A Template for Simulating the Urban Dynamics of a City

Tyler J. Frazier, AICP, PhD
Senior Scientist, Department of Transportation System Planning and Telematics
Fakultät V, Technische Universität Berlin
Sekr. SG12, SG 4.1
Salzufer 17-19
10587 Berlin
+49 (0)30-314 22418 (office)
+49 (0)30-314-26269 (fax)
http://www.vsp.tu-berlin.de/frazier/
http://www.vsp.tu-berlin.de/gauss/
tyler.j.frazier@tu-berlin.de

Amongst all species, humans have a highly sophisticated, remarkably unique and individual set of rules which govern our behavioral patterns. The ultimate aggregate exhibition of these behavioral patterns is the phenomenon of urban dynamics and these interactions within the larger biosphere where the city resides. Millions of activities comprise this complex urban system, implemented as individual choices taking place along different temporal scales as spatially distributed throughout the city. Some location choices involve decisions with short term implications, such as travelling to work, school or shopping, while other decisions have a more lasting and long term impact, such as accepting a job, moving into a new home, or deciding to start a business. These two fundamental components, activities related to temporary locations (transportation uses) and activities related to permanent locations (land uses) are the fundamental building blocks for modeling urban dynamics. By understanding these two basic types of activities which occur within the city, a template for simulating the urban dynamics of a city can be constructed. From this basic urban simulation system, models of infrastructures (transportation networks, electricity, water & sewer, solid waste, stormwater and parks and recreation) and institutions (public health, education and safety) can be implicitly incorporated.

An urban simulation system is comprised of several disaggregate data sets, integrated variables and sequential models, which are used to predict the choices of persons, households and businesses in order to spatially simulate the potential land development patterns forming throughout the region. Individual models forecast demographic and economic growth, predict the probability that a particular household, job or business will relocate, and if so, which available new location it will chose, project real estate prices and simulate potential new land developments. Primarily through the use of probabilities derived from surveys and data describing the physical geography of the city, an urban simulation system executes millions of predictions about the decisions of agents interacting with each other as well as within their spatial environment. The foundation of this system is these local interactions, which result in an emergent global structure representing some of the likely spatial land use and development scenarios which could occur in the future.

Central to the creation of an urban simulation system is the development of a number of data sets used

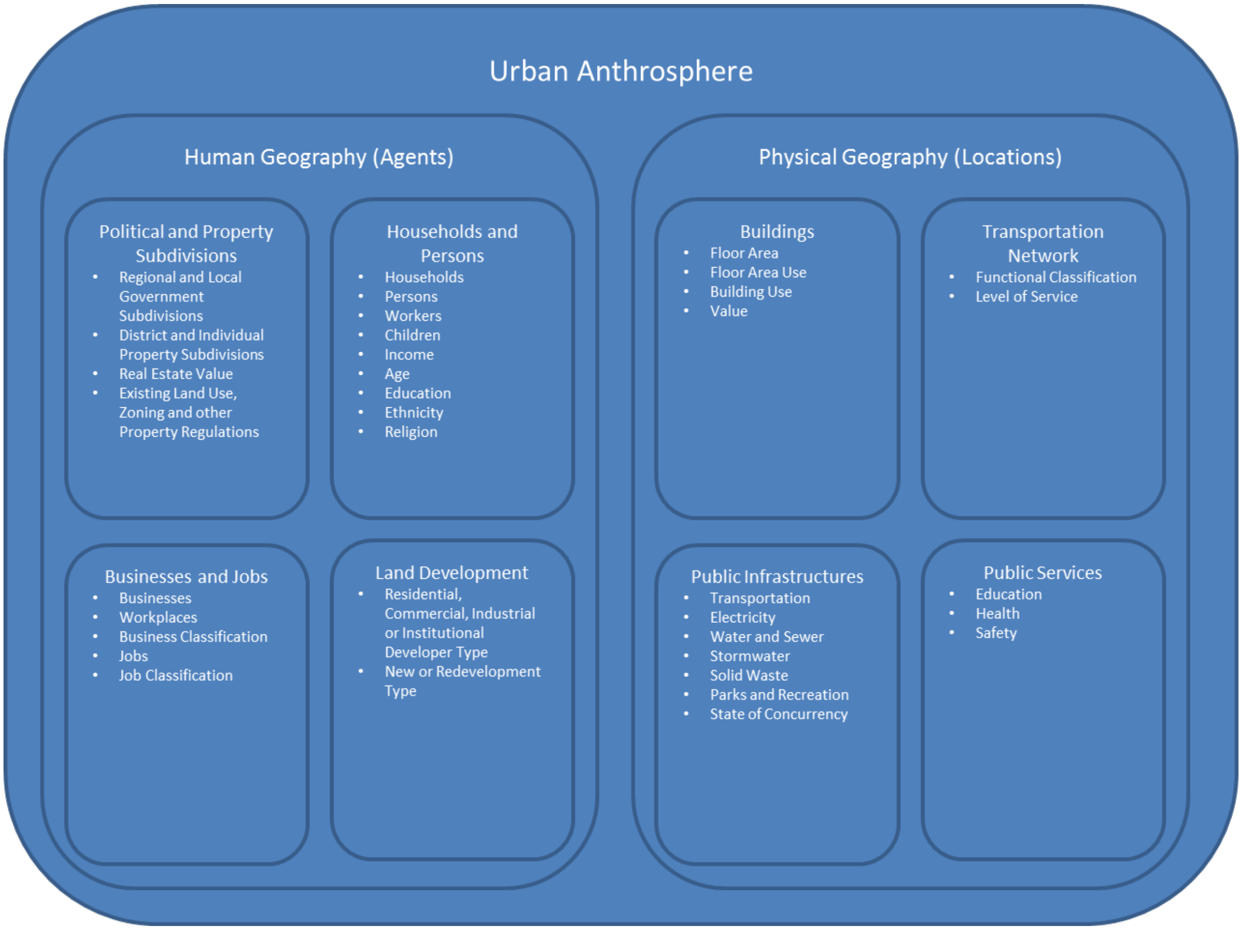

Figure 1 Data sets used in describing the Human and Physical Geography of a City

in the model system which represent the existing condition of the city at the beginning point in time of a scenario run. Agent and Location data sets spatially describe the physical, demographic, and economic geography in terms of each person and household, job and business, as well as building, parcel and zone. The person, household, job and business tables are typically synthetically generated datasets which describe agent attributes including how they choose where to live, work and conduct business. The buildings table describes every physical structure, which is typically available from high resolution aerial photography, while newer approaches enable recording buildings in three dimensions and improved inference of use as derived from the signature of the building envelope. The parcels table describes the fundamental unit of property subdivision and ownership where each building is located and generally is found at the property recording agency survey department but is generally neither digitized, geographically projected nor topologically defined. The zones table is used as a means to describe spatial data at higher levels of aggregation when higher resolution data is not currently available or needed. Agent and Location data sets serve as the fundamental input for the models as each one executes its series of functions, generates output and then updates these tables based on the projection results.

Exogenous allocations and endogenous probabilities provide the number of new agents projected for immigration to the city as well as the probability a new agent will be introduced (births), removed (deaths) or decides to move from its currents location. Allocation data sets include estimates for the total number of new households (by size) and businesses (by sector) projected in terms of a low, medium and high demographic and/or economic scenarios for each simulation year. Relocation data sets include the probability of a single household and business relocating within a given year, and are typically derived from a specialized survey. Statistical Model data sets specify how to spatially calculate a given variable used in a regression or location choice model, as well as the actual parameter used in that calculation.

An urban simulation begins its runtime environment by loading all base year data into a cache for use by each of the models. This model system for predicting land use and development patterns begins with household and business allocation models, continues with the household and business probability models, and is then followed by statistical models for determining real estate value, household and business location choice, and land development and infrastructure demand. As each model completes its initial run, the base year data sets within the cache are updated to reflect the results. This process continues through each annual cycle until the simulation has reached is scenario horizon date.

The model system begins with the Household Transition Model, which estimates allocations for how many new households will move to the urban area during the initial year. The household transition model compares the actual number of households residing in the city with the household control totals table, subtracts the difference, and queues this number of new households for subsequent input to the household location choice model. At the beginning of the simulation, the urban simulation exists as thousands of uninhabited residential locations waiting to be occupied by tens of thousands of households; therefore, all of the synthetically simulated households from the households table are queued for input to the household location choice model during the initial simulation year. Beginning the following year and with all subsequent years, the Household Transition Model also randomly selects households for removal, when relocating households have chosen to move outside of the Metropolitan Area.

Following the household transition model, the Business Transition Model estimates how many new businesses will move to the city during the initial simulation year. The business transition model compares the actual number of businesses with the business control totals table, subtracts the difference, and queues this number of new businesses for subsequent input to the business location choice model. As with the household transition model, at the beginning of the simulation, the city exists as thousands of unoccupied business locations waiting to be occupied by tens of thousands of potential businesses; therefore, all synthetically simulated businesses are initially queued for input to the business location choice model. As local data becomes more readily available and reliable, synthetically generated business data is less likely to be needed, and the focus will turn to describing workers occupying each workplace.

Once the business transition model has completed its run, the Household Relocation Model predicts the probability that a household will move from its current location or remain in place during that particular simulation year. The household relocation model determines the relocation probability for each household and then randomly selects a number to determine if each individual household has been chosen to move from its current location within that simulation year. As households are identified for relocation, they are combined with the new households moving to the region as allocated from the household transition model for input into the household location choice model. After the household relocation model, the Business Relocation Model is initiated in order to predict the probability that a job will move from its current location during that particular simulation year. The business relocation model determines which businesses will be scheduled for relocation based on the probabilities in the relocation rates table.

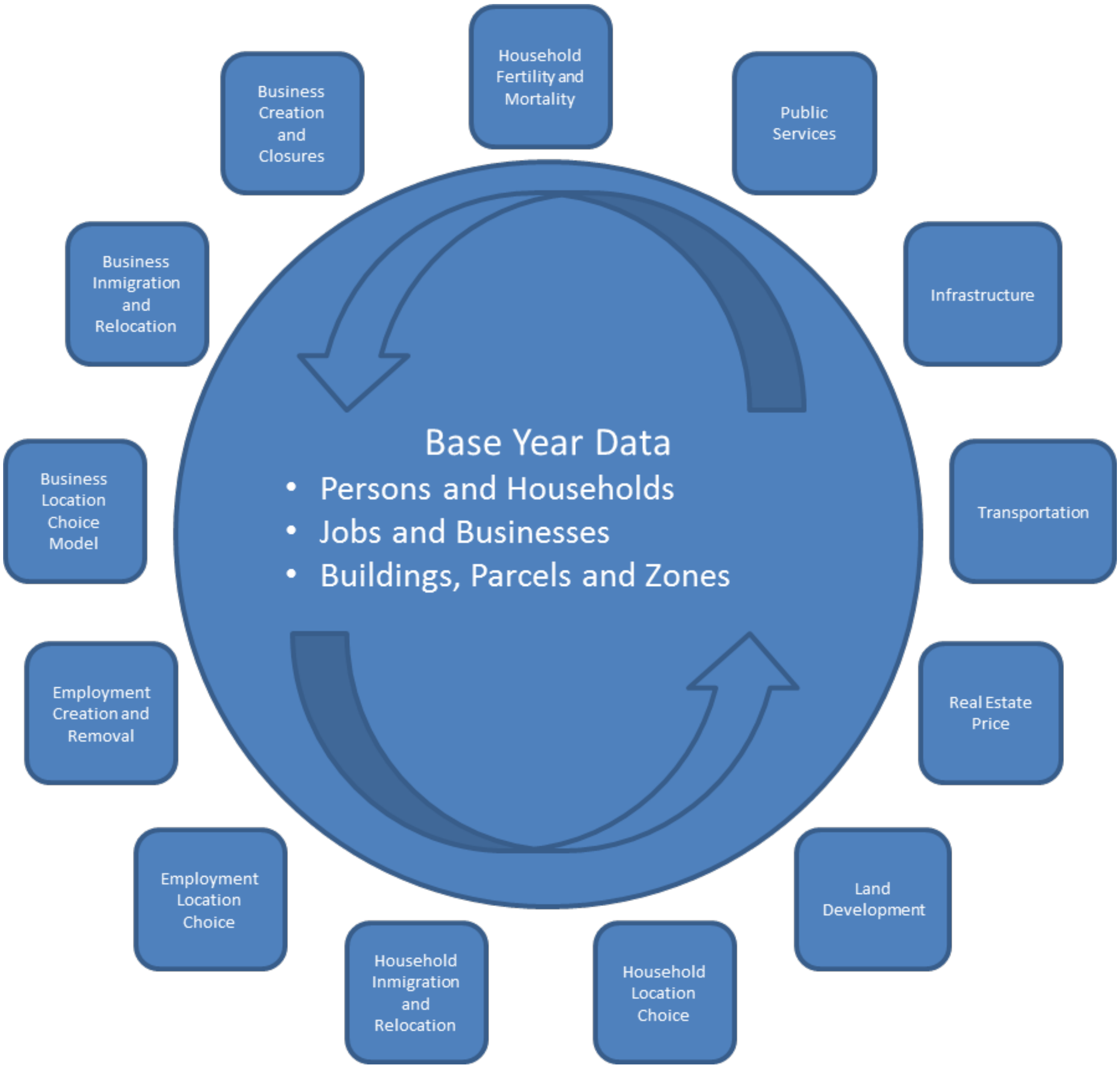

Figure 2 Sequential Models used in an Urban Simulation System

The Real Estate Price Model uses a linear regression model to predict real estate value as dependent to the spatial variables found in the real estate price model specifications table. These six variables have been specified to describe building, household, and employment attributes by zone for each building type and then used to calculate the average_value_per_unit attribute in the buildings table. Model coefficients found in the real estate price model coefficients table are generally estimated from tax assessor or real estate professional's data. The Household Location Choice Model predicts in which particular residential structure a new household (from the Household Transition Model) or an existing household (from the Household Relocation Model) will be located. The household location choice model uses the six spatial variables found in the household location choice model specifications table to calculate the probability of a household selecting a particular location from a set of 30 to 50 available dwelling units. Once a household chooses a location, the buildings table is updated to reflect occupancy of that particular residential unit. Model coefficients parameterizing building area, population density, job density, number of households, work travel time and household income are found in the household location choice model coefficients table.

The Business Location Choice Model predicts in which particular institutional, industrial or commercial structure a new job (from the employment transition model) or an existing job (from the employment relocation model) will be located. The employment location choice model uses the three spatial variables found in the employment location choice model specifications table to calculate the probability of a job selecting a particular location from a set of 30 to 50 available business locations. Once a job chooses a location, the buildings table is updated to reflect occupancy of that particular workplace. Model coefficients parameterizing population density, number of jobs, and work travel time are found in the employment location choice model coefficients table.

The Infrastructure demand model projects infrastructure consumption or demand per parcel from a linear regression model which is typically estimated from historic monthly consumption data. Aggregated demographic and economic attributes can be used to parameterize to projected demand, as dependent to variables describing the number of households and their demographic composition as well as the total number of businesses and jobs and their economic composition. Regression coefficients are applied to aggregated results from the Household and Employment Location Choice models to project total demand.

The result is a basic template for an urban simulation which can be used to project some of the future land development patterns found in a city. The template given here is only in an abstracted state and the specific details for developing each model will largely be effected by the specific location. While included in this generalized template is a basic discussion of transportation modeling, this template only includes data that describes the utility associated with travel times to and from each point within the urban area. A more sophisticated transportation model would include a high resolution simulation that involves daily temporal scales and typically also involves agent "learning algorithms." Additionally, improving synthetically generated data such that the created data sets are more "close to reality" is also an important step. Issues related to indigenous land uses or representing the "informal economy" also require further explanation.

Reference:

Frazier, Tyler. 2011. Powering Accra: Projecting Electricity Demand for Ghana's Capital City. Zentrum fur Entwicklungsforschung, Ecology and Development Series at Cuvillier Verlag, Göttingen.