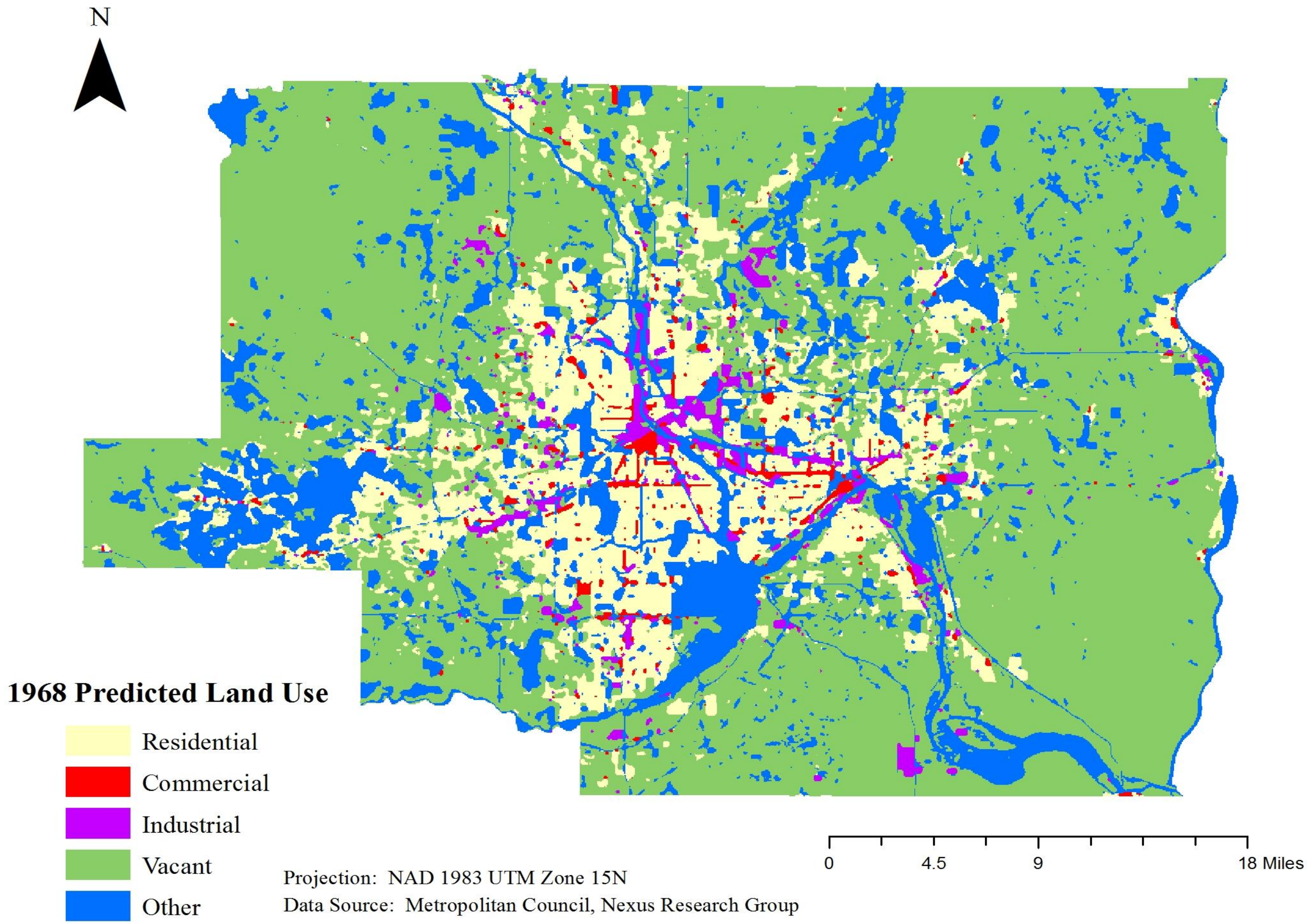


Figure 5-1: Predicted land use in 1968 (logistic regression)

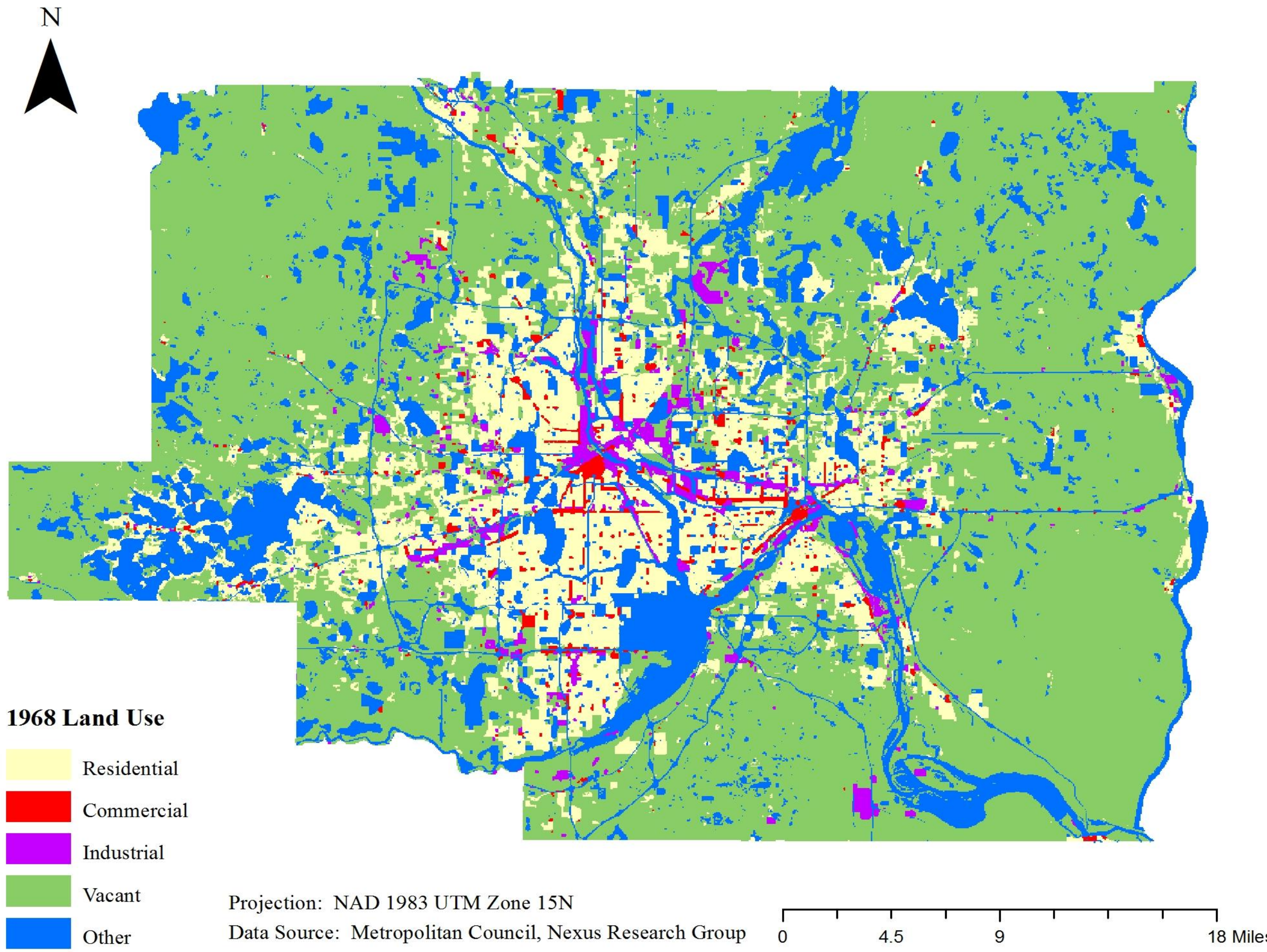


Figure 5-2: Actual land use in 1968

in suburban locations. The latter show up primarily as vacant or agricultural land. Residential land, while underpredicted as a whole, covers a disproportionate share of the central portions of the region, while failing to reflect the full extent of the urbanized fringe.

Table 5-10: Estimated logistic regression model of land use change, 1968–1978

Dependent	Independent	Coefficient	S.E.	t
Residential	Residential_{t-1}	2.03	0.03	67.95
	Commercial_{t-1}	2.27	0.07	32.09
	Industrial_{t-1}	2.45	0.07	32.91
	Vacant_{t-1}	1.96	0.03	66.46
	$\text{Residential}_{neighbor,t-1}$	0.47	0.00	98.46
	$\text{Commercial}_{neighbor,t-1}$	0.27	0.01	21.07
	$\text{Industrial}_{neighbor,t-1}$	0.07	0.01	5.69
	$\text{Vacant}_{neighbor,t-1}$	0.19	0.00	42.73
	Road_{t-1}	-1.13	0.04	-26.32
	$\text{Road}_{neighbor,t-1}$	-0.03	0.01	-3.63
	Constant	-3.44	0.02	-190.54
Commercial	Residential_{t-1}	2.31	0.07	34.28
	Commercial_{t-1}	2.19	0.09	24.56
	Industrial_{t-1}	2.70	0.10	26.22
	Vacant_{t-1}	2.46	0.07	35.80
	$\text{Residential}_{neighbor,t-1}$	0.30	0.01	30.32
	$\text{Commercial}_{neighbor,t-1}$	0.81	0.01	55.00
	$\text{Industrial}_{neighbor,t-1}$	0.39	0.02	23.51
	$\text{Vacant}_{neighbor,t-1}$	0.09	0.01	8.73
	Road_{t-1}	-1.02	0.07	-15.23
	$\text{Road}_{neighbor,t-1}$	0.19	0.01	14.80
	Constant	-5.45	0.05	-117.06
Industrial	Residential_{t-1}	2.72	0.07	37.22
	Commercial_{t-1}	2.77	0.11	24.59
	Industrial_{t-1}	2.93	0.08	36.70
	Vacant_{t-1}	2.80	0.06	44.45
	$\text{Residential}_{neighbor,t-1}$	0.07	0.01	6.27
	$\text{Commercial}_{neighbor,t-1}$	0.37	0.02	19.29
	$\text{Industrial}_{neighbor,t-1}$	0.75	0.01	59.08
	$\text{Vacant}_{neighbor,t-1}$	0.14	0.01	16.14
	Road_{t-1}	-1.17	0.07	-16.46
	$\text{Road}_{neighbor,t-1}$	0.15	0.01	11.84
	Constant	-5.35	0.04	-119.18
Vacant	Residential_{t-1}	1.81	0.03	56.08
	Commercial_{t-1}	2.00	0.08	24.21
	Industrial_{t-1}	2.37	0.07	36.04

Continued on next page

Dependent	Independent	Coefficient	S.E.	t
	Vacant_{t-1}	2.86	0.02	136.15
	$\text{Residential}_{neighbor,t-1}$	-0.03	0.01	-6.06
	$\text{Commercial}_{neighbor,t-1}$	-0.02	0.02	-1.43
	$\text{Industrial}_{neighbor,t-1}$	0.08	0.01	6.84
	$\text{Vacant}_{neighbor,t-1}$	0.30	0.00	90.67
	Road_{t-1}	-0.98	0.03	-28.76
	$\text{Road}_{neighbor,t-1}$	0.01	0.01	1.45
	Constant	-2.51	0.01	-199.33

N = 606,936; Likelihood ratio = 612,286.63; Log likelihood = -390,214.38; ρ^2 = 0.44

Table 5-11: Predicted land use change, 1968–1978 (logistic regression)

Land Use	Actual 1978	Predicted 1978	Difference	Percent
Commercial	12,718	5,706	-7,012	-55.1%
Industrial	16,917	10,578	-6,339	-37.5%
Other	133,059	112,136	-20,923	-15.7%
Residential	107,716	99,691	-8,025	-7.5%
Vacant	336,526	378,825	42,299	12.6%

The model parameter estimates, listed in Table 5-10, look similar to those estimated for the 1958-68 period. For each type of land use, the pre-existing, neighboring land uses continue to exert the strongest influence on land use change. The presence of roadway-related land in a cell has an almost uniformly negative influence on the probability of conversion to residential, commercial, industrial or vacant land relative to land uses in the 'other' category, while roads in neighboring cells have a modestly positive influence.

The 1978 to 1990 model produces results similar to the one estimated from the 1968-78 data. Just under 79 percent of cells in the data set are predicted correctly, with a ρ^2 value of 0.48 for the model (Table 5-12), but again the commercial and industrial land uses are severely underpredicted. Table 5-13 shows that over this period, not only are the commercial and industrial land uses underpredicted by more than 30 percent, but that residential cells are underpredicted by 22 percent. Again, the model predicts too little conversion to urban land uses (residential, commercial, industrial and others) and too much land remaining vacant or in agriculture.

Table 5-12: Estimated regression model of land use change, 1978–1990

Dependent	Independent	Coefficient	S.E.	t
Residential	Residential_{t-1}	1.26	0.02	54.25
	Commercial_{t-1}	1.30	0.06	23.38
	Industrial_{t-1}	1.62	0.06	25.50
	Vacant_{t-1}	1.04	0.02	46.67

Continued on next page

Dependent	Independent	Coefficient	S.E.	t
	$\text{Residential}_{neighbor,t-1}$	0.71	0.00	163.28
	$\text{Commercial}_{neighbor,t-1}$	0.38	0.01	39.04
	$\text{Industrial}_{neighbor,t-1}$	0.21	0.01	18.72
	$\text{Vacant}_{neighbor,t-1}$	0.36	0.00	102.43
	Road_{t-1}	-1.24	0.04	-32.65
	$\text{Road}_{neighbor,t-1}$	0.02	0.01	2.65
	Constant	-3.15	0.01	-229.98
Commercial	Land use (t-1)			
	Residential_{t-1}	1.74	0.06	29.95
	Commercial_{t-1}	2.01	0.07	29.03
	Industrial_{t-1}	2.16	0.09	24.81
	Vacant_{t-1}	1.75	0.05	33.49
	$\text{Residential}_{neighbor,t-1}$	0.32	0.01	33.08
	$\text{Commercial}_{neighbor,t-1}$	0.80	0.01	67.84
	$\text{Industrial}_{neighbor,t-1}$	0.52	0.01	35.55
	$\text{Vacant}_{neighbor,t-1}$	0.25	0.01	31.90
	Road dummy (t-1)	-0.92	0.06	-16.00
	$\text{Road}_{neighbor,t-1}$	0.36	0.01	31.65
	Constant	-5.44	0.04	-140.70
Industrial	Residential_{t-1}	1.59	0.06	26.27
	Commercial_{t-1}	1.67	0.09	18.55
	Industrial_{t-1}	1.84	0.06	29.22
	Vacant_{t-1}	1.43	0.04	32.06
	$\text{Residential}_{neighbor,t-1}$	0.10	0.01	9.16
	$\text{Commercial}_{neighbor,t-1}$	0.38	0.02	24.53
	$\text{Industrial}_{neighbor,t-1}$	0.92	0.01	85.02
	$\text{Vacant}_{neighbor,t-1}$	0.28	0.01	41.50
	Road dummy (t-1)	-1.02	0.06	-16.45
	$\text{Road}_{neighbor,t-1}$	0.19	0.01	15.74
	Constant	-4.74	0.03	-157.02
Vacant	Residential_{t-1}	0.62	0.03	19.31
	Commercial_{t-1}	0.88	0.07	11.85
	Industrial_{t-1}	1.31	0.06	20.34
	Vacant	1.93	0.02	92.91
	$\text{Residential}_{neighbor,t-1}$	0.32	0.01	59.34
	$\text{Commercial}_{neighbor,t-1}$	0.22	0.01	17.50
	$\text{Industrial}_{neighbor,t-1}$	0.34	0.01	31.47
	$\text{Vacant}_{neighbor,t-1}$	0.59	0.00	170.02
	Road dummy (t-1)	-1.00	0.03	-28.62
	$\text{Road}_{neighbor,t-1}$	0.16	0.01	21.29
	Constant	-3.81	0.02	-225.80

N = 606,936; Likelihood ratio = 727,785.39; Log likelihood= -402,268.62; ρ^2 = 0.48

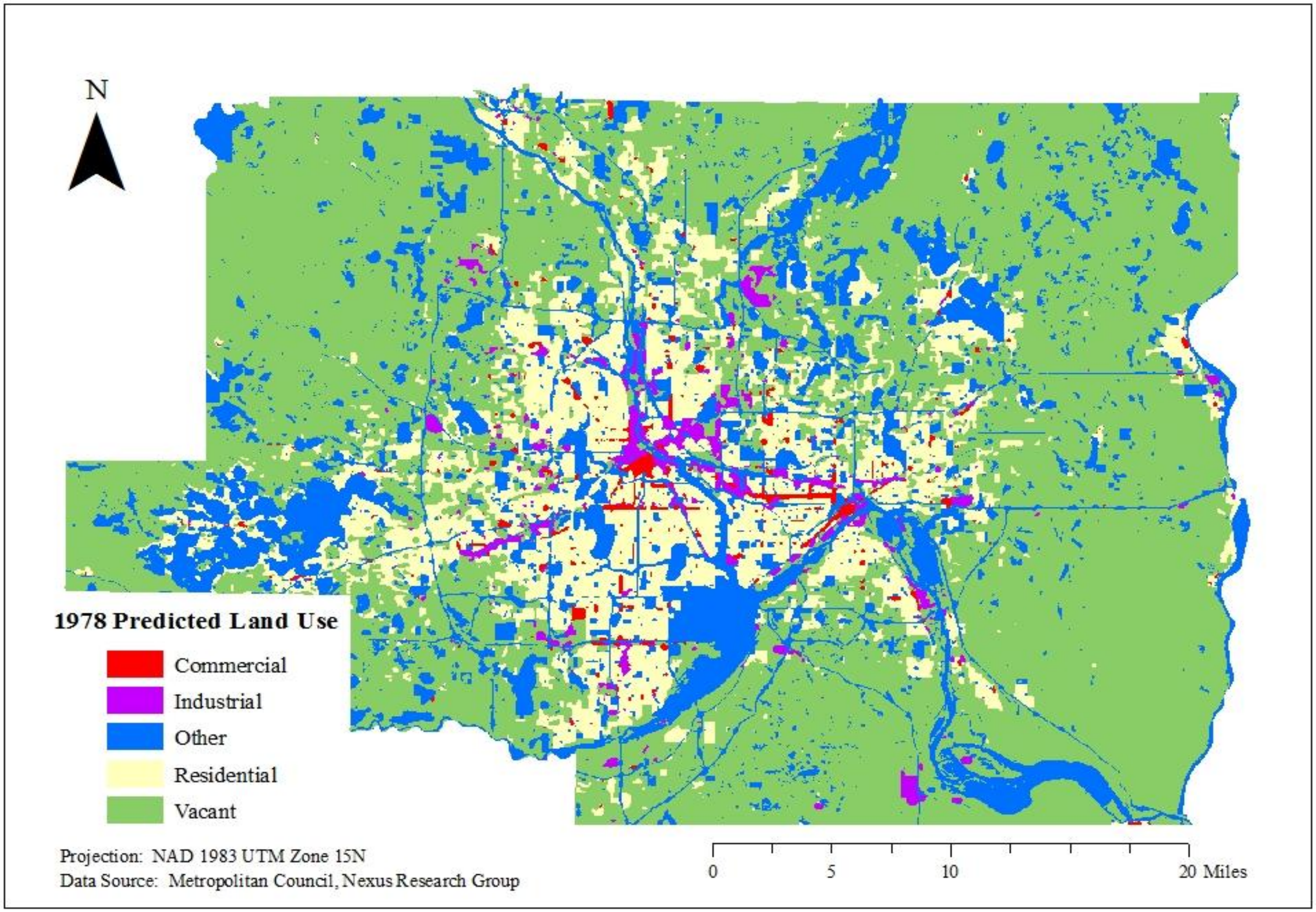


Figure 5-3: Predicted land use in 1978 (logistic regression)

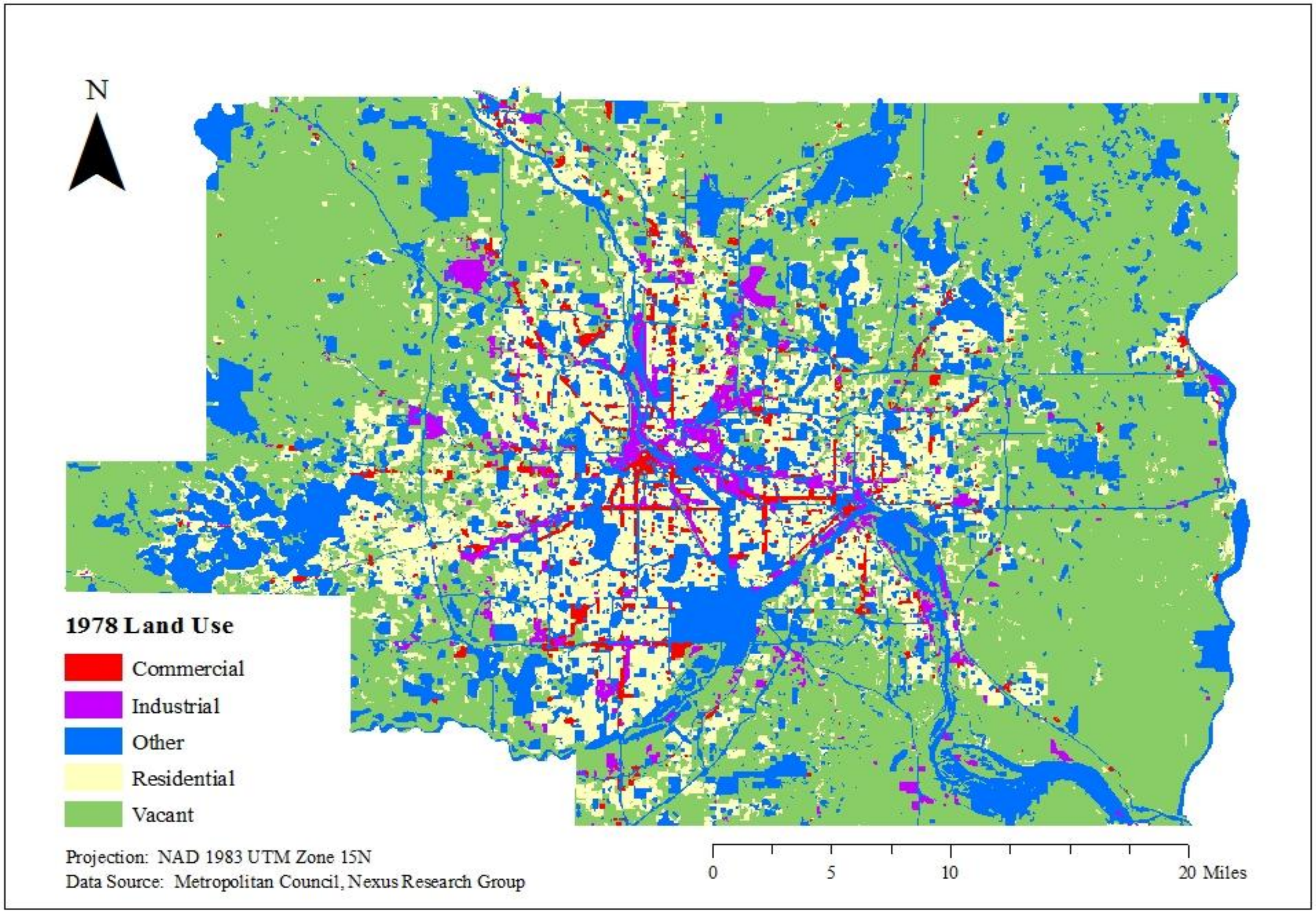


Figure 5-4: Actual land use in 1978

Table 5-13: Predicted land use change, 1978–1990 (logistic regression)

Land Use	Actual 1978	Predicted 1978	Difference	Percent
Commercial	14,132	7,714	-6,418	-45.4%
Industrial	22,939	15,907	-7,032	-30.7%
Other	142,282	132,214	-10,068	-7.1%
Residential	158,132	123,416	-34,716	-22.0%
Vacant	269,451	327,685	58,234	21.6%

Figures 5-5 and 5-6 show the spatial distribution of these predictions. The patterns strongly resemble those predicted for the previous decade. The predicted 1990 land use pattern again deviates from the actual pattern of land use in that the region is portrayed as more densely and contiguously developed. Land devoted to residential uses is predicted to be more centralized and less integrated with other land uses. Clusters of commercial and industrial land use are underestimated in terms of their spatial extent, not only in more central locations, but also especially along suburban highway corridors, such as Interstate 494.

The last regression model to be estimated for the region with available data covers the period from 1990 to 2000. In addition to the variables entered in the first three models, a separate variable is introduced to test for effects of access to regional employment on land use transition. Our hypothesis, stated previously, held that increases in accessibility to employment would increase the probability of transition to residential land use during the subsequent time period.

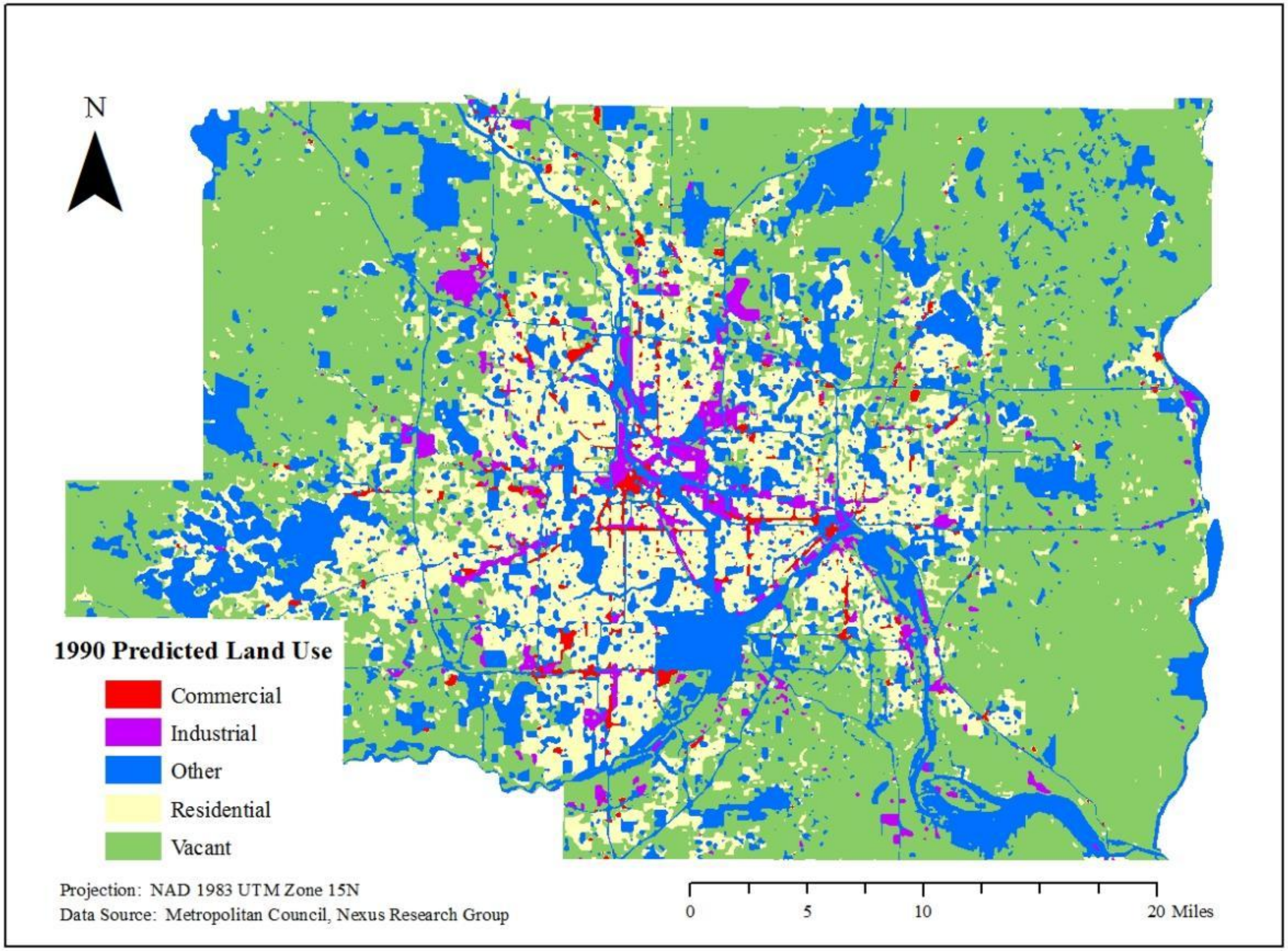


Figure 5-5: Predicted land use in 1990 (logistic regression)

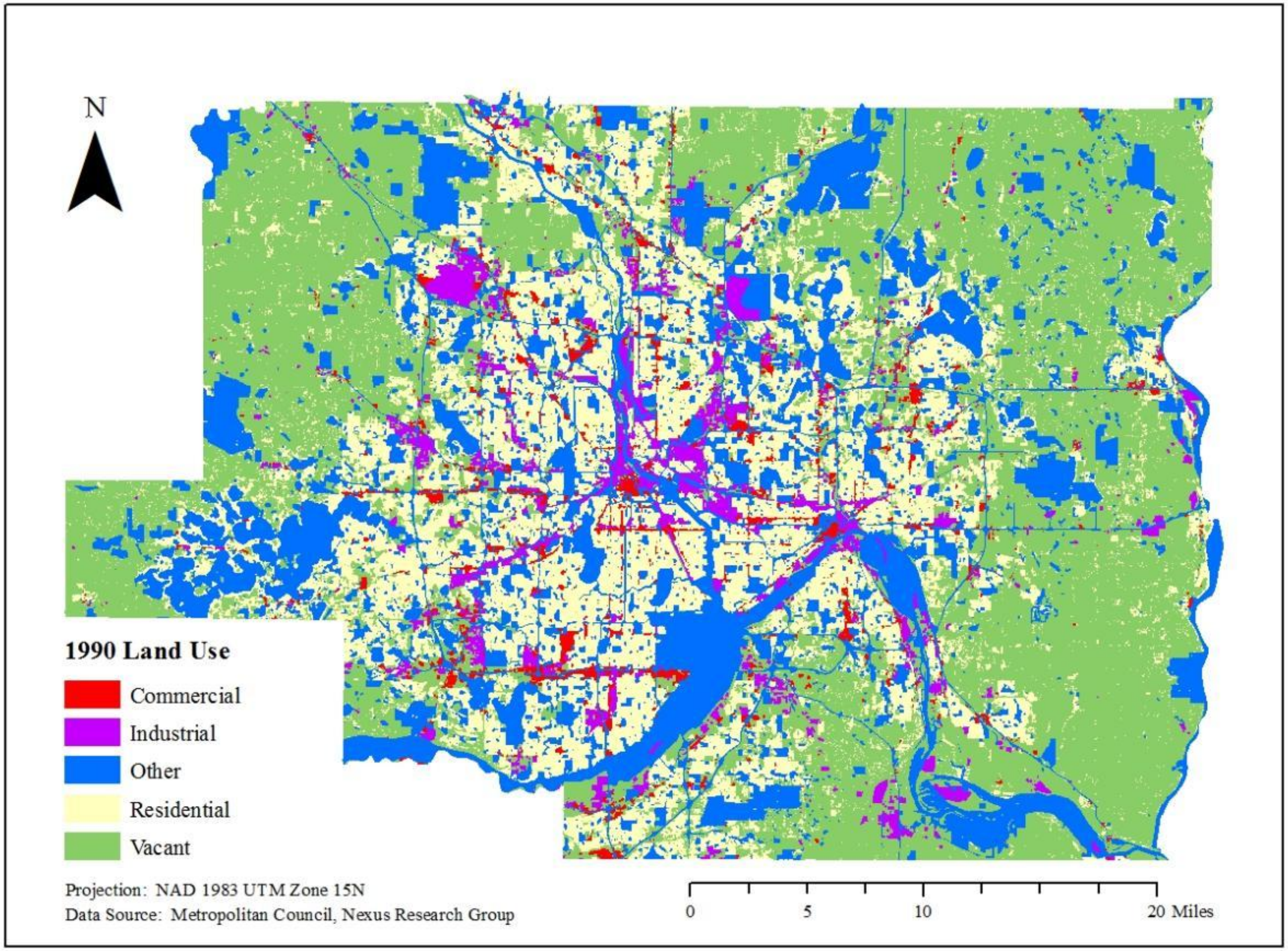


Figure 5-6: Actual land use in 1990

Table 5-14: Estimated regression model of land use change, 1990–2000

Variable	Coefficient	S.E.	t	
Residential	Residential$_{t-1}$	4.64	0.03	169.55
	Commercial$_{t-1}$	2.57	0.06	42.70
	Industrial$_{t-1}$	2.53	0.07	35.20
	Vacant$_{t-1}$	2.47	0.03	84.49
	Residential$_{neighbor,t-1}$	0.59	0.00	118.40
	Commercial$_{neighbor,t-1}$	0.18	0.01	14.95
	Industrial$_{neighbor,t-1}$	0.11	0.01	8.64
	Vacant$_{neighbor,t-1}$	0.30	0.00	64.28
	Road$_{t-1}$	-0.98	0.05	-20.61
	Road neighbor $_{neighbor,t-1}$	-0.11	0.01	-11.32
	Accessibility $(10^4)_{t-1}$	0.01	0.00	2.80
	Constant	-4.66	0.03	-173.57
Commercial	Residential$_{t-1}$	2.88	0.06	48.04
	Commercial$_{t-1}$	4.55	0.06	74.27
	Industrial$_{t-1}$	3.64	0.08	47.95
	Vacant	2.46	0.06	43.63
	Residential$_{neighbor,t-1}$	0.22	0.01	23.33
	Commercial$_{neighbor,t-1}$	0.75	0.01	63.19
	Industrial$_{neighbor,t-1}$	0.44	0.01	34.73
	Vacant$_{neighbor,t-1}$	0.21	0.01	24.70
	Road$_{t-1}$	-1.14	0.06	-18.51
	Road$_{neighbor,t-1}$	0.22	0.01	18.32
	Accessibility $(10^4)_{t-1}$	1.22	0.03	32.82
	Constant	-6.43	0.05	-128.04
Industrial	Residential$_{t-1}$	1.95	0.08	25.69
	Commercial$_{t-1}$	3.28	0.07	44.81
	Industrial$_{t-1}$	4.12	0.06	65.11
	Vacant	2.21	0.05	40.88
	Residential$_{neighbor,t-1}$	0.01	0.01	0.98
	Commercial$_{neighbor,t-1}$	0.47	0.01	34.23
	Industrial$_{neighbor,t-1}$	0.65	0.01	61.55
	Vacant$_{neighbor,t-1}$	0.22	0.01	27.02
	Road$_{t-1}$	-1.14	0.07	-17.11
	Road$_{neighbor,t-1}$	0.10	0.01	8.11
	Accessibility $(10^4)_{t-1}$	0.11	0.00	14.36
	Constant	-5.51	0.04	-130.81
Vacant	Residential$_{t-1}$	2.11	0.03	65.17
	Commercial$_{t-1}$	1.65	0.07	24.00
	Industrial$_{t-1}$	2.20	0.06	38.72
	Vacant	2.80	0.02	112.15

Continued on next page

Variable	Coefficient	S.E.	t	
	Residential$_{neighbor,t-1}$	0.20	0.01	40.53
	Commercial$_{neighbor,t-1}$	0.19	0.01	15.28
	Industrial$_{neighbor,t-1}$	0.49	0.01	52.57
	Vacant$_{neighbor,t-1}$	0.35	0.00	87.96
	Road dummy (t-1)	-0.92	0.04	-23.85
	Road$_{neighbor,t-1}$	0.01	0.01	1.20
	Accessibility $(10^4)_{t-1}$	-0.58	0.00	-103.04
	Constant	-2.36	0.02	-129.51

N = 606,936; Likelihood ratio = 944,398.68; Log likelihood = -326,876.65; ρ^2 = 0.59

The parameter estimates in Table 5-14 seem to bear this out. The effect is positive and significant, though it is difficult to determine from the table entry. Due to the units of accessibility being measured on a different scale than the other variables, the associated coefficient appears comparatively small, with a value of 0.00000806. However, given the large range of values that this variable takes in the data set (from 661 to over 55,000), its practical significance is on par with the remaining variables. Within the range of values identified, the accessibility variable could add between 0.005 to 0.44 units to the linear predictor for the residential alternative. In addition to residential land, the accessibility variable increases the probability of transition to commercial and industrial land uses, relative to the 'other' category. However, with respect to vacant land, the accessibility coefficient takes a negative sign. This is to be expected, since land that is vacant (or more importantly, agricultural) is less likely to be in a highly accessible location.

The remaining variables are again found to have the expected signs and magnitudes. For each land use, a cell is most likely remain in its same state during the next time period. Also, cells are increasingly likely to transition to land uses that were the predominant neighbors in the prior time period. Having roadway-related land in a cell reduces the probability of transition to that cell relative to the reference category, though having roads in a neighboring cell appears to have the opposite effect, albeit on a smaller scale.

The overall fit of the model appears to be moderately improved by the inclusion of the accessibility variable. Table 5-14 reports a ρ^2 value of 0.59 for the model, with land use in roughly 83 percent of cells being predicted correctly. Table 5-15 reveals that the model's predictions have the same downward bias for residential, commercial, industrial and 'other' land uses, but that the effect is somewhat attenuated relative to earlier modeled periods. Vacant and agricultural land are overpredicted by 19 percent, accounting for the difference.

The spatial patterns of predicted and actual land use, shown in Figures 5-7 and 5-8, reveal a familiar set of results. Residential land is underpredicted by the model and more centrally concentrated than the actual pattern of residential land. Commercial and industrial land, while less inaccurately predicted than in previous models (both in quantity and in spatial scope), are still underrepresented along certain highway corridors and in outlying suburban areas.

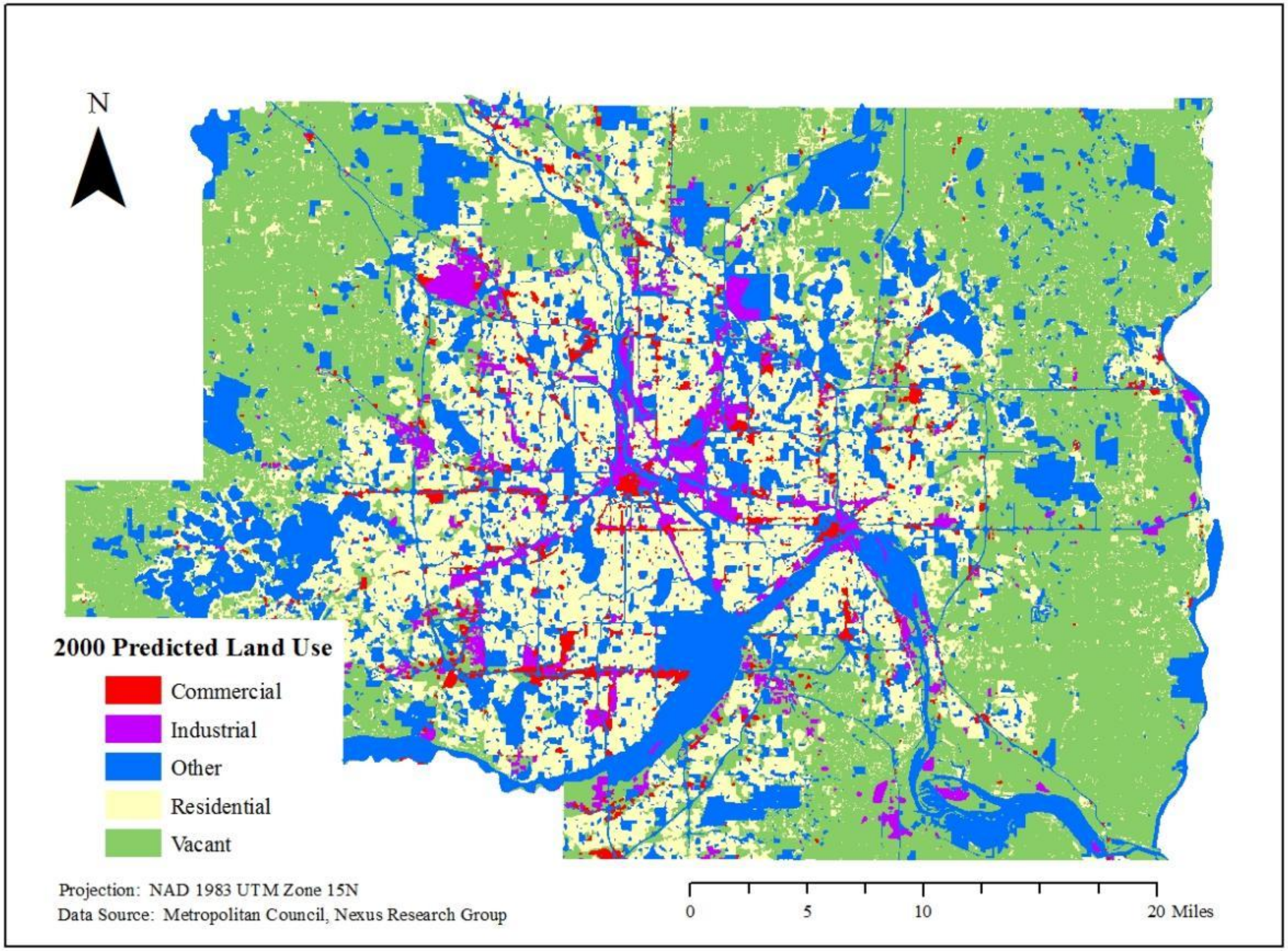


Figure 5-7: Predicted land use in 2000 (logistic regression)

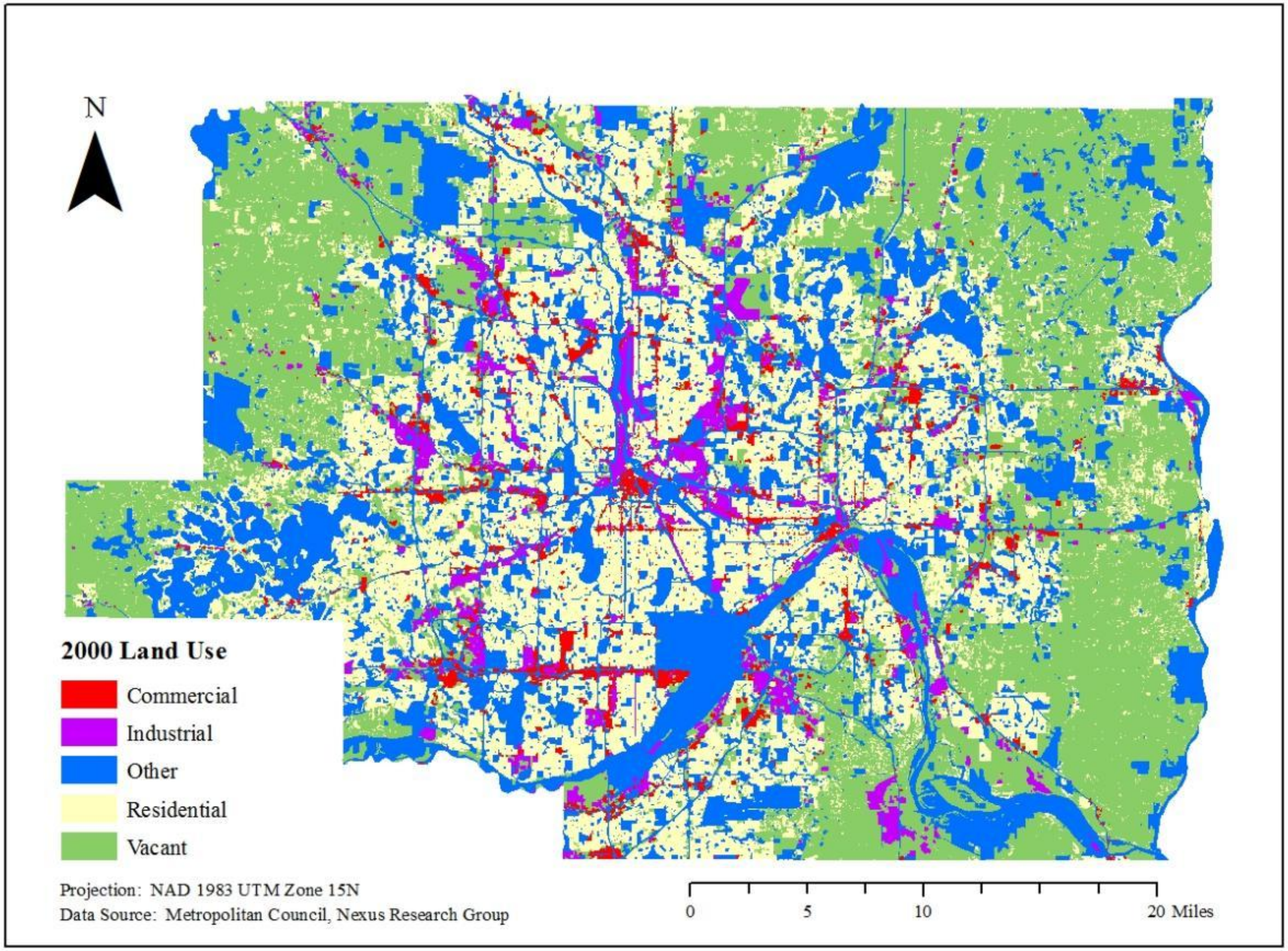


Figure 5-8: Actual land use in 2000

5.3 Markov Chain-Cellular Automata Model

The previous sections have outlined the validation and application of models of land use change based on probabilistic transitions between states (Markov chain) and on the influence of several neighbor characteristics, particularly neighboring land uses (logistic regression). This section introduces the Markov Chain-Cellular Automata (MCCA) model, which combines elements of both approaches.

As identified in Section 3.3, the MCCA model builds on the basic structure of the Markov chain model, with a transition matrix of land use change probabilities serving as the focal point. In our model, the two most common neighboring land uses surrounding a cell, along with its current land use, define a state. The addition of the neighboring land uses helps to define more homogeneous neighborhoods of certain land use types, particularly residential use, which are more resistant to change. The outcome of this process is demonstrated through application of the MCCA model to backcast previous land use change and to predict future land use to the year 2030.

The first historical forecast (backcast) for which we have results is the prediction of 1978 land use using a transition matrix estimated from 1958 and 1968 data. Table 5-16 presents the totals of predicted and actual land use by land use type. Many of the land use types are predicted within 10 percent of their actual values, with the exceptions being highways, parks, and public land. Highways and parks are overpredicted by more than 30 percent, while public lands exceeded predicted levels by more than 40 percent. Residential land use is predicted at slightly higher levels than what was actually observed, and this was largely offset by an underprediction of vacant and agricultural land. Table 5-17, which evaluates the overall accuracy of the model predictions in terms of the percentage of cells correctly predicted, indicates that just under 79 percent of cells were accurately identified. Comparing the maps in Figures 5-9 and 5-10, which depict predicted and actual land use in 1978, the most noticeable differences are the underprediction of commercial land uses along major transportation corridors and of industrial land at major industrial sites, as well as the preponderance of predicted vacant and agricultural land.

Table 5-15: Predicted land use change, 1990–2000 (logistic regression)

Land Use	Actual 2000	Predicted 2000	Difference	Percent
Commercial	18,779	13,949	-4,830	-25.7%
Industrial	23,777	21,248	-2,529	-10.6%
Other	160,646	145,022	-15,624	-9.7%
Residential	186,347	168,030	-18,317	-9.8%
Vacant	217,387	258,687	41,300	19.0%

Table 5-16: Predicted land use change, 1968–1978 (MCCA Model)

Land Use	Predicted 1978	Actual 1978	Difference	Percent
Airport	4,735	4,845	-110	-2.3%
Commercial	11,915	12,718	-803	-6.3%
Highway	17,348	12,860	4,488	34.9%
Industrial	15,733	16,917	-1,184	-7.0%
Parks	55,278	41,750	13,528	32.4%
Public	10,661	18,575	-7,914	-42.6%
Railway	2,260	2,228	32	1.4%
Residential	117,419	107,716	9,703	9.0%
Vacant	318,835	336,526	-17,691	-5.3%
Water	52,752	52,801	-49	-0.1%

Table 5-17: Accuracy of MCCA land use model predictions, 1978–2000

Prediction Year	Percent Correct
1978	78.9
1990	66.4
2000	68.1

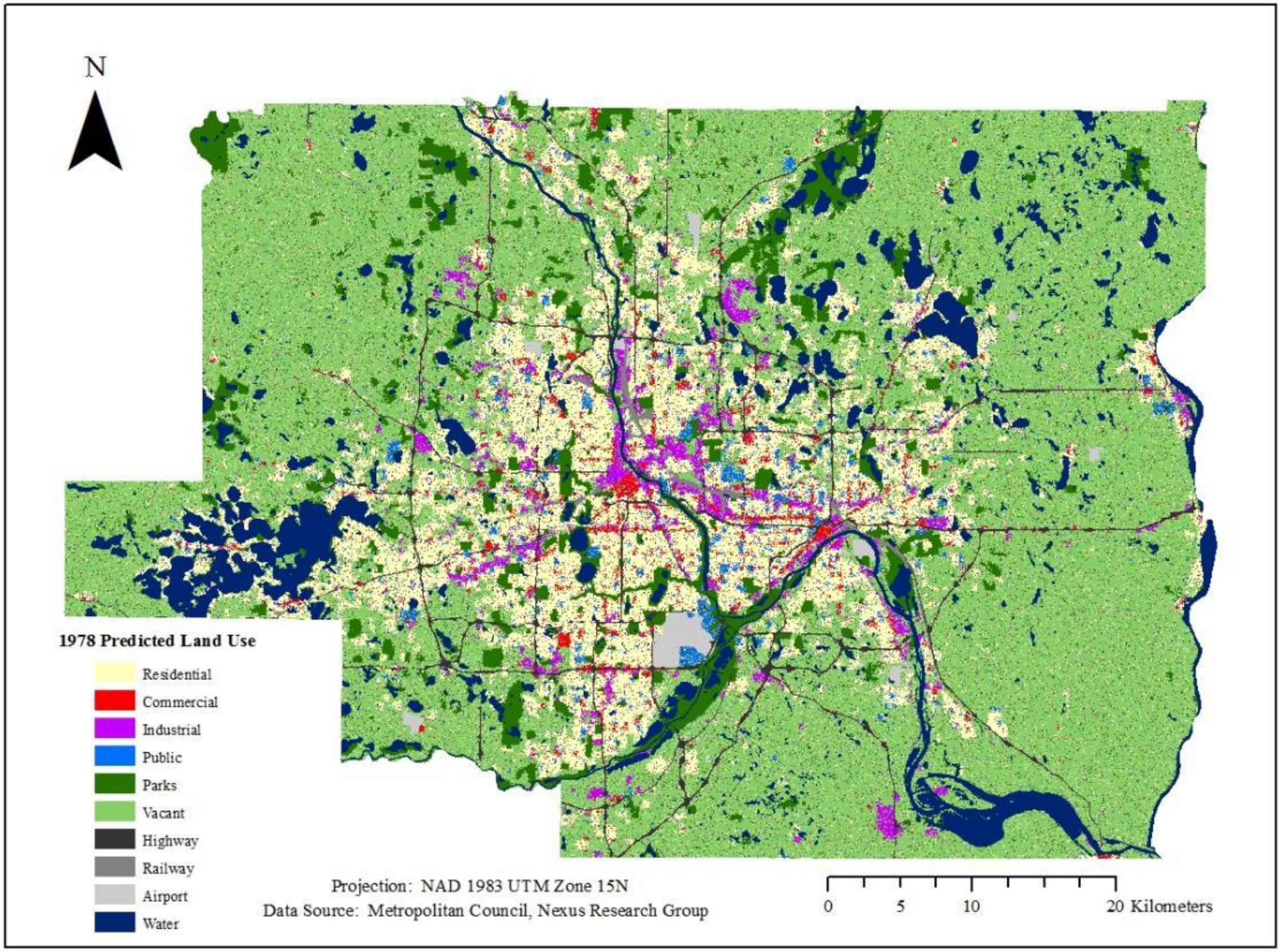


Figure 5-9: Predicted land use in 1978 (MCCA Model)

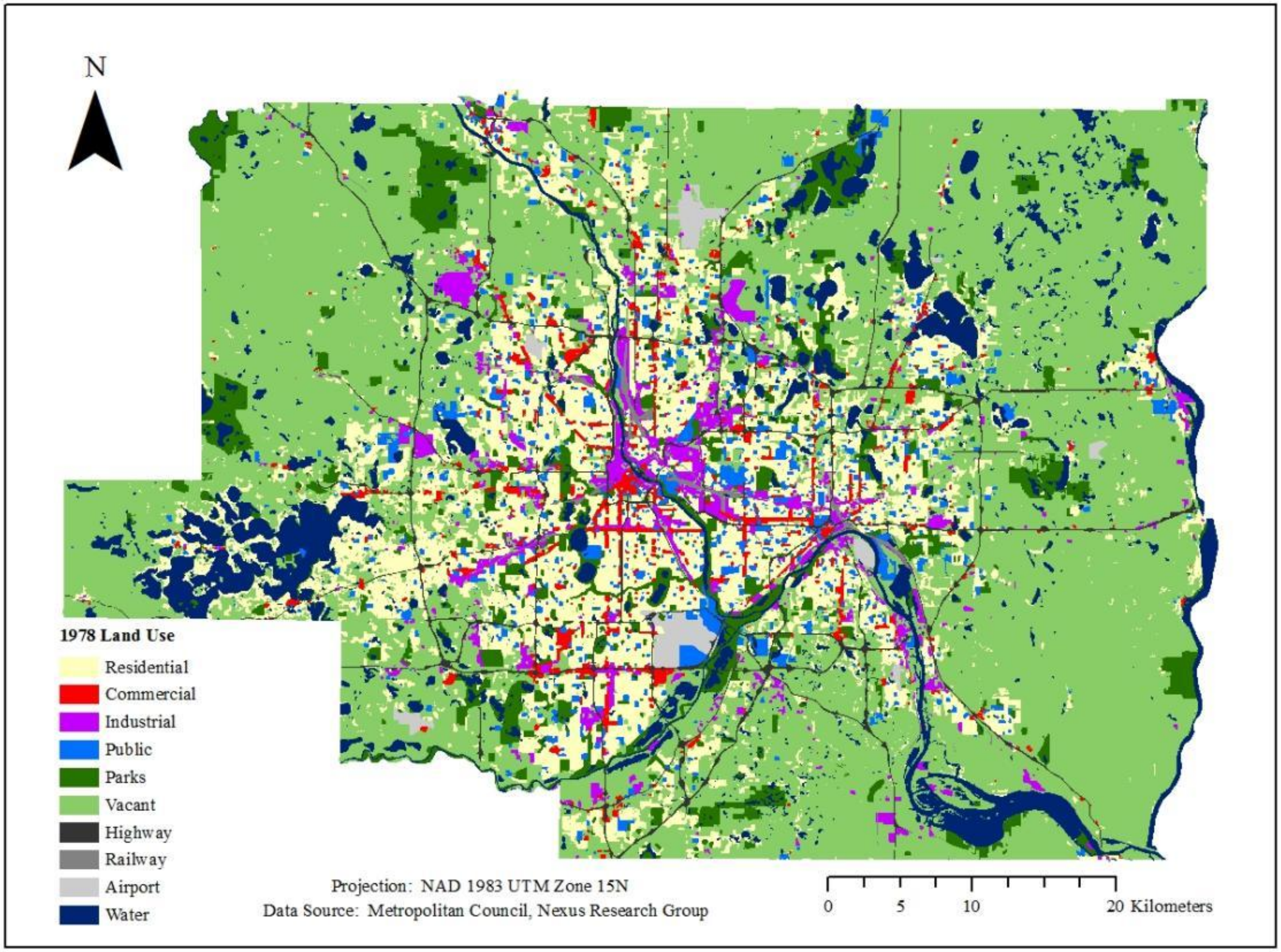


Figure 5-10: Actual land use in 1978

The next historical forecast uses the 1968 to 1978 matrix of land use change to predict land use in 1990. This represents a 12-year period rather than 10 years, however land use data were not available in 1988. The 12-year transition period is considered a reasonable approximation. Table 5-18 breaks down the totals of land use change by type. Of interest, there appears to be greater variation in land use change than in the previous period, with several land uses exceeding their forecast levels by 40 percent or more. However, several of these land uses (e.g. airport, commercial, highway) represent only a small share of total land use in the region. The major exception appears to be vacant land, which is underpredicted by more than 20 percent. The maps in Figures 5-11 and 5-12 show that not only was the spatial extent of residential land growth underpredicted, but also its organization into contiguous residential neighborhoods. Public and industrial land are predicted more consistently during this period, though the decentralization of industrial sites is not fully capture by the model. Overall, about two-thirds of land use cells in the region were accurately predicted.

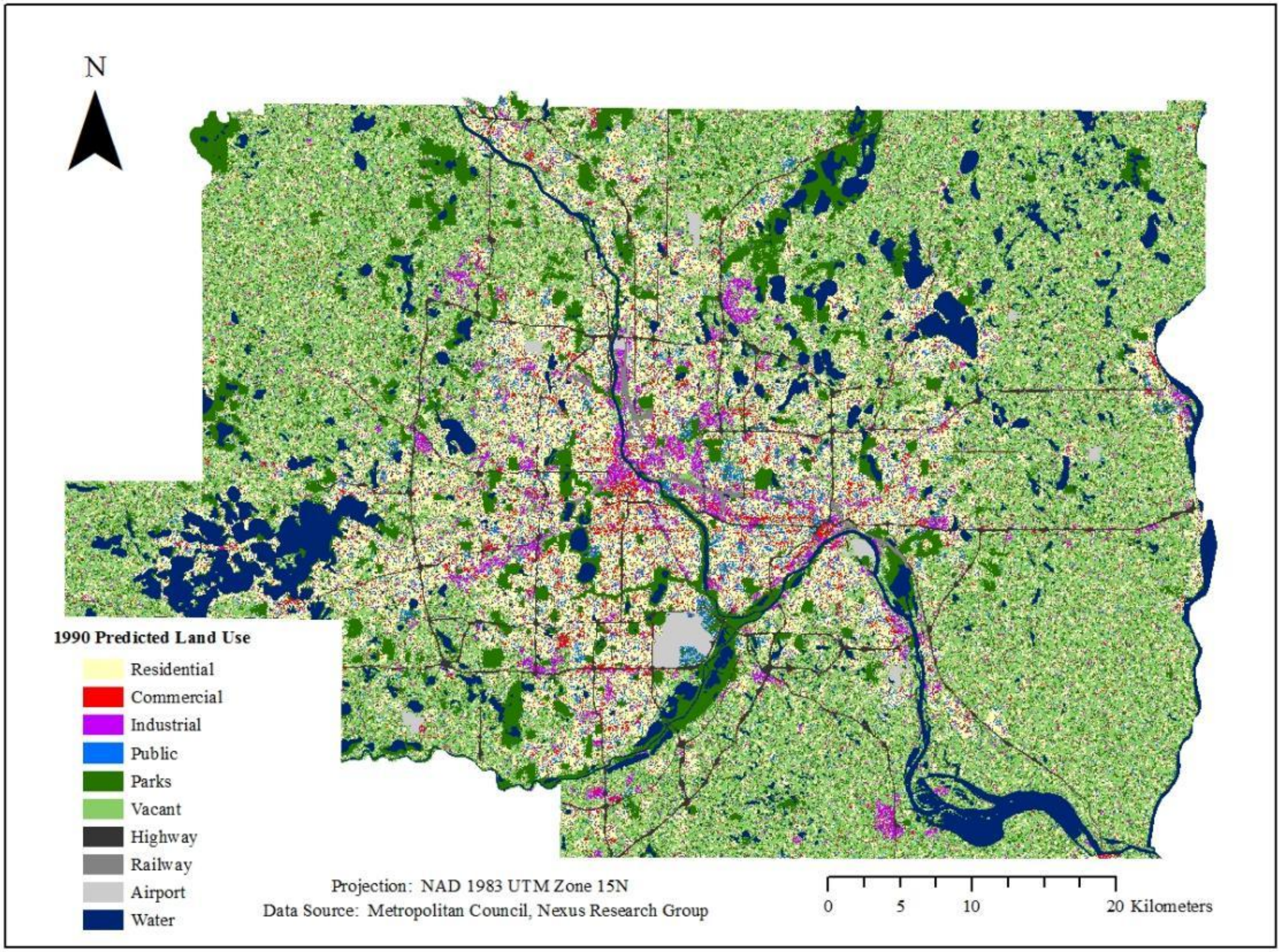


Figure 5-11: Predicted land use in 1990 (MCCA Model)

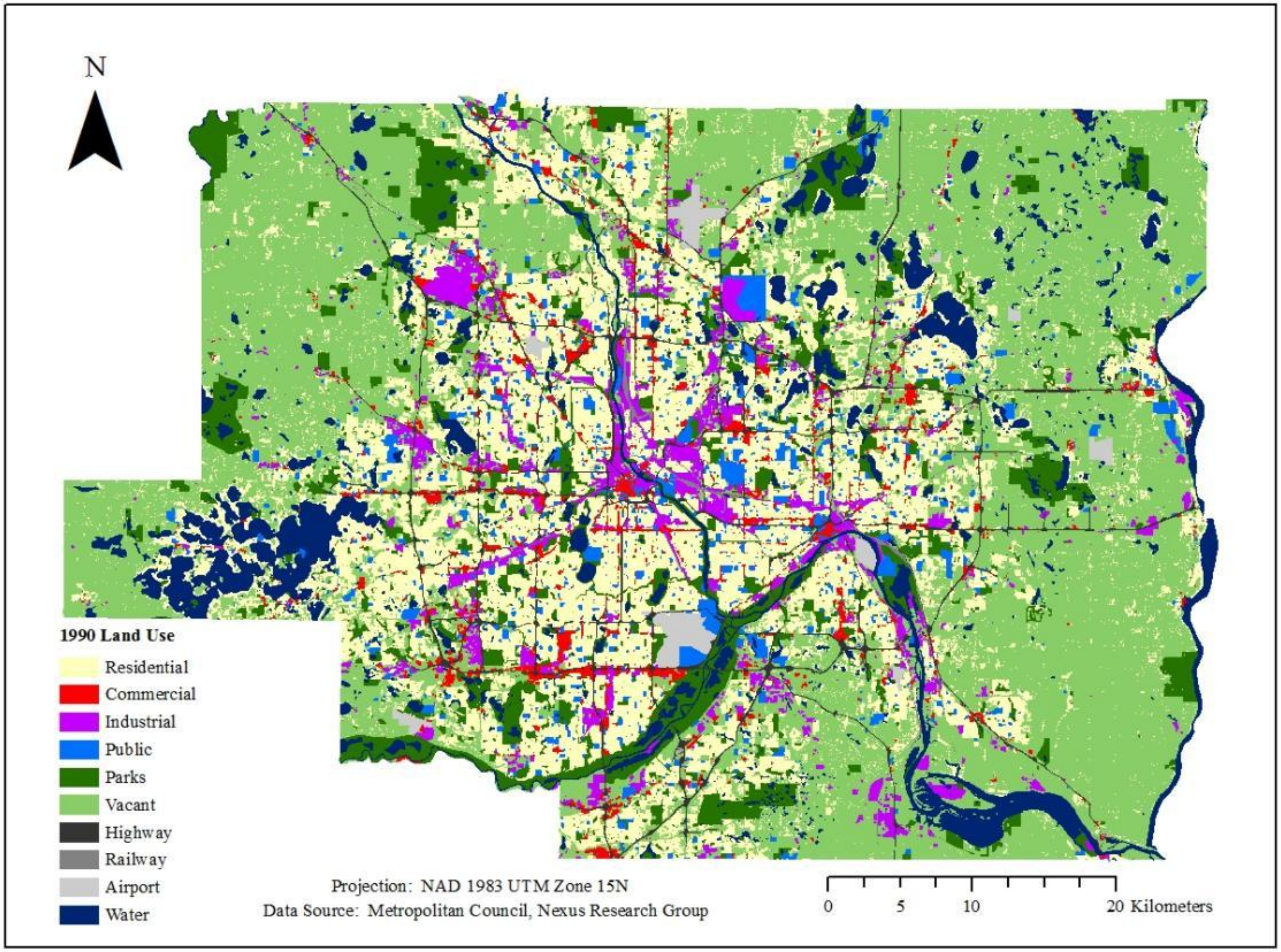


Figure 5-12: Actual land use in 1990

Table 5-18: Predicted land use change, 1978–1990 (MCCA Model)

Land Use	Predicted 1990	Actual 1990	Difference	Percent
Airport	7,287	5,068	2,219	43.8%
Commercial	19,627	14,132	5,495	38.9%
Highway	24,692	12,730	11,962	94.0%
Industrial	21,719	22,939	-1,220	-5.3%
Parks	89,886	53,760	36,126	67.2%
Public	18,754	19,679	-925	-4.7%
Railway	2,404	1,496	908	60.7%
Residential	155,591	158,132	-2,541	-1.6%
Vacant	214,224	269,451	-55,227	-20.5%
Water	52,752	49,549	3,203	6.5%

The last period for which a backcast was conducted was from 1990 to 2000. Over this period there was a large increase in new residential land, partly coinciding with the housing boom of the late 1990s. This change had major implications for regional land use. The MCCA model responded by adding roughly 20,000 cells of residential land to the region during this period, even though this about five percent below actual observed residential land use. The underforecast of vacant land seems to suggest that prediction errors often arise (at least in part) in miscalculating the amount of land at the urban fringe that is converted to housing. Table 5-19 shows that the prediction error in the MCCA model appears to be compounded over several prediction periods. There appears to be a lack of stability in certain types of land uses over time. Growth in Highway, airport and park lands in the 1950s and 1960s did not carry over at the same rate to the 1990s. This seems reasonable, as the infrastructure building booms associated with the early growth in air travel and the development of urban highway networks took place largely during the postwar period, but have not kept pace during more recent decades. Figures 5-13 and 5-14 display maps of predicted and actual changes. While the dispersion of land use in the MCCA model is among the most prominent contrasts, there is also substantial variation in the prediction of residential land at the urban fringe.

Using the transition matrix that was estimated for data from 1990 and 2000, the MCCA model can be applied to forecast land use change during the coming decades. Here we apply the model to forecast land use in 10-year time steps from 2000 to 2030. Table 5-20 provides a summary of the land use forecasts, in both absolute and percentage terms with the actual land use distribution in 2000 serving as a control point. Park land use is forecast to grow most rapidly, increasing by over 90 percent between 2000 and 2030. However, in absolute terms, residential land use accounts for most of the growth, with commerical and highway uses growing by smaller amounts. Industrial land use is predicted to grow over this period, though only very slowly, while vacant land is forecast to decline by more than 35 percent, representing a loss in vacant land between 2000 and 2030 that is equivalent to more than one-seventh of the land in the entire study area. The dynamics of these changes are shown in the series of maps listed as Figures 5-15, 5-16 and 5-17.

Table 5-19: Predicted land use change, 1990–2000 (MCCA Model)

Land Use	Predicted 2000	Actual 2000	Difference	Percent
Airport	10,016	4,616	5,400	117.0%
Commercial	23,131	18,779	4,352	23.2%
Highway	31,124	14,849	16,275	109.6%
Industrial	24,102	23,777	325	1.4%
Parks	123,222	68,565	54,657	79.7%
Public	25,245	18,322	6,923	37.8%
Railway	2,636	1,523	1,113	73.1%
Residential	177,809	186,347	-8,538	-4.6%
Vacant	136,899	217,387	-80,488	-37.0%
Water	52,752	52,771	-19	0.0%

Table 5-20: Forecast land use change, 2000–2030 (MCCA Model)

Land use	2000	2010	2020	2030	Change 2000-2030	Change (%)
Residential	186,347	205,658	224,234	237,125	50,778	27.2%
Commercial	18,779	21,894	24,993	26,984	8,205	43.7%
Industrial	23,777	22,871	23,256	22,179	-1,598	-6.7%
Public	18,322	17,089	16,851	16,625	-1,697	-9.3%
Parks	68,565	87,348	109,936	131,793	63,228	92.2%
Vacant	217,387	171,517	119,887	78,093	-139,294	-64.1%
Highway	14,849	17,824	21,347	24,684	9,835	66.2%
Railway	1,523	1,594	1,689	1,789	266	17.5%
Airport	4,616	4,852	5,086	5,320	704	15.3%
Water	52,771	56,289	59,657	62,344	9,573	18.1%
Total	548,026	544,201	540,504	537,483	-10,543	-1.9%

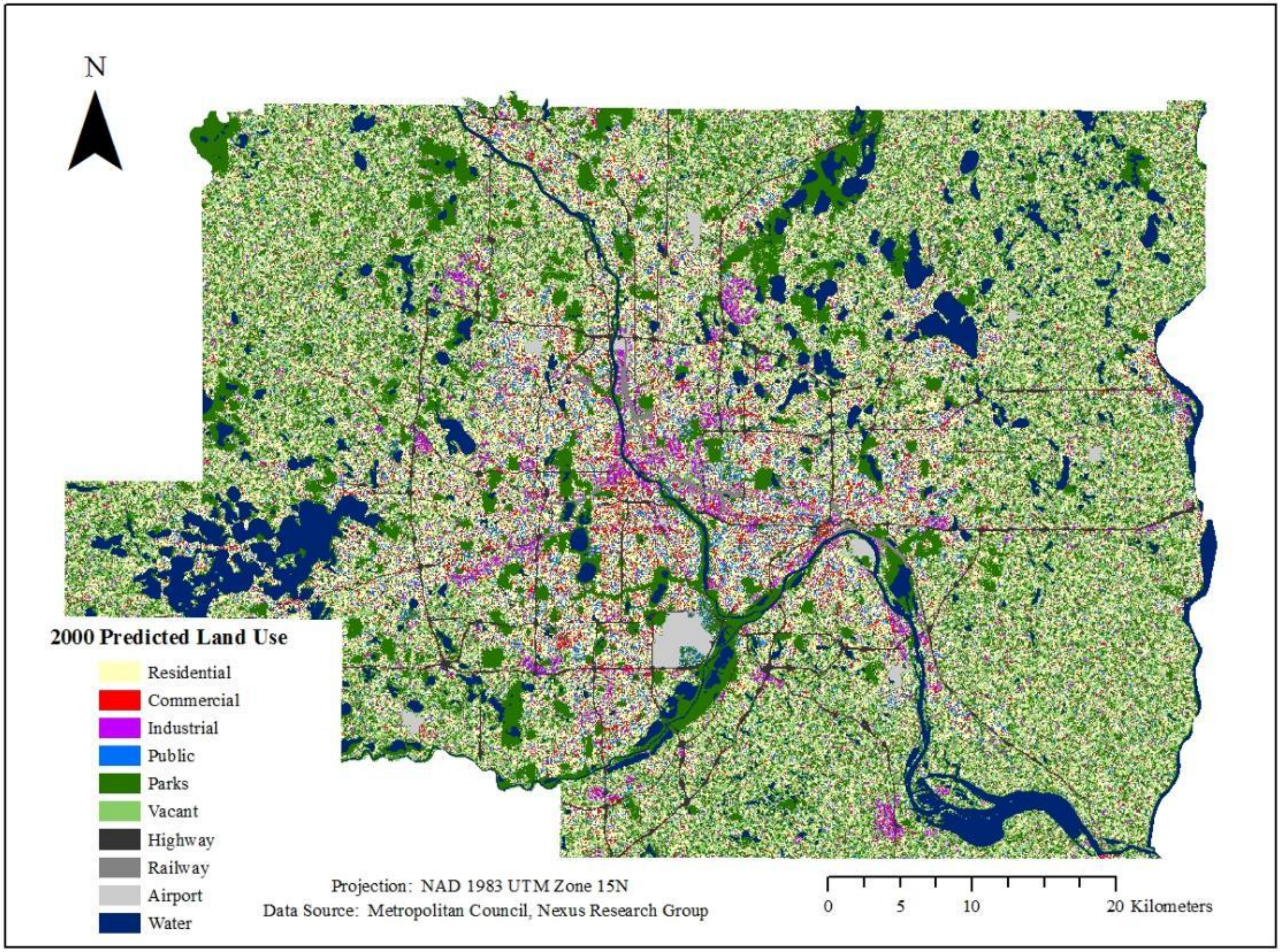


Figure 5-13: Predicted land use in 2000 (MCCA Model)

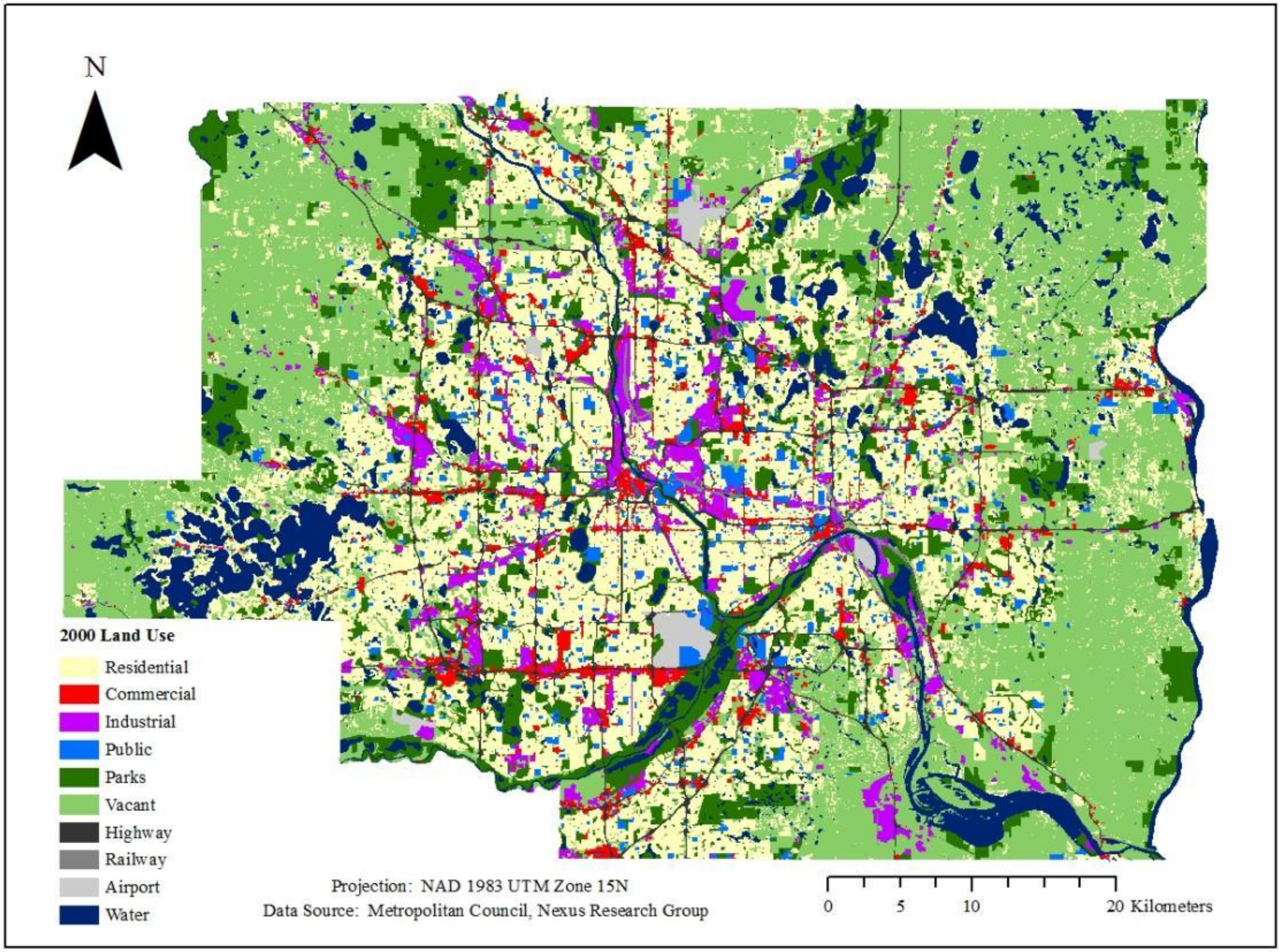


Figure 5-14: Actual land use in 2000

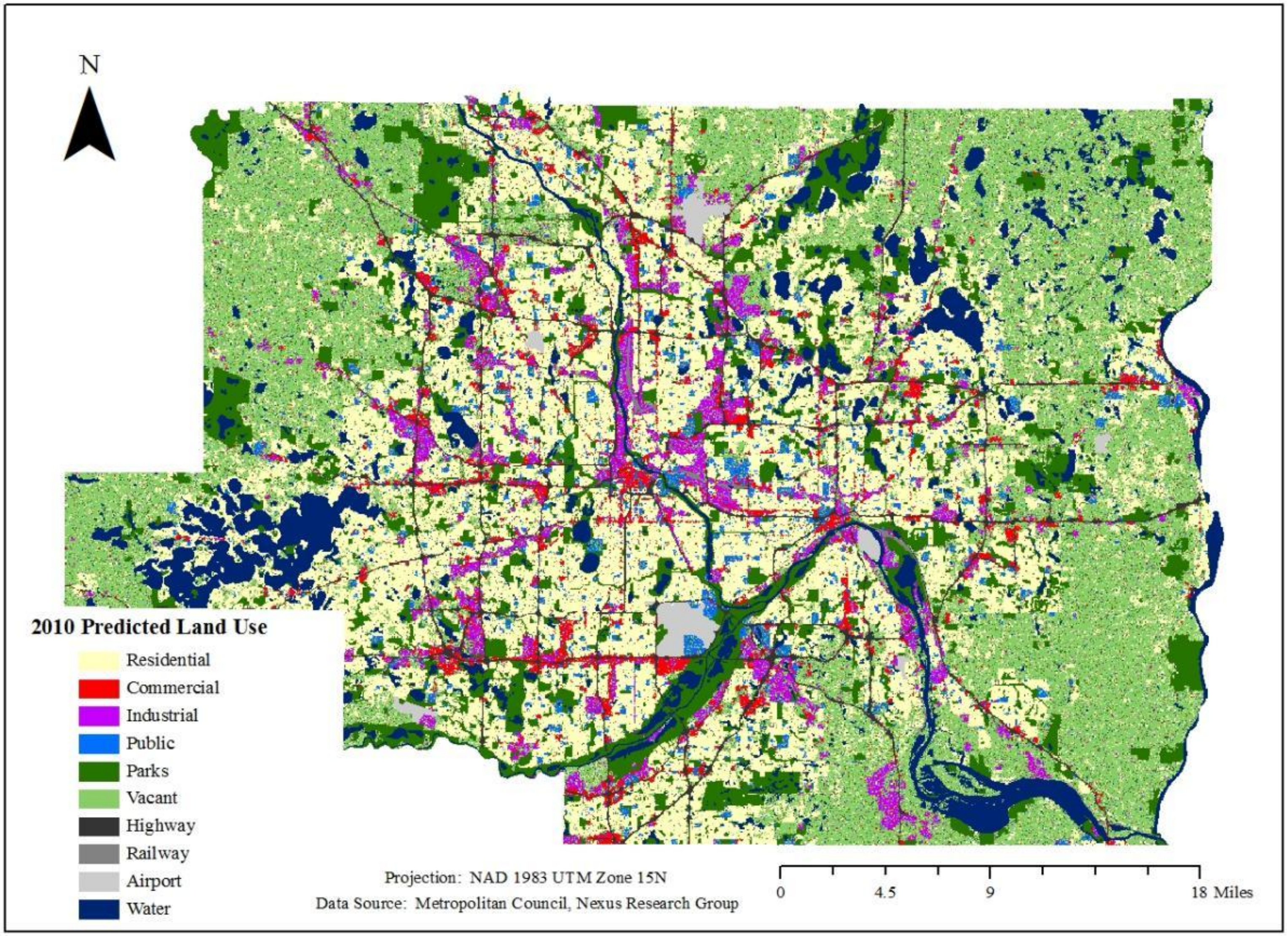


Figure 5-15: Predicted land use in 2010 (MCCA Model)

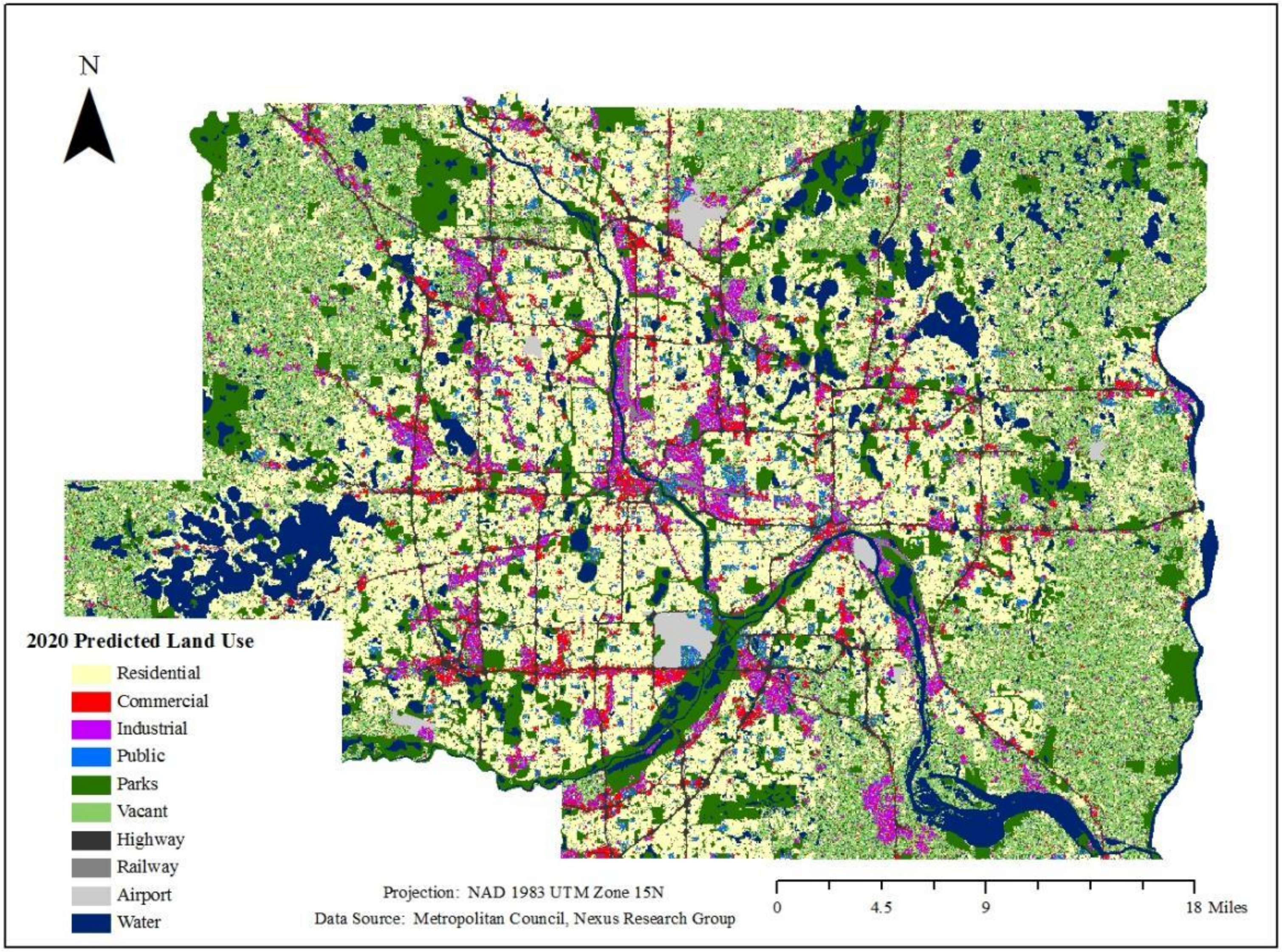


Figure 5-16: Predicted land use in 2020 (MCCA Model)

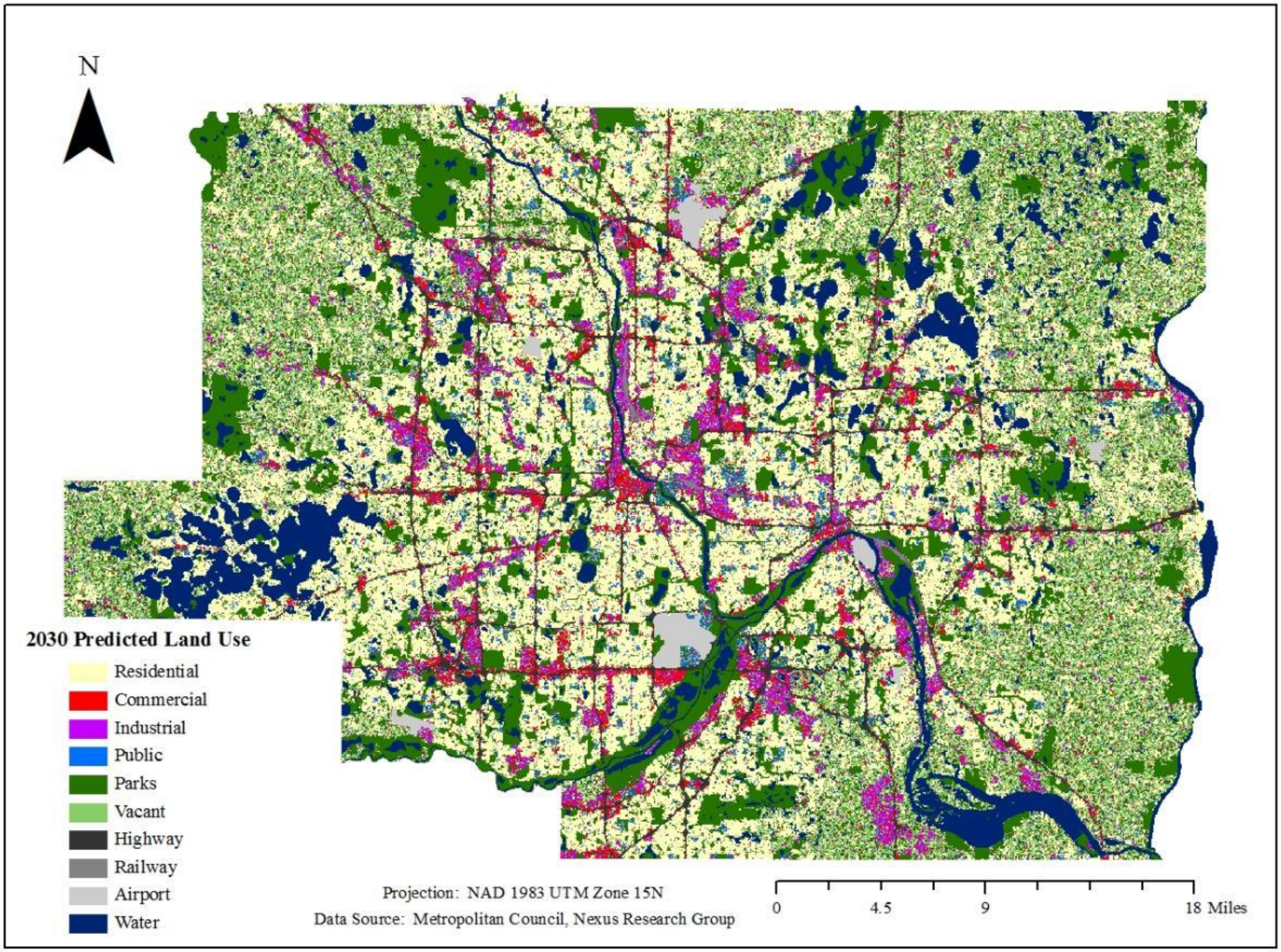


Figure 5-17: Predicted land use in 2030 (MCCA Model)

Chapter 6

Policy Simulation

In the process of applying and validating the models of land use change we dealt with a large study area covering much of the core seven counties of the Twin Cities region. The results described in the previous section are averaged out across different parts of the region, some of which have little in common (e.g. developed central city parcels and vacant agricultural land at the urban fringe). It would be helpful to test the sensitivity of the models by applying them to a part of the region that is experiencing relatively strong growth, and where important improvements to the transportation network are occurring.

We do so by applying the models to a part of the region near Minnesota State Highway 610, a new four-lane, limited-access highway in the northern Minneapolis suburbs of Brooklyn Park and Maple Grove. The highway is currently being built in sections (the first having been completed in the late 1990s), with the last 5 mile (8 km) stretch between U.S. Highway 169 and Interstate 94 awaiting completion. A corridor is defined for the Highway 610 by including all land within two miles (3.2 km) of the highway. This area still has a significant amount of vacant land, some of which local officials expect to see developed following the highway's construction. The models are applied to forecast land use change, as was done in the previous section, assuming the completion of the last section of highway.

6.1 Markov Chain Model

The base case for analysis of land use change in the Highway 610 corridor is the Markov Chain model. Unlike the other two models, the MC model does not incorporate any neighbor effects. Hence, any effects of the new highway construction will primarily appear as land being converted to highway uses. Land use growth or decline in other classes will reflect growth occurring during the 1997-2005 period, during which growth in the corridor began to respond to the availability of part of the new Highway 610 link.

This application of the MC model will also differ from the region-wide application in that it will incorporate *absorbing states*. These are defined as land use classes that can receive additional growth from other land use classes, but from which no new transitions will originate. Highways and railroads, given their rather fixed nature, will be considered as absorbing states, along with water and park lands (airports are absent from this study area). This modification should provide for more consistency in the transition matrices, by eliminating the possibility of some transitions

Table 6-1: Forecast land use change in Highway 610 Corridor, 2005-2029 (MC Model)

Land use	2005	2013	2021	2029	Change 2005-2029	Change (%)
Residential	8,705	9,611	10,220	10,599	1,894	21.8%
Commercial	719	839	869	891	172	23.9%
Industrial	1,396	1,695	1,876	1,985	589	42.2%
Public	802	772	761	716	-86	-10.7%
Parks	3,802	4,122	4,343	4,531	729	19.2%
Vacant	7,156	5,133	3,761	2,857	-4,299	-60.1%
Highway	1,012	1,359	1,646	1,857	845	83.5%
Railway	55	55	56	56	1	1.8%
Water	991	1,052	1,106	1,146	155	15.6%
Total	24,638	24,638	24,638	24,638	0	0

that merely represent coding or other types of classification error. It also provides for a more orderly representation of land use in the corridor, reducing some of the random scattering that generally appears in the output of the MC model.

Forecasts are made in eight-year increments to the year 2029, based on the estimated transition matrix for the period 1997-2005. Results of the forecasts are summarized in Table 6-1, which displays the land use distribution in each of the forecast years, along with the absolute and relative changes during the entire forecast period. Maps showing land use in the corridor at each of the forecast years are contained in Figures 6-1, 6-2 and 6-3. The MC model predicts a continuation of strong residential land use growth in the Corridor. According to the MC forecasts, by 2029 more than two-fifths of the land in the corridor could be consumed by residential uses. Significant growth is also forecast for industrial uses, which already comprise about five percent of land use in the corridor. The large growth in Highway land mostly reflects the major construction activity during the first phase of Highway 610 during the late 1990s, and is carried forward into future periods. Much of the growth in the corridor is forecast to occur on previously vacant land, as the amount of vacant land forecast for 2029 is less than half of its actual total in 2005.

Figures 6-1 through 6-3 indicate that the model predicts much of the growth to occur on vacant land in the northwestern and central parts of the corridor, along Highway 610 and Interstate 94. The random nature of the MC model, lack of neighbor effects and high probability of transition from vacant to other states largely explains the high degree of mixing of different land uses in new growth areas. Commercial and industrial land uses appear to be complements, as some of the commercial growth is forecast to occur within the bounds of formerly exclusive industrial sites.

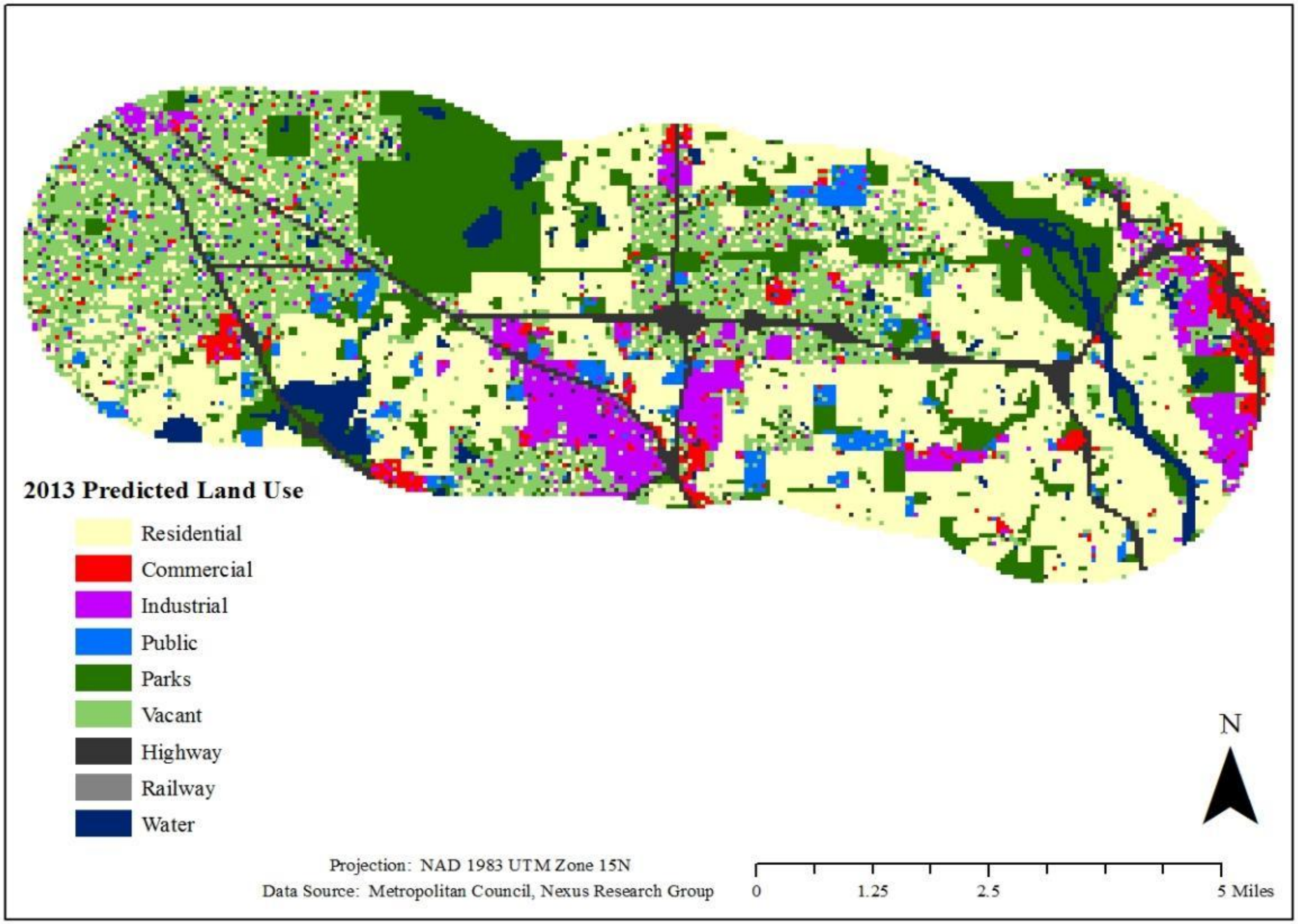


Figure 6-1: Predicted land use for Highway 610 Corridor in 2013 (MC Model)

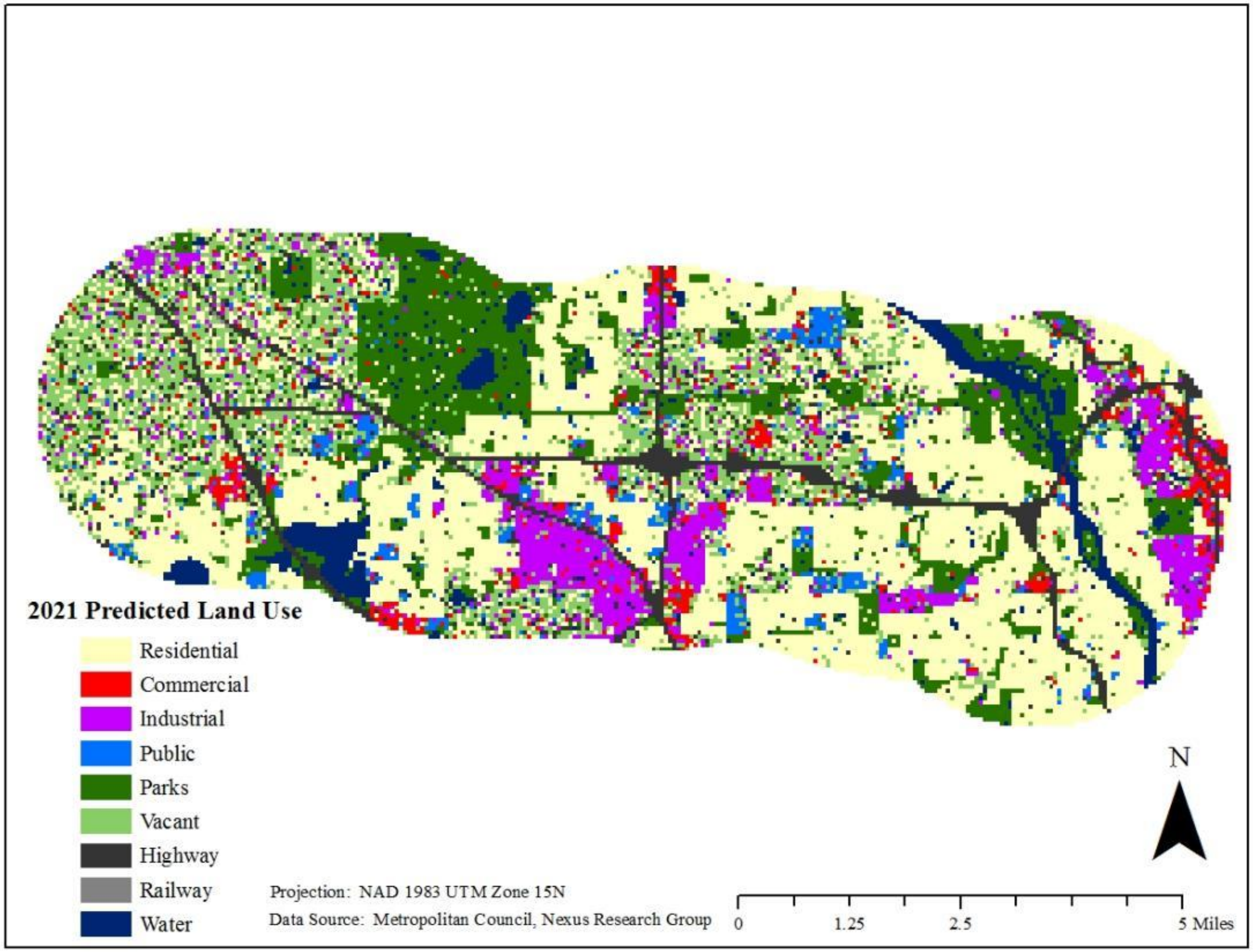


Figure 6-2: Predicted land use for Highway 610 Corridor in 2021 (MC Model)

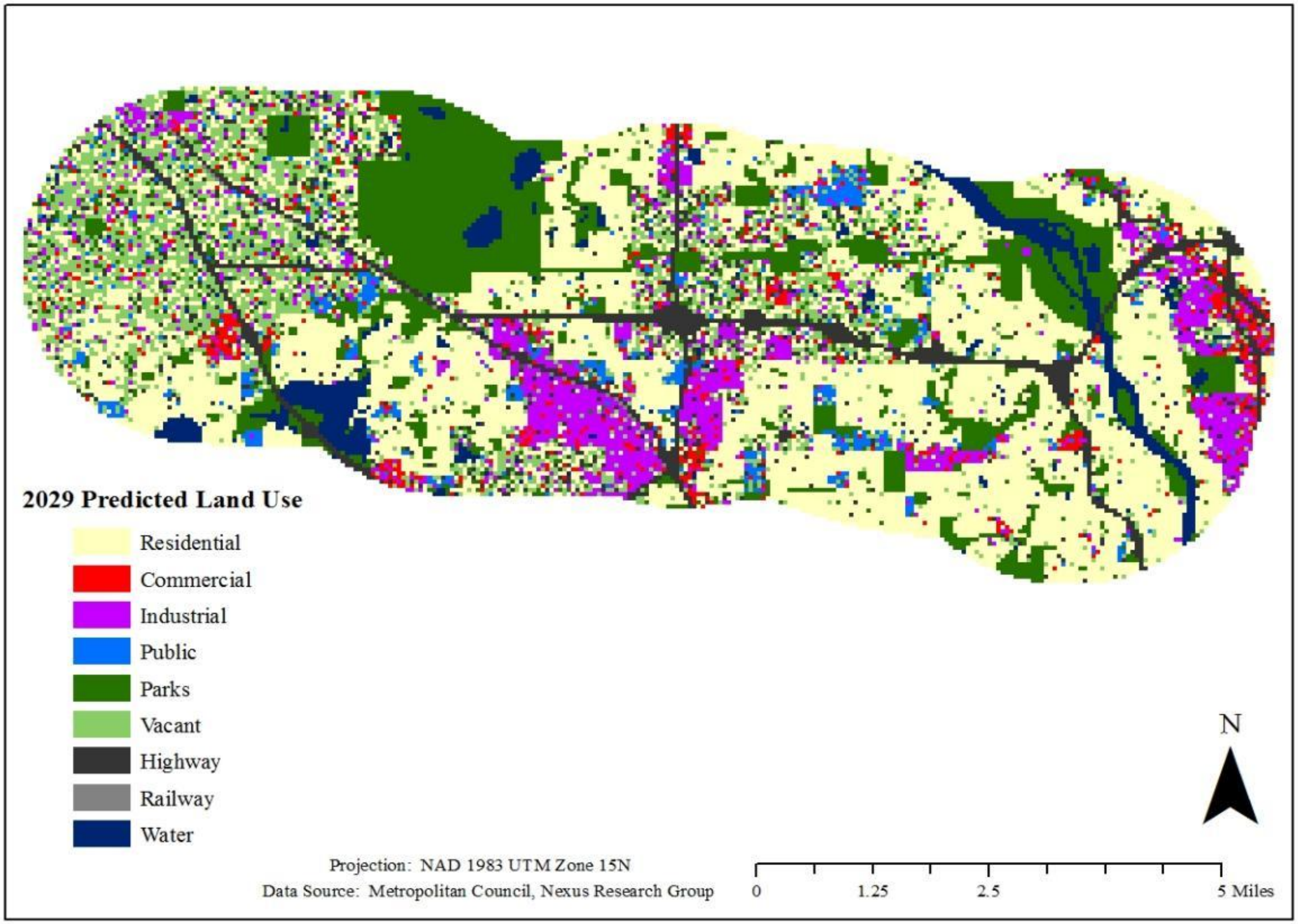


Figure 6-3: Predicted land use for Highway 610 Corridor in 2029 (MC Model)

Table 6-2: Forecast land use change in Highway 610 Corridor, 2000-2030 (logistic regression)

Land use	2000	2010	2020	2030	Change 2000-2030	Change (%)
Residential	8,089	8,332	8,421	8,460	371	4.6%
Commercial	661	747	787	817	156	23.6%
Industrial	1,169	1,136	1,120	1,118	-51	-4.4%
Vacant	8,382	8,003	7,817	7,704	-678	-8.1%
Other	6,337	6,420	6,493	6,539	202	3.2%
Total	24,638	24,638	24,638	24,638	0	0.0%

6.2 Logistic Regression Model

In order to predict land use change for the Highway 610 corridor using the logistic regression framework, a separate model was fitted to the land use data representing the corridor for the period 1990 to 2000. This model was used to forecast in 10-year increments between 2000 and 2030. Since land use was the only type of variable that could be reliably forecast for future periods, other variables were held at their 2000 values. This accounts for the more modest forecasts of change shown in Table 6-2. Among the five land use classes listed, only commercial land use was forecast to grow by more than 10 percent between 2000 and 2030. Residential land is forecast to grow by less than five percent. However, even this modest increase is enough to make residential land use account for more than one-third of the land in the corridor by 2030. Industrial land use is forecast to decline slightly by 2030.

As with the other forecasts, most growth is expected to occur at the expense of vacant land. Figures 6-4 through 6-6 show maps of land use in the corridor in 2010, 2020 and 2030. Much of the residential growth appears to take place along the western edge of the corridor, west of I-94, and in the central sections of the corridor near Highway 610. Commercial nodes expand from their previous locations near the major highways. Overall, the contiguous sections of different land uses are well preserved by the regression model.

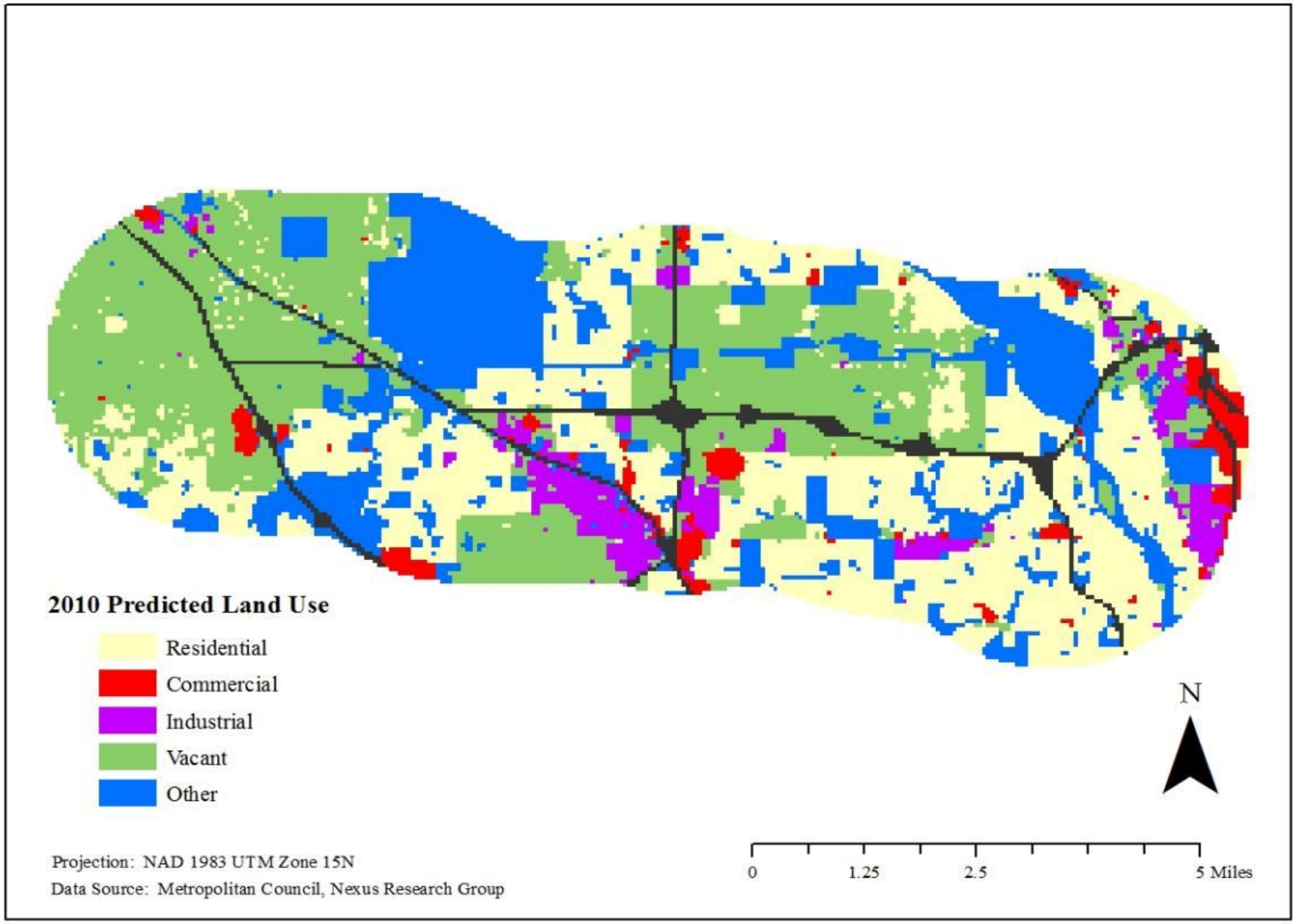


Figure 6-4: Predicted land use for Highway 610 Corridor in 2010 (logistic regression)

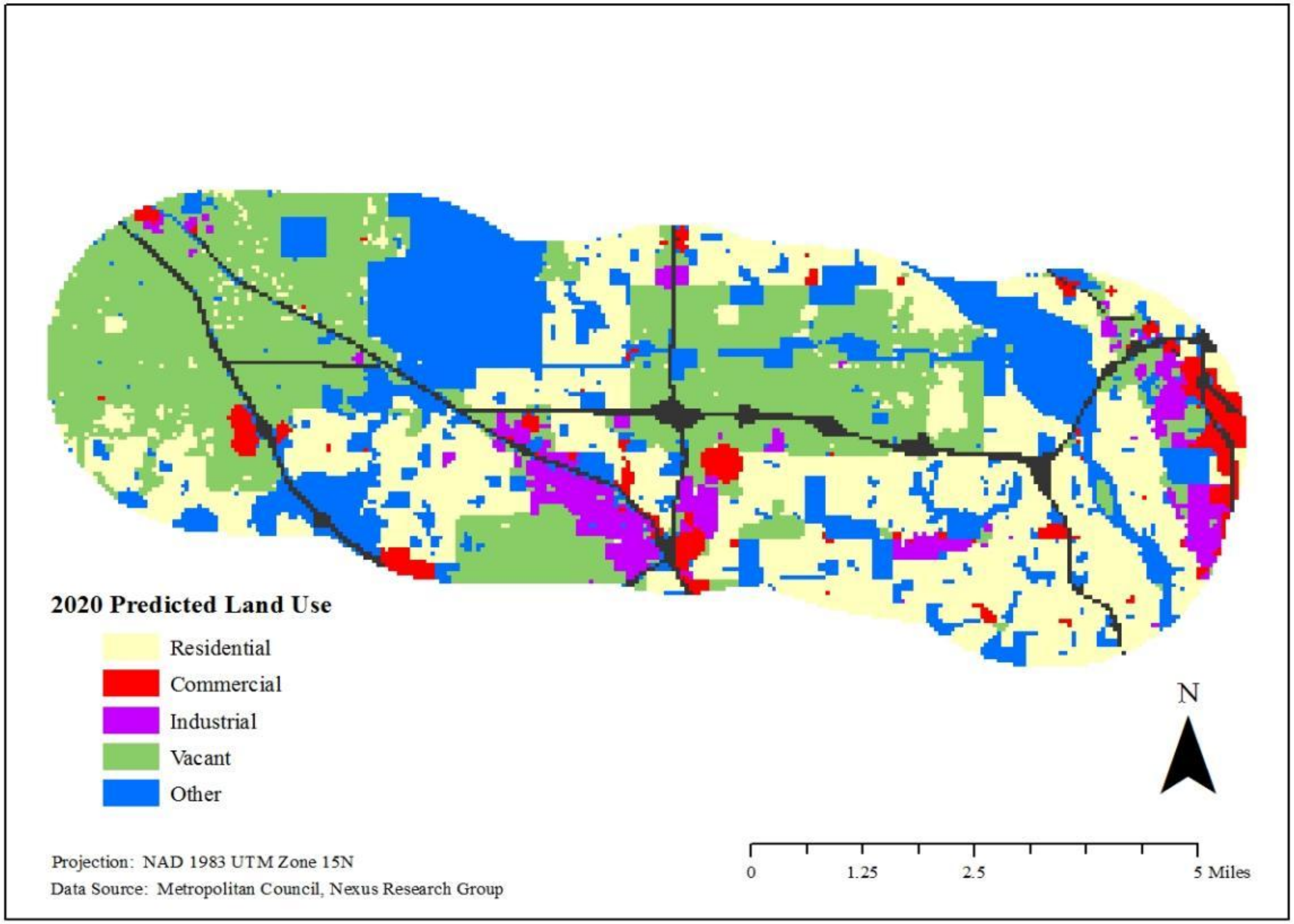


Figure 6-5: Predicted land use for Highway 610 Corridor in 2020 (logistic regression)

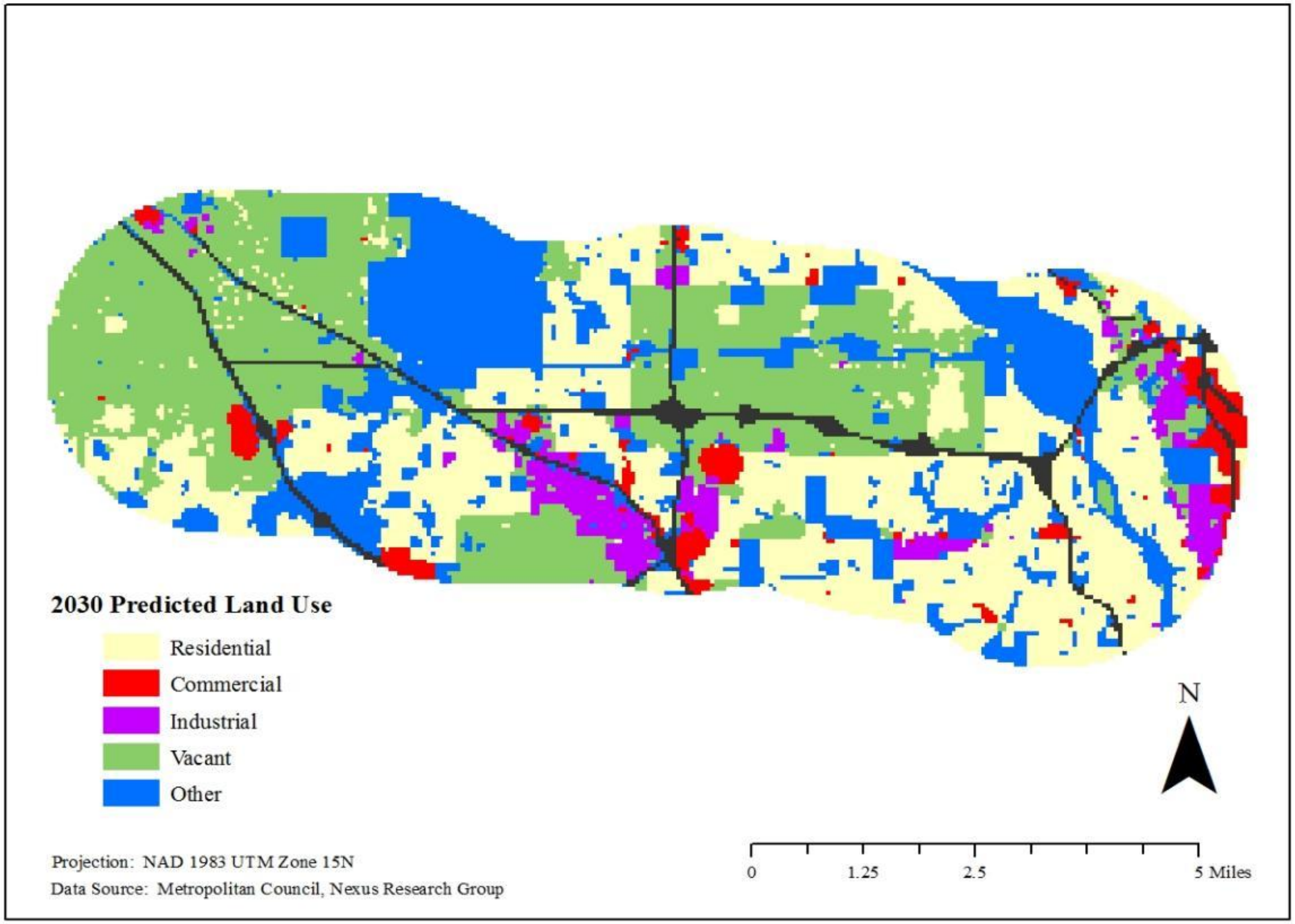


Figure 6-6: Predicted land use for Highway 610 Corridor in 2030 (logistic regression)

Table 6-3: Forecast land use change in Highway 610 Corridor, 2000-2030 (MCCA Model)

Land use	2005	2013	2021	2029	Change 2005-2029	Change (%)
Residential	8,705	9636	10,293	10,706	2,001	23.0%
Commercial	719	849	939	957	238	33.1%
Industrial	1,396	1679	1,934	2,072	676	48.4%
Public	802	779	742	666	-136	-17.0%
Parks	3,802	4171	4,533	4,808	1,006	26.5%
Vacant	7,156	5032	3,274	2,101	-5,055	-70.6%
Highway	1,012	1405	1,795	2,166	1,154	114.0%
Railway	55	56	59	66	11	20.0%
Airport	0	0	0	0	0	0.0%
Water	991	1031	1,069	1,096	105	10.6%
Total	24,638	24,638	24,638	24,638	0	0

6.3 Markov Chain-Cellular Automata Model

The Markov Chain-Cellular Automata (MCCA) model is the third model to be applied to simulate land use change in the corridor. For the purpose of forecasting, all ten land use classes are defined, with five of them (highways, railways, airports, parks and water) defined as absorbing states (they can expand but not contract). The aggregate forecasts of land use change in the Highway 610 corridor using the MCCA model are somewhat similar to those produced by the Markov Chain model, as is shown in Table 6-3. Residential land use growth of around 20 percent is forecast for the corridor, while commercial growth exceeds 30 percent. Highway-related land is again forecast to grow rapidly in the corridor, more than doubling in coverage. An educated guess would suggest that this prediction is probably too high, given that conversion of land to highway uses will probably slow considerably following the completion of Highway 610. The MCCA and MC models both predict significant growth in industrial uses, in contrast to the prediction of decline by the logistic regression model. The growth in cells classified as water most likely reflect classification errors in the land use data set between 1997 and 2005, the period for which a transition matrix was estimated for the study area.

The pattern of growth implied by the MCCA model forecasts are depicted in Figures 6-7 through 6-9. Similar to the Markov Chain model, much of the new growth is accommodated in the large tracts of vacant land in northern Maple Grove and Brooklyn Park. There is still a high degree of mixing of land uses, but new growth appears to be more concentrated in the MCCA model, both near existing clusters and along major transportation corridors. The MCCA model seems to be capable of preserving existing, contiguous tracts of residential and park land, but seems to perform less well at sorting out future land use growth on formerly vacant or agricultural land. Additional information about spatial location, proximity to urban infrastructure and other constraints might be needed to better handle land conversion at the urban fringe.

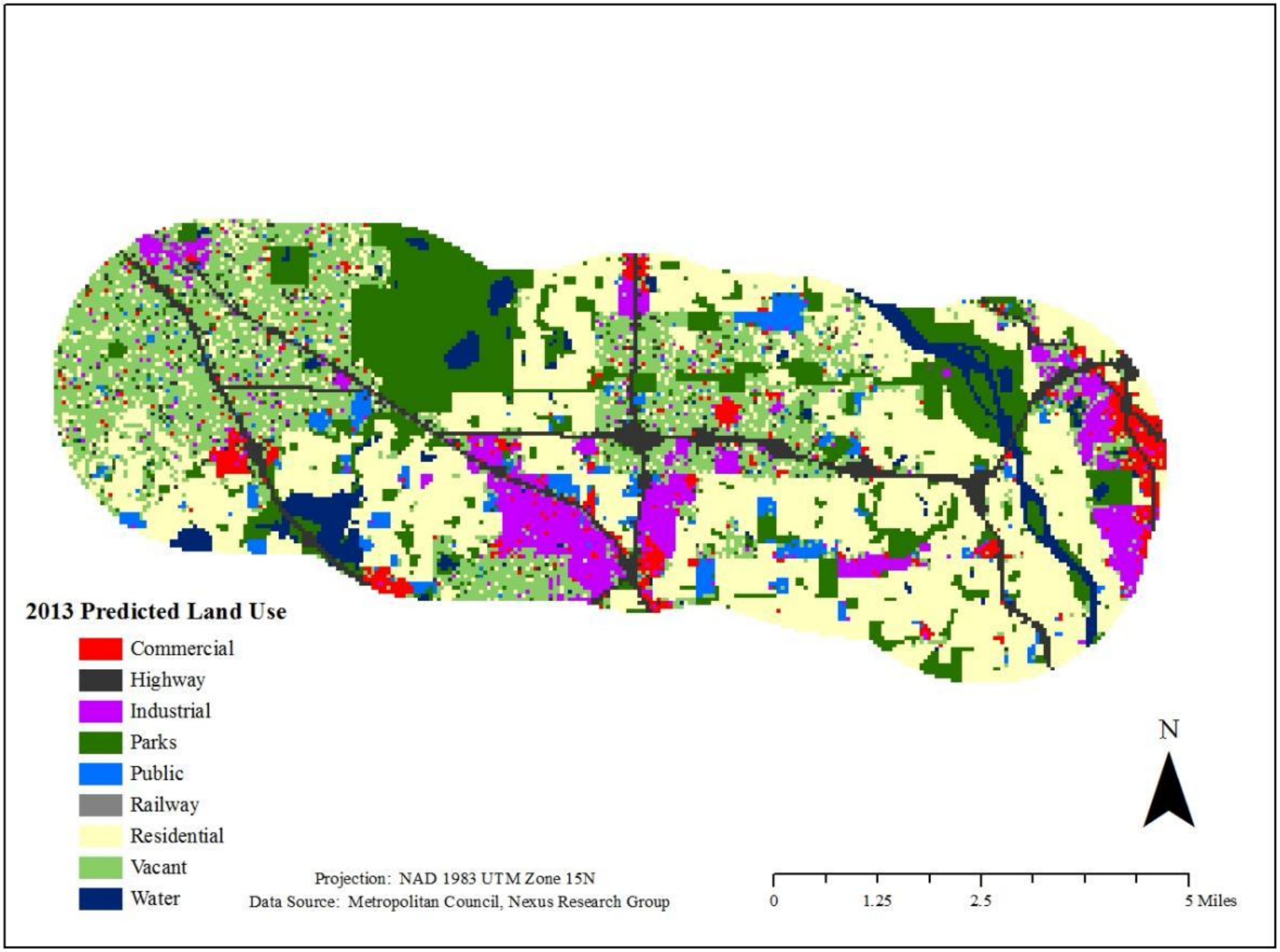


Figure 6-7: Predicted land use for Highway 610 Corridor in 2010 (MCCA Model)

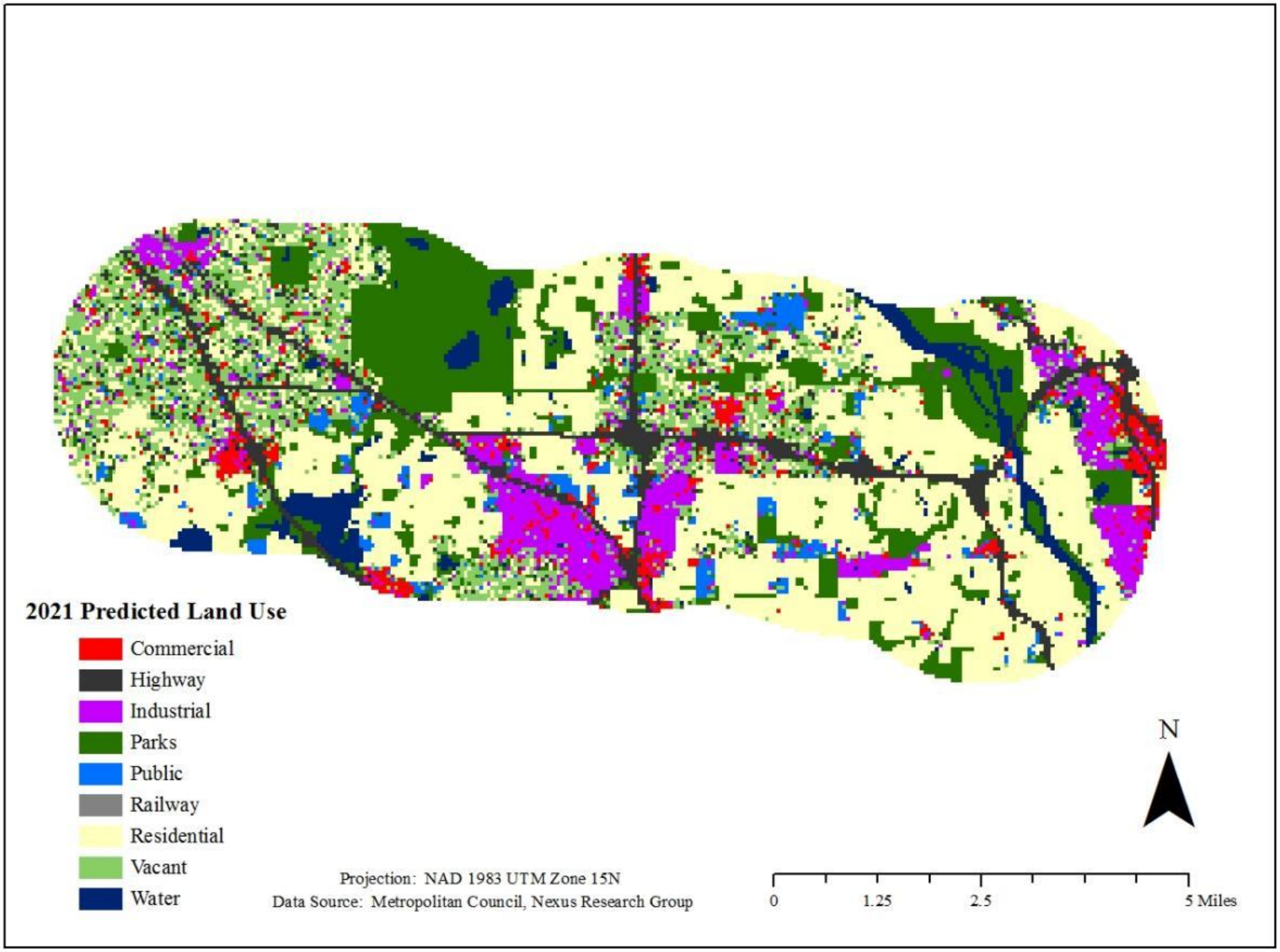


Figure 6-8: Predicted land use for Highway 610 Corridor in 2020 (MCCA Model)

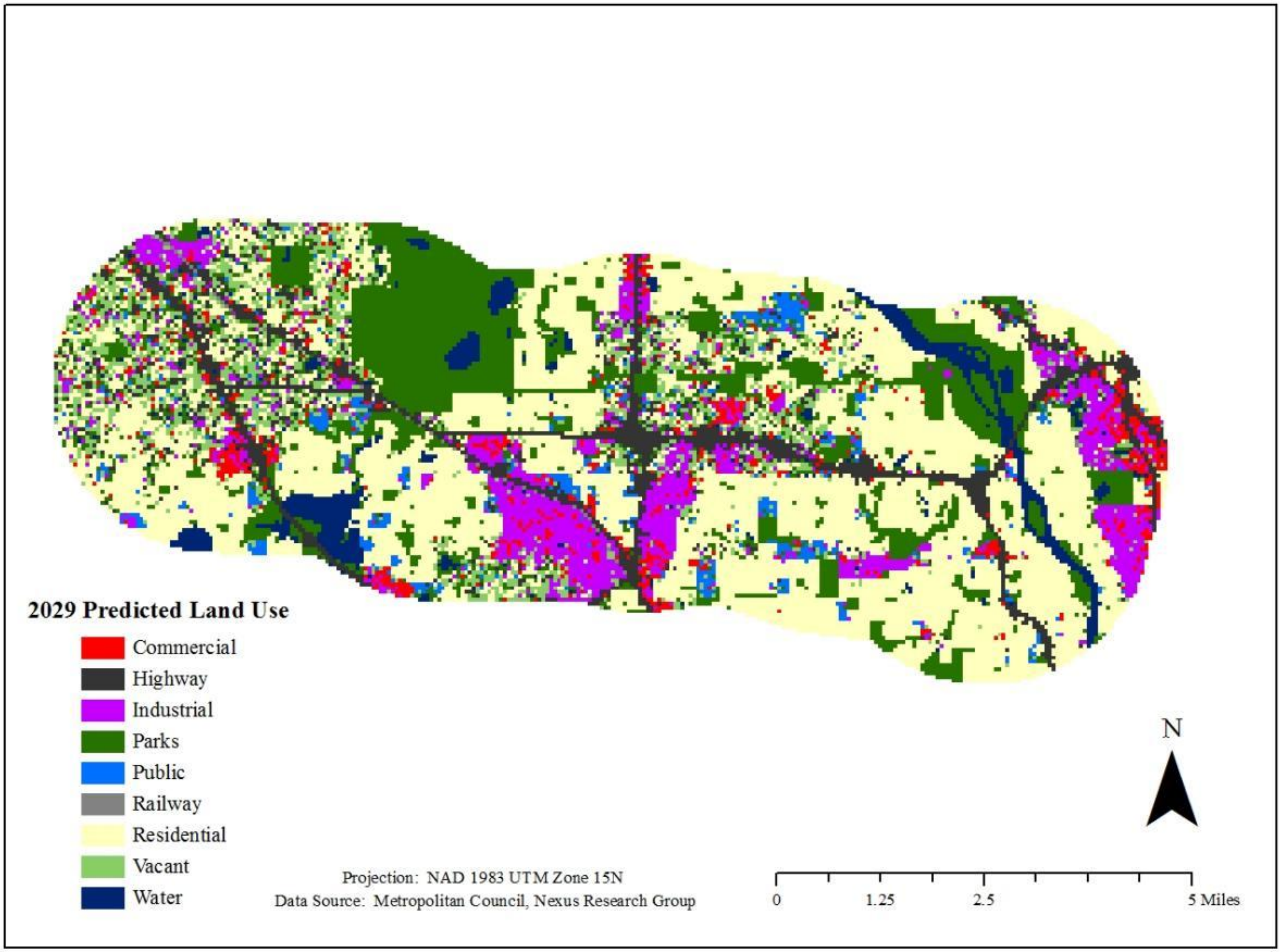


Figure 6-9: Predicted land use for Highway 610 Corridor in 2030 (MCCA Model)

Chapter 7

Conclusion

In this study we have formulated and applied a set of cellular models for tracking and predicting urban land use change. Each of the models retains a relatively simple structure, but is able to incorporate important elements of land use change processes, such as historical dependence, stochasticity, and neighbor effects. Using a uniquely fine-grained, cell-level land use data set as an input, these models were validated at the regional level, then applied in the context of a high-growth suburban environment near a new highway corridor. Each of the models showed certain strengths, but also some limitations as practical planning tools.

The Markov Chain model represents the most basic formulation of the land use change process. While it does a reasonable job of forecasting aggregate levels of land use change, its lack of spatial relationships and detail make its results less policy-sensitive and more difficult to interpret.

Some of these problems are alleviated by incorporating neighbor effects into the model design. The logistic regression model of land use change took account of both previous land use states and states of neighboring cells, in addition to other important factors such as proximity to highway networks and level of accessibility. As we saw, models specified with only these variables were able to produce reasonably accurate forecasts of past land use change. The ability to forecast these policy variables well into the future makes using this model as a forecasting tool difficult, however. Also, the land use patterns predicted by the regression model tended to be too centralized and contiguous, particularly in the case of residential development. This characteristic makes predicting low-density residential growth at the urban fringe difficult.

The Markov Chain-Cellular Automata Model combined the features of the previous two models to provide a projection model that incorporates spatial detail. The more refined definition of land use states, including neighboring land uses, allows for a better organization of space than the Markov Chain model. Despite this advantage, it did not seem to predict much more accurately than the basic MC structure. The MCCA model is also more difficult to apply at smaller scales, since using a larger number of land use classes can imply very large sets of possible transitions, and lower probabilities for each type of transition.

The design of the cell-based models is such that they can remain relatively simple and transparent, like the models described in this study. This is an important feature, since it allows the effects of different scenarios to be traced rather easily. However, this simplicity does present a tradeoff. The cell-based models described in this study are essentially projection models, and lack a strong theoretical structure. The individual cells are treated as agents, which makes the establishment of behavioral principles difficult.

While cellular models may not be the best available tools for testing theories about the behavior of actors in the land development process, they may hold more value in simple sketch planning applications. Models like the MCCA model can be built upon to provide greater realism, such as the inclusion of land use intensity in defining states, while trading this additional realism off against data collection and computational costs. Also, since cellular models generally are developed from assumptions about self-organizing behavior, they can provide an interesting counterfactual or control forecast for forecasts based on strong assumptions about responses to established land use plans. In this way, they should be seen as complementary tools to existing land use planning methods, as opposed to complete replacements.

References

Anas, A. (1982), *Residential location markets and urban transportation: economic theory, econometrics and policy analysis with discrete choice models*, Academic Press, New York.

Anas, A. (1983), "Discrete choice theory, information theory and the multinomial logit and gravity models", *Transportation Research, Part B* , Vol. 17B, pp. 13–23.

Anas, A. (1984), "Discrete choice theory and the general equilibrium of employment, housing and travel networks in a Lowry-type model of the urban economy", *Environment & Planning A* , Vol. 16, pp. 1489–1502.

Anas, A. (1987), *Modeling in urban and regional economics*, Fundamentals of pure and applied economics, v. 26, Harwood Academic Publishers, Chur, Switzerland.

Anas, A. (1998), *NYMTC transportation models and data initiative: the NYMTC Land Use Model*, technical report, Alex Anas and Associates, Williamsville, NY.

Anas, A. and Arnott, R. J. (1994), "The Chicago prototype housing market model, with tenure choice and its policy implications", *Journal of Housing Research* , Vol. 5, pp. 73–129.

Anderson, T. and Goodman, L. (1957), "Statistical inference about Markov chains", *Annals of Mathematical Statistics* , Vol. 28, pp. 89–110.

Arentze, T. A. and Timmermans, H. J. (2003), "Modeling agglomeration forces in urban dynamics: a multi-agent system approach". Paper presented at the 8th International Conference on Computers in Urban Planning and Urban Management, May 27-29, Sendai, Japan.

Arentze, T. A. and Timmermans, H. J. (2004), "A learning-based transportation-oriented simulation system", *Transportation Research, Part B: Methodological* , Vol. 38B, pp. 613–633.

Axhausen, K. W. (1990), "A simultaneous simulation of activity chains and traffic flow", *in* P. M. Jones, ed., '*Developments in dynamic and activity-based travel analysis*', Oxford studies in transport, Avebury, Brookfield, VT, pp. 206–225.

Axhausen, K. W. and Gärling, T. (1992), "Activity-based approaches to travel analysis: conceptual frameworks, models and research problems", *Transport Reviews* , Vol. 12, pp. 323–341.

Baker, W. (1989), "A review of models of landscape change", *Landscape Ecology* , Vol. 2, pp. 111–133.

Barra, T. d. l. (1989), *Integrated transport and land use modeling: decision chains and hierarchies*, Cambridge University Press, Cambridge, UK.

Barrett, C. (1995), *An operational description of TRANSIMS*, Technical Report LA-UR-95-2393, Los Alamos National Laboratory, Los Alamos, NM.

Batty, M. (1979), "Progress, success and failure in urban modeling", *Environment & Planning A* , Vol. 11, pp. 863–878.

Batty, M. (1997), "Cellular automata and urban form: a primer", *Journal of the American Planning Association* , Vol. 63, pp. 264–274.

Batty, M. (2005), *Cities and complexity: understanding cities with cellular automata, agent-based models and fractals*, MIT Press, Cambridge, MA.

Benenson, I. (1998), "Multi-agent simulations of residential dynamics in the city", *Computers, Environment and Urban Systems* , Vol. 22, pp. 25–42.

Berechman, J. and Small, K. A. (1988), "Modeling land use and transportation: an interpretive review for growth areas", *Environment & Planning A* , Vol. 20, pp. 1285–1309.

Bourne, L. (1971), "Physical adjustment processes and land use succession: a conceptual review and central city example", *Economic Geography* , Vol. 47, pp. 1–15.

Bowman, J. L. and Ben-Akiva, M. E. (2001), "Activity-based disaggregate travel demand model system with activity schedules", *Transportation Research, Part A: Policy and Practice* , Vol. 35A, pp. 1–28.

Cervero, R. (1984), "Light rail transit and urban development", *Journal of the American Planning Association* , Vol. 50, pp. 133–147.

Chang, J. (2006), "Models of the relationship between transport and land-use: a review", *Transport Reviews* , Vol. 26, pp. 325–350.

Chapin, F. S. (1974), *Human activity patterns in the city: things people do in time and in space*, Wiley, New York.

Clarke, K. and Gaydos, L. (1998), "Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore", *International Journal of Geographic Information Science* , Vol. 12, pp. 699–714.

Clarke, K., Hoppen, S. and Gaydos, L. (1997), "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area", *Environment & Planning B: Planning and Design* , Vol. 24, pp. 247–261.

Clay, M. J. and Johnston, R. A. (2006), "Multivariate uncertainty analysis of an integrated land use and transportation model: MEPLAN", *Transportation Research, Part D: Transport and Environment* , Vol. 11D, pp. 191–203.

Crecine, J. (1964), *TOMM: Time-Oriented Metropolitan Model*, Technical report, Department of City and Regional Planning, Pittsburgh, PA.

Dong, X., Ben-Akiva, M. E., Bowman, J. L. and Walker, J. L. (2006), “Moving from trip-based to activity-based measures of accessibility”, *Transportation Research, Part A: Policy and Practice* , Vol. 40A, pp. 163–180.

Echenique, M. H. (2004), “Econometric models of land use and transportation”, *in* D. A. Hensher, K. J. Button, K. E. Haynes and P. R. Stopher, eds, ‘*Handbook of transport geography and spatial systems*’, Handbooks in Transport, v.3, Pergamon, Amsterdam, The Netherlands, pp. 185–202.

Echenique, M. H., Crowther, D. and Lindsay, W. (1969), “A spatial model of urban stock and activity”, *Regional Studies* , Vol. 3, pp. 281–312.

Echenique, M. H., Flowerdew, A., Hunt, J., Mayo, T., Skidmore, I. and Simmonds, D. (1990), “The Meplan models of Bilbao, Leeds and Dortmund”, *Transport Reviews* , Vol. 10, pp. 309–322.

Ellickson, B. (1981), “An alternative test of the hedonic theory of housing markets”, *Journal of Urban Economics* , Vol. 9, pp. 56–79.

Ettema, D. and Timmermans, H. J. (1997), “Activity-based approaches: an introduction”, *in* D. Ettema and H. J. Timmermans, eds, ‘*Activity-based approaches to travel analysis*’, Pergamon, Oxford, UK, pp. 1–36.

Garin, R. A. (1966), “A matrix formulation of the Lowry model for intra-metropolitan activity”, *Journal of the American Institute of Planners* , Vol. 32, pp. 361–364.

Goldner, W. (1971), “The Lowry model heritage”, *Journal of the American Institute of Planners* , Vol. 37, pp. 100–110.

Hägerstrand, T. (1970), “What about people in regional science?”, *Papers of the Regional Science Association* , Vol. 24, pp. 7–21.

Harris, B. (1994), “The real issues behind Lee’s ‘Requiem’”, *Journal of the American Planning Association* , Vol. 60, pp. 31–34.

Hess, S., Train, K. E. and Polak, J. W. (2006), “On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice”, *Transportation Research, Part B: Methodological* , Vol. 40B, pp. 147–163.

Horowitz, A. J. (2004), “Lowry-type land use models”, *in* D. A. Hensher, K. J. Button and K. E. Haynes, eds, ‘*Handbook of transport geography and spatial systems*’, Handbooks in Transport, v. 3, Pergamon, Amsterdam, The Netherlands, pp. 167–183.

Hunt, J. D. and Abraham, J. E. (2005), “Design and implementation of PECAS: a generalised system for allocating economic production, exchange and consumption quantities”, *in* M. E. Lee-Gosselin and S. T. Doherty, eds, ‘*Integrated land-use and transportation models: behavioural foundations*’, Elsevier, Amsterdam, The Netherlands, pp. 253–273.

Hunt, J. D., Miller, E. J. and Kriger, D. S. (2005), “Current operational urban land-use-transport modelling frameworks: a review”, *Transport Reviews* , Vol. 25, pp. 325–350.

Jantz, C. A., Goetz, S. J. and Shelley, M. K. (2004), “Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land in the Baltimore/Washington metropolitan area”, *Environment & Planning B: Planning and Design* , Vol. 31, pp. 251–271.

Kelly, E. D. (1994), “The transportation land-use link”, *Journal of Planning Literature* , Vol. 9, pp. 128–145.

Kii, M. and Doi, K. (2005), “Multiagent land-use and transportation model for the policy evaluation of a compact city”, *Environment & Planning B: Planning and Design* , Vol. 32, pp. 485–504.

Krishnamurthy, S. and Kockelman, K. M. (2003), “Propagation of uncertainty in transportation land use models: Investigation of DRAM/EMPAL and UTPP predictions in Austin, Texas”, *Transportation Research Record* , Vol. 1831, pp. 219–229.

Kwan, M.-P. and Weber, J. (2003), “Individual accessibility revisited: implications for geographical analysis in the twenty-first century”, *Geographical Analysis* , Vol. 35, pp. 341–353.

Lee, D. B. (1973), “Requiem for large scale urban models”, *Journal of the American Institutue of Planners* , Vol. 39, pp. 163–178.

Lerman, S. R. (1976), “Location, housing, automobile ownership and mode to work: a joint choice model”, *Transportation Research Record* , Vol. 610, pp. 6–11.

Levinson, D. M. and Chen, W. (2005), “Paving new ground: a Markov Chain model of the change in transportation networks and land use”, *in* D. M. Levinson and K. J. Krizek, eds, ‘*Access to destinations*’, Elsevier Science Publishers, Amsterdam, The Netherlands, chapter 12, pp. 243–266.

Levinson, D. M. and Karamalaputi, R. (2003), “Induced supply: a model of highway network expansion at the microscopic level”, *Journal of Transport Economics and Policy* , Vol. 37, pp. 297–318.

Levinson, D. M. and Yerra, B. (2006), “Self-organization of surface transportation networks”, *Transportation Science* , Vol. 40, pp. 179–188.

Lowry, I. S. (1964), *A model of metropolis*, Rand Corporation, Santa Monica, CA.

Mackett, R. L. (1983), *The Leeds Integrated Land-use Transport model (LILT)*, Technical report, Transport and Road Research Laboratory, Crowthorne, UK.

Mackett, R. L. (1990), “Exploratory analysis of long-term travel demand and policy impacts using micro-analytical simulation”, *in* P. M. Jones, ed., ‘*Developments in dynamic and activity-based approaches to travel analysis*’, Oxford studies in transport, Avebury, Brookfield, VT, pp. 384–405.

Mackett, R. L. (1991), “LILT and MEPLAN: a comparative analylsis of land-use and transport policies for Leeds”, *Transport Reviews* , Vol. 11, pp. 131–154.

Martinez, F. J. (1992), “The bid-choice land use model: an integrated economic framework”, *Environment & Planning A* , Vol. 24, pp. 871–885.

Martinez, F. J. (1996), “MUSSA: A land-use model for Santiago city”, *Transportation Research Record* , Vol. 1552, pp. 126–134.

McFadden, D. L. (1978), “Modelling the choice of residential location”, *in* A. Karlqvist, ed., ‘*Spatial interaction theory and planning models*’, North-Holland, Amsterdam, The Netherlands, pp. 75–96.

McNally, M. G. (1998), “Activity-based forecasting models integrating GIS”, *Geographical Systems* , Vol. 5, pp. 163–184.

McNally, M. G. (2000), “The activity-based approach”, *in* D. A. Hensher and K. J. Button, eds, ‘*Handbook of transport modelling*’, Handbooks in Transport, v. 1, Pergamon, Amsterdam, The Netherlands, pp. 53–69.

Miller, E. J. (2003), “Microsimulation”, *in* K. G. Goulias, ed., ‘*Transportation systems planning: methods and applications*’, New directions in civil engineering, CRC Press, Boca Raton, FL, chapter 12, pp. 12.1–12.22.

Miller, H. J. (2005), “Place-based versus people-based accessibility”, *in* D. M. Levinson and K. J. Krizek, eds, ‘*Access to destinations*’, Elsevier Science Publishers, Amsterdam, chapter 4, pp. 63–89.

Moeckel, R., Spiekermann, K., Schürmann, C. and Wegener, M. (2003), “Microsimulation of land use, transport and environment”. Paper presented at the 8th International Conference on Computers in Urban Planning and Urban Management, May 27-29, Sendai, Japan.

Ottensmann, J. R. (2005), “Accessibility in the Luci2 urban simulation model and the importance of accessibility for urban development”, *in* D. M. Levinson and K. J. Krizek, eds, ‘*Access to Destinations*’, Elsevier Science Publishers, Amsterdam, The Netherlands, chapter 14, pp. 297–324.

Pas, E. I. (1990), “Is travel analysis and modelling in the doldrums?”, *in* P. M. Jones, ed., ‘*Developments in dynamic and activity-based approaches to travel analysis*’, Oxford studies in transport, Avebury, Brookfield, VT, chapter 1, pp. 3–33.

Pendyala, R. M., Kitamura, R., Chen, C. and Pas, E. I. (1997), “An activity-based microsimulation analysis of transportation control measures”, *Transport Policy* , Vol. 4, pp. 183–192.

Pradhan, A. and Kockelman, K. M. (2002), “Uncertainty propagation in an integrated land use-transportation modeling framework: output variation via UrbanSim”, *Transportation Research Record* , Vol. 1805, pp. 128–135.

Putman, S. H. (1974), “Preliminary results from an integrated transportation and land use models package”, *Transportation* , Vol. 3, pp. 193–224.

Putman, S. H. (1983), *Integrated urban models: policy analysis of transportation and land use*, Pion, London.

Putman, S. H. (2001), "The METROPILUS planning support system: urban models and GIS", *in* R. K. Brail and R. E. Klosterman, eds, '*Planning support systems: integrating geographic information systems, models and visualization tools*', ESRI Press, Redlands, CA, pp. 99–128.

Recker, W. W., McNally, M. G. and Root, G. S. (1986*a*), "A model of complex travel behavior (I): theoretical development", *Transportation Research, Part A: General* , Vol. 20A, pp. 307–318.

Recker, W. W., McNally, M. G. and Root, G. S. (1986*b*), "A model of complex travel behavior (II): an operational model", *Transportation Research, Part A: General* , Vol. 20A, pp. 319–330.

Rohr, C. and Williams, I. (1994), "Modelling the regional economic impacts of the Channel Tunnel", *Envrionment & Planning B: Planning and Design* , Vol. 21, pp. 555–568.

Roorda, M. J., Doherty, S. T. and Miller, E. J. (2005), "Operationalising household activity scheduling models: addressing assumptions and the use of new sources of behavioural data", *in* M. E. Lee-Gosselin and S. T. Doherty, eds, '*Integrated land-use and transportation models: behavioural foundations*', Elsevier, Amsterdam, The Netherlands, pp. 61–85.

Roorda, M. J. and Miller, E. J. (2006), "Assessing transportation policy using an activity-based microsimulation model of travel demand", *ITE Journal* , Vol. 76, pp. 16–21.

Ryan, S. (1999), "Property values and transportation facilities: finding the transportation-land use connection", *Journal of Planning Literature* , Vol. 13, pp. 412–427.

Salvani, P. and Miller, E. J. (2005), "ILUTE: an operational prototype of a comprehensive microsimulation model of urban systems", *Networks and Spatial Economics* , Vol. 5, pp. 217–234.

Schelling, T. C. (1978), *Micromotives and macrobehavior*, Norton, New York.

Simmonds, D. C. (1999), "The design of the DELTA land-use modelling package", *Environment & Planning B: Planning and Design* , Vol. 26, pp. 665–684.

Strauch, D., Moeckel, R., Wegener, M., Gräfe, J., Mühlhans, H., Rindsfüser, G. and Beckmann, K.-J. (2003), "Linking transport and land use planning: the microscopic dynamic simulation model". Paper presented at the 7th International Conference on Geocomputation, Sept. 8-10, Southampton, UK.

Timmermans, H. J. (2003), "The saga of integrated land use-transport modeling: how many dreams before we wake up?". Paper presented at the 10th International Conference on Travel Behaviour Research, Aug. 10-15, Lucerne, Switzerland.

Torrens, P. M. and O'Sullivan, D. (2001), "Cellular automata and urban simulation: where do we go from here?", *Environment & Planning B: Planning and Design* , Vol. 28, pp. 163–168.

Veldhuisen, K. J., Timmermans, H. J. and Kapoen, L. L. (2000), "Ramblas: a regional planning model based on the microsimulation of daily travel patterns", *Environment & Planning A* , Vol. 32, pp. 427–443.

Veldhuisen, K. J., Timmermans, H. J. and Kapoen, L. L. (2005), "Simulating the effects of urban development on activity-travel patterns: an application of Ramblas to the Randstad North Wing", *Environment & Planning B: Planning and Design* , Vol. 32, pp. 567–580.

Vovsha, P., Bradley, M. A. and Bowman, J. L. (2005), "Activity-based travel forecasting models in the United States: progress since 1995 and prospects for the future", *in* H. J. Timmermans, ed., '*Progress in activity-based analysis*', Elsevier, Amsterdam, The Netherlands, pp. 389–414.

Waddell, P. A. (2000), "A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim", *Environment & Planning B: Planning and Design* , Vol. 27, pp. 247–263.

Waddell, P. A. (2002*a*), "Design of an integrated land use and activity-based travel model system for the Puget Sound region", *Transportation Research Record* , Vol. 1805, pp. 108–118.

Waddell, P. A. (2002*b*), "UrbanSim, modeling urban development for land use, transportation and environmental planning", *Journal of the American Planning Association* , Vol. 68, pp. 297–314.

Waddell, P. A. and Borning, A. (2004), "A case study in digital government: developing and applying UrbanSim, a system for simulating urban land use, transportation, and environmental impacts", *Social Science Computer Review* , Vol. 22, pp. 37–51.

Waddell, P. A., Borning, A., Noth, M., Freier, N., Becke, M. and Ulfarsson, G. (2003), "Microsimulation of urban development and location choices: design and implementation of UrbanSim", *Networks and Spatial Economics* , Vol. 3, pp. 43–67.

Wegener, M. (1982), "Modeling urban decline: a multi-level economic-demographic model of the Dortmund region", *International Regional Science Review* , Vol. 7, pp. 21–41.

Wegener, M. (1994), "Operational urban models: state of the art", *Journal of the American Planning Association* , Vol. 60, pp. 17–30.

Wegener, M. (2004), "Overview of land use transport models", *in* D. A. Hensher, K. J. Button, K. E. Haynes and P. R. Stopher, eds, '*Handbook of transport geography and spatial systems*', Handbooks in Transport, v. 3, Pergamon, Amsterdam, The Netherlands, pp. 127–146.

Williams, H. C. (1977), "On the formation of travel demand models and economic evaluation measures of user benefit", *Environment & Planning A* , Vol. 9, pp. 285–344.

Wilson, A. G. (1967), "A statistical theory of spatial distribution models", *Transportation Research* , Vol. 1, pp. 253–269.

Wilson, A. G. (1970), *Entropy in urban and regional modelling*, Pion, London.

Wilson, A. G. (1971), "A family of spatial interaction models, and associated developments", *Environment & Planning* , Vol. 3, pp. 1–32.

Yamins, D., Rasmussen, S. and Fogel, D. (2003), "Growing urban roads", *Networks and Spatial Economics* , Vol. 3, pp. 69–85.

Yeh, A. G. and Li, X. (2002), “A cellular automata model to simulate development density for urban planning”, *Environment & Planning B: Planning and Design* , Vol. 29, pp. 431–450.

Yerra, B. and Levinson, D. M. (2005), “The emergence of hierarchy in transportation networks”, *Annals of Regional Science* , Vol. 39, pp. 541–553.

Zhang, L. and Levinson, D. M. (2007), “The economics of transportation network growth”, *in* P. Coto and V. Inglada, eds, ‘*Essays on transportation economics*’, Springer Science, pp. 317–340.

Appendix A

A Review of Models of Land Use and Transportation Change

A-1 Introduction

Models are the basic tool of analysis for planners working in the fields of transportation and land use forecasting. Current practice in these fields generally accepts the notion of some type of reciprocal relationship between transportation and land use. For more than four decades now, urban researchers have sought to formalize this relationship using mathematical, statistical and logical methods, and to produce models capable of predicting changes to transportation and land use systems as the result of policy measures.

This paper reviews some of the important theoretical frameworks adopted by researchers to represent the complex relationship between transportation and land use. Each framework has guided the development of a number of different operational models, that is, models that have been applied using data from real-world metropolitan regions. Several of these models are described in some detail to illustrate how each modeling framework is used to represent the processes of urban change[1]. Before turning to the models however, some background is provided on the transportation-land use relationship and the chronological development of transportation and land use modeling.

The first two modeling frameworks to be discussed are those based on aggregate models of spatial interaction and econometric models. These two modeling frameworks provide the vast majority of current operational models that are used in planning practice. We might refer to these first two frameworks as "top-down" modeling frameworks, since they specify the interaction between transportation networks and location as a set of aggregate relationships based on the behavior of a representative individual, usually the mean calculated from a representative sample of the population. The third class of models to be introduced falls under the general category of microsimulation models. These models cover a number of different approaches to representing the dynamics of land use change and travel behavior, but generally share the common focus of attempting to disaggregate the population and to simulate changes from the "bottom up", redefining the nature of actors in the model. Models of activity-based travel are discussed here, along with multi-agent models and cell-based models, a special type of multi-agent model that offers an alternative mechanism for representing the dynamics of land use change. Some examples of prototype urban models that are being developed entirely within a microsimulation framework are described. The later sections of the paper review some of the common criticisms directed toward land use and transportation models and note how these criticisms have (or have not) been addressed in the most recent generation of models. Some outstanding issues are discussed and suggestions offered as to important future research directions. A concluding section follows with some general remarks on the state of transportation and land use modeling and its relationship to planning as a discipline.

A-2 The Transportation-Land Use Relationship

Transportation networks and the spatial patterns of land use they serve are assumed to mutually influence each other over time. Changes to transportation networks, such as the construction of a new link or expansion of an existing one, eventually influence the location of investment in land, which in turn influences the demand for travel to and from a particular location. This relationship

[1] Reviews that cover a large number of models, including some that have seen less commercial application, are provided in recent papers by Timmermans (2003) and Wegener (1994, 2004). Chang (2006) also provides a review of models based on mathematical programming formulations, which are not discussed here.

is sometimes referred to as the transportation-land use "link" or "cycle", emphasizing a feedback relationship (Kelly, 1994). The mediating factor in determining changes in the location of activities and the demand for travel is $accessibility$, which measures the situation of a location relative to other activities or opportunities (work, shopping, etc.) distributed in space. Changes in relative accessibility are measured indirectly when researchers attempt to identify the influence of new infrastructure, such as a highway link or transit station, on local land markets. In these cases, accessibility is usually approximated by some measure of access to the transportation network, such as travel time or distance (Ryan, 1999). Generally, the degree of land market impact is related to the impact of the new transportation link on regional accessibility, and so is roughly proportional to the increase in speeds (and reduction in travel time) permitted by the new link (Cervero, 1984).

In order to operationalize the transportation-land use relationship within models of transportation and land use, measures of accessibility are incorporated in determining the location of activities. It is typically assumed that households wish to locate in areas with higher accessibilities to opportunities such as employment or shopping, while firms are assumed to locate in areas with higher accessibility to labor markets, perhaps stratified by occupational type. In models where land and floor space markets are considered explicitly, these accessibility factors can be important determinants of price. Since most models of transportation and land use contain a land use component that is integrated with, or at least loosely coupled with, a travel demand model containing a network assignment component, congested network travel times can be fed into the calculation of accessibility, thus providing a measure of the impact of congestion on regional accessibility and activity location. In order to simulate these changes within models of metropolitan regions, the region is typically broken down into a set of small geographic zones, similar (or in many cases identical) to the set of zones used for regional travel forecasting. Accessibility is typically calculated from each zone to all other zones in the region via the regional transportation network. Changes to the travel network that alter zone-to-zone travel times thus impact the relative accessibility of a location.

A-3 Chronology of Model Development

The history of simulation models of transportation and land use is dated back to the late 1950s (Batty, 1979). While models of regional travel demand had been established as far back as the early 1950s and some early experiments with transportation and land use models were carried out in the following years, it wasn't until the early 1960s that the first operational land use simulation model was built. The *Model of Metropolis* developed by Lowry (1964) is widely considered to be the first operational simulation model of urban land use. Lowry's model was the first of a generation of models based on theories of spatial interaction, including the gravity model that was popular in quantitative geography at the time. Models based on a spatial interaction framework continued to be developed through the early to mid-1980s, when they became largely replaced by models grounded in random utility theory and econometric methods.

Figure 2-1 from the text describes this process and gives an approximate timeline for the adoption of various modeling frameworks within transportation and land use research. Several of the models that follow an econometric framework continue to be used today, although some, like the UrbanSim simulation system (Waddell, 2002*b*; Waddell et al., 2003) are being redeveloped within a microsimulation design. The broad class of transportation and land use models that could fall under the title of 'microsimulation' began to be developed in the early 1990s, in parallel with major improvements in computational power that allowed for their operation. These included prototype models of activity-based travel, cell-based models land use change and the introduction of multi-agent models for urban simulation. More recently, some researchers have begun to devote effort to developing comprehensive urban microsimulation models that fully reflect the dynamics of changes in the population and the urban environment within which they make choices.

A-4 Spatial Interaction Models

The earliest class of land use and transportation simulation models are a set of highly aggregate models based on principles of spatial interaction that were popular in the regional science and quantitative geography fields in the 1950s and 1960s. There were many different formulations of this type of model, though most revolved around variations of the gravity model, an adaptation from Newtonian physics. The derivation of the gravity model from principles of entropy maximization (Wilson, 1967, 1970) was a major accomplishment and formed the basis for many of the allocation mechanisms within spatial interaction models. A general form of the gravity model can be expressed as:

$$T_{ij} = A_i B_j O_i D_j \exp(-\beta c_{ij}) \tag{A.1}$$

where T_{ij} represents trips (or other measures of interaction) between two zones, O_i represents origins at zone i, D_j represents destinations to zone j, and A_i and B_j are balancing factors to ensure that total origins equal total destinations. The exponential term in the model is used to capture the effect of decreasing interaction as a function of travel cost, including travel time.

As mentioned previously, the first operational land use simulation model was the model developed by Lowry (1964) for the Pittsburgh region. This model has great importance, since many of the other land use and transportation models that follow a spatial interaction framework have similar structures. A detailed review of this model and its variations are provided in Horowitz (2004).

The Lowry Model and Derivatives

The land use model developed by Lowry was a spatial interaction model designed to simulate patterns of residential and service location in the Pittsburgh, Pennsylvania region. The impetus for building the model was to be able to simulate the effects of urban renewal and slum clearance programs on the distribution of activities within the region. The model borrowed from economic base theory, which divides a region's employment into basic and non-basic services. Basic industries are assumed to export much of their product outside the region, generating additional income which can then support additional non-basic services. Non-basic industries then serve households (e.g. retail activities) and other industries within the region.

Lowry's model assumed that the location of basic industries was fixed. This required an initial allocation of basic employment to zones within the region. Households were then allocated to zones $from$ the initial basic employment locations, using a function similar to the deterrence function used in the trip distribution step of most trip-based travel forecasting models (Horowitz, 2004):

$$f\left(t_{ij}\right) = \exp\left(-\beta t_{ij}\right) \tag{A.2}$$

where $f\left(t_{ij}\right)$ is a deterrence function value representing the inverse of the likelihood of workers working in zone i and living in zone j, and t_{ij} is a measure of the disutility of travel between zones, typically defined as travel time, and $-\beta$ represents the marginal disutility per unit of time. This functional form implicitly assumes that workers choose to locate near their workplace and that only one household member is employed outside the home. Lowry chose to define this measure of disutility as the airline distance between zones. He did this partially because of the difficulty of generating matrices of trips between zones using the travel models that existed at the time, but also because he noted a high degree of correlation between observed airline and network distances in his study region (Lowry, 1964). Using the deterrence function described above, $f\left(t_{ij}\right)$, the number of workers working in zone i and living in zone j (defined here as T_{ij}) could be calculated by using a modified expression that included a value of attractiveness for each residential zone (w_j):

$$T_{ij} = \frac{e_i w_j f\left(t_{ij}\right)}{\sum\limits_{j} w_j f\left(t_{ij}\right)} \tag{A.3}$$

where e_i is the employment in zone i. The residential attractiveness measure as used in this formulation simply relates to the amount of land available for residential development in a particular zone. Deleting the variable for zonal employment in the above expression yields an expression for the probability of residing in a zone given a fixed workplace location that is very similar to the probability expression in the multinomial logit model. This relationship is important, since it is used extensively in transportation and land use models that derive from random utility theory, as will be discussed in the next section.

The process of worker/household allocation is followed by a similar process in which the locations of non-basic industries serving households and other (basic) industries are allocated assuming fixed locations for these quantities. Once these activities have been allocated, it is possible to couple the land use model with a conventional, trip-based travel forecasting model to produce a set of network flows. These new flows and travel times can be used to modify the deterrence function and produce a new allocation of households and non-basic employment.

Several models extended the basic Lowry framework in new directions. Table 2-1 lists some of these models along with their distinguishing features. For example, the Time Oriented Metropolitan Model (TOMM) described by Crecine (1964) disaggregated the population into socio-economic groups in order to improve the model's representation. It also differed from the Lowry model in that only some of the non-basic activities in a region would be reallocated between model iterations, reflecting a certain degree of inertia in location. Garin (1966) recast the original Lowry model by proposing a matrix representation for the model's components and substituting a production-constrained, gravity-type interaction model as the basis for allocation. Garin's version also allocated all activities at each iteration, an improvement over Lowry's formulation since it improved the coupling between allocation and generation (Timmermans, 2003). Another land use model designed by Goldner (1971) allocated activities according to an intervening opportunity model, a special case of the gravity model (Wilson, 1971). The design of the model also sought to improve realism by using different dispersion parameters for each of the nine counties of the San Francisco Bay area, where it was calibrated and tested.

ITLUP/METROPILUS

Building on the Lowry-Garin framework, Putman (1974, 1983) developed the Integrated Transportation and Land Use Package (ITLUP), widely considered to be the first fully operational transportation-land use modeling software package. ITLUP has been applied in over a dozen locations within the U.S., and has been calibrated over 40 times (Hunt et al., 2005). Designed in the mold of the Lowry model, ITLUP initially contained a land use model that was similar to Goldner's PLUM model. ITLUP offered a network representation that allowed for the incorporation of congested travel times in the distribution of activities. At the core of ITLUP were two allocation submodels: a household allocation submodel called DRAM, and an employment allocation submodel, EMPAL. Trip generation and distribution functions for the travel forecasting model are developed within DRAM, simultaneously with household location, while mode choice and trip assignment are handled with separate submodels. Travel times from runs of the travel model are fed forward to calculate new activity distributions.

More recently, the ITLUP model framework has been updated to incorporate modifications to some of its submodels and new data and visualization tools (Putman, 2001). The new package, called METROPILUS, is housed within a geographic information system (GIS) environment that permits improved visualization of output. Other important features of METROPILUS include multivariate, multiparametric attractiveness functions that include lag terms to better capture location dynamics. The addition of zonal constraints can limit allocation of activities to zones where land is not available. Land supply in the model is managed by a land supply function that translates the location demands from employers and households from DRAM and EMPAL into land uses and intensities.

LILT and IRPUD

Two other spatial interaction-based models merit attention, since they have been extensively applied and tested. The first is the Leeds Integrated Land Use (LILT) model, developed by Mackett (1983, 1991). LILT combines a Lowry-type land use model with a conventional, four-step travel model. Forecasts of change in population are allocated to zones according to accessibility functions derived from work trips and zonal attractiveness functions. Other salient features of LILT include the ability to handle demolition, changing occupancy rates and vacancies, and a car ownership submodel, which estimates vehicle ownership as a function of network travel times and costs (Timmermans, 2003).

The IRPUD model (Wegener, 1982) was developed by Wegener and colleagues at the University of Dortmund in Germany. IRPUD is quite complex and contains seven interlinked submodels of aging, firm relocation, residential and non-residential construction, rehabilitation and demolition, change of job, change of residence and car ownership/travel demand. IRPUD is somewhat unusual in that it contains a microsimulation model of land use, in which land uses are allowed to change through aging. Another desirable feature of IRPUD's design is that it allows different submodels to take place at different spatial scales (intra-regional location takes places at a meso-scopic scale, while land development takes places at a micro/tract level). These features are emulated in some of the newer, emerging urban microsimulation models.

The first generation of land use and integrated transportation and land use models based on spatial interaction formulations produced a multitude of models that were tested and applied in numerous settings. Some models, such as the METROPILUS planning support system package, continue the legacy of these models to the present. However, very few examples of this type of model framework remain. The shortcomings of these models were numerous: most were static equilibrium models incapable of capturing the dynamics of urban systems, none of models actually represented land markets with explicit prices, zones were highly aggregate and lacked spatial detail, and the models were inadequately supported by theory. Inadequate theory may have also been a reason that many of the models forecasted so poorly. There were many high-profile failures in terms of using the models for policy analysis purposes (Batty, 1979). Some of these were seized upon by Lee (1973) in a critique which highlighted some of the mistakes of the first generation of models. Lee characterized them as being too complicated, overly aggregate, data hungry, wrong-headed, extraordinarily complicated, too mechanical and expensive. Many of these criticisms informed the next generation of models, which took their cue from developments in econometric modeling based on random utility theory.

A-5 Econometric Approaches

As noted previously, one of the major shortcomings of the aggregate spatial interaction models was the absence or use of inappropriate theory to describe the behavior captured in the model. Developments in the use of random utility theory to describe choices among discrete alternatives, such as the choice of travel mode, provided the impetus for a new generation of models based on the study of disaggregate behavior. When it was shown that discrete choice models could be applied to problems such as residential location (Lerman, 1976; McFadden, 1978), researchers began to look for ways to model the interrelated choices individuals made in terms of location and travel behavior.

Land use and transportation models that follow econometric frameworks can be thought of as comprising two types of models: regional economic models and land market models. In these two types of simulation models the economic model and the land market model each form the core of a simulation system that includes the prediction of transportation flows. Both types tend to have improved representation of land markets that include endogenously-determined (determined within the model) prices and market clearing mechanisms. A summary of these models and their characteristics are provided in Table 2-2.

Regional Economic Models

Two of the most important and widely-used transportation and land use models grounded in econometric modeling approaches, MEPLAN and TRANUS, are largely built around a core of a regional economic model. MEPLAN (Echenique, 2004; Echenique et al., 1990) is a model that began as a more simple model of urban stock and activity (Echenique et al., 1969) and expanded into a more comprehensive urban simulation model. Similar to other types of models, MEPLAN has a zone-based structure. In contrast to spatial interaction models though, the activities in zones are determined by a spatial input-output model which predicts trade flows by sector between zones of a region, driving the demand for space. Production and consumption are linked in the spatial input-output model, replacing the trip generation and distribution steps in trip-based travel forecasting models. The trade flows are converted to demand for commercial and passenger traffic through the application of scaling constants. The generated traffic is then fed into models of mode and route choice. Congestion and travel times from the transportation model are then fed back into the land use and economic model, yielding time-lagged measures of accessibility, which affect location choice. The structure of MEPLAN, including its spatial economic model, makes it appropriate for modeling not only at an $intraurban$ scale but also at an $interurban$ scale. It has been used in a variety of major applications, including modeling the regional impacts of the Channel Tunnel between England and France (Rohr and Williams, 1994).

The TRANUS model (Barra, 1989) is similar to MEPLAN in that it incorporates a spatial input-output model as the basis of its generation and allocation of activities. The regional economy is disaggregated into sectors, with the demand for each zone and sector generated and then allocated to production zones and sectors via a multinomial logit model. A land supply model is also available to simulate the behavior of developers, who choose where to build (new land vs. existing sites), what type of space to build, and at what density. This choice process is governed by explicit prices or rents for new or replacement stock, demolition and building costs. Another unique feature of TRANUS is its relatively advanced trip-based travel forecasting model. Similar to MEPLAN, flows of traffic between zones are generated from input-output matrices. Personal travel is estimated by time of day by mode as a function of cost. Trips are assigned to the network according to distinct mode-path combinations. Accessibility is calculated as a logsum composite utility measure from the mode choice model and input directly to the land use model to generate a new set of spatial flows.

A third model system that takes as its centerpiece a regional economic model is the PECAS system (Production, Exchange, and Consumption Allocation System), developed by Hunt and Abraham (2005). PECAS is a generalization of the spatial input-output modeling approach used in MEPLAN and TRANUS. The model system is based on a quasi-dynamic equilibrium structure with flows of exchanges, including goods, services and labor, from production to consumption based on technical coefficients. Flows of exchanges from production to zones of exchange and from exchange zones to consumption are based on nested logit models that take into account exchange prices and transport disutilties. Similar to other spatial input-output models, trade flows are converted to transport demands and loaded onto networks in order to calculate congested travel times (disutilities). Exchange prices for space drive changes in available space, simulating developer actions. The model system is run in one-year time steps, with travel disutilities and changes in space in a given year influencing the flows of exchanges in the next year (Hunt and Abraham, 2005).

PECAS now features activity-based travel modules, as well as microsimulations of land development, with land parcels as the unit of analysis. While PECAS is run at the scale of a metropolitan region, it can, like other input-output models, be adapted to larger-scale applications. Recent versions of the model system have been applied in statewide models of land use and transportation for Ohio and Oregon, as well as metropolitan-level applications in Sacramento, CA and Calgary and Edmonton, Alberta, Canada.

Land Market Models

Improved land market representation is a distinguishing characteristic of many of the econometric approaches to transportation and land use models. In fact, several them have at their core markets for residential and commercial real estate, with transportation models linked into the overall model structure. Some of these models, such as those developed by Anas (1982, 1984), seized on theoretical advances in linking the related strands of gravity-based models with those based on the multinomial logit specification (Anas, 1983; Williams, 1977).

Anas and colleages developed a series models (Anas, 1982, 1998; Anas and Arnott, 1994) designed to simulate the effects of transportation improvements on land markets and overall social welfare. The first such model, CATLAS, emphasizes a discrete choice framework to describe both the supply and demand sides of the housing market. The supply side of the model contains vacancy-occupancy, construction and demolition submodels that respond to factors such as construction costs, land prices, taxes and operating costs, and expected future resale values. Developers are assumed to be profit maximizers, and so select the location and type of construction to maximize profit. The demand side of the model takes a nested logit choice model form, assuming that households have a fixed workplace location and choose a residential location and travel mode to maximize their utility. Only two workplace locations are considered in the model (CBD and non-CBD), though commuters have a variety of modes available (auto, bus, heavy rail and commuter rail), depending on their residential location. The model is calibrated with Census data and can predict changes in mode splits, house prices and rents, demolitions, and new construction activity (Anas, 1987). The economic evaluation component of the model estimates changes in economic welfare due to changes in modal utility arising from investment in different modes. The changes in utility are captured in an inclusive value (logsum) accessibility measure and are capitalized into housing prices or rents.

The original CATLAS framework was modified in an enhanced model called METROSIM (Anas and Arnott, 1994), designed for the New York City metropolitan region. METROSIM incorporated a dynamic model of metropolitan housing markets (Anas and Arnott, 1994), along with a model of commercial floor space markets. The full modeling system combined models of employment, residential and commercial real estate, vacant land, households, work and non-work travel and traffic assignment, which was absent in the CATLAS system. A recent extension of this system is the NYMTC-LUM model (Anas, 1998), a simplification of METROSIM designed to facilitate the evaluation of changes in transit policies for the New York City transit system. The model is slightly refined, adding a local labor market submodel and using very small zones to better model transit and auto network flows. The combined model determines housing prices and floor space rents endogenously (within the model), and uses modal utilities from the mode choice model as accessibility inputs to the land use model.

A similar framework was adopted by Simmonds (1999) in developing DELTA, a land use model designed to form the basis of a dynamic model system of land use and transportation interaction. The model system is divided into processes which represent spaces and those which represent activities. Processes dealing with activities include household formation and dissolution, employment growth or decline, location and property markets and the employment status of individuals. Processes representing the change in spaces predict the quantity and quality of floor space available. The model system is designed to be run over a series of short steps of no more than one or two years, and was originally coupled with START, a transportation model developed for the city of Leeds, UK . A distinguishing feature of DELTA is that attempts to add a quality variable to the prediction of location choices. In the case of residential location, the quality variable relates to local income and vacancy rates. Hence, the quality of development can change over time. The DELTA model has seen several applications in the UK and parts of Western Europe and is currently being developed as a microsimulation model system.

An alternative framework for modeling land markets in transportation and land use models was provided by Martinez (1992, 1996), who built an integrated model called MUSSA for the city of Santiago, Chile. MUSSA adopted a modified version of the ”bid-rent” framework for land markets, first articulated by Ellickson (1981). The ”bid-choice” framework used by Martinez combines bid-rent and discrete choice approaches to land markets by dealing simultaneously with both sides of an auction in a bi-level framework. The MUSSA system provides an equilibrium model of building stock supply and demand, where buyers maximize their surplus, sellers maximize price, and builders maximize profits. Building stock prices are then endogenously determined in the model.

The MUSSA system also includes a rather sophisticated four-step travel forecasting model that is linked to the land use component. The travel model features a detailed transit network representation and the ability to forecast demand for 11 separate alternatives, including road, transit and mixed modes. The combined transportation and land use models are referred to as 5-LUT (indicating a 5-step forecasting procedure), and are able to provide equilibrated forecasts of land use and travel demand. A notable feature of MUSSA is that the model uses smaller-than-average zones as units of analysis in order to achieve a higher level of spatial disaggregation. Likewise, there is an effort to disaggregate the treatment of households within the model, with the Santiago application containing 65 different household types. This is an important step in the development of transportation and land use models, and one that is being replicated in the current generation of transportation and land use models based on microsimulation techniques, as will be discussed in the following section.

Another transportation and land use simulation model that adopts this highly disaggregate structure is the UrbanSim model developed by Waddell and colleagues (Waddell, 2000, 2002*b*). Like MUSSA, UrbanSim is primarily a model of land markets, though extensions have been considered to add an activity-based travel forecasting model (Waddell, 2002*a*), as well as an environmental analysis module (Waddell and Borning, 2004). Like MUSSA, UrbanSim initially contained a highly disaggregated household treatment, with 111 distinct household types identified in an early calibration of the model (Waddell, 2000). Demographic transition in population and household formation are microsimulated within a separate submodel. Residential mobility of households is characterized by a two-stage process in which households decide whether to search and then whether to move. Location choice of households and firms are represented by a multinomial logit model considering all zones in the region within the choice set. While UrbanSim makes extensive use of econometric models in its structure, predictions are based on Monte Carlo simulation methods, indicating that it also has the characteristics of a microsimulation model system.

UrbanSim's structure is also unusual in that it operates in $disequilibrium$ from year to year, with no general equilibrium in land markets assumed at the end of a time step, though market clearing does occur at the transportation analysis zone (TAZ) level. This feature sets it apart from all of the preceding models that incorporate land markets, which are typically static within each time step of a simulation. Researchers in the field of urban modeling have previously commented on the importance of modeling different elements of urban systems at the time scales in which they operate (Miller, 2003; Wegener, 1994). Since urban areas do not really ever reach a general equilibrium in land and travel markets, this disequilibrium structure will likely be adopted in many future attempts to model land markets.

UrbanSim's model of land markets also estimates supply at the parcel level, using parcel databases within a GIS. Demand for housing and floor space are calculated at the TAZ level in the original version of the model, though subsequent versions are attempting to reconcile the spatial scale of the supply-demand relationship. Land markets are simulated using the bid-choice framework, similar to the MUSSA model (Waddell, 2000). Land prices are estimated from hedonic regressions containing building unit and neighborhood characteristics, and regional accessibility to work and shopping. The neighborhood characteristics are determined by partitioning the region into 150 by 150 meter grid cells, each containing information about neighborhood composition and nearby land uses.

Further work on UrbanSim is focusing on converting it to a comprehensive microsimulation modeling system (Waddell et al., 2003). Many of the elements of the original model lent themselves to this treatment, including the high level of household type disaggregation and demographic transition submodel. The land market simulation is already highly disaggregated and requires only further refinement of developer behavior. The structure of the model system suggests that modified transportation sub-models, such as an activity-based travel model, could be coupled with the other elements in the model system. Long-term goals of the project include developing the software architecture to support an agent-based simulation version of the modeling system and the exploration of new model structures.

The experience with the generation of transportation and land use models based on econometric frameworks has been valuable and addressed one of the most pointed criticisms of the previous generation of spatial interaction models, that of lack of theory. The use of random utility theory and advancements in discrete choice modeling of individual behavior have allowed for the inclusion of economic evaluation components in several of the models, as well as improved accessibility measures based on utility functions. Also, the introduction of model systems built around a regional economic model allowed for the inclusion of commercial travel in forecasts and the general treatment of travel as a derived demand. Despite these advancements, many of the econometric models retained a number of problems left over from the previous generation of models. For example, most of the models remained highly aggregate, despite the use of disaggregate calibration methods. This became one source of bias in the model forecasts. Also, with the exception of UrbanSim, all of the models were essentially static in nature. Their structure forced them to reach a general equilibrium between each time step in the model; this was especially true of the models focusing on land markets. Furthermore, little advancement was made in the transportation component of the model. Most models continued to use trip-based, four-step forecasting procedures, where all submodels except mode choice were run at an aggregate level. Much of the current research into microsimulation methods is attempting to address this issue, along with other pressing research questions in the design of comprehensive simulation models of transportation and land use.

A-6 Disaggregate and Microsimulation Models

Since the late 1980s, advances in computing power and efficiency of data storage have allowed researchers to begin to build models that address many of the shortcomings associated with previous large-scale modeling efforts and represent important change processes in cities with the detail they require. Examples of these include activity-based models of travel behavior, multi-agent models of urban land use and transportation, and cell-based models of urban land use. The common conceptual underpinning of each of these models is that they attempt to represent processes of change from the bottom up, that is, they account for the behavior of individual agents in space and/or time, along with interactions between agents. The use of the term *microsimulation* can be applied to each of these types of models, though it requires some definition. As defined by Miller (2003), microsimulation relates to "a method or approach (rather than a model *per se*) for exercising a disaggregate model over time." All of the types of models identified above are what would be considered disaggregate models and all have a significant temporal element. Microsimulation methods are particularly effective for modeling systems that are dynamic and complex, which urban systems invariably are.

Activity-Based Travel Models

The literature on activity-based approaches to travel analysis is quite extensive and dates to the 1970s. Thus, a comprehensive review of this literature is not possible here. Instead, the focus will be on covering a few of the models that have been tested using real-world data at least once. The interested reader is directed to papers by Axhausen and Gärling (1992), Ettema and Timmermans (1997), McNally (2000), Vovsha et al. (2005), and the collection of papers in the August 1996 issue of the journal *Transportation*, which describes the early results of research work funded through TMIP.

Research into the foundations of travel behavior dating back to the 1970s has identified many shortcomings in the use of sequential, trip-based travel demand forecasting models (Chapin, 1974; Hägerstrand, 1970). However, there was little incentive until this time to attempt to recast travel forecasting procedures. Oil crises during the 1970s precipitated research into various energy use reduction strategies, including demand management measures and transportation system management techniques. It was then that the inability of existing forecasting models, which were mostly static and aggregate, to predict behavioral responses to such policy measures became apparent (McNally, 2000).

A combination of factors brought about a resurgence in interest in reconceptualizing travel behavior for modeling purposes during the 1990s. The completion of the interstate system and the difficulty of expanding existing urban road networks led many regional planning organizations to emphasize preservation and management of transportation systems through such policies as flexible working hours, travel information provision, traffic flow improvements and diversion of some travel to alternate modes. The potential changes in travel behavior implied by these types policies cannot be forecast using existing methods, since trip-based models separate travel decisions from their broader context of activity participation and temporal constraints. At the same time, improvements in computing power and the use of geographic information systems have allowed for the formalization and testing of models that previously only existed at conceptual or limited empirical levels. Support from the Federal Highway Administration in the form of the Travel Model Improvement Program (TMIP), which attempted to improve the state of practice in transportation modeling and facilitate development of a new generation of travel demand models, has also had a significant impact.

The first demonstration of an operational model of activity-based travel preceded the TMIP, and was conducted by Recker et al. (1986*a*,*b*). The STARCHILD model was developed to investigate dynamic ridesharing, but was designed for research purposes only and required collection of data that is still not commonly available (McNally, 2000). Models of activity chains and travel behavior were coupled with a mesoscopic traffic simulation in work by Axhausen (1990). Pendyala et al. (1997) developed an activity-based simulation model capable of predicting activity scheduling changes in response to transportation control measures. They demonstrate their model with an application to evaluate the impacts of control measures in the Washington, D.C. metropolitan region. Activity-based forecasting models incorporating GIS applications have also been developed by McNally (1998). Bowman and Ben-Akiva (2001) structured a model of activity participation within a nested logit framework to predict travel tours (clusters of chained trips). Their model was calibrated using travel survey data from the Boston region. A model system developed by Arentze and Timmermans (2004) attempts to simulate learning behavior by agents within the context of activity scheduling and travel behavior. Perhaps the most ambitious effort to date in the U.S. has been the research program associated with the TRANSIMS modeling system, which is designed to combine an activity-based forecasting model with a region-wide traffic microsimulation system (Barrett, 1995).

Activity-based models are necessarily disaggregate and attempt to simulate travel behavior within the limits of time and space. Due to spatial and temporal interdependencies, this process cannot be modeled within a framework that treats trips as independent and generates trips at an aggregate level. An alternative, agent-based approach is typically adopted in formal travel forecasting applications. This focus on the behavior of individual agents and addition of temporal elements makes activity-based travel models a natural complement to microsimulation models of transportation and land use that focus on the activity of agents at an individual or household level.

Agent-Based Microsimulation Models

The state-of-the art in transportation and land use modeling is defined by current research efforts aimed at building comprehensive microsimulation systems of urban areas, with representation at the level of individual agents (persons, households, firms, etc.) and simulations of the behavior of the entire population of interest. The advantages of adopting such a modeling approach for urban systems are many (Miller, 2003):

- Urban systems are dynamic, with a significant time element and components changing at different temporal scales

- The behavior of these systems is complex, with interacting agents, complex decision-making processes, and significant probabilistic elements

- Closed-form mathematical and statistical representations of urban systems often introduce large amounts of bias and lead to poor forecasts

The seeds of comprehensive microsimulation models had been sown in a number of earlier models, where one or more elements of the system were governed by a microsimluation process. For example, Wegener's IRPUD model contained microsimulations of population and building stock. Mackett (1990)'s MASTER model simulated location choices and travel decisions, and MUSSA and UrbanSim disaggregated households at a level sufficient to operate them in a static microsimulation format, where a representative sample is used within a microanalytic framework for short-run applications. However, for long-term forecasts, which most transportation and land use models are designed for, the population must be synthesized or updated to represent the dynamics of individuals and the environments within which they make choices.

An overview of some of the comprehensive microsimulation systems currently under development are presented in Table 2-3. The UrbanSim system was the only simulation model to transition from a static simulation format to a dynamic microsimuation model. As noted previously, the original version of UrbanSim contained a number of microsimulation submodels within its structure, thus eliminating the need for as radical a redesign as would be needed for many of the static, equilibrium models.

The ILUMASS simulation system (Moeckel et al., 2003; Strauch et al., 2003), being developed by a research team at the University of Dortmund, builds on the experience of Wegener and others with the IRPUD model in the 1980s. The design of ILUMASS embeds a microscopic dynamic simulation model of urban traffic flows within a comprehensive model system incorporating changes in land use and building stock.

The microsimulation modules of ILUMASS include models of demographic change, household formation, firm lifecycles, residential and non-residential construction, labor mobility in a regional labor market, and residential mobility in a regional housing market. These modules are linked with models of daily activity participation and travel, as well as goods movement. The activity-travel module uses data collected via a hand-held survey instrument. This innovation in data collection allows for near-real-time information on activity and travel behavior, obviating the need for respondents to recall their activities later on. The GIS component of ILUMASS combines raster-based and vector-based representations, allowing for the advantages spatial disaggregation in land use representation and efficient network algorithms for the transportation network model.

The ILUTE model (Salvani and Miller, 2005), being developed by researchers at a number of Canadian universities, chiefly the University of Toronto, represents the most complete microsimulation model to date. The product of a long-term effort to design an 'ideal' simulation model of transportation and land use, ILUTE centers around a behavioral core consisting of four interrelated components: land use, location choice, auto ownership and activity/travel patterns. The model system is highly integrated with feedback mechanisms whereby higher-level (longer-term) decisions, such as residential mobility, affect lower-level (shorter-term) decisions, such as activity participation and travel. ILUTE is not based on a single modeling technique (e.g. random utility), but rather uses a variety of modeling approaches to represent the behavior of agents in the model, such as state transition models, random utility models, computational rule-based models, learning models, and hybrids of previous approaches.

ILUTE's treatment of land markets explicitly assumes a constant disequilibrium framework, indicating that a particular house could be on the market for several months without selling, since no market clearing is assumed. The time steps in the model are brought down to the level of $months$, rather than years, to provide greater temporal detail. The disequilibrium framework and absence of market clearing also means that projects with extended construction periods (e.g. greater than one year) can be accommodated. The housing market submodel within ILUTE assumes a three-step process to describe residential mobility, involving a mobility decision, a search process, and bidding and search termination.

The transportation component of ILUTE is quite sophisticated and includes submodels for automobile transactions and activity scheduling. The activity scheduling submodels characterizes activities as occurring in time and space, with various scheduling dependencies to represent temporal constraints (Roorda et al., 2005). Future plans include adding a network model, which is needed to provide travel times and costs by mode, along with a formal model of activity participation. Like most comprehensive microsimulation models, ILUTE is still in the process of calibrating some of the submodels in the system, and has yet to be used in a full forecasting application, though the travel demand component has been applied in a policy simulation (Roorda and Miller, 2006).

Another agent-based simulation model that merits attention is the Ramblas model (Veldhuisen et al., 2000, 2005). While it is not as comprehensive as the other models described here, Ramblas is designed to simulate the effects of land use and transportation planning policies, with an emphasis on the prediction of activity participation and traffic flows. An unusual aspect of Ramblas is that it is designed to simulate the effects of policies on the entire Dutch population (estimation at over 16 million). The model also distinguishes itself by being entirely rule-based, rather than adopting a formal theoretical framework to guide the behavior of agents. These aspects of the model derive from its stated purpose of being a practical planning tool to assess the impact of various transportation and land use scenarios.

Ramblas is run by selecting households, stratified according to size and structure. Individuals are classified according to one of 24 population segments, defined on the basis of age, gender, education and employment status. An activity agenda and transportation mode are drawn at random, with seven activity types available. Destinations are randomly drawn from a choice set, sometimes delimited by a given action space or distance constraint. Origin-destination pairs are generated from the activity and mode allocations and traffic flows are then microsimulated, calculating travel times via a speed-flow method. Output from the microsimulation of traffic is used to forecast changes in land use, dwelling stock and road construction.

Cellular Models

The representation of land use in integrated models of transportation and land use change has been one of the less satisfactory elements of these models (Chang, 2006). Until recently, land use had generally been represented by zones that served as convenient areal units for the location of activities, and coincided with zonal designations for transportation models. Models that provide greater simplicity and a clearer representation of the dynamics of land use change using cell-based representations of regions have emerged within the past two decades as an increasingly attractive land use modeling alternative.

Cell-based models, and particularly those based on cellular automata (CA) theory, arise from the application of *complexity theory* to cities (Batty, 1997, 2005). Complexity theory conceptualizes systems, such as urban systems, as being too complex to synthesize using closed-form, predetermined mathematical representation. Rather, these systems arise from the collective interaction and self-organization of large numbers of individual agents which generate the observed macro-level states (Benenson, 1998). Cell-based models of land use can range from simple state transition models in which cells change states (land uses) according to some observed probability, to the more general form of CA, in which cell states are also a function of states in neighboring cells. CA models can be seen as extension of agent-based microsimulation models, in which individual cells are the agents, rather than persons or households.

CA models generally require four basic elements: a lattice of regular spaces or cells, a set of allowed states, neighborhoods that are defined by the lattice, and a set of transition rules governing the evolution of individual cells in the system. Many CA models also add a fifth, temporal element. CA models are basically deterministic, rule-based models, using ”if-then-else” logical statements to build their transition rules, though stochastic elements can be added to transition rules using probabilistic expressions and random number generation. Other types of modifications to CA models intended to introduce complexity include changes to the structure and dimension of the lattice of cells, expansion of allowable cell states, expanded neighborhood definitions to include action at a distance, and changes to temporal elements, such as Markov chains Torrens and O'Sullivan (2001).

A precursor to many of the contemporary cellular models being used to describe the dynamics of urban systems is the model of self-forming neighborhoods presented by Schelling (1978). As part of a larger exposition of self-organizing principles, Schelling demonstrated how "individually motivated" forms of segregation could arise through the interaction of many agents (households) pursuing their own objectives. Preferences for individuals of a different race, income, or any other form of social stratification were shown to lead to highly segregated outcomes under a variety of initial conditions and preference structures.

The compatibility of CA models with GIS, remote sensing data and associated visualization capabilities make them particularly suitable for land use modeling applications (Torrens and O'Sullivan, 2001). It is here that they have received the most attention. One example is the model of urban land use developed by Clarke et al. (1997) to estimate the regional impact of urbanization on the San Francisco Bay Area's climate. This model is an example of a self-modifying CA, in which the CA can adapt to the circumstances it generates. Clarke and Gaydos (1998) applied the same model to the Baltimore-Washington region to generate long-term urban growth predictions. Jantz et al. (2004) also studied growth in the Baltimore-Washington region using CA, with the objective of simulating the effects of different patterns of land use on the Chesapeake Bay watershed. Levinson and Chen (2005) describe the development of a Markov Chain model of land use change for the Minneapolis-St. Paul region. Their model adopts the discrete-time version of a Markov Chain and predicts the evolution of transportation networks and land use patterns over the period from 1958 to 1990. A next step for this model would be to add neighbor effects, which would move it to a CA-Markov Chain framework.

Other applications of CA include simulating land use density conditions, as in the model developed by Yeh and Li (2002). Their model incorporates a density gradient in the simulation of urban development for different urban forms. The transition rules of their model specify a density, obtained from a distance-decay function, to be applied to cells as they are converted to developed cells. Kii and Doi (2005) provide a similar application to demonstrate the effects of compact city form and mixed land use on total trip length, energy consumption and social welfare in Takamatsu, Japan. The model they present, MALUT, is a multi-agent model of transportation and land use, where a CA model of land use is coupled with a microsimulation model of travel. Accessibility can be incorporated into a CA model of land use change, as demonstrated by Ottensmann (2005)'s LUCI2 model. LUCI2 was designed to predict employment and land conversion change over a 44-county region of Central Indiana, consisting of eight separate metropolitan statistical areas. The model found access to employment to be an important determinant of residential development and density.

CA models appear to be growing more complex. Their many applications reflect the relative ease and flexibility with which they can be modified to describe processes of change. CA models are not without their weaknesses, though. Their simplicity, which is one of their most desirable attributes, is also a significant limitation. In most cases, they are inappropriate for modeling systems with complex interactions. For example, processes like land development represent the interaction between human and physical systems, but CA models cannot capture both. Also, CA models are not designed to be forecasting tools. Since they are calibrated on historical data and lack a strong behavioral interpretation, most forecasts have little meaning. Rather, CA are better suited to idealized principles of cities and urban design applications than large-scale simulations or strategic planning (Batty, 1997).

A-7 Resolved and Ongoing Modeling Issues

Old Issues

The models currently being developed to describe change in transportation and land use systems look very different than those that existed a few decades ago. One might question then, to what extent these newer models have overcome the deficiencies of earlier generations of models, such as the criticisms lodged against the first generation of spatial interaction-based simulation models.

Reflecting on the earlier experience, some modelers claimed in the early 1990s that advances in computer processing power and data storage would obviate many of the problems identified by Lee (1973) in his critique of the early modeling experience (Harris, 1994). While these advances have undoubtedly reduced some of the costs of building, operating and maintaining transportation and land use models, concomitant expansions in the scope of these models, as exemplified by the current generation of urban micrsosimulation models, ensures they will continue to be a resource-intensive effort. These models also remain highly complex, with many interacting submodels. Calibration is still a daunting task, even for models that are available as commercial packages. Data requirements are still large, especially for dynamic models that require synthesis of a population or continual updating of a sample.

It must also be recognized though, that a number of problems identified with earlier models have been, at least partially, resolved. Most microsimulation models are no longer static, and can simulate changes in transportation network performance and land use through time. Nearly all models now are able to model land markets with explicit prices and the ability to simulate the behavior of various agents in the land development process. The level of aggregation of agents is being reduced, especially in comprehensive microsimulation models. The size of zones in most models is now much smaller, and should continue to decrease as computing power permits, though spatial detail in many models could certainly improve. Perhaps most importantly, the theoretical basis of models has improved, especially in ongoing efforts to reconceptualize the relationship between individual activity patterns and travel choices for travel demand forecasting.

New Directions

The development of advanced models of transportation and land use change brings about opportunities for exploring some important topics related to the models themselves and their representation of real-world urban regions. The following are some issues worthy of more attention.

The Use of Theory Some researchers question the continuing use of broad theoretical frameworks to guide agent behavior in model systems. Timmermans (2003) points to the use of random utility theory to describe a wide range of spatial choices in many models. Noting that utility is a concept that must be built up over several repetitive choice situations, he questions the applicability this concept to rare decisions such as mobility and residential location. Also, it is questionable whether discrete choice methods and random utility theory are applicable to entities such as firms, which comprise collective choice situations, as opposed to individual agents. Models like Ramblas and ILUTE, which use rule-based or hybrid modeling approaches, suggest that tailoring the right tool to each model component can overcome this issue. Timmermans also noted that most models are consumers of theory rather than producers, indicating that model development ought to coincide with the process of theory development.

Forecast Accuracy Recent studies that have sought to explore the propagation of uncertainty through transportation and land use models (Clay and Johnston, 2006; Krishnamurthy and Kockelman, 2003; Pradhan and Kockelman, 2002) have identified a continuing trend of large variation in output from these models. Presumably, the addition of better model dynamics and disaggregation of population groups within microsimulation models will reduce some of the bias present in earlier, more aggregate models. However, long-term forecasting models of many types necessarily retain significant amounts of irreducible uncertainty, and the lack of available forecasting results from applications of newer models leaves some room for concern.

Treatment of Supply Side In most forecasting applications the supply side of transportation, as represented by the extent and capacity of networks, is held fixed or treated as a policy variable. The limited available evidence on the evolution of networks over time (Levinson and Yerra, 2006; Yamins et al., 2003; Yerra and Levinson, 2005; Zhang and Levinson, 2007) suggests that scenarios such as alternative ownership regimes and their impact on transportation-land use systems are a topic worthy of exploration with more comprehensive models.

Agglomeration Effects Previous reviews of operational models of transportation and land use (Berechman and Small, 1988) identified the absence of agglomerative effects as a major weakness of the land use component of these models. Recent work using multi-agent systems (Arentze and Timmermans, 2003) suggests that modeling this effect is possible, and it is deserving of further exploration.

Person-Based Accessibility Since accessibility is still seen as an important component of location choice in transportation and land use models, especially for residential location, it makes sense to pursue measures of accessibility that recognize the importance of treating travel behavior as a process constrained in time and space, as is reflected in activity-based travel models. Examples have been provided in work by Kwan and Weber (2003) and Miller (2005).

Future work in these key areas holds some promise to improve the validity of land use and transportation models. Many of the suggested actions can, and in some cases have, been incorporated into existing models. Recent versions of UrbanSim have attempted to simulate agglomeration effects by including a variable in the utility expression for firm location choice reflecting the existence of employment in the same industry. As a proxy for agglomeration effects, this feature should improve location choice models by providing a complement to traditional accessibility measures as determinants of employment location.

Incorporation of person-based accessibility measures also seems feasible, and has been demonstrated by Dong et al. (2006) in an application of activity-based accessibility using an activity-based travel modeling system developed by Bowman and Ben-Akiva (2001). While this approach represents a definite improvement to the modeling of travel demand, it is unclear whether the use of activity-based measures of accessibility will greatly affect longer-term location decisions as currently structured within land use models.

Explicit treatment of the supply side in transportation models would also improve their capacity to reflect the dynamics of land use and network growth. In a previous application, Levinson and Karamalaputi (2003) demonstrated how factors such as existing traffic demand, demographic characteristics, and present network conditions could be used to predict network expansion. Incorporation of a separate sub-model within current operational models seems a viable and useful alternative to treating the supply side of transportation systems as fixed.

Perhaps the most important line of inquiry for future work in land use and transportation modeling relates to understanding the accuracy and level of uncertainty inherent in existing and proposed operational models. As indicated, some of this work has already begun. Future work may continue to use simulation methods, such as Monte Carlo or Latin Hypercube sampling (Hess et al., 2006) in order to relate changes in inputs and model parameters to various model outputs. The outcome of this line of work will provide important feedback to model users about the acceptability of forecasts based on current operational model systems.

A-8 Conclusion

Models of transportation and land use change have evolved significantly since their early applications more than four decades ago. In the search to design models that capture the recursive relationship between transportation and land use, there has been a general trend toward the disaggregation of the representation of people and space. Newer models represent in greater detail the dynamics of the transportation-land use change process. Experiments with bottom-up approaches to modeling urban systems, especially those that recognize the interactions between agents, provide an alternative means for understanding their complexity. Yet, the ability to forecast these processes for policy applications remains and important goal. Most of the newer generation of microsimulation models are designed with the objective of making them more policy sensitive. Unfortunately, few of them have yet reached a point where they can be fairly evaluated on this criterion, and the older operational models still raise important questions about the utility of such complex tools.

One must be more circumspect though, in evaluating the transportation and land use modeling experience more generally. In reflecting on the experience with the first generation of models nearly three decades ago, Batty (1979) noted that models should be evaluated in terms of their contribution to both science and design (i.e. policy). Many of the earliest models were failures on both accounts, though there has arguably been some success on the science side since then.

Models continue to represent an important means of testing theories and developing knowledge about the behavior of urban systems. For example, land use and transportation models have emphasized the role of accessibility in location choices from their earliest origins. They have provided a method for formally and quantitatively understanding this link and its effects on urban structure. Another contribution that modeling efforts have made has been to treat cities has living, dynamic systems. While virtually all operational models include some type of feedback effect, the increasing inclusion of dynamic effects in the form of lagged responses to transportation network or land use changes has increased the realism and applicability of many models. Further, recent modeling efforts are beginning to incorporate principles of self-organization in representing urban growth and change. Agent-based and microsimulation models demonstrate the importance of individual action, as opposed to top-down decision-making, in representing dynamics of urban growth and change. Recent microsimulation models such as UrbanSim have been able to capture unique spatial effects, such as neighborhood effects, in residential location choices. Agglomeration effects in firm location choice have also been added to reflect concentrations of an industry type in specific locations.

Another observation by Batty related to the status of planning as a science. He argued that planning was (at the time) an 'immature' science, marked by poor theoretical development, continuing controversy about methods and results, and the tendency to follow 'fashions.' In many respects this is true of the field today, as in the continuing controversy over the influence of land use patterns on travel behavior. Batty suggested that this status may be inherent to planning, which is considered a 'policy' science, and hence subject to the dictates of short-term policy needs, albeit at the expense of long-term theory development. This trend continues to the present and will likely do so in the future, as the needs for policy-oriented analysis (design) work continue to dominate planning practice. He noted though, that periodic reflection and critical review by those engaged in research can be seen as a sign of maturity. There has been much of this in the field of transportation and land use modeling, as in the related field of travel demand analysis (see, for example Pas (1990)). Continued reflection, along with a commitment to developing models that reflect the relevant theoretical constructs of the behavior or system being studied, are seen then as the most promising paths toward developing transportation and land use modeling toward a more 'mature' state and building more practically useful tools.