

Dynamic Simulation of Real Estate Development and Land Prices within an Integrated Land Use and Transportation Model System

Paul Waddell^a

Gudmundur F. Ulfarsson^b

^aDaniel J. Evans School of Public Affairs
University of Washington, Box 353055
Seattle, WA 98195
Fax: (206) 543-1096

^bDepartment of Civil and Environmental Engineering
University of Washington, Box 352700
Seattle, WA 98195
Fax: (206) 543-1543

Paper length: 7500 words. (Text: 5,750 words; 5 Tables, 2 Figures: 1,750 words)
Paper submitted: July 31, 2002

Telephone numbers and email addresses:
P. Waddell: (206) 221-4161, pwaddell@u.washington.edu
G. F. Ulfarsson: (206) 221-4161, gfu@u.washington.edu

ABSTRACT

This paper reports on the design, specification and empirical results of a discrete choice real estate development model and a hedonic regression land price model for the Greater Wasatch Front Area of Utah, using 150 by 150 meter grid cells as the units of analysis. The real estate development and land price models are components of the UrbanSim model system, in which, they interact with demand model components for residential location and employment location. Parcel data were geo-coded to grid cells, classified into composite development types, and the year of construction used to generate a chronology of development within each cell. The real estate development model simulates the annual probability that a cell will remain in its current state or experience a development event that involves the addition of housing or nonresidential floor-space, and a potential transition to one of 24 development types. The land price model simulates the land price for each cell, given its locational characteristics and accessibility. Spatial measures describing the real estate and land use composition of the neighborhood, recent development patterns, distance from existing development, site characteristics such as proximity to highways and arterials, and regional access measures incorporating the composite utility of travel to population, are used as independent variables. Estimation results indicate the importance of both localized and regional accessibility on real estate development and prices, and reflect the effects of comprehensive land use plan constraints on development outcomes and prices.

INTRODUCTION

The Greater Wasatch Front Region of Utah, home of Salt Lake City and recent host of the winter Olympics, has been experiencing rapid economic and population growth, with predictable consequences for spatial expansion and travel demand. Concerns over the effects of such rapid growth precipitated the launching of a unique regional visioning process known as Envision Utah, an initiative that convened community leaders and residents to attempt to forge consensus about vision for the region over the next 20 to 30 years. Having adopted a preferred strategy that attempts to reconcile the range of values and perspectives within the community, the task now at hand is to determine the most effective ways to bring the vision to reality. To assist in doing this, the Governor's Office of Planning and Budget, in coordination with the Wasatch Front Regional Council (WFRC) and the Mountainlands Association of Governments (MAG), chose to support the analysis of alternative land use and transportation policy strategies for achieving the vision by using the UrbanSim model system in conjunction with the regional travel model system operated by WFRC and MAG.

This paper describes the real estate development and land price components of UrbanSim, and the results of estimation of these models for the Greater Wasatch Front region. The description of the real estate development and land price models is combined, since these are highly interdependent and share most explanatory variables. These models are also interdependent with the demand-side models of residential location and employment location, which we describe separately due to space limitations and differences in the structure and specification of those models. A description of the residential location model and estimation results from the Greater Wasatch Front application is reported in (1), and the design and application of the employment location model is reported in (2).

An overview of the UrbanSim model and validation results of its application in Eugene-Springfield has been described in (3) and further references provide empirical results from the original specification of UrbanSim (4), description of the data development process (5), detailed specifications of the current UrbanSim implementation (6), analysis of its relationship to land supply monitoring (7), description of its theoretical foundations (8, 9), its application as a decision support system (10), and the underlying software infrastructure (11).

In the following sections, we review the theory and approach used in these model components, and then describe the methodology and results from estimation using data for the Greater Wasatch Front Region.

THEORY

The theoretical foundations of the model components described here draw heavily on Random Utility Maximization (RUM) models pioneered by McFadden (12, 13), on bid-rent theory of land markets (14, 15), and on hedonic price theory (16). The work represents an ongoing effort to integrate these strains of theory and methodological developments into an operational urban simulation framework.

Our approach in modeling real estate prices assumes that individual consumers and suppliers are too small in scale to manipulate prices directly, making them exogenous to these individual actors. Whereas this assumption could be argued in the event of oligopolistic behavior by large-scale developers or large corporations seeking sites, it is a relatively weak assumption to impose and avoids complications arising from modeling prices as endogenous to the interaction between consumers and sellers, such as having to simulate search and auction processes, imperfect information, and oligopolistic market behavior. A second assumption is that the

advantages of location, such as neighborhood amenities and accessibility, are capitalized into land values. This assumption follows from a wide consensus of theoretical and empirical work in urban economics that has consistently found that in competitive land markets, the quasi-unique characteristic of land (they aren't producing any more of it, every location is unique, and housing or commercial buildings are tied to their location) implies that consumers bid for location based on their willingness to pay for locational attributes, and the highest bidder wins the use of the site and sets the market price for it (14, 17).

Rosen (16) developed the approach of hedonic price analysis, which attempts to disentangle the implicit prices for the components of the bundle of services provided by housing (the same theory applies to nonresidential space). By regressing the sale price of housing on characteristics of the housing structure and location, we obtain estimates of the implicit prices of individual characteristics—holding other characteristics constant—despite us observing only the single price of the bundle for any individual property. These implicit prices do not, strictly speaking, represent either demand functions (willingness to pay) or supply functions (reservation prices), but rather, the composite of all of the willingness to pay and reservation price functions of all consumers and sellers in the market. Given our assumption that prices are exogenous to individual consumers or sellers, this provides a reasonable way to estimate the land price function within a given market.

Following DiPasquale and Wheaton (18), we interpret market prices of land within a metropolitan market as consisting of two parts. The first component is a mean price level, which fluctuates around long-term trends that are driven by short-term imbalances between supply and demand of real estate, by interest rates and other development costs, and in the longer-term by overall expansion and contraction of the metropolitan economy, population, and changes in income. The second component is the relative price of land across sites within the metropolitan market. These relative prices are based on relative advantage and abundance of sites with characteristics that are valued or avoided by consumers. As these underlying characteristics and the resulting relative advantage change, so to do relative prices, as these advantages are capitalized into land values. This paper focuses principally on the characteristics influencing relative prices, since these will have the greatest influence on intra-metropolitan variation in real estate development and consumer location choices.

Real estate development is a collection of choices made by individual developers on individual sites, about whether, when, and how to develop or redevelop those sites. We assume their behavior is motivated by profit (they attempt to maximize their profits), within constraints imposed by their resources, the physical environment, and by public land use regulations. The main influences on their choices will then be factors influencing prices of different types of real estate at different locations, the costs of producing those development projects, and the constraints relevant at those sites. There are two general approaches that developers consider in making development choices. The first is known as the site looking for a use, and corresponds to a specialized developer who has a specific project in mind, and attempts to find the most profitable site for the project. The second general approach is known as the use looking for a site, and corresponds more closely to the landowner's problem of sorting out which type of developer to sell the property to, that will generate the highest return. In the real world, both approaches occur. We have structured the current model as a discrete choice model from the perspective of the site looking for a use—the landowner's perspective. This approach lends itself to formulation as a standard multinomial logit model, where an individual landowner considers alternative uses, or developments, for a particular site.

Taking these elements together, we propose the modeling of land prices as a hedonic regression, and the real estate development model as a multinomial logit model of development of a site into alternative uses over a specific time frame.

DATA AND METHODOLOGY

Development of the UrbanSim Database

The data used by UrbanSim to construct the model database includes parcel data from tax assessor offices, business establishment files from the state unemployment insurance database or from commercial sources, census data, GIS overlays representing environmental, political, and planning boundaries, and a location grid. These data are diagnosed and analyzed for missing, erroneous, or inconsistent values. Each household in the metropolitan area is represented as an individual entity, with the primary characteristics relevant to modeling location and travel behavior: household income, size, age of head, presence of children, and number of workers. Employment is represented as individual records for each job and its employment sector. Locations are represented using grid cells of 150 by 150 meters (just over 5.5 acres), whose size can be modified. This location grid allows explicit cross-referencing of spatial features such as planning and political boundaries, including city, county, traffic zones, urban growth boundaries; and environmental features such as wetlands, floodways, stream buffers, steep slopes, or other environmentally sensitive areas.

The parcel data are collapsed into the 150 by 150 meter cells to generate composite representations of the mix and density of real estate at each location, labeled development types. These development types are somewhat analogous to the development typology developed by Calthorpe (19), in that they represent at a local neighborhood scale the land use mix and density of development. Table 1 provides the rules for classifying grid cells into development types, based on the combination of housing units, nonresidential square footage, and the principal land use of the development. Cells containing some housing and almost no nonresidential square footage are considered residential in character. Those containing a diverse mixture of housing and nonresidential floor-space are considered mixed-use, and those cells containing principally nonresidential square footage are further classified into commercial, industrial or governmental types. For a more detailed description, see (3, 6).

Model Design—Real Estate Development

The purpose of the real estate development model is to simulate discrete developer choices about whether to develop particular sites within a given year, what type of construction to undertake, and the quantity of construction. The construction of real estate can be either new development (sometimes referred to as Greenfield development) or the intensification or conversion of existing development (referred to as infill and redevelopment, respectively). The model takes a bottom-up view, i.e. from the vantage point of a developer or a land-owner at a single location (grid cell) making choices about whether to develop, and into what type of real estate. This bottom-up view is tempered by market information that reflects the state of the market as a whole, such as vacancy rates.

The model is designed in terms of discrete alternatives that represent development events, including the base case of no development on a particular site within a given year. In addition, there are development alternatives that represent transitions between the different development

types defined in Table 1, including the alternative of increasing the density of the current cell without changing its development type (where this is feasible).

The probability of each alternative (the no development, the increasing density of current cell within its development type, and transitions to other development types) being chosen is calculated using a discrete choice model. We draw on discrete choice theory and random utility maximizing models, following the work of McFadden (12, 13), to design a multinomial logit model. Similar approaches have been developed to model land cover change (20) and land use change (21), although none of these models interact with disaggregate demand-side models of residential and employment location choice as is done in UrbanSim.

To arrive at a choice model for development we assume that 1) each cell has a developer, 2) that each development alternative, indexed by i , has attached to it some utility, U_i , for the developer, based principally on profit expectations, and 3) that the development event with the highest utility has occurred (maximization of utility). We only observe events that have occurred and do not observe the utility. We proceed by assuming that the utility of alternative i for a particular cell developer can be separated into a systematic part and a random part:

$$U_i = u_i + \varepsilon_i, \quad (1)$$

where $u_i = \beta_i \cdot x$ is a linear-in-parameters function, β_i is a vector of k estimable coefficients, x is a vector of k observed, exogenous, independent variables, and ε_i is an unobserved random error that is assumed to be distributed with a Gumbel distribution (Type I extreme value distribution), which leads to the familiar multinomial logit model (12, 13):

$$P_i = \frac{e^{\beta_i \cdot x}}{\sum_{\forall i'} e^{\beta_{i'} \cdot x}}. \quad (2)$$

The probability, P_i , represents the probability of a particular cell developer choosing development alternative i . The estimable coefficients, β_i , of (2) are estimated with the method of maximum likelihood (see for example (22)).

We estimate one choice model (i.e. one set of coefficients) for each development type, since the types are very different and the development alternatives open to each development type vary. To estimate the model coefficients we need data for cells experiencing no development and for cells development events of all types.

Data and Estimation Process for Real Estate Development Model

The estimation data are derived from the parcel and grid data for a base year of 1997. The year-built values of the existing development in the assessor's records are the foundations of the process. Year-built values are imputed for records for which they are missing by examining the surrounding cells of the same type and drawing from the distribution of observed values. Historical development 'events' are identified in the data for a user-specified period of time. Events, within this framework, are any changes in the real estate development within a cell that is identified by examining the year built values within the data.

The procedure is capable of identifying any new construction that has a year-built occurring within the specified time frame. However, the procedure does not identify events that involve the demolition of buildings at some time in the past, since normally there is no record of demolitions within the current assessor database. This procedure could be augmented with data derived from building demolition and permit records.

The result is a set of cells experiencing development events that represent all observed transitions between any pairs of development types, including increases in density that didn't result in a development type change, within each year of the specified historical time frame. The time slice for determining the existence of an event is annual, since this is the limit of the information on the vintage of real estate. For further explanations of this process, see (6).

To form the estimation data we take all the development event cells, i.e. cells with a known development event and look up the values for a set of independent variables from the grid cell database. The independent variables in the real estate development model include characteristics of the grid cell (current development, land use plan, environmental constraints, policy constraints, land and improvement value), characteristics of the site location (proximity to highways, arterials, existing development, and recent development, neighborhood land use mix and property values, local accessibility measures), and regional accessibility (access to population and employment, travel time to central business district and airport). The local accessibility measures correspond to activities that can be reached by walking, over a distance of 600 meters (approximately 1/3 mile) and they are calculated using spatial queries on the network of grid cells.

We now need to take into account the much larger set of cells that didn't experience a development event. We take a random sample of these cells to generate a set of similar size as the development event set. This gives us a choice-based sample of cells. Choice-based sampling only biases the alternative-specific constants but other coefficients remain consistent (23). We adjust the alternative-specific constants after estimation to account for this bias.

Given this data, we estimate multinomial logit, discrete choice models for real estate development for each development type, with alternatives of no development, increasing density of cell without changing the development type, and observed transitions to other development types. The models are estimated using maximum likelihood. See, for example, (22) for a description of maximum likelihood estimation and multinomial logit models.

Simulating Real Estate Development

UrbanSim simulates real estate development using the presented model on a yearly basis. Each year, the model iterates over all grid cells on which development is allowed and creates a list of possible development alternatives for each cell. Development constraints may reduce the number of alternatives from the estimation stage.

Constraints on development outcomes are included through a combination of user-specified spatial overlays and decision rules about specific types of development allowed in different situations. Each cell is assigned a series of overlays through spatial preprocessing using GIS overlay techniques. These overlays can be used to assign user-specified constraints on the type of development that is allowed to occur within each of these overlay designations. The constraints are indicated as allowed conversions between each land use plan designation and each development type. Currently, if users wish to examine the impact of these constraints, they would need to relax a particular constraint and compare the results to the results for more restrictive policy. For example, the plan designation of 'agricultural' may not allow conversion

to any developed urban category under restrictive interpretation of the land use plan, or may allow conversion to rural density single-family residential under a less restrictive interpretation. The overlays used in the Greater Wasatch Front Region of Utah model application include the following features: a) water cover, b) flood plane, c) steep slope, d) open space, e) public space, f) roads, g) land use plan designation. Constraints are implemented by eliminating the constrained development alternatives from the choice set for any cell affected by the constraint. These constraints are therefore interpreted as binding constraints, and not subject to market pressure.

The estimated logit model is used to calculate the probabilities of each allowed alternative, i.e. the no development alternative, and one or more development alternatives. Development is then simulated using a Monte Carlo sampling process. Actual implementation of development takes place by using a development template, which gives the most likely characteristics of the resulting development project within the cell. The development template has defined probability distributions for development changes, including the number of housing units, square feet of commercial, industrial and government space, improvement value, and construction schedule. These development events are then added to the ‘development event’ queue in UrbanSim, to be built as scheduled.

Land Price Model

The land value for each cell, taken as the aggregation of the land value of the parcel fragments that lie within the cell, and originating from the tax assessor’s estimates of the land value of each parcel, is used as the basis for the dependent variable of the land price model. The independent variables used as predictors—essentially the same as for the real estate development model—are the characteristics of the cell, its surrounding environment, and its accessibility. A semi-log specification is used, with the log of land price as the dependent variable, as is common in hedonic price studies since it generally provides a more robust specification.

The model is a linear multiple-regression of the log of land prices, $\ln(P_i)$, for each cell i , on an array of housing structural (S_i), neighborhood (N_i), and accessibility (A_i) characteristics:

$$\ln(P_i) = \alpha + \beta S_i + \delta N_i + \gamma A_i + \varepsilon_i, \quad (3)$$

where α is the estimable intercept term; β , δ , and γ are the estimable coefficient vectors on the housing structural, neighborhood, and accessibility characteristics, respectively; ε_i is an unobserved error term, assumed to be normally distributed with mean zero and variance σ^2 .

The full set of grid cells in the study area is used in model estimation, using base year (1997) characteristics and values. As such, this is a cross-sectional estimation of the market hedonic price function, rather than an estimation of a dynamic price function. Dynamics are introduced through the process of annual changes in the characteristics of grid cells due to simulated results from the real estate development, residential location and employment location models, and the external transportation model system, all of which combine to change the characteristics of grid cells on an annual basis. Each year, after all other model components have executed, the land price model simulates end-of-year prices of land, based on the updated cell characteristics. These become the land prices that influence location choice and developer behavior in the subsequent year.

RESULTS AND CONCLUSIONS

The real estate development model is estimated separately for cells of each development type, representing a total of 24 models. Due to space limitations, only the results of one of the 24 models, representing cells initially classified as vacant land, are reported in Table 2. Although this model represents the predominant mode of new real estate construction, conversion of vacant land to urban uses (greenfield development), it is important to recognize that infill and redevelopment do contribute substantially to the real estate inventory, and moreover, change the spatial structure of urban areas in quite significant ways. The dynamics of infill and redevelopment are represented in the other 23 real estate development models not reported here.

The results shown in Table 2 provide the estimation results for conversion of cells classified initially as vacant land to one of the alternative 23 development types. The alternatives not reported as outcomes (8, 10–16, 18, 19, 22, 23) are omitted because there were too few historical transitions of this type observed in the 10 years over which these events were compiled. The first seven alternatives are residential development, ranging from very low density (one residential unit in a cell of just over 5.5 acres), to relatively high density (31 to 75 units in the same area). Alternative 9 is a low-density mixed use development type, alternative 17 is low-density commercial development, and alternatives 20 and 21 are low- and moderate-density industrial development, respectively. Coefficients in many cases are constrained to be the same across related alternatives.

Among the key patterns in the results are the following:

- 1) Higher density of residential units in a cell increases the utility of further development, particularly for the higher density residential outcomes.
- 2) Proximity to existing development, a measure of leapfrogging or spatial dispersion of new development from existing development, was found to be consistently significant for both residential and nonresidential development.
- 3) There is an observable pattern of commercial and industrial development following recent development of housing, while recent commercial development significantly increased the utility of mid- and high-density residential and mixed use development, but decreased the utility of further commercial development.
- 4) Patterns of land use segregation appeared in the results, due to some combination of enforced segregation due to zoning and market forces, as evidenced by the positive coefficients on percent of residential cells in the surrounding area for further residential development, but negative coefficients for percent mixed use, percent commercial, and percent governmental cells.
- 5) Higher land values encourage further development, with differential patterns of response across alternative development types.
- 6) Proximity to a highway reduced the utility of residential development but increased the utility of mixed use and commercial development.
- 7) Higher percentages of cell area covered by water, floodplain, steep slope, or open space decreased the utility of most development.

These coefficients can only provide limited direct interpretation, since they affect the utility of alternatives but the resulting probability is affected by the relative utilities of all the alternatives. To provide more directly interpretable results, we compute the elasticities of the continuous independent variables, reported in Table 3. All elasticities in this model have the

same sign as the coefficients, note that elasticities in multinomial logit models do not necessarily have the same sign as the coefficients but it holds in this case. The elasticities, contrary to the coefficients, are directly comparable to each other, reflecting the percentage change in the probability of a development alternative, as a function of a one percent change in an independent variable, evaluated at the mean of all continuous variables. Elasticities on the percentage of cells within 600 meters that are of a commercial development type, show large positive impacts on the probability of all types of development.

Figure 1 depicts the dynamic trends in the number of development events of each type of transition from vacant land to each simulated development outcome. The period from 1985 to 1997 is observed, and the period post 1997 is simulated based on one set of default assumptions, including a 2 percent annual rate of growth in population and employment, roughly comparable to a long term historical average. The simulated results are fairly consistent with historical patterns, with one notable exception. The lowest density residential category, representing one housing unit within 5.5 acres, drops off markedly after the transition from observed to simulated events, with a slight resurgence late in the simulation period. This would tend to indicate underproduction of low-density housing compared to historical patterns. This may be due to a combination of two main causes. First, a significant number of the cells classified as lowest density residential may have originated due to an artifact of the spillover of one housing unit from a higher density subdivision that is mostly contained in other, adjacent cells—meaning that these are not planned as low-density residential. Second, the imposition of development constraints in the simulation may overly inhibit this level of low-density residential development.

Results of the land price model show that the model explains approximately 75% of the variation in the log of land value of cells. In these results, the coefficients reported are all significant at the 95% level, and the coefficients are directly interpretable. The coefficients on the continuous independent variables that are nominal show the percentage effect on land value in a cell associated with a one-unit change in the independent variable (multiply the coefficient with 100 to arrive at the percentage change). Coefficients on variables that are log-transformed are directly interpretable as elasticities. For dummy, or categorical variables, the coefficients indicate the percentage change in the land value of a cell associated with a change from a value of 0 to a value of 1 on the dummy variable (multiply the coefficient with 100 to arrive at the percentage change). Each coefficient must be interpreted holding all other variables constant.

The first set of variables reflects the influence of the current development type of a cell, as compared to the omitted category of vacant. The largest percentage effects are for mid- and high-density commercial, with price premiums of 74% and 95%, respectively, and for high-density industrial (63%) and high-density mixed-use (61% premium as compared to vacant). The second set of variables show the effects of land use plan designations on land values, compared to agricultural land use plan designation, with premiums of up to 166% for plan type 8.

Improvement value reflects the influence of the quality and value of construction of real estate, and this has an elasticity of .07 on land value. Residential density within a cell, measured by log of number of units, has an elasticity of .25. Density within the surrounding neighborhood of cells within 600 meters had an elasticity of .16. These would seem to clash with expectations that households prefer lower densities, since the locations with higher densities command higher land values. But one should note that we are describing land values here, and at higher land values, developers will tend to use land more intensively, and build higher density development to bring the price per unit of development down, and to maximize total profit from development.

Regional access to population and regional access to employment are fairly correlated, but access to employment is more concentrated in and near employment centers. The coefficients on access to population and employment were both significant, but in opposite directions, with access to population positive and of larger magnitude. Local land use mix also proved to be significant, with mixed use development lowering land values, and commercial and industrial composition increasing it, all else being equal. Proximity to a highway increased land values by 9%. Land values were reduced as the percentage of a cell increased in the following categories: water, animal sanctuary, floodplain, wetland, steep slope, public space, and roads. By contrast, an increase by one percent of land within a cell in open space increased land values by .52%.

Figure 2 depicts land values per acre simulated by the model for each 150 by 150 meter grid cell in 1997, averaged across the cells within a 3 by 3 cell neighborhood to provide a more generalized representation. The results show the composite effects of all the influences incorporated into the model: current and planned land use, density, nearby land use mix and recent development activity, regional accessibility, and relationship to existing development and to highways.

These results reveal a complex but systematic pattern of influences on land prices and the probability of alternative development outcomes. Regional transportation accessibility interacts with localized site characteristics and neighborhood-scale characteristics to texture the patterns of prices and development. The models presented here interact with demand-side models of residential and employment location on an annual basis, to reflect the ongoing dynamics of demand, supply, and prices of real estate and the effects of transportation and land use policies and investments. These results are subject to change as further testing of the model system proceeds. An extensive testing of the interaction of the UrbanSim model system with the combined regional transportation model system operated by the Wasatch Front Regional Council is now in its early stages.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grants CMS-9818378, EIA-0090832, BCS-0120024, and EIA-0121326, and by the Governor's Office of Planning and Budget (GOPB), the Wasatch Front Regional Council (WFRC), the Mountainlands Association of Governments (MAG). In particular, we wish to acknowledge the assistance of John Britting at WFRC, Carl Johnson at MAG, Peter Donner at GOPB, Natalie Gochnour (formerly at GOPB), and Stuart Challender (formerly with Utah Automated Geographic Reference Center) for their assistance.

REFERENCES

1. Waddell, P. and F. Nourzad. Incorporating Non-Motorized Mode and Neighborhood Accessibility in an Integrated Land Use and Transportation Model System. *Transportation Research Record*, forthcoming.
2. Waddell, P. and G. F. Ulfarsson. Accessibility and Agglomeration: Discrete-Choice Models of Employment Location by Industry Sector. 2002. <http://www.urbansim.org/papers>. Accessed Aug. 1, 2002.
3. Waddell, P. UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. *Journal of the American Planning Association*, Vol. 68, No. 3, 2002, pp. 297-314.

4. Waddell, P. A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim. *Environment and Planning B: Planning and Design*, Vol. 27, 2000, pp. 247-263.
5. Waddell, P. *Exploiting parcel level GIS in land use models*. in *ASCE Conference on Land Use, Transportation and Air Quality: Making the Connection*. 1998. Portland, OR: American Society of Civil Engineers.
6. Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke and G. F. Ulfarsson. UrbanSim: A Simulation System for Land Use and Transportation. *Networks and Spatial Economics*, forthcoming.
7. Waddell, P. Monitoring and Simulating Land Capacity at the Parcel Level. In *Monitoring Land Supply with Geographic Information Systems: Theory, Practice and Parcel-Based Approaches*, Vernez-Moudon, A. and M. Hubner, Editors. John Wiley & Sons, Inc. New York. 2000, pp. 201-217.
8. Waddell, P. Towards a Behavioral Integration of Land Use and Transportation Modeling. In *The Leading Edge in Travel Behavior Research*, Hensher, D., Editor. Pergamon Oxford. 2001, pp.
9. Waddell, P. and T. Moore. Forecasting Demand for Urban Land. In *Land Market Monitoring for Smart Urban Growth*, Knaap, G., Editor. Lincoln Institute for Land Policy Cambridge, MA. 2001, pp. 187-217.
10. Waddell, P. Between Politics and Planning: UrbanSim as a Decision Support System. In *Planning Support Systems: Integrating Geographic Information Systems, Models, and Visualization Tools*, Brail, R. a. K., Richard, Editor. ESRI Press and Center for Urban Policy Research Redlands, CA. 2001, pp. 201-228.
11. Noth, M., A. Borning and P. Waddell. An Extensible, Modular Architecture for Simulating Urban Development, Transportation, and Environmental Impacts. *Computers, Environment and Urban Systems*, forthcoming.
12. McFadden, D. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, Zarembka, P., Editor. Academic Press New York. 1974, pp.
13. McFadden, D. Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*, Manski, C. F. and D. McFadden, Editors. MIT Press Cambridge, MA. 1981, pp. 198-272.
14. Alonso, W. *Location and Land Use*. Harvard University Press, Cambridge, 1964.
15. Wheaton, W. C. A bid rent approach to housing demand. *Journal of Urban Economics*, Vol. 4, No. 2, 1977, pp. 200-217.
16. Rosen, S. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, Vol. 82, 1974, pp. 34-55.
17. Mills, E. S. An Aggregative Model of Resource Allocation in a Metropolitan Area. *American Economic Review*, Vol. 57, 1967, pp. 197-210.
18. DiPasquale, D. and W. Wheaton. *Urban economics and real estate markets*. Prentice Hall, Englewood Cliffs, NJ, 1996.
19. Calthorpe, P. *The Next American Metropolis: Ecology, Community and the American Dream*. Princeton Architectural Press, New York, 1993.
20. Turner, M. and R. Gardner. *Quantitative Methods in Landscape Ecology*. Springer-Verlag, New York, 1991.

21. Landis, J. and M. Zhang. The Second Generation of the California Urban Futures Model Part I: Model Logic and Theory. *Environment and Planning B: Planning and Design*, Vol. 25, 1998, pp. 657-666.
22. Greene, W. *Econometric Analysis*. 4th ed. Prentice Hall, 1999.
23. Manski, C. F. and D. Mcfadden. Alternative Estimators and Sample Designs for Discrete Choice Analysis. In *Structural Analysis of Discrete Data with Econometric Applications*, Manski, C. F. and D. Mcfadden, Editors. MIT Press Cambridge, MA. 1981, pp. 2-50.

LIST OF TABLES

TABLE 1 Development Type Classification

TABLE 2 Estimation Results for Real-Estate Development of Vacant Land

TABLE 3 Direct Elasticities of Continuous Variables in Estimation Results for Real-Estate Development of Vacant Land

TABLE 4 Estimation Results for Land Price

LIST OF FIGURES

FIGURE 1 Quantity of development (y-axis) per year (x-axis) for each development type. The results for the years 1985-1997 represent observed data, whereas the results from then on represent model predictions.

FIGURE 2 Simulated land value per acre for 1997, averaged over 450 by 450 meter neighborhood.

TABLE 1 Development Type Classification

Dev. Type	Name	Min. Units	Max. Units	Min. Sqft.	Max. Sqft.	Primary Use
1	R1	1	1	0	999	Residential
2	R2	2	4	0	999	Residential
3	R3	5	9	0	999	Residential
4	R4	10	14	0	2,499	Residential
5	R5	15	21	0	2,499	Residential
6	R6	22	30	0	2,499	Residential
7	R7	31	75	0	4,999	Residential
8	R8	76	65,000	0	4,999	Residential
9	M1	0	9	1,000	4,999	Mixed Use
10	M2	10	30	2,500	4,999	Mixed Use
11	M3	10	30	5,000	24,999	Mixed Use
12	M4	10	30	25,000	49,999	Mixed Use
13	M5	10	30	50,000	9,999,999	Mixed Use
14	M6	31	65,000	5,000	24,999	Mixed Use
15	M7	31	65,000	25,000	49,999	Mixed Use
16	M8	31	65,000	50,000	9,999,999	Mixed Use
17	C1	0	9	5,000	24,999	Commercial
18	C2	0	9	25,000	49,999	Commercial
19	C3	0	9	50,000	9,999,999	Commercial
20	I1	0	9	5,000	24,999	Industrial
21	I2	0	9	25,000	49,999	Industrial
22	I3	0	9	50,000	9,999,999	Industrial
23	GV	0	99,999	0	9,999,999	Government
24	Vacant Developable	0	0	0	0	Vacant Developable
25	Undevelopable	0	0	0	0	Undevelopable

TABLE 2 Estimation Results for Real-Estate Development of Vacant Land

New development type	1	2	3	4	5	7	9	17	20	21
Constant	-12.680 (0.444) [‡]	-20.875 (0.641) [‡]	-25.948 (0.879) [‡]	-28.078 (1.358) [‡]	-27.077 (1.758) [‡]	-23.022 (2.329) [‡]	-10.004 (0.689) [‡]	-16.502 (1.988) [‡]	-21.011 (1.746) [‡]	-21.481 (1.748) [‡]
Ln(Total residential units in cell)	–	0.169 (0.0401) [‡]	0.267 (0.0534) [‡]	0.371 (0.0623) [‡]	0.474 (0.0875) [‡]	1.647 (0.162) [‡]	–	0.263 (0.0739) [‡]	0.450 (0.0901) [‡]	0.450 (0.0901) [‡]
Scaled distance to development (1000-distance)/1000	3.495 (0.30) [‡]	3.495 (0.30) [‡]	3.495 (0.30) [‡]	3.495 (0.30) [‡]	3.495 (0.30) [‡]	3.495 (0.30) [‡]	2.994 (0.538) [‡]	1.794 (0.699) [‡]	1.794 (0.699) [‡]	1.794 (0.699) [‡]
Transitions to residential, within 600 m, in last 3 years	–	–	–	–	–	–	–	0.120 (0.0241) [‡]	0.0953 (0.0104) [‡]	0.0953 (0.0104) [‡]
Transitions to mixed use, within 600 m, in last 3 years	-0.0732 (0.0431) [‡]	-0.0732 (0.0431) [‡]	-0.0732 (0.0431) [‡]	-0.0732 (0.0431) [‡]	-0.0732 (0.0431) [‡]	-0.0732 (0.0431) [‡]	–	–	–	–
Transitions to commercial, within 600 m, in last 3 years	0.121 (0.0547) [‡]	0.121 (0.0547) [‡]	0.121 (0.0547) [‡]	0.121 (0.0547) [‡]	0.121 (0.0547) [‡]	0.121 (0.0547) [‡]	–	–	0.591 (0.0824) [‡]	0.591 (0.0824) [‡]
Ln(Residential units added, within 600 m, in last 3 years)	–	–	0.0601 (0.0231) [‡]	0.0601 (0.0231) [‡]	0.132 (0.0556) [‡]	0.216 (0.0848) [‡]	0.293 (0.0415) [‡]	-0.376 (0.147) [‡]	–	–
Percentage of residential cells, within 600 m, before transition	0.0233 (0.00159) [‡]	0.0131 (0.00248) [‡]	0.00909 (0.00299) [‡]	–	–	-0.030 (0.00818) [‡]	–	–	-0.0896 (0.0184) [‡]	-0.0896 (0.0184) [‡]
Percentage of mixed use cells, within 600 m, before transition	-0.0125 (0.00651) [‡]	-0.0125 (0.00651) [‡]	-0.0125 (0.00651) [‡]	-0.0125 (0.00651) [‡]	-0.0125 (0.00651) [‡]	-0.0125 (0.00651) [‡]	0.0838 (0.00648) [‡]	0.0984 (0.0110) [‡]	0.0521 (0.0127) [‡]	0.0521 (0.0127) [‡]
Percentage of commercial cells, within 600 m, before transition	-0.0389 (0.00837) [‡]	-0.0509 (0.00884) [‡]	-0.0679 (0.0109) [‡]	-0.0679 (0.0151) [‡]	-0.0647 (0.0198) [‡]	–	0.0388 (0.00790) [‡]	–	–	–
Percentage of industrial cells, within 600 m, before transition	-0.0541 (0.0119) [‡]	-0.0456 (0.0119) [‡]	-0.0369 (0.0134) [‡]	-0.0308 (0.0150) [‡]	-0.0308 (0.0150) [‡]	-0.0603 (0.0368)	–	–	–	–
Percentage of governmental cells, within 600 m, before transition	–	–	–	–	–	–	-0.0526 (0.0102) [‡]	-0.0317 (0.0163) [‡]	-0.0713 (0.0172) [‡]	-0.0713 (0.0172) [‡]
Ln(Land value)	0.717 (0.0335) [‡]	1.331 (0.0524) [‡]	1.661 (0.0723) [‡]	1.738 (0.114) [‡]	1.557 (0.149) [‡]	0.747 (0.202) [‡]	0.385 (0.0472) [‡]	0.631 (0.147) [‡]	1.105 (0.139) [‡]	1.105 (0.139) [‡]
Highway within 300 m	-0.499 (0.136) [‡]	-0.970 (0.178) [‡]	-1.219 (0.243) [‡]	-0.732 (0.307) [‡]	-0.913 (0.435) [‡]	–	1.033 (0.176) [‡]	1.074 (0.264) [‡]	–	–

(Continued)

TABLE 2 Estimation Results for Real-Estate Development of Vacant Land (Continued)

New development type	1	2	3	4	5	7	9	17	20	21
Percent cell covered by water	-4.630 (2.142) [‡]	-4.630 (2.142) [‡]	-4.630 (2.142) [‡]	-4.630 (2.142) [‡]	-4.630 (2.142) [‡]	-4.630 (2.142) [‡]	–	–	–	–
Percent cell covered by flood plane	-1.668 (0.667) [‡]	-1.668 (0.667) [‡]	-1.668 (0.667) [‡]	-1.668 (0.667) [‡]	-1.668 (0.667) [‡]	-1.668 (0.667) [‡]	-2.313 (1.345) [†]	–	–	–
Percent cell covered by slope	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	-1.565 (0.215) [‡]	–	–
Percent cell covered by open space	-2.542 (0.184) [‡]	-2.542 (0.184) [‡]	-2.542 (0.184) [‡]	-2.542 (0.184) [‡]	-2.542 (0.184) [‡]	-2.542 (0.184) [‡]	-2.580 (0.686) [‡]	–	–	–
Percent cell covered by public space	–	–	–	–	–	–	–	-6.194 (3.556) [†]	–	–
Percent cell covered by roads	2.464 (0.209) [‡]	2.464 (0.209) [‡]	2.464 (0.209) [‡]	2.464 (0.209) [‡]	2.464 (0.209) [‡]	2.464 (0.209) [‡]	-1.830 (0.558) [‡]	–	–	–
Plan type 2	0.975 (0.0802) [‡]	0.975 (0.0802) [‡]	0.975 (0.0802) [‡]	0.975 (0.0802) [‡]	0.975 (0.0802) [‡]	0.975 (0.0802) [‡]	–	2.280 (1.024) [‡]	–	–
Plan type 3	–	–	–	–	–	–	–	3.786 (1.031) [‡]	2.311 (0.599) [‡]	2.311 (0.599) [‡]
Plan type 4	-0.893 (0.182) [‡]	-0.893 (0.182) [‡]	-0.893 (0.182) [‡]	-0.893 (0.182) [‡]	-0.893 (0.182) [‡]	-0.893 (0.182) [‡]	0.551 (0.145) [‡]	3.576 (1.030) [‡]	4.001 (0.487) [‡]	4.001 (0.487) [‡]
Plan type 7	0.616 (0.172) [‡]	0.616 (0.172) [‡]	0.616 (0.172) [‡]	0.616 (0.172) [‡]	0.616 (0.172) [‡]	0.616 (0.172) [‡]	–	–	–	–
Plan type 8	-1.143 (0.445) [‡]	-1.143 (0.445) [‡]	-1.143 (0.445) [‡]	-1.143 (0.445) [‡]	-1.143 (0.445) [‡]	-1.143 (0.445) [‡]	–	–	–	–
Number of observations	16,323									
Log-likelihood at 0	-40,561.1									
Log-likelihood at convergence	-19,931.2									
ρ^2	0.50861									
Corrected ρ^2	0.50837									

Logit model coefficient estimates with standard errors in parentheses.

The no build option is the base case. The coefficients represent a difference in utility from the base case.

[†] Indicates the coefficient is significant at the $p = 10\%$ level, with a two-tailed t-test.

[‡] Indicates the coefficient is significant at the $p = 5\%$ level with a two-tailed t-test.

– Indicates the coefficient was restricted to zero because of lack of significance.

TABLE 3 Direct Elasticities of Continuous Variables in Estimation Results for Real-Estate Development of Vacant Land

New development type:	No build	1	2	3	4	5	7	9	17	20	21
Ln(Total residential units in cell)	-0.0003	-0.0003	0.3981	0.6277	0.8736	1.1152	3.8748	-0.0003	0.6187	1.0575	1.0575
Scaled distance to development (1000-distance)/1000	-0.0093	2.5966	2.5966	2.5966	2.5966	2.5966	2.5966	2.2231	1.3282	1.3282	1.3282
Transitions to residential, within 600 m, in last 3 years	-2.E-05	-2.E-05	-2.E-05	-2.E-05	-2.E-05	-2.E-05	-2.E-05	-2.E-05	0.6505	0.5156	0.5156
Transitions to mixed use, within 600 m, in last 3 years	0.0001	-0.0161	-0.0161	-0.0161	-0.0161	-0.0161	-0.0161	0.0001	0.0001	0.0001	0.0001
Transitions to commercial, within 600 m, in last 3 years	-5.E-05	0.0147	0.0147	0.0147	0.0147	0.0147	0.0147	-5.E-05	-5.E-05	0.0720	0.0720
Ln(Residential units added, within 600 m, in last 3 years)	-0.0001	-0.0001	-0.0001	0.0979	0.0979	0.2147	0.3527	0.4779	-0.6131	-0.0001	-0.0001
Percentage of residential cells, within 600 m, before transition	-0.0015	0.4880	0.2741	0.1895	-0.0015	-0.0015	-0.6313	-0.0015	-0.0015	-1.8850	-1.8850
Percentage of mixed use cells, within 600 m, before transition	3.E-05	-0.0241	-0.0241	-0.0241	-0.0241	-0.0241	-0.0241	0.1617	0.1897	0.1004	0.1004
Percentage of commercial cells, within 600 m, before transition	0.0001	-0.0417	-0.0546	-0.0728	-0.0728	-0.0694	0.0001	0.0419	0.0001	0.0001	0.0001
Percentage of industrial cells, within 600 m, before transition	0.0001	-0.0438	-0.0369	-0.0298	-0.0249	-0.0249	-0.0488	0.0001	0.0001	0.0001	0.0001
Percentage of governmental cells, within 600 m, before transition	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	-0.3329	-0.2009	-0.4519	-0.4519
Ln(Land value)	-0.0286	7.1060	13.2228	16.5027	17.2708	15.4720	7.4072	3.8062	6.2529	10.9747	10.9747
Highway within 300 m	0.0001	-0.0246	-0.0479	-0.0602	-0.0361	-0.0451	0.0001	0.0512	0.0532	0.0001	0.0001
Percent cell covered by water	0.0002	-0.0545	-0.0545	-0.0545	-0.0545	-0.0545	-0.0545	0.0002	0.0002	0.0002	0.0002
Percent cell covered by flood plane	0.0001	-0.0336	-0.0336	-0.0336	-0.0336	-0.0336	-0.0336	-0.0467	0.0001	0.0001	0.0001
Percent cell covered by slope	0.0002	-0.0498	-0.0498	-0.0498	-0.0498	-0.0498	-0.0498	-0.0498	-0.0498	0.0002	0.0002
Percent cell covered by open space	0.0003	-0.0845	-0.0845	-0.0845	-0.0845	-0.0845	-0.0845	-0.0858	0.0003	0.0003	0.0003
Percent cell covered by public space	9.E-06	9.E-06	9.E-06	9.E-06	9.E-06	9.E-06	9.E-06	9.E-06	-0.3320	9.E-06	9.E-06
Percent cell covered by roads	-0.0006	0.2048	0.2048	0.2048	0.2048	0.2048	0.2048	-0.1532	-0.0006	-0.0006	-0.0006

TABLE 4 Estimation Results of Hedonic Land Price Regression Model

Variable	Coefficient	Standard error
Constant	-6.911	(0.110) [‡]
Residential 1	-0.0749	(0.0187) [‡]
Residential 2	0.231	(0.0196) [‡]
Residential 3	0.372	(0.0229) [‡]
Residential 4	0.484	(0.0265) [‡]
Residential 5	0.433	(0.0285) [‡]
Residential 6	0.259	(0.0331) [‡]
Residential 7	-0.114	(0.0396) [‡]
Residential 8	-0.305	(0.0675) [‡]
Mixed use 2	0.247	(0.0631) [‡]
Mixed use 3	0.255	(0.0436) [‡]
Mixed use 4	0.219	(0.0889) [‡]
Mixed use 5	0.417	(0.0797) [‡]
Mixed use 8	0.614	(0.119) [‡]
Commercial 1	0.438	(0.0272) [‡]
Commercial 2	0.738	(0.0349) [‡]
Commercial 3	0.946	(0.0370) [‡]
Industrial 1	0.307	(0.0338) [‡]
Industrial 2	0.303	(0.0403) [‡]
Industrial 3	0.627	(0.0399) [‡]
Governmental	0.111	(0.0146) [‡]
Plan type 1	-0.202	(0.0210) [‡]
Plan type 2	0.523	(0.0109) [‡]
Plan type 3	0.729	(0.0165) [‡]
Plan type 4	0.597	(0.0139) [‡]
Plan type 5	1.069	(0.128) [‡]
Plan type 6	0.580	(0.0202) [‡]
Plan type 7	0.502	(0.0222) [‡]
Plan type 8	1.659	(0.0259) [‡]
Ln(Total improvement value + 1)	0.0701	(0.00137) [‡]
Ln(Number of units + 1)	0.247	(0.00990) [‡]
Ln(Number of units within 600 m + 1)	0.158	(0.00254) [‡]
Ln(Total employment within 600 m + 1)	0.0727	(0.00298) [‡]
Ln(Access to population)	2.989	(0.0521) [‡]
Ln(Access to employment)	-1.812	(0.0488) [‡]
Ln(Percent mixed use within 600 m + 1)	-0.0217	(0.00346) [‡]
Ln(Percent commercial within 600 m + 1)	0.0168	(0.00426) [‡]
Ln(Percent industrial within 600 m + 1)	0.0572	(0.00471) [‡]
Ln(Percent governmental within 600 m + 1)	0.0686	(0.00259) [‡]
Distance to highway <= 300 m	0.0924	(0.0148) [‡]
Ln(Distance to highway + 1)	0.0411	(0.00281) [‡]
Percent cell covered by water	-0.512	(0.0377) [‡]
Percent cell covered by animal sanctuary	-0.890	(0.0841) [‡]
Percent cell covered by flood plane	-0.607	(0.0320) [‡]
Percent cell covered by wetland	-0.929	(0.0367) [‡]
Percent cell covered by slope	-1.375	(0.0231) [‡]
Percent cell covered by open space	0.516	(0.0210) [‡]
Percent cell covered by public space	-0.440	(0.0211) [‡]
Percent cell covered by roads	-0.231	(0.0240) [‡]
Number of observations	107,208	
R ²	0.746	
Adjusted R ²	0.746	

[‡] Indicates the coefficient is significant at the $p = 5\%$ level with a two-tailed t-test.

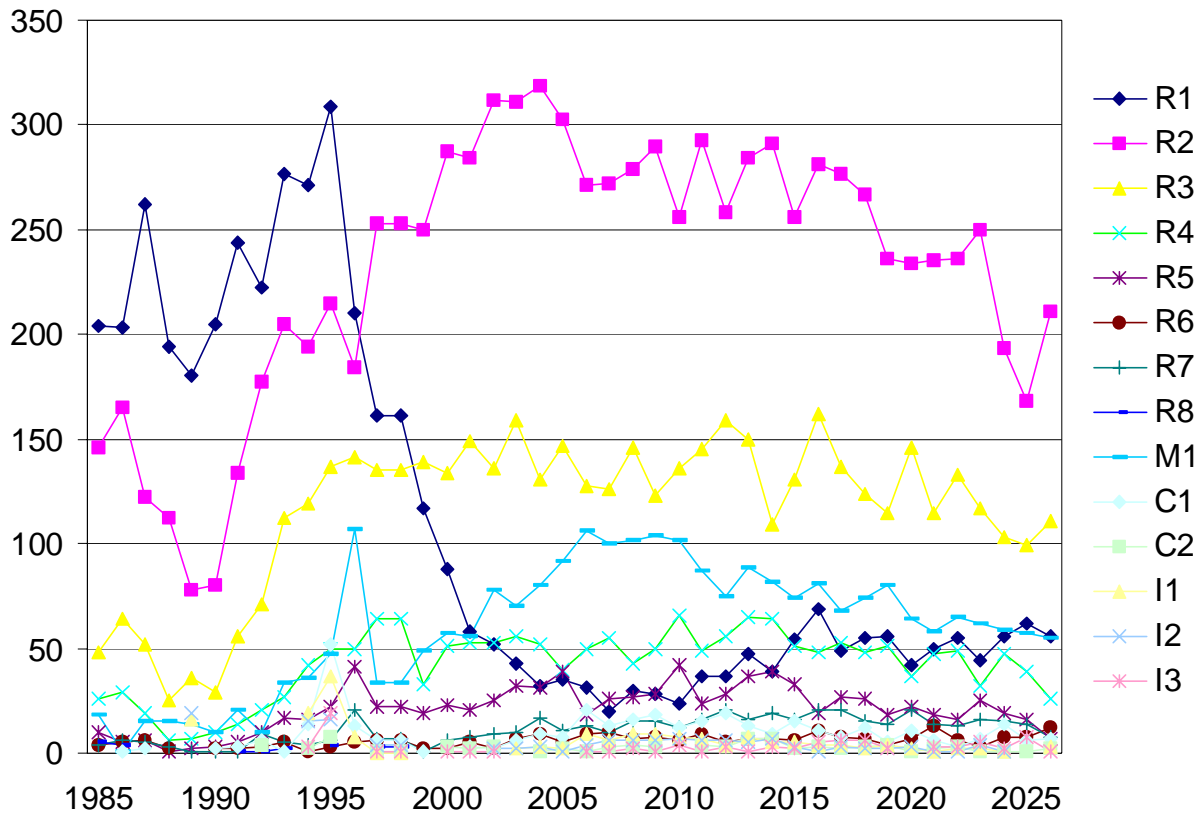


FIGURE 1 Quantity of development (y-axis) per year (x-axis) for each development type. The results for the years 1985-1997 represent observed data, whereas the results from then on represent model predictions.

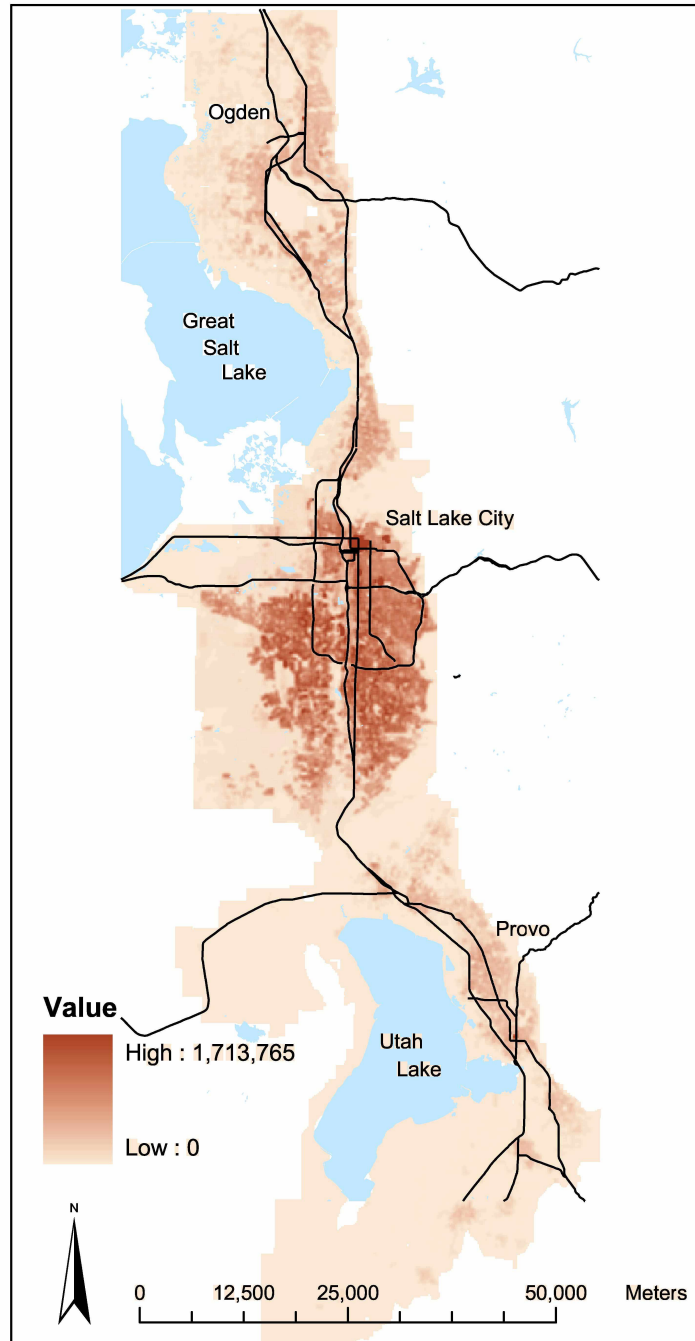


FIGURE 2 Simulated land value per acre for 1997, averaged over 450 by 450 meter neighborhood.