

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13

## Tracking Land Cover Change in a Mixed Logit Model: Recognizing Temporal and Spatial Effects

14  
15  
16  
17  
18  
19  
20  
21

**Xiaokun Wang**

Graduate Student Researcher  
The University of Texas at Austin  
6.9 ECJ, Austin, TX 78712-1076  
wangxk@mail.utexas.edu

**Kara M. Kockelman**

Clare Boothe Luce Associate Professor of Civil, Architectural and Environmental Engineering  
The University of Texas at Austin  
6.9 ECJ, Austin, TX 78712-1076  
kkockelm@mail.utexas.edu  
Phone: 512-471-0210  
FAX: 512-475-8744  
(Corresponding Author)

22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

The following paper is a pre-print and the final publication can be found in  
*Transportation Research Record No. 1977:112-120*, 2006.  
Presented at the 85th Annual Meeting of the Transportation Research Board,  
January 2006

### Abstract

42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

As an essential part of integrated land use-transport models, prediction of land cover changes and illumination of the many factors behind such change are always of interest to planners, policy makers, developers and others. Using a mixed logit framework, this paper studies land cover evolution in the Austin, Texas region, recognizing distance-dependent correlations -- both observed and unobserved -- over space and time, in a sea of satellite image pixels. The paper describes the computational methods used for model estimation and application, including generalized Cholesky decomposition and likelihood simulation. Results indicate that neighborhood characteristics have strong effects on land cover evolution: Clustering is significant over time, but high residential densities can impede future development. Model application produces graphic predictions, allowing one to visually confirm these results and appreciate the variability in potential urban futures.

**Key Words:** land cover change, satellite imagery, mixed logit model, temporal and spatial correlation, dynamic unobserved effect model, simulation, generalized Cholesky decomposition, randomly shuffled Halton sequence, predictive uncertainty

## Introduction

Over the years, researchers have put considerable effort into understanding relationships between populations, travel and land cover (e.g., Fujita [1989], Nelson and Hellerstein [1997], Clarke [1997] and Munroe et al. [2001], and several land use-transport models have resulted (e.g., de la Barra's TRANUS [1989], Waddell's UrbanSim [2002] and Klosterman's Whatif? [1999]). However, due to the complexity of the real world, no existing models are perfectly satisfactory in application. There is growing interest in more disaggregate data sets as well as the temporal and spatial correlations that underpin urban patterns. Fortunately, developments in remote sensing via satellites provide opportunities for modeling improvements through temporally consistent and spatially detailed information.

As part of an integrated land use-transport model's development, this paper presents a model of land cover change based on data derived from satellite imagery. The derived panel data allows direct incorporation of temporal and spatial correlation into a model's specification. In addition to the smaller spatial units (measuring 30 m x 30 m), another advantage of satellite data is that it offers much more precise information on vegetation, which can be critical to air quality, due to biogenic sources of (and sinks for) various chemicals of interest. This implies that using satellite data may be an optimal choice for integrated land use-transport-environment (ILUTE) models, which are required in many regions in order to demonstrate compliance with air quality-related planning standards. Moreover, with further development in the satellite image precision and classification techniques, such data may provide a very promising future. It is important that researchers and planners begin working with such data sets now, in order to begin harnessing their potential for future planning and prediction efforts.

In many cases, proximity produces correlation. In contrast to time-series data sets, spatial relationships are two dimensional – and time adds an important third dimension. Obviously, knowing what land covers presently exist – and where – is very helpful in formulating future year predictions. A rigorous recognition of spatial and temporal correlations across observational units, both latent and observed, can be critical.

In this study, a model incorporating the effects of both time and space is used. It can be viewed as an extension to Wooldridge's (2002) dynamic unobserved effect model because the lagged (both temporally and spatially) dependent variable is used as an explanatory variable and individual specific effects are correlated across observations (in both space and time). Such a model also falls within the mixed-logit (ML) framework, as described by Greene (2000) and McFadden and Train (2000), where slope parameters can be random, and a correlation structure can be specified. The ML model is normally estimated using maximum simulated likelihood estimation (MSLE), and in this study a randomly shuffled sequence of Halton draws is used.

The following sections review related works, present the methodology and applications of this research, and draw some conclusions. The data come from the Austin, Texas region and model results are applied to a 2.1 km square area for predictions of pixel-level predictions of land cover in the year 2005.

## Previous Research and Motivation

In a review of existing land use/land cover models Parker et al. (2003) concluded that no one approach yet "dominates this nascent field". However, some approaches are more common than others and have already been used in the land use/land cover part of integrated land use-transport models. For example, Waddell's (2002) UrbanSim has been rather broadly applied in

1  
2  
3 the U.S. It offers a module for simulating land development at the 150m x 150m grid cell level.  
4 Explanatory variables include regional, neighborhood and cell characteristics.

5 The recently designed PECAS model (Hunt and Abraham, 2003) has a structure similar to  
6 TRANUS (de la Barra, 1989). Like UrbanSim, PECAS' land development model relies on very  
7 standard MNL specifications. But PECAS recognizes inter-industry trade flows via a separate  
8 input-output-based sub-model. Finally, the cellular automata (CA)-based SLEUTH model  
9 (Candau et al., 2000) is also receiving much attention. SLEUTH is one of the few models that is  
10 designed to work with satellite image data. It consists of Clarke et al.'s (1996, 1997) Urban  
11 Growth Model (UGM) and Deltatron Land Use/Land Cover Model. Though SLEUTH simulates  
12 land cover change recognizing the temporal and spatial context of each cell, it can only share  
13 information across immediate cells, so that more dispersed interactions and correlations are  
14 largely ignored. In addition, SLEUTH is not (yet) designed to flexibly accommodate the effects  
15 of many influential and related human factors (such as land prices, employment and population  
16 density.) Finally, it relies on rule-driven algorithms, rather than more behavioral or statistical  
17 models.

18 Most present models are not adept at analyzing satellite data, which its highly disaggregate  
19 units and large sample size, and clear potential for various inter-observation correlations.  
20 Therefore, the goal of this study is to a rigorous econometric method to appreciate the many facets of this  
21 particular data source. However, few studies have used econometric methods to address both  
22 temporal and spatial effects. As Anselin (1988, 1999) explains, spatial econometrics allows such  
23 interactions. He suggests three principal methods for addressing the spatial effects that exist in  
24 land use, home price, air quality, and other data sets: use of spatial stochastic processes, a direct  
25 representation of correlations, and a non-parametric framework.

26  
27 The relatively common spatial autoregressive (SAR) and spatial moving average (SMA)  
28 specifications (see, e.g., Besner [2002], Miyamoto et al. [2004] and Frazier and Kockelman  
29 [2005]) are examples of spatial stochastic processes. In contrast to their time-series counterparts,  
30 where a single serial correlation parameter is used, SAR and SMA use a spatial lag operator (or  
31 weight matrix), which essentially produces a weighted average of the neighbors' values. A  
32 frequently encountered issue in spatial stochastic process is the lack of stationary covariance  
33 terms, which can violate asymptotic property assumptions established based on the central limit  
34 theorem and laws of large numbers. (Anselin, 1999). By using a "direct representation", the  
35 covariance matrix among unobservable components can be directly expressed as an inverse  
36 function of distances. In this way, it also will meet stationary requirements (Anselin 1999).  
37 However, the formulation of this inverse function is often restricted, because flexible expressions  
38 may suffer from estimation and identification problems. The third approach, using non-  
39 parametric methods, is seldom found in practice because it requires a long panel of data, with a  
40 time dimension ( $T$ ) that is much greater than the cross-sectional dimension ( $N$ ) (Anselin  
41 1999). None of these three methods is simple; and, as far as the authors are aware, only the first  
42 has been applied in a discrete-response context (see Beron and Vijverberg 2004). In such cases,  
43 the model calibration becomes intractable and requires approximation via simulation of the  
44 associated distributions.

45 Many studies recognizing spatial effects have tried to either remove all spatial correlation  
46 through strategic sampling (to provide a dispersed sample, with minimal interactions and/or  
47 spatial effects) or constructing and controlling for a variety of neighborhood attributes. For  
48 example, Nelson and Hellerstein (1997) sampled selectively and created exogenous variables  
49 from neighboring land covers when studying the effect of roadways on deforestation. Wear and  
50 Bolstead (1998) controlled for prior land uses in the neighborhood of each data cell.

1  
2  
3 Munroe et al. (2001)'s paper on land cover change is one of the few that also considers  
4 temporal correlations. They attempted to filter out spatial correlations through sampling and then  
5 removed the residual spatial dependence through a "trend surface" approach (Cliff and Ord  
6 1981). Mohammadian et al. (2004)'s study on housing type choice also attempted to control for  
7 neighborhood attributes in a mixed logit framework. However, issues remain in previous  
8 approaches because spatial correlation generally cannot be perfectly removed through sampling  
9 and spatially lagged (in contrast to temporally lagged) variables are not strictly exogenous.

10 Frazier and Kockelman's (2005) land cover model probably follows Anselin's (1988, 1999)  
11 standard approach most closely, by using an SAR specification in their continuous models of  
12 population and logistic model of land development (using 300 m x 300 m aggregates of satellite  
13 pixels in the Austin region). Miyamoto et al.'s (2004) work is most similar to this paper in nature:  
14 Relying on relatively few zones (a sample of just 163), Miyamoto et al. used a SAR formulation  
15 and Monte Carlo MSLE of a logit model for location choice. However, standard SAR  
16 specifications require information on all correlated units, which is problematic with very large  
17 data sets, such as those based on satellite images. In such situations, inversion of the weight  
18 matrix becomes practically infeasible and a sampling method must be used, rendering SAR  
19 unsuitable. For this reason, the specification pursued here relies on a direct representation of the  
20 correlation structure. 100 randomly selected samples of 1,000 observation units (pixels) were  
21 used, rather than the roughly 3 million pixels available in the satellite images for the Austin  
22 region.  
23

## 24 **Data Description**

25 The data used for the following analysis comes from two sources: 1) land cover  
26 information derived from satellite images and 2) Census of Population data. These two datasets  
27 are processed and combined to the same spatial and temporal coordinates.  
28

29 The satellite images used for deriving land cover information come from Landsat 4, 5 and 7  
30 systems and cover the urban area of Austin, Texas. (See Frazier and Kockelman (2005) for more  
31 details.) The processing of remote sensing data is computationally intensive work, so only four  
32 years' data were available, for 1983, 1991, 1997 and 2000. The data was originally classified by  
33 students supervised by Dr. Barbara Parmenter at the University of Texas of Austin, using  
34 ERDAS Imagine's ISODATA classifier (Jensen, 1996). For each year, the dataset covers a 48.5  
35 km x 55.8 km area that contains around 3 million 30 m x 30 m grid cells. Each of these cells is  
36 indexed as one of the nine land-cover types: water, barren, forest/woodland, shrubland,  
37 herbaceous natural/semi-natural, herbaceous planted/cultivated, fallow, residential, or  
38 commercial/industrial/transportation. To focus on human development of land, the nine cover  
39 types were aggregated into three categories: undeveloped, residential and  
40 commercial/industry/transportation.

41 Since pixels were classified based on their spectral qualities (i.e., reflected light), rather  
42 than information of how humans actually "use" the land, the terms residential and  
43 commercial/industrial/transportation are best interpreted as lands developed at low and high  
44 intensities, respectively (i.e., exhibiting only moderate or little to no vegetation). Therefore, the  
45 term "land cover" will be used throughout this paper, rather than "land use". A comparison of these land  
46 cover designations, aerial photos and parcel-based land use data suggests that the classes generally  
47 correspond to residential or commercial/industrial/transportation uses, respectively. However, residential  
48 land is sometimes coded as intensely developed (if the entire 30 m cell is covered by concrete or roofing  
49 materials, without any vegetation). Moreover, land visibility is affected by a number of factors,  
50 including cloud cover and humidity. As a result, image classification techniques are imperfect  
51  
52

1  
2  
3 and errors can (and do) occur in dataset overlays. For these and possibly other reasons, about 5%  
4 of the region's 3 million grid cells present unrealistic changes. (For example, one cell's  
5 classification changes from undeveloped to residential in the year 1991 but changes back to  
6 undeveloped in the following years). To neutralize this data imperfection, the model allows land  
7 to change from developed to undeveloped, recognizing the uncertainty inherent in the  
8 classifications. If desired, restrictions could be imposed in the model's specification, in order to  
9 prevent such land cover changes.

10 The U.S. Census of Population also provided key data, in years 1990 and 2000. The typical  
11 census block encompasses dozens of 30 m grid cells. In order to combine the datasets (across  
12 space and time), each block's population was assumed to increase exponentially with time,  
13 according to its own growth parameter. Block populations outside the decennial years (1990 and  
14 2000) were interpolated/extrapolated, and these inter-decennial block group populations were  
15 distributed uniformly across all residentially-classified cells within the block.

16 The Euclidean distance to the nearest highway and that to the central business district  
17 (CBD) also were computed for each grid cell (and then controlled for in the land cover change  
18 models). These distances were computed in TRANSCAD (Caliper, 2004) based on a network  
19 composed of Interstate, U.S., and state highways. The network in the study area contains a total  
20 length of 199 centerline miles.  
21

## 22 **Methodology**

23  
24 As discussed in the literature review, land use/land cover change is discrete in nature.  
25 Meanwhile, to incorporate the temporal and spatial correlations across observations units, a  
26 specification more flexible than the basic MNL may be required. This leads to the choice of  
27 mixed (or random parameter) logit model, in which the parameters are assumed to be random  
28 across individual cells. Furthermore, the total number of observations in this study is too large  
29 and a set of randomly selected samples has to be used, which abolishes the necessary condition  
30 for using SAR. Therefore, the direct representation method is used to deal with the spatial  
31 correlation across observations. The following sections describe the model specification and the  
32 related estimation considerations, including generalized Cholesky decomposition and the choice  
33 of simulation methods.

## 34 **Model Specification**

35  
36 The development of urban land is essentially the outcome of decisions made by households,  
37 developers and planners, under a fair amount of uncertainty. It is helpful to apply random utility  
38 theory, in which the land cover type of a cell is determined by the following equation:

$$39 \quad y_{it}^m = \begin{cases} 1 & \text{if } U_{it}^m > U_{it}^n \quad \forall n \in M \\ 0 & \text{else} \end{cases} \quad (1)$$

40  
41  
42 where  $y_{it}^m$  is an indicator of observed land cover type  $m$  for cell  $i$  in time period  $t$ ;  $U_{it}^m$  is its  
43 latent variable (utility) and  $M$  is the set for all land cover alternatives.  $m$  (or  $n$ ) =1 denotes  
44 undeveloped land, 2 denotes residential (i.e., less intensively developed land), and 3 denotes  
45 commercial/industrial/transportation (i.e., more intensively developed land).  $t=0$  in 1983,  $t=1$  in  
46 1991,  $t=2$  in 1997, and  $t=3$  in 2000.

47 As usual, the latent variable  $U$  is composed of two parts: a deterministic part  $V$  and a  
48 random part  $\varepsilon$ :

$$49 \quad U_{it}^m = V_{it}^m + \varepsilon_{it}^m \quad (2)$$

50  
51  
52

In order to allow for both temporal and spatial correlations across observations, the random part  $\varepsilon$  is assumed to have two independent components:

$$\varepsilon_{it}^m = u_i^m + \xi_{it}^m \quad (3)$$

where  $u_i^m$  is an individual-specific random effect (constant over time) and  $\xi_{it}^m$  is assumed to be iid Gumbel across land cover types and observations. This results in the logit model:

$$P(y_{it}^m = 1 | x_{it-1}, u_i^m; \pi_{it}^m) = \frac{\exp(x_{it-1} \pi_{it}^m + u_i^m)}{1 + \sum_{n=2}^M \exp(x_{it-1} \pi_{it}^n + u_i^n)} \quad (4)$$

where  $\pi_{it}^m$  and  $x_{it-1}$  are the vectors of parameters and explanatory variables comprising  $V_{it}^m$ , respectively. Based on data availability and inspired by previous studies, the explanatory variables include distance to the nearest highway  $d_{HW,t-1}$ , distance to the CBD  $d_{CBD,t-1}$ , neighborhood population, and neighborhood land cover types. That is,

$$V_{imt} = \beta_t^{m,HW} d_{HW,t-1} + \beta_t^{m,CBD} d_{CBD,t-1} + \sum_{j \in S_i} \rho_{ijt}^{m,POP} POP_{jt-1} + \sum_{n=2}^M \sum_{j \in S_i} \rho_{ijt}^{m,n} y_{jt-1}^n \quad (5)$$

In other words,  $x_{it-1} = (d_{HW,t-1}, d_{CBD,t-1}, (POP_{jt-1}, y_{jnt-1})_{j \in S_i})$  and  $\pi_{it}^m = (\beta_t^{m,HW}, \beta_t^{m,CBD}, \rho_{ijt}^{m,POP}, \rho_{ijt}^{m,n})$ , where  $POP_{jt-1}$  is the population in cell  $j$  at time  $t-1$  and  $y_{jt-1}^n$  indicates the land cover type of cell  $j$  at time  $t-1$  (e.g.,  $y_{jt-1}^2 = 1$  if cell  $j$  is residential land at time  $t-1$ , and 0 otherwise).  $S_i$  is the 6.03 km x 6.03 km square neighborhood centered on cell  $i$  (which consists of 201x201, or 40,401 cells).

The fact that time  $t-1$  is used in this model instead of time  $t$  helps ensure exogeneity of control variables. In this way, each neighborhood's control-variable values can be preprocessed and stored. Thus, unlike Frazier and Kockelman's (2005) and Miyamoto's (2004) methods, the necessary neighborhood information is not lost through sampling. Finally, because the time intervals between the panel's four data years differ, a time adjustment coefficient is included in all cases. The coefficients  $\rho_{ijt}^{POP}$  and  $\rho_{ijt}^n$  also are assumed to be related to distance, via the following expression:

$$\rho_{ijt}^{m,POP} = \frac{\theta_0^{m,POP} + \theta_1^{m,POP} \Delta t}{(d_{ij})^2} \quad (6)$$

where  $\Delta t$  is the time interval between two periods (e.g.,  $\Delta t = 1997 - 1991$ , or 6 years, for  $t=2$ ); and  $d_{ij}$  is the distance between cells  $i$  and  $j$ , using the pixel coordinates as scale (i.e., 1 unit is 30m and the distance between cell  $i$  and its furthest "neighbor" is  $\sqrt{2} \times 200$ ).  $(d_{ij})^2$  was chosen as the denominator because it yielded the highest likelihood ratio among  $d_{ij}$ ,  $(d_{ij})^2$ ,  $\exp(d_{ij})$  and  $\exp((d_{ij})^2)$ .

From Equation (5), one finds that  $y_{it-1}$  serves as an explanatory variable, which means this model is in fact a dynamic, unobserved-effects model. As Wooldrige (2002) discusses, in such a model, treating  $u_i^m$  as a single parameter does not result in consistent estimators. Instead, the unobserved effect should be integrated out by proposing a density for  $u_i^m$ , given  $y_{i0}$ . Here,  $u_i^m$  is assumed to be normally distributed with mean  $\alpha_{0m} + \alpha_{1m} (y_{i0}^{m=2} + y_{i0}^{m=3})$  and a standard deviation  $\sigma_m$ :

$$u_i^m = \alpha_0^m + \alpha_1^m (y_{i0}^{m=2} + y_{i0}^{m=3}) + \sigma^m \zeta_i^m \quad (7)$$

In Equation (7), if  $\zeta_i^m$  is iid across cells, this model is the same as an MNL model for panel data, with random effects (Greene, 2000). However, in this study this random effect also needs to capture the spatial correlation among cells, which means the covariance matrix of  $\zeta_i^m$  is no longer an identity matrix.

As mentioned earlier, the method of direct representation is used here, in order to avoid the problem caused by sampling. This also has the advantage of making the relationships between spatial correlation and distance more obvious and ensuring stationarity of the covariance matrix. Here, the correlation between cells  $i$  and  $j$  is assumed to be inversely proportional to the distance between them:

$$\text{corr}(\zeta_i^m, \zeta_j^m) \propto (d_{ij})^{-1} \quad (8)$$

Several other correlation patterns also were examined (e.g.,  $(d_{ij})^{-1.5}$  and  $(d_{ij})^{-2}$ ), but  $(d_{ij})^{-1}$  offered the highest log likelihood value, as well as more intuitive coefficients.

Overall then, the model's final likelihood function is as follows:

$$\ln L_{\pi} = \sum_{i=1}^N \ln L_{i|\pi} = \sum_{i=1}^N \sum_{t=1}^T \sum_{m=1}^M y_{it}^m \ln \left( P(y_{it}^m = 1 | x_{it-1}; u_i^m, \pi_i^m) f(u_i^m | y_{i0}, \zeta_i^m) \right) \quad (9)$$

which can be estimated using MSLE (as described by Greene [2000], Bhat[2001] and Train [2003]). The random variable values  $u_i^m$  were generated according to Equations (7) and (8).

## Simulation Methods and Other Computational Issues

Simulation can be computationally intensive work. In order to recognize the correlation across observations in time and space, one must generate 2 (for  $m=2,3$ )  $N$ -dimensional normally distributed random vectors with the correlation matrix given by Equation (8). Computational issues encountered in the simulation process, and the methods to address these, are discussed below.

### Generalized Cholesky Decomposition

Normally, the multidimensional normal distribution is generated using a Cholesky decomposition, which decomposes the covariance matrix into a triangular matrix. The triangular matrix is then multiplied by  $N$  iid normal values in order to obtain the required  $N$ -dimensional normal vector. However, in the study of spatial effects, the correlation matrix between observations can only be guaranteed to be symmetric, not positive definite, which can make it impossible to apply a Cholesky decomposition. In this study, 60 of the 100 samples did not offer

1  
2  
3 positive definite covariance matrices. Therefore, Gill and King's (2004) generalized Cholesky  
4 decomposition method was used. In this method, when a symmetric matrix is not positive  
5 definite, a nonnegative diagonal matrix with element values as small as possible is added to the  
6 original matrix, in order to produce a positive definite matrix. Gill's (2004) GAUSS code was  
7 modified and integrated into Train's MSLE code (1999). The method performed well with the  
8 samples in this study: the largest element in the added diagonal matrix is less than 0.01 and 100  
9 sample estimates (rather than just 40) resulted. With generalized Cholesky decomposition  
10 providing the pseudo triangular matrices needed, the MSLE focus turned to generating  $N$  iid  
11 normally distributed random values.

## 12 **Randomly Shuffled Halton Sequence**

13  
14 Four commonly used methods for drawing random numbers were considered for the  
15 simulation: pseudo Monte Carlo (PMC), Halton sequence, scrambled low discrepancy (SLD)  
16 method, and randomly shuffled Halton sequence. While PMC has been most broadly used, in  
17 recent years researchers have sought better efficiency via quasi Monte Carlo methods (such as  
18 Halton draws), in order to cover the number space more uniformly. (Bhat [2001] and Train  
19 [2003]). However, the correlation between successive Halton draws results in a very periodic  
20 correlation across observations that may not be desirable. (Usually, when the correlation across  
21 observations is not specified, Halton draws are favored for low dimension integration because  
22 correlation patterns cause a relatively low asymptotic variance of coefficients and the mean  
23 squared error is still less than that from PMC. However, when using such draws to generate  
24 multivariate normal distributions, it is very likely that this potential correlation may be confused  
25 with the specified correlation pattern.)

26 Another popular method, the SLD approach evolves from a standard Halton sequence. The  
27 basic idea of this method is that instead of radical inverse function for the base prime, this  
28 method uses a scrambled radical function. Many have suggested different scrambling methods –  
29 for example Faure (1992), Tuffin (1996), Braaten and Weller (1979), Atanassov (2004) and  
30 Mascagni and Chi (2004). However, successive draws from the same scrambled sequence still  
31 generate periodic correlation. Figure 1 compares these methods. (Each observation uses 250  
32 numbers in sequence based on the prime number 17. For SLD, Faure's (1992) permutation was  
33 used.) Figure 1 shows that when the draws come from the same sequence, the Halton and its  
34 scrambled version make the first observation periodically correlated with all successive  
35 observations (shown through 40 here). The highest correlation between observations reaches 1,  
36 at multiples of 17.

37 It seems that to generate uncorrelated random terms, in order to avoid introducing  
38 unintended covariance in the model, the PMC method may be best among these. However, PMC  
39 methods lose the efficiency given by the coverage of Halton draws. To avoid correlation  
40 problems yet maintain the coverage of a Halton draw, one may wish to treat each observation as  
41 a dimension (rather than as a separate draw). Unfortunately, when the dimension increases, most  
42 Scrambled Low Discrepancy sequences are hard to generate because the permutation requires a  
43 higher dimension matrix. (For example, with Faure's permutation method, if 1000 dimensions  
44 are required, a 7917 x 7917 triangular matrix will have to be used for calculating the  
45 permutation.) In light of all these considerations, a random shuffled Halton sequence (Hess et al.,  
46 2003) was finally chosen. The idea of generating a random shuffled Halton sequence is very  
47 simple: (1) Generate Halton sequence to cover the space; (2) For each dimension, randomly  
48 order the numbers in the original Halton sequence. In this way, the numbers still cover the space  
49 uniformly so it should be more efficient than PMC. Moreover, this method requires less memory  
50 and computation time than a scrambled Halton sequence. The correlation between observations  
51  
52



1  
2  
3 also is very low <sup>1</sup>, as shown in Figure 1. (The original Halton sequence contains 250 draws based  
4 on prime number 17.)

5 In this study, each sample uses 1000 observations. Hence, 2000 shuffled Halton sequences  
6 (m = 2, 3 for each observation) were generated. The original Halton sequence contains 300  
7 draws based on the prime number 3.<sup>2</sup> The model's estimation proves that this random shuffled  
8 Halton sequence is more efficient than PMC. (Ten samples were run using these two different  
9 methods, based on the same convergence criteria, and only one PMC sample converged with  
10 fewer iterations.)

## 11 Results

12 As mentioned, the study area contains about 3 million cells. However, cells around the  
13 region's edge are not usable since they lack information on their 6.03km x 6.03km  
14 neighborhoods. This results in a population size of 2.3 million cells, which is still too large for  
15 model estimation. Instead, this study uses cell sampling to reduce the burden: 100 samples were  
16 randomly selected, each containing 1000 cells. Of these 100 samples, 64 converged while others  
17 resulted in singular Hessian matrices for their likelihood functions. One of the reasons for  
18 causing these singular Hessian matrices is that the commercial/industrial/transportation land  
19 accounts for only 8 percent of the whole area. In some samples, the proportion of developed land  
20 may be too small for model estimation.

21 The coefficients and t-statistic values shown in Table 1 are the mean values of the estimates  
22 from these 64 samples. As Greene (2000) indicates, the means of sample coefficients are  
23 consistent estimates of the corresponding coefficients for the whole population. Though means  
24 of the t-statistics are not consistent estimates of the population's t-statistic values, they still  
25 provide some information of the significance level across samples.

26 Elasticities of land cover change probabilities were computed for 100,000 randomly  
27 selected observations using average parameter values, as follows:

$$28 \text{Elasticity}(X_{i,k}) = \frac{\partial P(Y_i = 1 | X_i)}{\partial X_{i,k}} \cdot \frac{X_{i,k}}{P(Y_i = 1 | X_i)} \quad (10)$$

29 The resulting elasticities were averaged, using land cover probabilities as weights (as in  
30 Hensher and Johnson (1981)). The results of these elasticity calculations are discussed below.

31 The neighborhood's land cover type is estimated to have a strong effect on future  
32 development. As the share of neighboring land cover becomes more residential or  
33 commercial/industrial/transportation, so does the likelihood of that cell's future cover. Moreover,  
34 these effects are estimated to increase with time. They suggest a strong clustering effect in land  
35 development, as discussed by Fujita (1983), especially for commercial/industrial/transportation  
36 land, since the estimated elasticity for this variable is more than 9. Everything else constant,  
37 however, neighboring population densities reduce the likelihood of such transitions, particularly  
38 for commercial/industrial/transportation land, with an elasticity estimate of nearly -3. Higher  
39 population densities may imply higher land prices as well as stronger zoning (and neighborhood  
40  
41  
42  
43  
44  
45  
46

47 <sup>1</sup> If the original Halton sequence has R draws (normally, R is at least 100), there are R! possible ways to shuffle the  
48 sequence. Even when there are millions of observations, there is still little chance of seeing high correlation between  
49 observations because  $10^6 \ll 100!$ .

50 <sup>2</sup> With fewer than 300 sample draws, sequences based on other relatively small prime numbers were found to leave  
51 large, uncovered spaces. Thus, such sequences were avoided here.

1  
2  
3 resistance to development) in order to preserve what undeveloped space remains. These results  
4 suggest that development happens sooner in neighborhoods that are less populated but more  
5 developed.

6 As expected, as the distance to the nearest highway increases, the probability of  
7 development falls, though this effect diminishes over time (i.e., as the time interval increases).  
8 The results also suggest that development is more likely to emerge away from the CBD, where  
9 land development restrictions are likely to be fewer, land values lower, and construction costs  
10 lower. Evidently (and somewhat intuitively), infill is a slower process than green-field  
11 development.

12 Finally, the results show that it is necessary to consider the individual-specific error terms  
13 associated with each cell; the mean and the standard deviation on both alternatives are highly  
14 statistically (& practically) significant. The standard deviation for residential land is somewhat  
15 smaller than that of commercial land, indicating lesser variability in development potential.

## 17 Prediction and Validation

18  
19 Prediction using the calibrated coefficients can provide helpful insights into the future land  
20 cover patterns. In the case of unequal time intervals, this dynamic model's explicit accounting  
21 may produce more reliable results. Prediction into year 2005 and the consequent comparison  
22 with actual land cover information also provide a credible way for evaluating model performance.

23 A 2.1km x 2.1km squared area (i.e., a sample of 70 x 70 = 4900 cells) in north Austin was  
24 selected as an example site for model application. The correlation matrix  $\Omega$  was first generated  
25 based on the distance between cells, resulting in a 4900x4900 matrix. This was decomposed  
26 using generalized Cholesky methods. And 9800 randomly shuffled Halton sequences (each  
27 containing 300 values) were generated and multiplied by the triangular matrix, producing the  
28 random vectors  $\zeta_i^m$  ( $i = 1, 2, \dots, 4900; m = 2, 3$ ).

29  
30 As mentioned above, in order to validate the model, the prediction year was selected to be  
31 the current year (2005), and the explanatory variables come from the year 2000 values (resulting  
32 in a  $\Delta t = 5$  year time interval). For each draw and for each cell, the utilities of all three land cover  
33 types were calculated, and the type offering the highest utility became the cell's 2005 predicted  
34 land cover. 300 such draws were made, and the most common land cover type for each cell is  
35 shown in Figure 2.

36  
37 Of course, Figure 2's "most likely" pattern will not occur with a high likelihood. The  
38 predictions are really a set of 300 simulations' averages, or highest likelihoods, so more dramatic  
39 clustering and the lack of a fine-grained diversity in cover is expected, as shown. In fact, there is  
40 great flexibility and uncertainty in the future of this complex site, with its 4900 cells. It is very  
41 helpful for planners to quantify and visualize such uncertainty. As in SLEUTH (Candau et al.  
42 2000), uncertainty can be calculated based on the predicted frequencies of different land covers  
43 taking hold. This study relies on an entropy statistic to quantify predictive uncertainty. (See, e.g.,  
44 McKay [1995] and Kotz and Johnston [1982].) That is, the uncertainty associated with the set of  
45 potential land covers in cell  $i$  is formulated as follows:

$$46 \text{uncertainty}_i = \frac{-1}{\ln(3)} \sum_{m=1}^3 P_{im} \ln(P_{im}) \quad (11)$$

47  
48  
49 This formulation generates a value between 0 and 1 for each cell. The higher the value, the  
50 more uncertain is the prediction for that cell. When all three future land cover alternatives have  
51  
52

1  
2  
3 equal probabilities ( $P_{im} = 0.33 \forall m$ ), the uncertainty index equals 1, indicating maximum  
4 uncertainty. When the land cover type remains the same for all 300 simulations, this index equals  
5 0. The land cover uncertainty associated with this site and this set of predictions can be  
6 visualized, as shown in Figure 2. Higher uncertainty appears at the border of two different land  
7 cover types and decreases with land cover uniformity.

8  
9 The 2005 predictions were compared to detailed aerial photos and parcel-based land use  
10 maps of the site, as a form of validation. Figure 2 shows the 2000 data alongside the actual 2005  
11 land cover classes. It seems the 2000 data were 25% misclassified, resulting in most of the “mis-  
12 prediction” evident in the 2005 image. Partly as a result, undeveloped land was seriously under-  
13 predicted while lightly and intensely developed lands were over-predicted by 17% and 26%,  
14 respectively. This does not mean that the model is not performing well, but rather, the start data  
15 was not as reliable as it should be. Improvements in satellite image quality and data classification  
16 are expected to resolve some of these issues. Anyone wishing to project forward can classify the  
17 area of interest manually, and then apply the model, in order to avoid such issues with current  
18 classifications. Another possible reason for over-prediction of 2000-2005 land development is  
19 the economic slowdown that began in 2000, particularly for the Austin region. With model  
20 parameters calibrated using 1983 through 2000 data, the estimated rate of development is  
21 probably faster than the actual speed between 2000 and 2005, contributing to the over-prediction.

## 22 **Conclusions**

23  
24 Satellite imagery presents a promising future to those interested in land cover change  
25 dynamics. However, the data points are profuse, land cover classifications are discrete, and  
26 spatial and temporal effects are fundamental. For these reasons, new specifications and  
27 computational methods are very much needed. In light of these challenges, this work undertakes  
28 several relatively unusual techniques using a series of Austin, Texas images. Sampling of  
29 observations is central to these techniques. Halton sequences are randomly shuffled and  
30 Cholesky decomposition of non-positive definite matrices is achieved. Likelihood values are  
31 simulated (for parameter estimation via maximization), as are images of Austin’s future. Lagged  
32 variables are key to the model’s specification, as is explicit recognition of unequal time intervals  
33 between each of the three satellite images. A series of neighborhood variables, computed for  
34 each cell in turn, control for population and local development. And uncertainties in predictions  
35 are quantified via an entropy statistic.

36 Model results suggest strong clustering effects, by developed land cover type. After  
37 controlling for neighboring land covers, future development is most likely in areas offering lower  
38 residential densities. Model application to a 4900-cell area in the year 2005 confirms these  
39 results, via simulation of 300 futures for each cell. Uncertainty indices (i.e., entropy terms)  
40 indicate greatest land development uncertainty at the “edges” of distinct land cover types where  
41 land cover transition tends to occur.

42 It is hoped that the methods and conclusions of this work will offer valuable tools and  
43 insights for researchers, planners, policy makers, developers and others seeking to predict the  
44 future. New techniques should allow us to tap the potential of satellite databases.

45  
46  
47  
48  
49  
50  
51  
52

## Acknowledgements

The authors are grateful to Kenneth Train (who also shared his GAUSS code for the mixed logit model) and Chandra Bhat for providing useful suggestions regarding the simulation techniques, and to Annette Perrone for editorial assistance.

This research is sponsored by the Environmental Protection Agency (EPA) under its STAR grant program. The full project is titled "Predicting the Relative Impacts of Urban Development Policies and On-Road Vehicle Technologies on Air Quality in the United States: Modeling and Analysis of a Case Study in Austin, Texas."

## References

- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Press.
- Anselin, L. (1999). *Spatial Econometrics*. Working paper. Accessed July 10, 2005: [http://www.csiss.org/learning\\_resources/content/papers/baltchap.pdf](http://www.csiss.org/learning_resources/content/papers/baltchap.pdf).
- Atanassov, E.I. (2004). "On the discrepancy of the Halton sequences". *Mathematica Balkanize*, 18(1-2): pp. 15-32.
- Beron, K. and Vijverberg, W. (2004). "Probit in a spatial context: a Monte Carlo analysis". In L. Anselin, R. Florax and S. Rey (eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag. Ch. 8.
- Besner C. (2002) "A Spatial Autoregressive Specification with a Comparable Sales Weighting Scheme". *Journal of Real Estate Research* 24: pp. 193-211.
- Bhat, C. R. (2001) "Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model", *Transportation Research: Part B*, 35: pp. 677-693.
- Braaten, E. and Weller, G. (1979). "An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration". *Journal of Computational Physics*, 33: pp. 249-258.
- Caliper Corporation. (2004). TransCAD GIS Software, version 4.7. Caliper Corporation, Newton, Massachusetts.
- Candau, J., Rasmussen, S. and Clarke, K. C. (2000). "A coupled cellular automaton model for land use/land cover dynamics". Accessed July 10, 2005: <http://www.geog.ucsb.edu/~kclarke/ucime/banff2000/533-jc-paper.htm>.
- Cliff, A.D. and Ord., J. K. (1981). *Spatial Processes: Models & Applications*. London: Pion.
- Clarke, K.C., Hoppen and S., Gaydos, L. (1996). "Methods and techniques for rigorous calibration of a cellular automaton model of urban growth". Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January 21-25, 1996.
- Clarke, K. C. (1997). "Land transition modeling with deltatrons". Paper presented at the NCGIA Land Use Modeling Workshop. Accessed July 10, 2005: <http://www.ncgia.ucsb.edu/conf/landuse97>.
- de la Barra, T. (1989). *Integrated Land Use and Transport Modeling*. Cambridge: Cambridge University Press.

- 1  
2  
3 Faure, H. (1992). "Good permutations for extreme discrepancy". *Journal of Number Theory*, 42:  
4 pp. 47-56.
- 5 Frazier, C. and Kockelman, K. (2005). "Spatial econometric models for panel data: Incorporating  
6 spatial and temporal data". Forthcoming in *Transportation Research Record*. Accessed July 10,  
7 2005: [http://www.ce.utexas.edu/prof/kockelman/public\\_html/TRB05satdata.pdf](http://www.ce.utexas.edu/prof/kockelman/public_html/TRB05satdata.pdf)  
8
- 9 Fujita, M.(1989). *Urban Economy Theory*. Cambridge University.
- 10 Gill, J. and King, G.. (2004) "What to do when your Hessian is not invertible: Alternatives to  
11 model respecification in nonlinear estimation". *Sociological Methods and Research*, 32(1): pp.  
12 54-87.
- 13 Gill, J. and King, G.. (2004). "Generalized Cholesky decomposition". GAUSS code. Accessed  
14 July 10, 2005: <http://psblade.ucdavis.edu/papers/gmchol.g>  
15
- 16 Greene, W. (2000). *Econometric Analysis*. Upper Saddle River: Prentice-Hall.
- 17 Hensher, D. A. and Johnson, L. W. (1981). *Applied Discrete-Choice Modelling*, John Wiley &  
18 Sons.
- 19 Hess, S., Polak, J.W. and Daly, A.J. (2003). "On the performance of shuffled Halton sequences  
20 in the estimation of discrete choice models". Presented at the European Transport Conference,  
21 Strasbourg, France.
- 22 Hunt, J.D. and Abraham, J.E. (2003). "Design and application of the PECAS land use modeling  
23 system". *Proceedings of the 8th International Conference on Computers in Urban Planning and*  
24 *Urban Management*, Sendai, Japan.
- 25  
26 Jensen, R.J. (1996). *Introductory Digital Image Processing*. Prentice Hall.
- 27  
28 Klosterman, R. E. (1999). "What if?: Collaborative Planning Support System". *Environment and*  
29 *Planning B* 26: pp.393-408.
- 30 Kotz, S. and Johnson, N. (1982). *Encyclopedia of Statistical Sciences*. John Wiley & Sons, New  
31 York.
- 32 Mascagni, M. and Chi., H. (2004). "On the scrambled Halton sequence". *Monte Carlo Methods*  
33 *and Applications* 10(3): pp. 435-442.
- 34 McFadden, D. and Train, K. (2000) "Mixed MNL models for discrete response". *Journal of*  
35 *Applied Econometrics*, 15(5): pp. 447-470.
- 36  
37 McKay, M. D. (1995). "Evaluating prediction uncertainty". Technical Report NUREG/CR-  
38 6311, U.S. Nuclear Regulatory Commission and Los Alamos National Laboratory.
- 39  
40 Miyamoto, K., Vichiensan, V.; Shimomura, N. and Paez, A. (2004). "Discrete choice model with  
41 structuralized spatial effects for location analysis" *Transportation Research Record* 1898: pp.  
42 183-190.
- 43 Mohammadian, A. K., Haider, M. and Kanaroglou, P. S. (2004) "Incorporating spatial  
44 dependencies in random parameter discrete choice models" Presented at 84th Annual  
45 Transportation Research Board Meeting. Accessed July 10, 2005:  
46 [http://www.milute.mcgill.ca/Research/Senior/Spatial-Mixed-TRB\\_rev.pdf](http://www.milute.mcgill.ca/Research/Senior/Spatial-Mixed-TRB_rev.pdf)  
47
- 48 Munroe, D., Southworth, J. and Tucker, C. M. (2001). "The dynamics of land-cover change in  
49 western Honduras: Spatial autocorrelation and temporal variation". *Conference Proceedings*.
- 50  
51  
52

- 1  
2  
3 *American Agricultural Economics Association*. AAEA-CAES 2001 Annual Meeting. Accessed  
4 July 10, 2004: [http://agecon.lib.umn.edu/cgi-bin/pdf\\_view.pl?paperid=2611](http://agecon.lib.umn.edu/cgi-bin/pdf_view.pl?paperid=2611)
- 5 Nelson, G. C., and Hellerstein, D. (1997). "Do roads cause deforestation: Using satellite images  
6 in econometric analysis of land use". *American Journal of Agricultural Economics* 79: pp. 80-88.
- 7  
8 Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J. and Deadman, P. (2003). "Multi-  
9 agent systems for the simulation of land-use and land-cover change: A review". *Annals of the*  
10 *Association of American Geographers*, 93(2): pp. 314-340.
- 11 Train, K. (2003). *Discrete Choice Methods with Simulation*. New York: Cambridge University  
12 Press.
- 13 Train, K., Revelt, D. and Ruud, P. (1999). "Mixed logit estimation routine for panel data".  
14 GAUSS code. Accessed July 10, 2005:  
15 <http://elsa.berkeley.edu/Software/abstracts/train0296.html>
- 16  
17 Tuffin, B. (1996), "On the use of low-discrepancy sequences in Monte Carlo methods," *Monte*  
18 *Carlo Methods and Applications*, 2: pp. 295-320.
- 19 Waddell, Paul. (2002). "UrbanSim: Modeling urban development for land use, transportation,  
20 and environmental planning". *The Journal of the American Planning Association* 68(3): pp. 297-  
21 314.
- 22  
23 Wear, D. N., and Bolstad, P. (1998). "Land-use changes in southern Appalachian landscapes:  
24 Spatial analysis and forecast evaluation." *Ecosystems* 1: pp. 575-594.
- 25  
26 Wooldrige, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press,  
27 Cambridge.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

## LIST OF TABLES

TABLE 1 Variable Summary Statistics for All Variables

TABLE 2 Model Estimation Results

## LIST OF FIGURES

FIGURE 1 Correlation between Observations Generated by Different Simulation Methods

FIGURE 2 Predicted Land Cover and Predictive Uncertainty for Year 2005

**Table 1. Summary Statistics for All Variables**

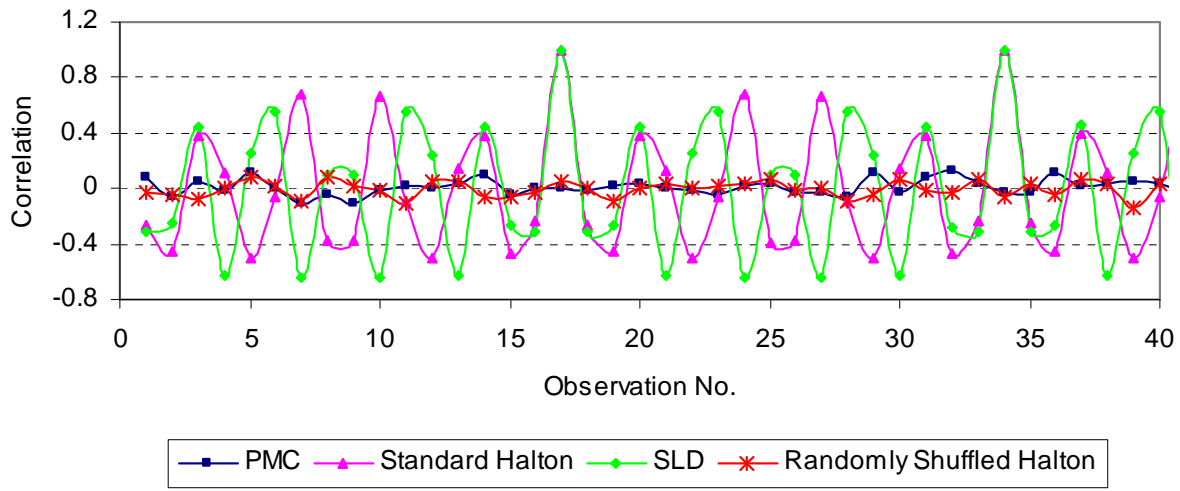
Variable	Year	Variable Statistics			
		Min	Max	Mean	S.D.
<i>Dependent Variables</i>					
Indicator for residential (1 if residential; 0 otherwise) $y_{it}^{m=2}$	1983	0	1	0.052	0.223
	1991	0	1	0.138	0.345
	1997	0	1	0.216	0.412
	2000	0	1	0.237	0.425
Indicator for commercial/industrial/transportation (1 if commercial, industrial or transportation; 0 otherwise) $y_{it}^{m=3}$	1983	0	1	0.089	0.285
	1991	0	1	0.080	0.271
	1997	0	1	0.087	0.281
	2000	0	1	0.106	0.308
<i>Explanatory Variables</i>					
Distance to the nearest highway (km) (constant over time) $d_{HW,t-1}$	---	0.020	20.60	5.088	3.963
Distance to CBD (km) (constant over time) $d_{CBD,t-1}$	---	0	42.65	20.24	9.414
Weighted sum of the neighborhood's population $\sum_{j \in S_i} \frac{POP_{jt-1}}{d_{ij}^2}$	1983	0	230.8	4.977	11.13
	1991	0.033	384.7	8.567	15.19
	1997	0.117	447.4	11.560	18.81
	2000	0.045	590.2	12.797	20.01
Weighted sum of the neighborhood's residential land $\sum_{j \in S_i} \frac{y_{jt-1}^{m=2}}{(d_{ij})^2}$	1983	0	25.34	1.760	3.75
	1991	0.094	27.04	4.560	4.952
	1997	0.305	25.13	7.209	5.505
	2000	0.077	28.44	7.889	7.001
Weighted sum of the neighborhood's commercial/industrial/transportation land $\sum_{j \in S_i} \frac{y_{jt-1}^{m=3}}{(d_{ij})^2}$	1983	0.027	26.92	2.941	4.084
	1991	0.007	26.15	2.678	4.212
	1997	0.020	26.05	2.855	3.735
	2000	0.027	27.35	3.555	4.599

Note: Neighborhood is defined as a 6.03 km x 6.03 km square area centered on the observation, which consists of 201x201, or 40,401 cells.

**Table 2. Model Estimation Results**

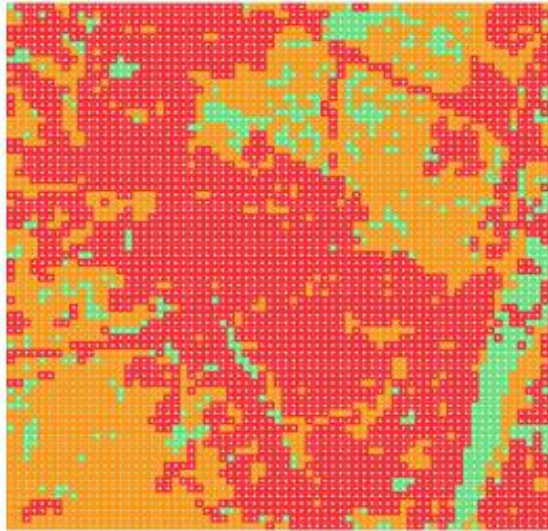
Parameters	Coef.	t-stat.	Elasticity	Coef. Stat. across 100 Samples			
				Max	Min	S.D	
<i>Attributes for Residential Land Cover (m=2)</i>							
Distance to the nearest highway $\beta^{m,HW} = \beta_0^{m,HW} + \beta_1^{m,HW} \Delta t$	$\beta_0^{m,HW}$	-1.035E-02	0.768	-0.043	1.390E-01	-1.569E-01	6.533E-02
	$\beta_1^{m,HW}$	5.801E-03	0.891	0.140	2.669E-02	-3.447E-02	1.278E-02
Distance to CBD $\beta^{m,CBD} = \beta_0^{m,CBD} + \beta_1^{m,CBD} \Delta t$	$\beta_0^{m,CBD}$	2.079E-02	0.946	0.359	5.993E-02	-1.555E-02	1.843E-02
	$\beta_1^{m,CBD}$	5.954E-04	0.567	0.060	7.726E-03	-7.546E-03	3.401E-03
Population in the neighboring cell $\rho_{ijt}^{m,POP} = \frac{\theta_0^{m,POP} + \theta_1^{m,POP} \Delta t}{(d_{ij})^2}$	$\theta_0^{m,POP}$	-3.528E-02	1.563	-0.159	1.349E-02	-8.190E-02	1.884E-02
	$\theta_1^{m,POP}$	-1.005E-02	1.976	-0.253	-4.154E-04	-2.146E-02	4.929E-03
Neighboring cell is residential (n=2) $\rho_{ijt}^{m,n} = \frac{\theta_0^{m,n} + \theta_1^{m,n} \Delta t}{(d_{ij})^2}$	$\theta_0^{m,n}$	5.096E-01	6.947	1.093	6.814E-01	3.215E-01	7.582E-02
	$\theta_1^{m,n}$	2.490E-02	2.048	0.248	5.520E-02	-1.385E-03	1.329E-02
Neighboring cell is commercial/industrial/transportation (n=3) $\rho_{ijt}^{m,n} = \frac{\theta_0^{m,n} + \theta_1^{m,n} \Delta t}{d_{ij}^2}$	$\theta_0^{m,n}$	1.932E-01	1.954	0.241	4.229E-01	-6.462E-02	1.086E-01
	$\theta_1^{m,n}$	1.406E-02	1.084	0.106	5.626E-02	-3.412E-02	1.797E-02
Individual Specific Effect $u_i^m = \alpha_0^m + \alpha_1^m (y_{i0}^{m=2} + y_{i0}^{m=3}) + \sigma^m \zeta_i^m$	$\alpha_0^m$	-6.860	11.565	---	-5.514	-9.383	5.750E-01
	$\alpha_1^m$	-6.525E-02	0.689	---	1.213	-3.615	6.422E-01
	$\sigma^m$	1.380	5.614	---	2.065	0.828	2.212E-01
<i>Attributes for Commercial/Industrial/Transportation Land Cover (m=3)</i>							
Distance to the nearest highway $\beta^{m,HW} = \beta_0^{m,HW} + \beta_1^{m,HW} \Delta t$	$\beta_0^{m,HW}$	-3.265E-02	0.796	-0.148	1.471E-01	-2.988E-01	9.809E-02
	$\beta_1^{m,HW}$	7.228E-03	0.900	0.228	4.974E-02	-3.521E-02	1.663E-02
Distance to CBD $\beta^{m,CBD} = \beta_0^{m,CBD} + \beta_1^{m,CBD} \Delta t$	$\beta_0^{m,CBD}$	3.846E-03	0.701	0.057	9.701E-02	-7.533E-02	3.394E-02
	$\beta_1^{m,CBD}$	6.751E-03	1.303	0.716	1.741E-02	-7.194E-03	5.023E-03
Population in the neighboring cell $\rho_{ijt}^{m,POP} = \frac{\theta_0^{m,POP} + \theta_1^{m,POP} \Delta t}{(d_{ij})^2}$	$\theta_0^{m,POP}$	-6.166E-02	1.984	-2.872	2.646E-02	-1.672E-01	3.264E-02
	$\theta_1^{m,POP}$	-5.355E-03	1.152	-1.426	9.664E-03	-2.279E-02	6.469E-03
Neighboring cell is residential (n=2) $\rho_{ijt}^{m,n} = \frac{\theta_0^{m,n} + \theta_1^{m,n} \Delta t}{(d_{ij})^2}$	$\theta_0^{m,n}$	3.491E-01	3.696	1.138	5.371E-01	1.171E-01	9.037E-02
	$\theta_1^{m,n}$	-1.521E-02	1.143	-0.288	3.492E-02	-5.043E-02	1.673E-02
Neighboring cell is commercial/industrial/transportation (n=3) $\rho_{ijt}^{m,n} = \frac{\theta_0^{m,n} + \theta_1^{m,n} \Delta t}{d_{ij}^2}$	$\theta_0^{m,n}$	8.547E-01	6.506	9.741	1.195	6.173E-01	1.268E-01
	$\theta_1^{m,n}$	2.483E-03	0.818	0.181	5.862E-02	-3.318E-02	1.811E-02
Individual Specific Effect $u_i^m = \alpha_0^m + \alpha_1^m (y_{i0}^{m=2} + y_{i0}^{m=3}) + \sigma^m \zeta_i^m$	$\alpha_0^m$	-8.671	9.160	---	-6.974	-10.447	8.355E-01
	$\alpha_1^m$	-9.401E-02	0.761	---	1.854	-1.948	7.573E-01
	$\sigma^m$	2.000	5.532	---	2.570	1.022	3.333E-01



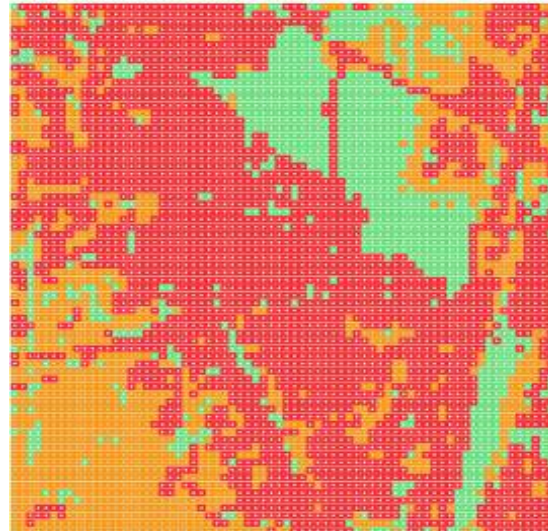


**Figure 1. Correlation between Observations Generated by Different Simulation Methods**

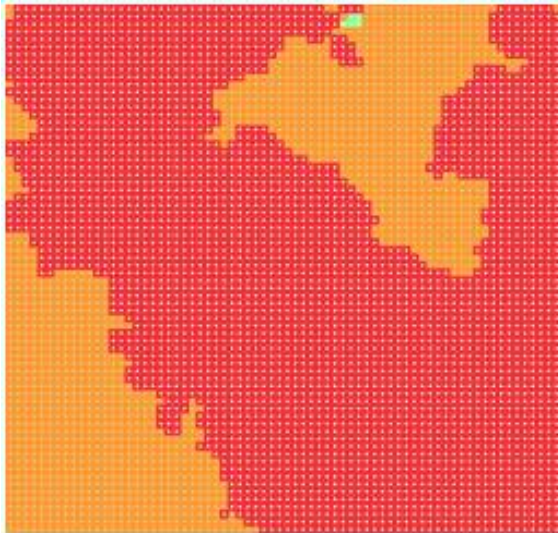
1  
2  
3  
4  
5  
6 **Land Cover in Year 2000**



**Actual Land Cover in Year 2005**



22  
23 **Predicted Land Cover in Year 2005**



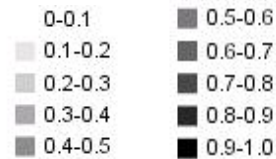
**Predictive Uncertainty in Year 2005**



40  
41 **Land Cover**



**Uncertainty**



46 **Figure 2. Predicted Land Cover and Predictive Uncertainty for Year 2005**

47  
48  
49  
50  
51  
52