# AN APPLICATION OF URBANSIM TO THE AUSTIN, TEXAS REGION: INTEGRATED-MODEL FORECASTS FOR THE YEAR 2030

Siva Karthik Kakaraparthi
Transportation Planner
Alliance Transportation Group Inc.,
11500 Metric Blvd., Building 1 Suite 150
Austin, Texas 78758
sivakarthik19@gmail.com

Kara M. Kockelman
(Corresponding author)
Professor and William J. Murray Jr. Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
6.9 E. Cockrell Jr. Hall
Austin, TX 78712
kkockelm@mail.utexas.edu

The following paper is a pre-print and the final publication can be found in *Journal of Urban Planning and Development,* Vol. 137 (3):238-247, 2011.

## ABSTRACT

Land use patterns are key determinants of traffic conditions, as well as energy consumption and greenhouse gas emissions. This work describes the modeling of year-2030 land use patterns of the Austin, Texas region using UrbanSim, an open-source model for microscopic simulation of land development, location choices and land values, at fine spatial resolution (typically 150 m x 150 m grid cells). An accompanying travel demand model was run every five years, resulting in accessibility indices for use in UrbanSim location choice models. A business-as-usual trend scenario was compared to urban growth boundary (UGB) and added transport-cost-sensitivity scenarios (TCS), expanded highway capacity scenario (EXPAN), and added state highway 130 (SH 130) scenario in order the appreciate UrbanSim's performance and the potential land use and travel impacts of such policies. As expected, several land use results (e.g., population densities), travel patterns and energy consumption results responded to scenario contexts. Local access variables (within 600-meter Euclidean distances) also enjoy significant relevance in this implementation of UrbanSim.

While UrbanSim specification limitations are multiple and its data requirements are serious (and may be impossible for almost any planning agency to meet – even after substantial effort), the model does run reasonably fast and may make good sense over the longer term for interested regions with sophisticated planning staff on board to pursue. Enhancements in the newer version of UrbanSim rendered the model to be more user-friendly.

## INTRODUCTION

With an annualized population increase of 3.5% per year over the past 10 years, Austin is one of the fastest growing mid-size regions in the U.S. Such shifts, coupled with major transportation investments (including several new tolled highways and a commuter rail line) and variations in transport, land use, and energy policies significantly impact the region's future land use patterns, traffic conditions, greenhouse gas emissions (GHG), housing affordability, environmental encroachment, and other key facets of community life. Such changes are evident in most regions, as population pressures, rising incomes, and evolving markets necessitate change. This change regularly sparks strong interest in land use forecasting, in synch with travel demand modeling and transportation planning.

The U.S. is the world's leading producer of greenhouse gases (emitting over 6 billion metric tons of $CO_2$-equivalents annually, and accounting for 22.2% of the world's emissions) (EIA, 2007). Home energy use accounts for 21% of the nation's GHG emissions, and transport accounts for 32% (EIA, 2006). Building energy use altogether accounts for 47% of the nation's GHG emissions (EIA 2007), but transport sector emissions are rising faster than the total (24% versus 13%, between 1990 and 2003, according to Brown et al. [2005]). While UrbanSim does not yet output estimates of all building sizes, the addition of modeling equations to anticipate square footage of various home types allows this work to anticipate energy demands as a function of both transport and buildings. This study evaluated the ease of UrbanSim's implementation in Austin, Texas and the model's sensitivity to various transport and land use policy scenarios.

## URBANSIM OVERVIEW

A variety of land use models (LUMs) now exist[1]. Most are mathematically and behaviorally based, while others are more normative in nature. While most LUMs rely on aggregations of space and agents (such as traffic analysis zones [TAZs] and all low-income households) (Dowling et al. 2000), UrbanSim emphasizes relatively small grid cells (and, more recently, parcels) while tracking the cell locations of individual households and jobs.[2] A 150 m x 150 m grid cell is 5.56 acres, while the average Austin TAZ is 1691 acres – or 300 times larger. Though various models ignore constraints on land use and built-space availability, UrbanSim emphasizes these key facets of urban form (Waddell et al. 2003).

Of course, land use modeling is a complex endeavor, and the UrbanSim modeling system, like any abstraction of reality, exhibits many limitations. For example, households and firms do not evolve, workers are not linked to job sites, jobs are not linked to firms, economic interactions are neglected, and job growth and economic conditions are exogenous. Moreover, the travel demand modeling process is external (with only a relatively weak link, through regional accessibility

[1] Relatively common LUMs include Putman's (1983, 2005) gravity-based ITLUP, Waddell's (2003) UrbanSim, and models based on spatial input-output specifications, like de la Barra's (1989, 1984) TRANUS, Hunt and Abraham's PECAS (2003), Kockelman et al.'s (2004) RUBMRIO, and Echenique's (1990, 1969) MEPLAN. Others include, Simmond's (2001, 1999) DELTA, Martinez's (1996) MUSSA, and Landis and Zhang's CUF (1998a, 1998b). Suitability models assign new development to highest scoring sites, and include *What If?* (Klosterman, 1996, 2007) and UPlan (Johnston et al., 2003).

[2] Other microsimulation-based LUMs are Landis and Zhang's (1998a, 1998b) California Urban Futures (CUF) model, Miller et al.'s (2001) Integrated Land Use, Transport and Environment (ILUTE) model, Timmermans' (2000) ALBATROSS model, and portions of Hunt and Abraham's (2003) PECAS model.

terms that may or may not make it into the location choice specification), the land price model is not generally fully estimated (since dynamic land price and vacancy data are so difficult to come by), and jobs and households are assigned one by one and do not compete for space in any given year[3]. In practice, UrbanSim's location choice models are typically calibrated using cross-sectional data sets (rather than those of recent movers) and population synthesis is up to the analyst. However, the package is evolving and many of these issues can be addressed in near-term and future versions.

There also are many computing challenges for new users of UrbanSim, and data set acquisition always poses a major issue for such detailed models. Nevertheless, UrbanSim pays attention to many important land market features; and, once it is running properly, it runs quickly. Moreover, UrbanSim permits estimation of most sub-model parameters within the Opus environment, rather than requiring that users enter the model with parameters in hand.

UrbanSim uses independent logit models to simulate the relocation decisions of existing households and firms, place households and jobs in grid cells[4], and anticipate grid cell-level changes in development type (Waddell 2004). Ordinary least squares (OLS) methods for parameters of continuous logistic expressions provide estimates of residential land shares across zones, and land price estimates are based on a hedonic regression equation (estimated using OLS)[5]. New transport infrastructure and local use restrictions are coded in at proper time points over the multi-year modeling process. Monte Carlo techniques are then used to simulate future year forecasts of location choices (by developers, workers, and households). Figure 1 of Waddell (2007) illustrates UrbanSim's sub-model and data set interactions, along with user-specified events (such as road building and changes in zoning policy) and scenario details.

## DATA SETS USED

As described below, calibration and application of the UrbanSim model is a highly data intensive process, particularly for a large multi-county region. UrbanSim can simulate land use patterns at any resolution (Waddell, 2001); a typical resolution is 150 m × 150 m and so was used here. Many assumptions had to be made, in placing and defining individual households and buildings at the grid cell level, in order to get the model to run. Few regions are likely to have such data and will need to resort to some sort of reasonable rules for data generation.

*Household and Employment Data*

UrbanSim's household data set consists of a list of all households, with current locations (by gridcell), household size (number of members), age of the household head, race, and number of workers, children and autos. Household data was synthesized using iterative proportional fitting techniques at the level of year-2000 Census block groups (as described in Lemp et al. [2007] and McWethy [2006]).

---

[3] Market response emerges via a land price shift in the following year, based on prior year vacancy rates in each cell.
[4] According to Waddell (2008), the number of sampled alternatives is user defined, and is available as an option in a new version of UrbanSim. They believe that allocation results generally stabilize by around 30 alternatives. However, in reality, locators are likely to consider alternatives far more strategically than simply 30 randomly-drawn sites (out of hundreds of thousands of cell alternatives).
[5] UrbanSim uses the predicted means of land prices and residential shares (rather than random, simulated values) (Waddell, 2008).

Exogenous regional household control totals were obtained from Capital Area Metropolitan Organization (CAMPO) and used as model inputs. Annual relocation probabilities for households and the vacancies in residential units were imported from Eugene data sets and assumed to hold for Austin. These range from 63 percent (for those under 24 years of age) to 2.4 percent per year.

The employment data consists of a list of all jobs (by sector), their cell locations, and building type occupied. This data set was generated from a file of firm point locations[6] provided by CAMPO. Annual relocation rates (for each job type) and non-residential vacancies also are required inputs, but had to be imported from Eugene data sets.

*Built Space and Transportation Data*
The Travis County Appraisal District (TCAD) provided residential unit locations (and year-built information), as well as square footage (and year-built information) of all commercial and industrial establishments within the region's central county. These data were used to estimate the proportions of each type of residential unit, square footage, and age distributions for buildings in the Hays and Williamson counties (which hold what 19.6% and 27.3% of the region's jobs and housing, respectively), using logistic and OLS regression equations

UrbanSim requires network travel times to the region's CBD and major airport from each TAZ centroid, along with Euclidean distances to the nearest arterial and freeway from TAZ and gridcell centroids.[7] These values were computed using CAMPO's 1997 network.

*Energy Data*
2005 Residential Energy Consumption Survey data (RECS, 2001) and 2003 Commercial Buildings Energy Consumption Survey data (CBECS, 2003) were obtained from the Energy Information Administration (EIA) and used to estimate energy per square foot , to apply to UrbanSim's year 2030 outputs.

## MODEL SPECIFICATION

Future land use patterns depend on household and job location choices, which in turn depend on the supply, quality, and price of built space, access to jobs and other destinations, household income, industry sector, and so forth. The following discussion describes the model estimation process for key sub-models.

### Household Location Choice Model (HLCM)

UrbanSim's household location choice sub-models are based on a transition model and a relocation model. The transition model generates a list of households to be added to or subtracted from the set of existing households. These shifts can result from various demographic processes, like aging, marriage, divorce, births, and deaths, and are provided as exogenous control totals for

[6] Point locations of employment data were obtained as GIS layers from the Texas Workforce Commission via the Texas DOT and then cleaned by CAMPO.
[7] Interestingly, UrbanSim does not call upon network files directly; all distances must be computed externally and provided (or coded internally by the analyst). This is an opportunity for relatively easy model improvement.

future years. The household relocation model generates vacant spaces when households are selected to move and adds all movers to the list of unplaced households. While the probability that a household moves should depend on the relative attractiveness of available alternatives, as compared to the current dwelling; this presently is controlled by exogenous relocation probabilities, as pre-defined by UrbanSim users.

The set of unplaced households are placed into grid cells via a multinomial logit (MNL) model of household location choice, based on the composite utility of 30 alternative grid cells (each having at least one vacant housing unit at the time the household is allocated). Explanatory variables affecting location choice can reflect elements of urban economic theory and sociology (CUSPA 2006), including regional and local accessibilities, race, incomes, and land rents (Waddell 2002).

Table 1 shows the HLCM's estimation results. These suggest that increases in housing and land costs and the share of land in residential use negatively impact a cell's residential location utility, everything else constant. In contrast, variables such as income of current cell residents, open space in the cell, proximity to arterials and highways, accessibility to regional jobs, local employment, population, and travel time to the region's CBD are predicted to have a positive impact on a cells' utility. Also, households with similar income levels and races tend to attract. The positive signs on distances to arterials and highways and on CBD travel times are not intuitive, but these are more than offset by the positive benefits of the regional jobs accessibility values of all almost all locations.

## Employment Location Choice Models (ELCMs)

UrbanSim's employment location choice model (ELCM) is analogous to its household location choice model. A transition model generates or removes the newly created jobs in each sector, depending upon the growth or decline of employment in that sector (as compared to the prior year). Such input assumptions are exogenously obtained from state economic forecasts and/or commercial and in-house sources (CUSPA 2006). An employment relocation model then determines which individuals will change jobs in any given year. Employment relocation probabilities are determined exogenously and given as inputs to UrbanSim.

Removed and relocating jobs are noted in the database of job-site vacancies, and jobs created via the employment transition and relocation models are added to the database of unplaced jobs. The unplaced jobs are allocated into the gridcells by sector, based on the composite utilities of each gridcell. UrbanSim assumes that every job is independent of the other and moves separately. Though, this assumption simplifies the modeling procedure, it is an important limitation of UrbanSim since most job relocation is based on relocation of firms. The employment location choice model is also based on a multinomial logit specification.

The only difference between household and employment location choice models is that the latter allocates jobs using sector-specific preference functions (MNL model estimates), while all households are allocated using a single MNL specification – with indicator variables for variables like race and income, to accommodate some forms of preference variation. The sectors considered in this study are industrial, commercial and home-based employment[8].

The *Industrial Employment Location Choice Model* locates new and relocating industrial jobs based on the relative values of all grid cells' composite utilities. Estimation results presented in Table 1 suggest that increases in home and work accessibility values, presence of industrial square feet, and the presence of high-income households have a positive effect on gridcells' locational utilities. In contrast, increases in distance to the nearest highway and in travel time to the CBD have negative impacts. All parameters have intuitive signs except that for the presence of higher-income households (perhaps both high income households and industries are attracted to similar properties).

*The Commercial Employment Location Choice Model* locates commercial jobs seeking locations. Table 1's estimates suggest that increases in commercial and industrial square footage, distance to the nearest highway, number of service sector jobs within "walking distance" (600 meters, Euclidean from gridcell centroid), access to population, and number of high-income households positively impact a cell's attractiveness. In contrast, the cost of land and the number of retail jobs have a negative impact. While the positive impact of highway distance and the negative impact of local retail jobs are not intuitive, these can be counteracted by semi-collinear variables (like square footage, service job access, and population access).

UrbanSim requires data on home-based jobs in order to run. However, Austin does not yet have such data, so this data set had to be manufactured (using a two-percent-of-jobs-per-zone assumption). Table 1 shows the results of this *Home-based Employment Location Choice Model*, which indicates a preference for older, more expensive residences in denser, low-vacancy locations, close to other home-based jobs. While many parameters appear reasonable, others may be at odd with actual trends, because these job sites were randomly selected. This model should not affect the overall forecasts much because only 2% of the modeled population works from home (Bayles, 2002).

## Modeling Land Prices

UrbanSim's *Land Price Model* provides a key input to the land development, household and employment location choice models. Land price is modeled using hedonic regression on attributes such as land use type (including 8 mixed use types, 8 residential, 3 commercial, 3 industrial, 1 government use, vacant and undeveloped land uses), site characteristics, access variables, and neighborhood and zoning characteristics. UrbanSim implicitly assumes that households, businesses and developers are all price takers, and annual price and development adjustments help match aggregate supply and demand over time. (CUSPA 2006) Moreover, as cell-level vacancy rates fall below a user-specified long-term structural vacancy rate, land price increases and vice versa (DiPasquale 1996). Land price estimates are updated annually, after all

---

[8] Buildings for workers are classified into industrial, commercial, governmental and home-based buildings. Industrial employment refers to jobs located in industrial buildings. Commercial employment refers to jobs located in commercial buildings. And home-based jobs are those where workers work from home. Government jobs are assumed to maintain their locations, so there is no choice model for these.

the construction and development is undertaken and vacancy rates have been computed, according to the following equation:

$$P_{ilt} = \alpha + \delta(\frac{V_i^s - V_{it}^c}{V_i^s}) + \beta X_{ilt} \quad (1)$$

where $P_{ilt}$ is the land price per acre of development type $i$ at location $l$ and time $t$, $V_i^s$ is the long-term structural vacancy rate, $V_{it}^c$ is the current vacancy rate (at time $t$), $X_{ilt}$ is the vector of site attributes, and $\alpha$, $\beta$, and $\delta$[9] are parameters to be estimated. However, grid-cell-level or building-level vacancy rates can be quite difficult for analysts to avoid, particularly vis-à-vis property prices over time, in order to estimate a parameter like $\delta$. Thus, users may often find themselves importing (or guessing at) this key parameter. This is a key concern for the model, though most users may not rely on UrbanSim for land price information. Either way, current year price estimates are used to estimate the next year's market activities, including land development and location choices (Waddell *et al*. 2003).

*Land Price* estimation results (Table 1) suggest that an increase in commercial square feet, total employment, population density, access to population, percentages of commercial and developed lands, and high-income households (within walking distance) are associated with higher land prices (as assessed by the Travis County Appraisal District), whereas the presence of industrial space, open space and residential land use within walking distance of a cell (600 meters) are associated with lower land prices, ceteris paribus. Such signs are consistent with behavioral expectations.

UrbanSim's *Residential Land Share Model* is used to compute the residential land share in a gridcell, as follows:

$$\ln\left(\frac{y}{1-y}\right) = X\beta + \in \quad (2)$$

where $y$ is residential land use shares in each grid cell. Essentially, the model assumes a logistic specification for the fractional shares, but ordinary least squares is used for estimation of parameters ($\beta$).

From Table 1, it can be observed that residential land share predictions tend to rise with the number of residential units in a cell, and job access, but fall with jobs counts, non-residential built space and travel time to the CBD, as expected.

UrbanSim also uses MNL models to place new development of residential, industrial and commercial structures, based on composite utilities. These model parameters, for development location choices, can be found in Kakaraparthi (2009).

[9] Due to a lack of occupancy rate data for two time points across zones, δ could not be estimated and was simply set equal to 1.0 here, to be able to run the model and evaluate performance.

## INCORPORATION OF A TRAVEL DEMAND MODEL (TDM)

According to Opus (2006) workshop documents, UrbanSim has been used with TDMs based in TP+, MinUTP, EMME/2, and other systems. Here, Caliper's TransCAD software was used, and outputs of both models manually transferred by the user.

Using 1996/1997 Austin Travel Survey (ATS) data, Lemp (2007) estimated the parameters of and coded (in TransCAD's GIS-DK) details of the fairly standard TDM employed here. Regression models are used for trip generation, at the household level for home-based trips and at the zonal level for non-home-based trips. An MNL model of destination choice is used for trip distribution, and includes a logsum parameter measuring the maximum expected utility achieved over all modes and times of day (TODs). A joint MNL model of mode choice (drive alone, shared ride, transit, and bike/walk) and four TODs is also used. Separate models were applied for each of four trip types (home-based work, home-based non-work, non-home-based work, and non-home-based non-work). Finally, deterministic network assignment routines were used in each TOD, and 25 feedback iterations (from network-equilibrium travel times and costs to trip distribution) were performed, in order to obtain estimates of interzonal travel times, trip distances and travel costs – specific to each of the four modes at each of the four TODs[10].

Using this information, logsum accessibility indices were computed and then input into UrbanSim, in order to anticipate the next five years' land use patterns. The logsum values are computed as follows (see, e.g., Ben-Akiva and Lerman [1985]):

$$L_{ij} = \ln\left(\sum_{m \in C} \exp(U_{nijm})\right) \tag{3}$$

$$U_{nijm} = \beta_{1nm} + \beta_{GC,n}\left(\frac{Cost_{nijm}}{VOTT_n}\right) + TT_{nijm} \tag{4}$$

And these are used in somewhat simplistic accessibility indices, as proscribed by UrbanSim documentation (CUSPA, 2006):

$$AI_i = \sum_{j=1}^{J} D_j e^{L_{ij}} \tag{5}$$

where $i$ and $j$ index origin and destination zones, $n$ indicates trip type (e.g., HBNW), $m$ indicates mode, $t$ denotes TOD, $\beta_{1nmt}$ is the alternative specific constant from the joint mode-TOD choice model, $\beta_{GC}$ is the coefficient of generalized trip costs, $TT$ is travel time, $VOTT$ is the assumed value of travel time ($9 per person-hour for work trips and $4.50 per person-hour for non-work trips), and $COST$ indicates trip cost (assumed to be $0.20 per vehicle-mile).

For purposes of TDM feedback to UrbanSim in this study, accessibility indices (AI's) were computed only for home-based work (HBW) trips made during the AM peak period, for zero-

[10] CAMPO's estimates of commercial and external trips were constant throughout the forecast period and added to TOD-specific trip tables before traffic assignment. While not ideal, such methods are not uncommon and provide a simple way for dealing with commercial travel, so the research team could focus on the land use model's operations.

vehicle and 1+-vehicle-owning households, separately[11]. As expected, these twin AI values were highly correlated, and would be highly correlated with other AIs at other times of day by other modes, so only the zero-vehicle-household AI values were controlled for in the location choice models described earlier.

As mentioned earlier, UrbanSim was run every year, for 30 years (2001 through 2030), the TDM was applied six times (2005 through 2030, at five-year intervals). The TDM on an average took about 4.5 hours and UrbanSim forecasted land use patterns at the rate of 0.5 hours/year on a 3.5 GB RAM, 2.66 GHz Dual Core processor. The runs provided household and job locations, and square footage of commercial establishments, in year 2030, along with network travel conditions. Such information was converted into energy estimates, as described in the next section.

## ENERGY CONSUMPTION ESTIMATES

The OLS regression model estimates developed using the nation's RECS and CBECS data sets to estimate energy consumption per square foot of built space in Austin can be found in Kakaraparthi (2009). Variables pertaining to household location and attributes, home type and age, price of electricity and gas, and number of heating/cooling degree days served as covariates. The resulting estimates were applied to UrbanSim's results, with proportions of different home types (single-family, mobile homes, and multi-family dwelling units [MFDUs]) based on Travis County Appraisal District shares, by zone, in the year 2000 (since UrbanSim does not predict home type). Logarithmic regression equations were used to estimate the proportions of each type of dwelling unit in Williamson and Hays counties, and the results of these can be found in Kakaraparthi (2009). The square footage per residential unit (by type) was allowed to pivot off year 2000 values (at the TAZ level), simply by adding 20 square feet/year (to recognize the nation's average home size growth.

## RESULTS AND DISCUSSION

Several scenarios were implemented to test the sensitivity of UrbanSim to various policies. These scenarios include (1) a No Travel Demand Model (NoTDM) scenario, in which UrbanSim ran continuously for 30 years, without TDM integration (so travel costs stay constant over the forecasting horizon and accessibility indices do not vary as much as they would with a TDM in place [though they do reflect new home and job land use distributions that come out of UrbanSim]), (2) a Business as Usual (BAU) scenario (with Austin's 1997 network held constant over the forecast period), (3) an Urban Growth Boundary (UGB) scenario (where new development was not permitted in zones outside Figure 2's boundary), (4) a doubled Travel-Cost Sensitivity (TCS) scenario (5) an expanded network (EXPAN) scenario, where capacities of three major arterials were doubled, and (6) addition of a 49.2-mile bypass freeway (SH130) to the network. Figures 3 and 4 present the plots of household and employment densities in all these scenarios.

---

[11] While the TDM was not segmented on the basis of vehicle ownership levels, transit and bike/walk mode choice model specifications were used for the zero-vehicle households' accessibility index calculations and drive times were used for the 1+ vehicle-owning households' index.

Figure 1 suggest a decentralizing behavior for households under the BAU scenario once the TDM is integrated with UrbanSim – relative to the NoTDM case. In other words, lower density population development becomes evident once travel times are permitted to rise (as the network congests, to accommodate population and jobs growth). To understand this somewhat unexpected result, the practical significance of all variables in the household location choice model was investigated by looking at the utility change corresponding to a one standard deviation change in each explanatory variable. Relatively low land prices in the peripheral TAZs attracted the development, resulting in this somewhat unexpected decentralizing behavior. However, travel times to the CBD and accessibility indices are key covariates for the jobs location models, so jobs were predicted to centralize (locate closer to the region's CBD) once the TDM was added to the modeling process. Better linkage of jobs and households (as the Puget Sound Regional Council is pursing with UrbanSim) may help slow this somewhat incompatible decentralization/centralization pattern that can emerge in applications of UrbanSim. Moreover, the fact that calibration of UrbanSim parameters is based on cross-sectional data sets is problematic, and can result in such predictions. (The model predicts current land use patterns, for movers and non-movers, rather than the location choice behaviors of new households, new jobs/firms, new developments and recent movers.)

Figure 1's UGB scenario results show unexpected and excessive land development predictions in the northern cities of Williamson County (i.e., Georgetown, Taylor and Florence), which are inside the UGB and just inside the three-county region's border. Though of low density in year 2000, these northern zones were contiguous with zones that met the UGB's 2-job-equivalents-per-acre density threshold. (It should be noted that 453 of the region's 1074 TAZs or 40% of the three-county area falls outside the UGB, which is a set of five zone clusters – the largest located centrally.) This behavior can be attributed to the lower land prices (which households prefer) and high land availability in those zones. However, jobs continued to prefer the region's center, thanks to higher accessibility indices and lower CBD travel times (which are key predictors of the cross-sectional data set's land use patterns in year 2000). Recognition of jobs-worker connections, as discussed above, may result in a different outcome. It also should be noted that odd model behaviors, in multiple disciplines, often emerge at border areas. Thus, a halo of zones/counties could be very useful for achieving a stronger sense of the future in these border zones. In reality, many of these households would likely just leave the region (by skipping over the county border) if growth restrictions were not in place elsewhere.

In the TCS scenario, travel cost sensitivity is doubled, so accessibility indices fall everywhere and households and jobs move closer towards the CBD (as compared to the BAU scenario), in order to reduce their transport costs. This behavior was expected and resembles the shifts emerging from increased gasoline prices. In the EXPAN and SH 130 scenarios, households and jobs shift towards the expanded corridors but at lower densities (than in the BAU case), thanks to improved travel conditions along the corridor. These are expected results, but pictures cannot tell the whole story. Useful summary metrics of all scenario results include region-level statistics, and estimates of GHG emissions, as discussed below.

Table 2 presents count-weighted household and employment densities for the region, city center accessibility indices, and annual VMT values. Interestingly, with the exception of the UGB scenario, percentage differences in these various indices across different scenarios are less than

5%, as compared to the BAU. Such outcomes seem unusually moderate, particularly when compared to results of other land use models for the Austin region.

Zhou and Kockelman (2009) applied a gravity-based land use model (G-LUM) for the region, and Tirumalachetty and Kockelman (2009) rely on microsimulation models for their results. Though calibrated with much of the same data and applied over essentially the same period to many of the same scenarios, their UGB results suggest 17% and 14% reductions in region-wide VMT (rather than the 10% found here). The increases in count-weighted household densities were roughly a striking 1900% (almost 20-fold) and 200%, rather than the 50% increase seen here, using UrbanSim's predictions.

It seems UrbanSim's results are unusually "stable" here, across distinctive scenarios – with the exception of the jump in population at the region's northern edge under the UGB scenario. Such stability is better than having wildly changing results (which can emerge quickly in unconstrained gravity model applications [see, e.g., Zhou and Kockelman 2009]). This property no doubt emerges from UrbanSim's tight connections between land, buildings, and space users (jobs or households). But it may be an indication that the model (as calibrated for Austin) is inadequately sensitive to policy changes.

As a result, energy consumption estimates also did not vary much across scenarios, though these were lowest in the UGB and TCS scenarios and highest in the EXPAN scenario, as expected. Centralizing tendencies evident in the UGB and TCS scenarios result in less travel and smaller dwelling units, due to a shift toward MFDUs. More MFDUs are built in the UGB and TCS scenarios due to reduced land availability in the more popular/accessible locations.

In contrast, population decentralization clearly observed in the EXPAN, BAU and SH 130 scenarios is predicted to encourage the development of single-family units over MFDUs, resulting in higher energy demands via housing and travel. Commercial energy estimates were highest in the UGB and TCS scenarios, because CBECS estimates increased energy consumption values with the increase in workers per square feet. Therefore, scenarios which predicted denser land use patterns showed increased commercial energy consumption. Also, shared walls in the multi-storied commercial structures which actually reduce the energy consumed were not captured by the model. Relatively little variation in industrial building energy consumption emerged across scenarios, in part because industrial jobs' location are largely independent of access considerations.

## CONCLUSIONS

This work describes the UrbanSim modeling results of 5 distinctive land use and transport scenarios for year-2030 land use patterns in Austin, Texas at fine (150 m cell) spatial resolution. As evident, UrbanSim is data intensive software. Key stages of UrbanSim use are data acquisition and assembly, model estimation, scenario development and forecasting; and these required roughly 60%, 20%, 5% and 15% of the research team members' time. It took approximately two person-years to apply UrbanSim to the Austin region, yet various data and model enhancements are still desired.

While the greatest challenges for UrbanSim users lie in acquisition, assembly and management of data, a variety of challenges also lie in the model estimation process. For example, multicollinearity in various access indices can result in a lack of statistical and practical significance, and/or odd behavioral implications. Analysts need to be wary of what covariates make it into the final model selections.

In the Austin case study, variations in scenario results were quite moderate, but potentially realistic as compared to results derived from other land use models developed for the same region (using highly similar data sets and scenarios). These moderate results also were reflected in energy implications of the various scenarios.

UrbanSim exhibits a variety of strengths and limitations. Key strengths include its freely available open-source code with some technical support forthcoming via a growing user listserv (frequented by UrbanSim developers). The program is designed to compute land use patterns at any level of resolution, and every household and job is tracked. It uses a dynamic disequilibrium approach but turns to cross-sectional data sets for parameter estimation (though dynamic behaviors could and probably should be used for calibration, where feasible). As multiple as its many sub-models are, it is efficiently programmed, requiring only a 25-minute run time per year for the Austin simulations.

Key computing issues include UrbanSim's installation, which requires numerous supporting packages (including Enthought Python, Numpy and Scipy) and a variety of optional packages (such as MySQL and dbfpy, along with household and jobs synthesizers and a travel demand model). UrbanSim use requires relatively high-end computers, with 3.5 GB RAM, a 2.66 GHz Intel Core 2 Duo processor running on Windows XP. Documentation does not yet specify the changes in the code needed for running UrbanSim with a csv database, as opposed to MySQL.

A key concern related to Austin's data assembly involves estimating grid-cell-level land use patterns from TAZ, block and block group data; but much of this may be avoided via a switch to parcel-level data. Acquiring residential and non-residential target vacancies in future years is also challenging, if not impossible.

UrbanSim presently calls for cross-sectional data for estimation of model parameters, which is an important limitation. The pace and nature of land use change, and the preference of recent movers (rather than the siting of all agents currently located in a region) are key behaviors that require additional or different data structures (e.g., a survey of recent movers/locators). Such modeling work should enhance future implementations and may not be difficult given the data sets users would have to have on hand (and the data acquisition efforts most MPOs already engage in).

Other concerns include the use of just 30 random choices (chosen from 328,000 alternatives in the Austin example) for estimating MNL models of location choice, and later evaluation of just 30 alternatives before placing a household or job in a suitable gridcell. Such defaults can and should be changed. Reliance on simply AM peak travel times and costs (from travel demand model outputs) can pose a serious limitation for certain analyses (e.g., when one seeks to appreciate the local land use effects of a new transit line). The lack of residential square footage forecasts, constant area-per-worker assumptions, lack of household and job evolution, neglect of firm (versus job) dynamics, and reliance on exogenously specified control total and standard, aspatial MNL models are also concerns. Of course, modeling urban systems is an incredibly complex endeavor, and no model can address all issues. The opportunity is simply to advance the state of the art and practice. This work offers a window into the challenges that lie ahead.

## REFERENCES


Arentze, T.A., and H.J.P. Timmermans. 2000. Albatross: A learning-Based Transportation Oriented Simulation System. *European Institute of Retailing and Services Studies.* Eindhoven, The Netherlands.

Bayles, B., R. Crawley, and D. Mass. 2002. The Texas Landscape: A Snapshot from the 2000 Census. URL: http://www.tracer2.com/admin/uploadedPublications/141_tlmr0210art.pdf. Accessed on July 28, 2009.

Brown M. A., F. Southworth, and T. K. Stovall. 2005. Towards a built climate friendly environment, Pew Center. URL: http://www.pewclimate.org/docUploads/Buildings_FINAL.pdf Accessed on 5th April, 2009.

Center for Urban Simulation and Policy Analysis (CUSPA). 2006. *Opus: The Open Platform for UrbanSim Version 4.0 Reference Manual and Users Guide*. URL: http://www.urbansim.org/opus/releases/opus-4-1-2/docs/opus-userguide.pdf. Accessed April 5, 2009.

EIA. 2007. Emissions of GHGs, 1980-2006. URL: http://www.eia.doe.gov/emeu/aer/consump.html. Accessed April 5, 2009.

De la Barra, T. 1989. Integrated Land Use and Transport Modeling. *Cambridge University Press*. New York.

De la Barra, T., B. Perez, and N. Vera. 1984. TRANUS-J: Putting Large Models into Small Computers, *Env. and Planning B*, Vol. 1, No.1 needed, pp. 87-101.

DiPasquale, D., and W. Wheaton. 1996. *Urban Economics and Real Estate Markets.* Prentice Hall, Englewood Cliffs, New Jersey.

Dowling, R., R. Ireson, A. Skabardonis, D. Gillen, P. Stopher,.A. Horowitz, J. Bowman, E. Deakin, and R. Dulla. 2000. Predicting Short-Term and Long-Term Air Quality Effects of Traffic-Flow Improvement Projects. National Cooperative Highway Research Program Report 535.

Echenique, Marcial H., D. Crowther, and W. Lindsay. 1969. A Spatial Model of Urban Stock and Activity. *Regional Studies*, Vol. 3, pp. 281-312.

Echenique, Marcial H., A.D.J. Flowerdew, John D. Hunt, T.R. Mayo, I.J. Skidmore, David C. Simmonds. 1990. The MEPLAN models of Bilbao, Leeds and Dortmund. *Transport Reviews*, Vol. 10, No. 4: pp. 309-322.

Handy, S. 1993. Regional Versus Local Accessibility: Implications for Non-work Travel. *Transportation Research Record*, No. 1400, pp. 58 – 66.

Hunt, J.D., and J. E. Abraham. 2003. Design and Application of the PECAS Land Use Modelling System. *Proceedings of the 8th Conference on Computers in Urban Planning and Urban Management*. Sendai, Japan.

Kockelman, K. M., L. Jin, Y. Zhao, and N. Ruiz-Juri. 2004. Tracking Land Use, Transport, and Industrial Production using Random-Utility-Based Multizonal Input-Output Models: Applications for Texas Trade. *Journal of Transport Geography*, Vol. 13, No. 3: 275-286.

Klosterman, R.E. 2007. *What if?* 2.0 User's guide. URL: http://www.whatifinc.biz . Accessed July 27, 2009.

Klosterman, R.E. 1999. The *What if?* Collaborative Support System. *Environment and Planning, B: Planning and Design*, Vol. 26, pp. 393-408.

Landis, J., and Zhang, M. 1998a. The second generation of the California Urban Futures model. Part 1: Model logic and theory. *Environment and Planning B: Planning and Design*, Vol. 25, pp. 657 - 666.

Landis, J., and Zhang, M. 1998b. The second generation of the California Urban Futures model. Part 2: Specification and calibration results of the land use change module. In: *Environment and Planning B: Planning and Design,* Vol. 25, pp. 795 – 824.

Lemp, J.D., McWethy, L., and Kockelman, K. 2007. From Aggregate Methods to Microsimulation: Assessing the Benefits of Microscopic Activity-Based Models of Travel Demand. *Transportation Research Record*, No. 1994, pp. 80-88.

Lowry, I.S. 1964. *A Model of Metropolis*, Rand Corporation. URL: http://www.rand.org/pubs/research_memoranda/2006/RM4035.pdf. Accessed July 27 2009.

Miller, E. J. 2001. Integrated Land Use, Transportation, Environment (ILUTE) Modelling System. URL: http://www.civ.utoronto.ca/sect/traeng/ilute/ilute_the_model.htm. Accessed June 28, 2008.

Parsons Brinckerhoff Quade and Douglas. 1999. Land Use Impacts of Transportation: A Guidebook. *National Cooperative Highway Research Program (NCHRP) Report 423A*. National Academy Press, Washington D.C.

Putman, S.H. 1983. *Integrated Urban Models: Policy Analysis of Transportation and Land Use*. Pion Press, London.

Tirumalachetty, S., K.M. Kockelman, S. Kumar. 2009. Micro-Simulation Models of Urban Regions: Anticipating Greenhouse Gas Emissions from Housing and Transport in Austin, Texas. *Proceedings of the 88th Annual Meeting of the Transportation Research Board*. Washington D.C.

Waddell, P. 2001. Analytical Tools for Land Use, Transportation and Growth Management. URL: http://www.urbansim.org/papers/Analytical_Tools.pdf. Accessed June 27, 2008.

Waddell, P. 2002. UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. Preprint of an article that appeared in the *Journal of the American Planning Association* 68: 297-314. URL: http://www.urbansim.org/papers/UrbanSim-JAPA.pdf. Accessed June 28, 2008.

Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke and F. Ulfarsson. 2003. Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Networks and Spatial Economics,* Vol. 3, No: 1, pp. 43-67.

Waddell, P., and G.F. Ulfarsson. 2004. Introduction to Urban Simulation: Design and Development of Operational Models. *Handbook in Transport, Volume 5: Transport Geography and Spatial Systems*, pp. 203-236.

Waddell, P., G. Ulfarsson, J. Franklin, and J. Britting. 2007. Incorporating Land Use in Transportation Planning**.** Preprint of paper published in *Transportation Research Part A*, Vol. 41, pp. 382-410.

Waddell, P. 2008. Phone Conversation with K.M., Kockelman and S.K., Kakaraparthi.

Zhou, B., K.M. Kockelman, and J. Lemp. 2009. Transportation and Land use Policy Analysis Using Integrated Transport and Gravity-Based Land use Models. Forthcoming in *Transportation Research Record*.

Table 1: Estimation Results of Sub-models Used in UrbanSim

| | HLCM | | ELCM - Industrial | | ELCM - Commercial | | ELCM – Home-based Jobs | | Land Price | | Residential Land Share | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Explanatory Variables** | **Coef.** | **t-stats** | **Coef.** | **t-stats** | **Coef.** | **t-stats** | **Coef.** | **t-stats** | **Coef.** | **t-stats** | **Coef.** | **t-stats** |
| Constant | -- | -- | -- | -- | -- | -- | -- | -- | 13.256 | 454.9 | 2.114 | 55.41 |
| Ln(housing cost in cell | -0.192 | -24.75 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Acres of open space (in cell | 0.175 | 9.54 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Average income ($ per household per year? Of HHs in cell | 3.18E-06 | 8.38 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Ln(miles to nearest arterial) | 0.12 | 6.69 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Ln(miles to nearest highway) | 0.0852 | 5.34 | -0.0328 | -1.64 | 0.0594 | 3.5 | -- | -- | -- | -- | -- | -- |
| Ln(home access to employment for households with one car) | 0.721 | 10.06 | 4.09E-04 | 1.04 | 0.0183 | 0.43 | -- | -- | 0.11 | 27.4 | 3.35E-03 | 5.22 |
| Ln (total land value in cell | -0.03 | -4.01 | -- | -- | -0.1902 | -5.38 | -- | -- | | | -- | -- |
| Number of jobs (in cell) | 5.25E-04 | 1.79 | -- | -- | -- | -- | -- | -- | 0.233 | 123.21 | -- | -- |
| Number of retail jobs (in cell) | -- | -- | -- | -- | -2.33E-03 | -22.32 | -- | -- | -- | -- | -- | -- |
| % residential land in cell | -0.0161 | -20.26 | -- | -- | -- | -- | -- | -- | -0.00322 | -28.02 | -- | -- |
| Population (in cell | 3.63 | 12.49 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| Number of high income households (in cell | -- | -- | -- | -- | 0.023 | 3.02 | -- | -- | -- | -- | -- | -- |
| Travel time to CBD (on network, in AM peak period | 0.029 | 7.27 | -0.028 | -2.94 | -- | -- | -- | -- | -- | -- | -0.016 | -7.55 |
| % high income households within 600 m if household is high income | 0.0252 | 20.64 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| % low income households within 600 m if household is low income | 0.0523 | 23.11 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| % mid income households within 600 m if household is mid income | 0.0243 | 21.41 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| % minority households within 600 m if household is non-minority | -0.045 | -10.32 | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ln(work access to jobs for households with 0 cars) | -- | -- | 0.244 | 2.77 | -- | -- | -- | -- | -- | -- | -- | -- |
| Indicator for cell's average income being high income | -- | -- | 0.209 | 2.83 | -- | -- | -- | -- | -- | -- | -- | -- |
| Ln(industrial SF in cell) | -- | -- | 0.0948 | 7.43 | 0.0132 | 3.79 | -- | -- | -- | -- | -- | -- |
| Ln(commercial SF in cell) | -- | -- | 0.0448 | 2.3 | -- | -- | -- | -- | 0.048 | 35.36 | -- | -- |
| Total non-residential SF in cell | -- | -- | -- | -- | 3.32E-08 | 5.09 | -- | -- | -- | -- | -- | -- |
| Ln(service sector employment within 600 m) | -- | -- | -- | -- | 0.1438 | 6.82 | -- | -- | -- | -- | -- | -- |
| Building age (years) | -- | -- | -- | -- | -- | -- | 0.102 | 2.47 | -- | -- | -- | -- |
| Ln(total value of the cell) | -- | -- | -- | -- | -- | -- | 0.2 | 1.77 | -- | -- | -- | -- |
| Residential density (HHs/acre) | -- | -- | -- | -- | -- | -- | 9.287 | 13.61 | -- | -- | -- | -- |
| Vacant home-based job space ( in cell?) | -- | -- | -- | -- | -- | -- | -1.69 | -13.41 | -- | -- | -- | -- |
| Vacant residential units in cell | -- | -- | -- | -- | -- | -- | -0.0102 | -5.58 | -- | -- | -- | -- |
| Ln(same sector employment within 600 m) | -- | -- | -- | -- | -- | -- | 0.329 | 3.36 | -- | -- | -- | -- |
| Ln(Industrial SF within 600 m) | -- | -- | -- | -- | -- | -- | -- | -- | -0.0163 | -23.12 | -- | -- |
| % commercial SF within 600 m | -- | -- | -- | -- | -- | -- | -- | -- | 0.0138 | 43.76 | -- | -- |
| % developed land within 600 m | -- | -- | -- | -- | -- | -- | -- | -- | 0.0103 | 117.64 | -- | -- |
| % high income households within 600 m | -- | -- | -- | -- | -- | -- | -- | -- | 0.00229 | 27.73 | -- | -- |
| % open space within 600 m | -- | -- | -- | -- | -- | -- | -- | -- | -0.0129 | -50.39 | -- | -- |
| Population density (Persons/acre) | 0.0876 | 96.88 | -- | -- | -- | -- | -- | -- | 0.0876 | 96.88 | -- | -- |
| Average residential value per housing unit within 600 m | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -1.31E-07 | -3.47 |
| Ln(basic sector jobs within 600 m) | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -0.0527 | -8.06 |
| Ln(retail sector jobs within 600 m) | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -0.0298 | -4.82 |
| Number of residential units (in cell) | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | 1.68E-03 | 13.35 |
| Sum of industrial & commercial | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -1.72E- | -3.24 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SF in cell | | | | | | | | | | | 07 | |
| Developable maximum commercial SF | -- | -- | -- | -- | -- | -- | -- | -- | -- | -- | -1.44E-07 | -23.11 |
| **Adj. likelihood ratio index** | 0.091 | | 0.00561 | | 0.03194 | | 0.521 | | -- | -- | -- | -- |
| **Adj. R- square** | -- | | -- | | -- | | -- | | 0.566 | | 0.0362 | |

Table 2: Overall Land Use and Transportation Results across All Scenarios

| **Scenario** | **Daily VMT (in millions)** | **Average Count-Weighted Household Density** | **Average Count-Weighted Jobs Density** | **Average Regional AI* for HHs** | **Average Regional AI* for Jobs** | **VMT Weighted v/c Ratios** | **VMT Weighted Average Speeds** |
|---|---|---|---|---|---|---|---|
| **NoTDM** | 71.64 | 1,317 | 9,356 | 94,814 | 309,642 | 0.5671 | 37.176 |
| **BAU** | 73.59 | 1,303 | 9,422 | 94,947 | 313,708 | 0.5759 | 36.746 |
| **UGB** | 67.18 | 1,992 | 10,237 | 116,030 | 334,821 | 0.5278 | 36.937 |
| **TCS** | 70.84 | 1,335 | 9,524 | 98,057 | 323,697 | 0.5866 | 36.347 |
| **EXPAN** | 77.67 | 1,290 | 9,412 | 95,079 | 314,203 | 0.5422 | 39.695 |
| **SH130** | 72.57 | 1,264 | 9,446 | 95,007 | 316,428 | 0.563 | 36.68 |

*Note: AI = Accessibility index = $\sum$*Count of Jobs or HHs*$_i$*/ Network distance from zone i to CBD*

Table 3: Energy Results

| Scenario | Energy Consumption (in kWh or ccf) | Energy Results | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | SFDU | Mobile | MFDU | Total Residential | Commercial | Industrial | Travel | $CO_2$ emitted (1000 metric tonnes) |
| **NoTDM** | kWh | 6,972 | 3,514 | 5,402 | 15,889 | 13,967 | 2,797 | 11,620 | |
| | ccf | 196.95 | 121.5 | 146.93 | 465.38 | 274.94 | 60.02 | | 31,274 |
| **BAU** | kWh | 7,395 | 3,647 | 5,084 | 16,127 | 13,962 | 2,797 | 11,946 | |
| | ccf | 209.71 | 126.05 | 139.58 | 475.35 | 274.82 | 59.99 | | 31,629 |
| **UGB** | kWh | 6,597 | 3,448 | 5,562 | 15,607 | 14,108 | 2,797 | 10,990 | |
| | ccf | 181.49 | 118.13 | 148.12 | 447.75 | 275.14 | 59.99 | | 30,799 |
| **TCS** | kWh | 7,027 | 3,881 | 5,217 | 16,126 | 13,977 | 2,797 | 11,510 | |
| | ccf | 198.14 | 134.29 | 143.01 | 475.45 | 274.8 | 60.02 | | 31,441 |
| **EXPAN** | kWh | 7,471 | 3,675 | 5,040 | 16,187 | 13,976 | 2,797 | 12,720 | |
| | ccf | 212.22 | 127.09 | 138.31 | 477.62 | 275.07 | 59.99 | | 32,043 |
| **SH130** | kWh | 7,473 | 3,705 | 5,028 | 16,208 | 13,977 | 2,797 | 11,744 | |
| | ccf | 212.19 | 128.21 | 138.06 | 478.46 | 275.07 | 60.00 | | 31,619 |

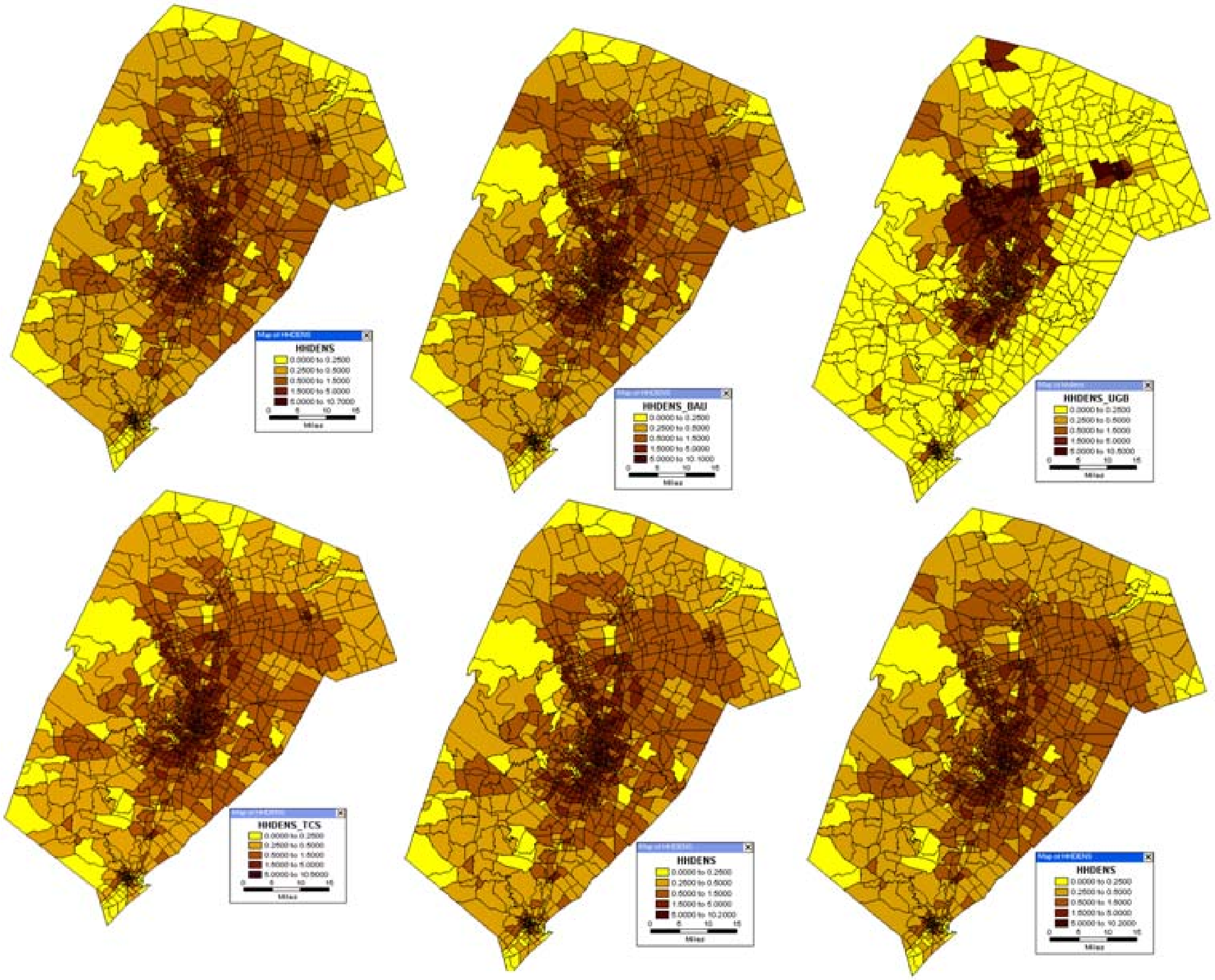


Figure 1: Household Densities in Year 2030 in NoTDM, BAU, UGB, TCS, EXPAN and SH 130 Scenarios

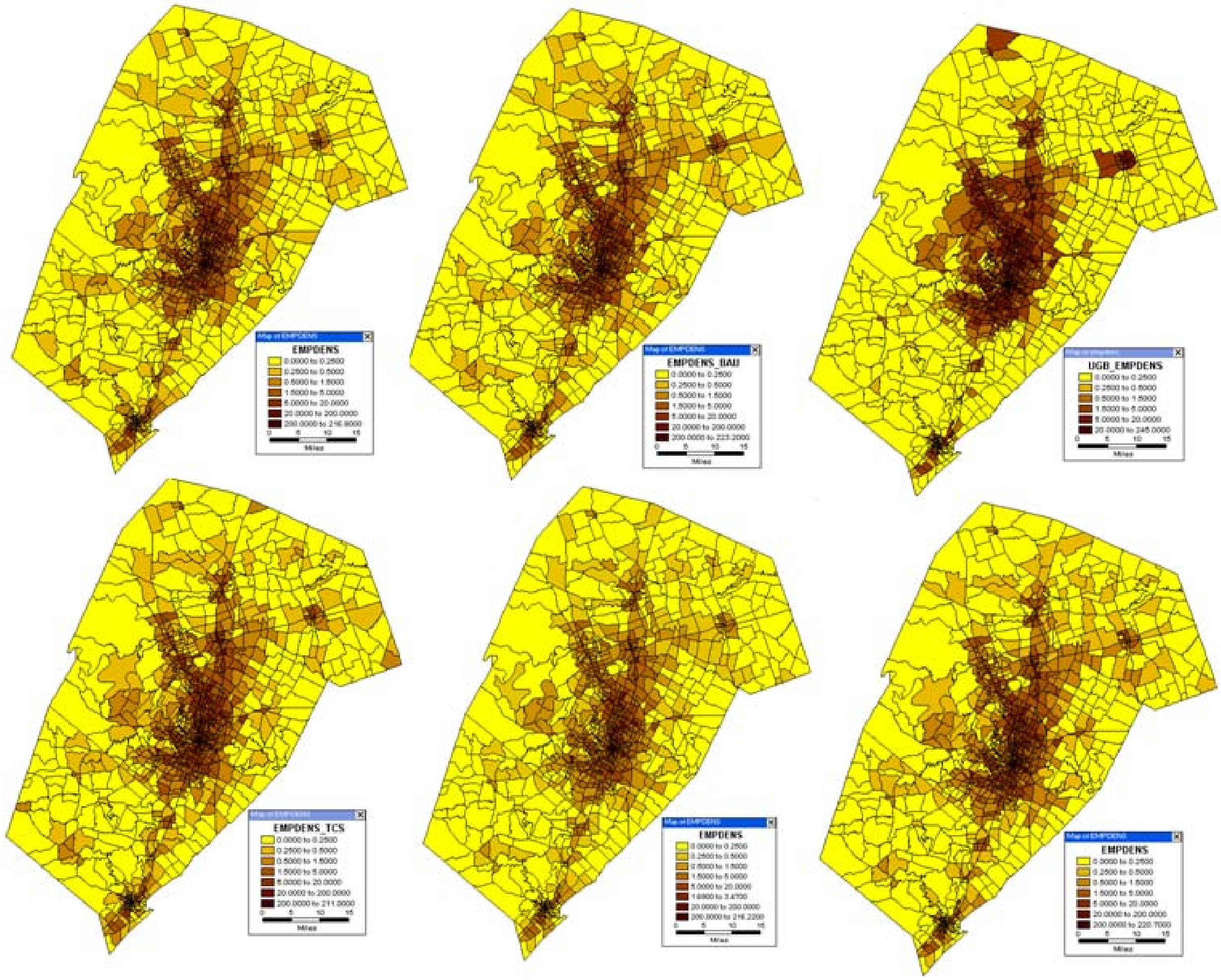


Figure 2: Employment Densities in Year 2030 in NoTDM, BAU, UGB, TCS, EXPAN and SH 130 Scenarios