Calibrating a Synthetic Built Form Generator

John E. Abraham
Institute for Advanced Policy Research
University of Calgary
Calgary, AB
jabraham@ucalgary.ca

Kristen Andersen
Auburn University

Michael Clay
Auburn University

JD Hunt
University of Calgary

Abstract

A system for assigning space (buildings) to parcels to establish a base-year parcel level description of built form is described. The system was applied repeatedly to Autauga County Alabama, where a land-use transport interaction model is being developed. The system sorts parcels according to suitability for different space types, with the details of the sorting processed controlled by user paramaters. Parameters were adjusted to achieve appropriate assignment in one county where target data were available to compare the assignment with observed data. Three map comparison techniques were applied. The resulting parameters will be used in the other counties in the MPO, and may be transferable to other areas in the USA. Major findings include the importance of an accurate zonal-level inventory, the usefulness of quantitative map comparisons, and the need for some information to identify vacant parcels.

Introduction

Integrated land-use and transportation models have significant data needs, requiring base-year information on population, land use, zoning, employment, housing, household income, travel accessibility, and floor space inventory and price. When data of sufficient quality or quantity cannot be found, a model's ability to generate usable results is compromised, and data weaknesses may derail a modeling project altogether. In particular, disaggregate data on built-form—describing the quantity, type, price, and cost of development of built space found in a metropolitan area—is often inconsistent with other regional data or not available at all. A built-form synthesizer will generate a synthetic built form representing real aggregate properties and statistical relationships, allowing modeling projects based on simulation to proceed.

Background

The contribution of synthetic built form to modeling efforts in Oregon, Oahu, and Sacramento, has been documented [1]. In these three modeling efforts, the characteristics of micro-level land use data was adjusted based on higher-level information. In Oregon, a small grid representation was used and top-down allocation approaches were applied. In Oahu, a larger grid representation was used and there was movement of quantities back-and-forth between adjacent grid cells. In Sacramento, a bottom-up approach was used where each parcel was scored and ranked for different floorspace types, and floorspace types were incrementally applied to parcels, their score and rank changing when they become full according to density measures. The work described here is based on the Sacramento algorithm described in [1].

The cartographic literature on "map generalization" describes visualization techniques where original data is modified for presentation purposes. For instance, Sester [6] describes a technique where building shapes and positions are adjusted to simplify shapes and align them with streets. The algorithm is guided by a least-squares difference between the synthetic data and the original data, subject to rules and constraints. These techniques may be useful for more than visualization: if there is reason to believe that the rules and constraints represent reality and the original data is subject to measurement error, then the techniques could be used as a Bayesian updating of prior information (the original data) due to new information (the rules and constraints).

Newkirk [11] describes a system of identifying vacant land inventory and development capacity by considering development constraints, zoning and cadastral datasets. He emphasizes the need for transparency in the process and algorithm since vacant land inventories can have a large impact on policy.

Various authors ([3],[6],[8]) describe methods of determining 3 dimensional models of built form using laser altimeter data in combination with 2 dimensional data. Such methods have not proven successful in land-use transport modeling, but do show promise. Laser altimeter data is not able to determine the type of building, however, so synthetic procedures to assign building type to parcels are still likely to be necessary.

Parcel level data from property tax assessment offices often have useful information on buildings, but information on non-assessed buildings (such as native reserves, government buildings, and churches) are rarely available, and commercial buildings are sometimes assessed only by value without systematic supporting documentation on type, size or age. Jarosz [10] describes a system of routine population estimates in San Diego, where the limitations of both assessor data and census data are described. For example, census structure-type information is based on a sample of information, and the categories do not match San Diego's needs. Thus a hybrid approach is used, with aerial photography resolving large conflicts.

Population synthesizers are commonly used in transportation demand modeling, and use an iterative proportional fitting technique [5], or, more recently, a simulated annealing technique [2]. A population synthesizer creates a list of households and corresponding individuals that reproduce certain aggregate characteristics, including totals and marginal distributions that are specified at different levels of geography. Typically, individual records from a sample are replicated, with each replica having its location changed, along with perhaps some other minor changes. The population synthesizer can be viewed as having two purposes.

1) It allows simulation procedures to be used, in which individual actions are sampled for individual agents. The agents are typically persons, households or jobs. Sampling individual actions from probabilities is much simpler mathematically and computationally than tracking joint probability distributions over many choice dimensions, so more complex problems can be tackled using what, in essence, are Monte Carlo techniques.
2) The relationships between attributes that are evident in the sample of individuals and households is maintained in the population, leading to richer cross relationships of attributes than the simpler marginals that feed the totals and marginal distributions specified to control the synthesizer.

Built form synthesis techniques, as described in this paper, share these two purposes with population synthesizers. To keep the simulation of each unit simple, each unit of land is described with single attributes for almost everything, including type of building, age of building, physical geography attributes and zoning regulations. In this way the diversity of built form is represented across a range of individual units of land. The complex joint distribution of type, age, physical geography and zoning regulations is represented by specifying some marginal totals at higher levels of geography, while other attributes are assigned to individual parcels or grid cells using Geographical Information Systems.

Purpose of study

The purpose of this study is to calibrate the behavior of the floor space synthesizer originally developed for the Sacramento project using data collected in one county of the Montgomery, Alabama region. The resulting parameters will then be of use in the other two counties in the Montgomery region, to improve the Montgomery PECAS ("Production Exchange Consumption Allocation System") land use and spatial economic model [7]. The parameters may also be applicable to other regions.

The initial synthetic built form inventory is required so that the "Space Development" module of PECAS can simulate, through time, the evolution of the region's built-form [9]. The module is a simulation model that represents the larger trends in built-form through a simulation of individual small units of land.

The Montgomery Area Metropolitan Planning Organization directs transportation planning in portions of three counties in central Alabama: Autauga, Elmore, and Montgomery (Figure 1). In the year 2000, the MPO area had approximately 300,000 residents, and approximately two-thirds of this population lived in Montgomery, the state capital.

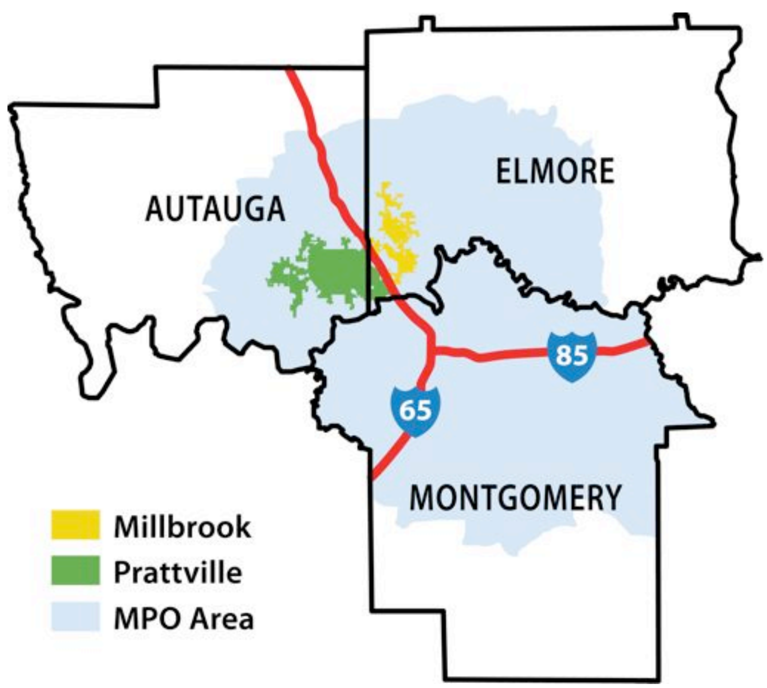

FIGURE 1 Montgomery MPO Area.

The Autauga County portion of the Montgomery Area MPO contains a cross section of urban and rural land uses covering 123,000 acres and 50,000 residents, 31,000 of whom reside in the city of Prattville. More importantly, data on built space in Autauga County has been obtained, making it possible to compare synthesizer results with actual conditions.

The study area is divided into 19,970 units, following actual parcel boundaries obtained from the Autauga County tax assessor. There is a wide range of parcel sizes, from .1 acre to 624 acres, with a mean parcel size of 6.2 acres.

The space development simulation module in PECAS requires small units of land – typically of the order of 1 acre or less [9], so that the random effects of simulating (re)development decisions on each unit of land are small compared to the aggregate or average effects. In the Sacramento model, land was pre-subdivided into "pseudo-parcels" before the synthetic floorspace was generated. In the Montgomery model, as in Baltimore [9], such subdivision is simulated alongside the development decision, negating the requirement of dividing larger parcels before generating the synthetic built-form.

DATA

Inputs

Data about attributes of parcels in the study area were used to influence the floor space synthesizer's space assignment decisions. Zoning classifications provided a key attribute used to guide the synthesizer. Due to geography and land development patterns, certain types of development are more suited to certain areas than others, and municipal zoning boundaries often reflect this. Including zoning discourages the synthesizer from assigning type and quantity of space in a way that violates existing zoning regulations. In some cases, however, there are prior-use developments that may have been legal when constructed but would not be legal under current regulations. If such "non-conforming uses" exist in a region, it is important that the synthesizer be allowed to assign them.

The zoning classifications are shown in Table 1. The set was created by consolidating the myriad zoning categories used by individual municipalities in the three-county area into a uniform set applied to the entire region. Prattville and Millbrook officials provided zoning maps in digital format, and municipal zoning ordinances were studied to determine which consolidated category best matched the regulations applicable in a particular municipal zone.

TABLE 1 Zoning Classifications Used as Floorspace Synthesizer Inputs

R-1	Low Density Residential (minimum lot size > 1 acre)
R-2	Med Density Residential (minimum lot size .25- 1 acre)
R-3	High Density Residential (minimum lot size < .25 acre)
R-4	Mobile Homes- Rural (minimum lot size > .5 acre)
R-5	Mobile Homes- Urban (minimum lot size < .5 acres)
R-6	Multi-family Dwellings and Apartments
O-1	Lower Density Office/Business
O-2	Higher Density Office/Business
I-1	Light Industrial/Manufacturing
I-2	Heavy Industrial/Manufacturing
PUD-2	Single and Multi-family Residential; Lower Density Office

The synthesizer steered synthetic built-form to parcels known to actually contain built space, and data from the Autauga County tax assessor's office was used to identify such parcels.

The total quantity of space-by-type was developed at a Transportation Analysis Zone (TAZ) level through a combination of census data, employment data, space-per-employee data and housing size information. Census data was used to determine the number of dwellings of each type in a zone, as well as the number of employees working in the TAZ. For each dwelling or employee, a certain amount of square feet was allocated to the TAZ for different types of development. The parcels were tagged with their TAZ, and the synthesizer was set up to work one TAZ at a time, assigning the space to parcels within the TAZ.

Four space type categories were used: single-family dwelling, apartment dwelling, mobile home dwelling, and non-residential space.

Comparison Map

To evaluate the synthetic built form in Autauga County, it was compared with county tax assessor data on type and size of building. The amount of built space on each parcel in the study area was provided by the Autauga County tax assessor. This data was disaggregated by building, with buildings not always identified as to type. Multiple buildings on the same parcel were added to produce a single figure representing the total square footage of built space for each parcel. Obtaining information on the type of built space in the study area was not as clear-cut, since manufactured homes and multi-family dwellings were not coded as separate types.

Synthesizer Structure

Scoring System Overview

The built-form synthesizer works through each TAZ in a region, examining various attributes of the parcels in a particular zone and evaluating each parcel's suitability for development. To influence the synthesizer's operation, rewards and penalties are set by the user to encourage or discourage the type or amount of space assigned to a particular parcel. The synthesizer works through the floorspace inventory, decrementing the inventory by small "chunks" and assigning the same chunks to the parcels that received the highest score for the type of space. As parcels approach capacity they receive a penalty, causing their score to drop and eventually causing them to drop below the top of the list, so that other parcels in the TAZ move to the top and receive space. Parcels ranked so low that they are not considered by the synthesizer before all space has been distributed are left vacant. The scoring algorithm works to replicate a real-world phenomenon: land competes for development, and each piece of land tends to attract the type and quantity of development to which it is best suited. When the synthesizer is calibrated properly, space is assigned to parcels in a way that approximates reality.

Scoring System and Match Coefficient Table

The information contained in the Match Coefficient table (Table 2) influences the score given to each parcel by the scoring algorithm. To change the likelihood that the synthesizer will assign space to parcels with a particular attribute, users modify the Match value associated with that attribute. Scores are increased by positive Match values, and decreased by negative Match values. If more than one attribute in the Match Coefficient table applies to a particular parcel, the Match values associated with applicable attributes are added together (e.g. In Table 2, if a Built parcel is located in zone R-1, its Match value for Single Family Dwellings is zero.)

TABLE 2 Portion of Match Coefficient Table

PECASTYPE	FIELDNAME	FIELDVALUE	FAR Target	Match
Single Family Dwelling	Built/Vacant	Built	0	5
Single Family Dwelling	Built/Vacant	Vacant	-0.25	-5
Apartment Dwelling	Built/Vacant	Built	0	5
Apartment Dwelling	Built/Vacant	Vacant	-0.25	-5
Mobile Home Dwelling	Built/Vacant	Built	0	5
Mobile Home Dwelling	Built/Vacant	Vacant	-0.25	-5
Non-residential space	Built/Vacant	Built	0	5
Non-residential space	Built/Vacant	Vacant	-0.25	-5
Apartment Dwelling	Zoning Category	PUD-2: Single and Multi-family Res; Lwr Density Office	1	-5
Apartment Dwelling	Zoning Category	R-6: Multi-family Dwellings and Apartments	1	-5
Mobile Home Dwelling	Zoning Category	R-1: Low Density Residential	0.3	0
Mobile Home Dwelling	Zoning Category	R-5: Mobile Hms- Urban	0.5	0
Non-residential space	Zoning Category	I-1: Light Industrial/Manufacturing	1	0
Non-residential space	Zoning Category	I-2: Heavy Industrial/Manufacturing	0.5	0
Non-residential space	Zoning Category	O-2: Higher Density Office/Business	0.5	0
Non-residential space	Zoning Category	PUD-2: Single and Multi-family Res; Lwr Density Office	0.5	0
Non-residential space	Zoning Category	O-1: Lower Density Office/Business	1.5	0
Single Family Dwelling	Zoning Category	R-1: Low Density Residential	0.3	-5
Single Family Dwelling	Zoning Category	R-2: Med Density Residential	0.5	-5
Single Family Dwelling	Zoning Category	R-3: High Density Residential	0.15	-5
Single Family Dwelling	Zoning Category	PUD-2: Single and Multi-family Res; Lwr Density Office	1	-5

To influence the quantity of space assigned to parcels, users provide Floor Area Ratio Targets ("FAR Targets") associated with particular attributes (e.g. the FAR Targets for non-residential space should be higher than those for single family dwellings, so that synthesized space will reflect the increased densities in which non-residential space is built). These values should approximate actual development densities in the area for which built space is being simulated. FARs are a widely available input, since maximum allowable FARs are published in municipal zoning ordinances, or actual FARs can be estimated by those with knowledge of the area. When assigning floor space, the synthesizer calculates a parcel's FAR—the quantity of space assigned to the parcel divided by the size of the parcel. When the quantity of floor space assigned to a parcel exceeds 70% of its FAR Target, its score is reduced in the amount of:

$$(FAR - farTarget_s * 0.7) / farTarget_s * 3.0 \qquad (1)$$

where:

$farTarget_s =$ the total of all FAR Targets applicable to a particular parcel (Like Match values, the effects of FAR Targets are cumulative if a particular parcel has more than one attribute included in the Match Coefficient table)

Within a particular TAZ, the parcel deemed most suitable for a particular type of space will be given all of that space, until the parcel fills to the point where the penalty causes it to drop down the sorted list.

When the synthesizer assigns space of a particular type within a TAZ, parcels assigned that space type earlier in the process are given a bonus to their score for that space type of 0.5. This increases the likelihood that parcels will continue to receive space of the same type, and that space will accumulate on parcels that have already been assigned space, instead of being assigned to new vacant parcels.

Conversely, if a parcel has previously been assigned a different space type, it receives a penalty of 5.0. However, if the penalized score for such a parcel still exceeds the scores of other parcels in the TAZ, its existing space type will be swapped with the conflicting type, so that all space of every type can be assigned without dramatically exceeding FAR targets on particular parcels, and without having more than one type of space on any one parcel.

The synthesizer assigns space in proportion to the total amount of space to be distributed within the TAZ. That is, when the synthesizer is halfway through distributing the total amount of space alotted to a particular TAZ within the Floorspace Inventory table, the synthesizer will have assigned 50% of each space type.

Floor Space Inventory Table

A portion of the Floor Space Inventory table is shown in Table 3.

TABLE 3 Portion of Floor Space Inventory Table

TAZ	Commodity	Quantity	Chunk Size
269	Single Family Dwelling	386716.69	500
269	Apartment Dwelling	84447.90	500
269	Mobile Home Dwelling	48584.80	500
269	Non-residential space	76772.50	150
270	Single Family Dwelling	206658.13	500
270	Apartment Dwelling	45128.24	500
270	Mobile Home Dwelling	25963.31	500
270	Non-residential space	45475.00	150
271	Single Family Dwelling	43991.58	500
271	Apartment Dwelling	9606.51	500
271	Mobile Home Dwelling	5526.84	500
271	Non-residential space	84262.50	150
272	Single Family Dwelling	41945.46	500
272	Apartment Dwelling	9159.69	500
272	Mobile Home Dwelling	5269.78	500
272	Non-residential space	53232.50	150

By manipulating values in the Chunk Size field, users change the square footage of built space the synthesizer allocates to parcels at a time. The use of larger space "chunks" allows the synthesizer to run more quickly, but smaller units of space increase the synthesizer's ability to assign space to other parcels whe a parcel's FAR target is approached..

Parcel Table

Information about the attributes of individual parcels is contained in the Parcel table, a portion of which is shown in Table 4. Field names referring to specific attributes must appear in both the Parcel and Match Coefficient tables to be considered by the floor space synthesizer.

TABLE 4 Portion of Parcel Table

Parcel Number	TAZ	City	Zoning Category	Built/ Vacant	Parcel Area (sq ft)	Actual Built Area (sq ft)	Assigned Built Area (sq ft)	Actual Type of Built Space	Space Type Assigned by Synthesizer
10093210000070840	305	Prattville	R-2: Med Density Residential	Built	2.93	2006	150	Single Family Dwelling	Non-residential space
10093240020030000	305	Prattville	R-2: Med Density Residential	Built	3.43	1928	500	Single Family Dwelling	Apartment Dwelling
10093240020040160	305	Prattville	R-2: Med Density Residential	Built	0.07	2080	150	Single Family Dwelling	Non-residential space
10093240020030080	305	Prattville	R-2: Med Density Residential	Vacant	4.10	0	0	None	None
10093240020040350	305	Prattville	R-2: Med Density Residential	Vacant	2.67	0	0	None	None
10093240020030160	305	Prattville	R-2: Med Density Residential	Built	0.03	1968	8500	Single Family Dwelling	Single Family Dwelling
10093240020030190	305	Prattville	R-2: Med Density Residential	Built	2.85	2163	500	Single Family Dwelling	Single Family Dwelling
10093240010070730	305	Prattville	R-2: Med Density Residential	Built	0.07	2557	500	Single Family Dwelling	Apartment Dwelling
10093210000070820	305	Prattville	R-2: Med Density Residential	Built	0.12	1842	150	Single Family Dwelling	Non-residential space
10093210000070030	305	Prattville	R-1: Low Density Residential	Vacant	3.61	0	0	None	None
19010210000450000	291	Prattville	R-3: High Density Residential	Built	4.86	1645	0	Apartment Dwelling	None
19010210000430000	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000450855	291	Prattville	R-3: High Density Residential	Built	4.86	1645	0	Apartment Dwelling	None
19010210000430463	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000450000	291	Prattville	R-3: High Density Residential	Built	4.86	1645	0	Apartment Dwelling	None
19010210000430894	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000450000	291	Prattville	R-3: High Density Residential	Built	4.86	1645	0	Apartment Dwelling	None
19010210000430008	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000430962	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000030000	291	Prattville	R-3: High Density Residential	Built	4.12	4495	0	Apartment Dwelling	None
19010210000030331	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	4495	0	Non-residential space	None
19010210000430000	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
19010210000430963	291	Prattville	O-2: Higher Density Office/Business	Built	1.66	1740	0	Non-residential space	None
10093240020040460	305	Prattville	R-2: Med Density Residential	Built	0.07	2110	150	Single Family Dwelling	Non-residential space
10093240020040390	305	Prattville	R-2: Med Density Residential	Built	0.10	1995	500	Single Family Dwelling	Apartment Dwelling
10093240020040420	305	Prattville	R-2: Med Density Residential	Built	1.59	2028	150	Single Family Dwelling	Non-residential space
19020330110090000	285	Prattville	R-2: Med Density Residential	Built	1.04	2617	150	Single Family Dwelling	Non-residential space
19010210000430000	291	Prattville	O-2: Higher Density Office/Business	Built	1.67	1740	0	Non-residential space	None
19011220010490000	293	Prattville	R-6: Multi-family Dwellings and Apartments	Built	1.14	1504	150	Apartment Dwelling	Non-residential space
09073500000130000	302	None	None	Vacant	3.64	0	0	None	None
10093240010040000	305	Prattville	R-2: Med Density Residential	Built	0.33	1823	500	Single Family Dwelling	Mobile Home Dwelling
10093240010040031	305	Prattville	R-2: Med Density Residential	Built	0.54	1823	500	Single Family Dwelling	Mobile Home Dwelling
10093240010040660	305	Prattville	R-2: Med Density Residential	Built	1.36	1823	500	Single Family Dwelling	Mobile Home Dwelling

Methods

Calibration efforts focused on three tasks: 1) directing synthesized development to actual built parcels; 2) directing the correct space type to as many of these developed parcels as possible; and 3) directing synthesized built space to developed parcels in amounts resembling actual quantities.

Visual Comparison

During this study, the floor space synthesizer was run over 100 times, using different inputs. After each run, the table of outputs was imported into a Geographical Information System, allowing results to be viewed spatially. The synthesizer-generated map was visually compared with the official map showing actual space development in the study area. This technique provided a powerful—though qualitative—view of how input changes had been processed by the synthesizer (see Figure 2, showing central Prattville in an early run of the system).

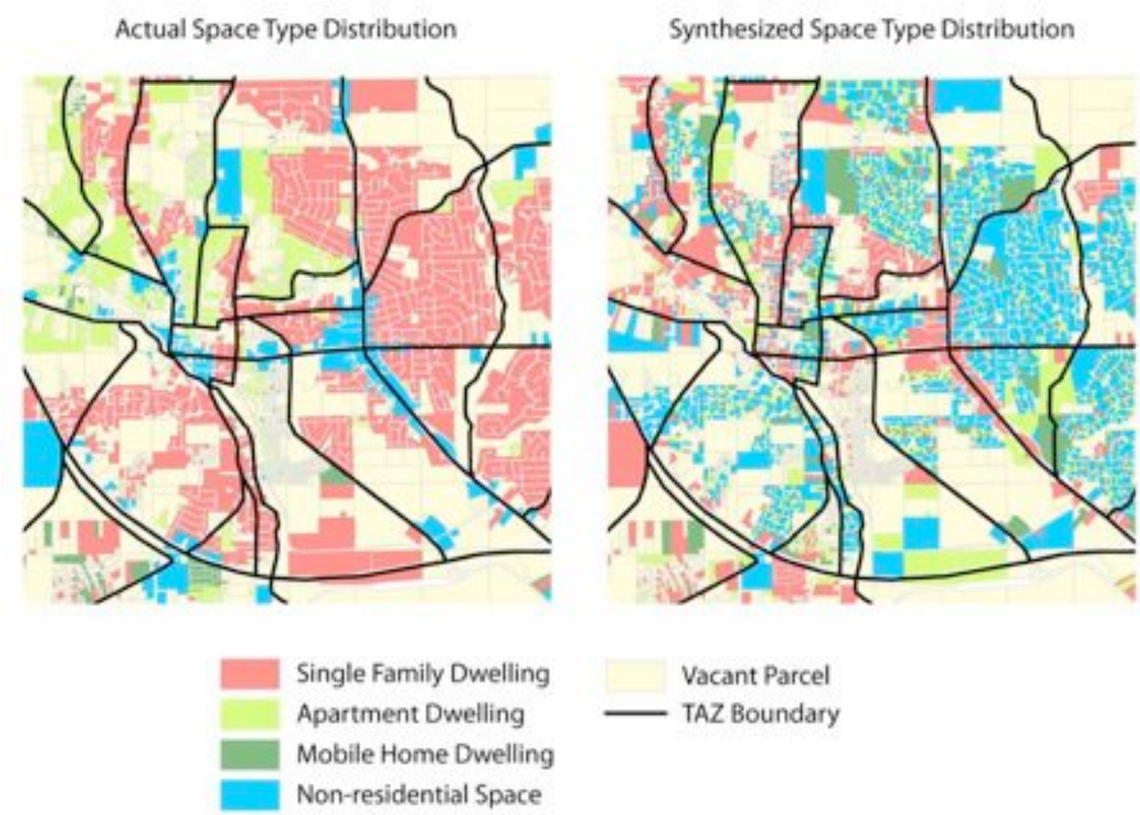

FIGURE 2 Example of visual map comparison technique: actual and synthesized space distribution in a portion of the study area.

Scoring system

A scoring system was developed to measure how well the amount and type of space assigned to parcels corresponded to development from the assessors data. The scoring system was relatively simple: for each parcel, if the assigned type is the same and the assigned floorspace quantity is within 10% of the actual sqft, the parcel score is 5. If within 20%, score is 4. If within 30%, score is 3. If within 50%, score is 2, if within 100%, score is 1. If greater than 100% (and type matches), score is 0.5. If type does not match score is zero, regardless of quantity assigned. The scores for each parcel were added together to get a total score for the synthetic space.

This scoring system is relatively simple, yet it proved valuable as a quantitative measure of the quality of the synthetic built form. More sophisticated map comparison techniques are described and compared in [12].

Summary table

In addition, two summary tables were created. The first showed the proportions of synthetic square feet that were assigned to parcels with the matching type. The second shows the proportions of parcels that had a synthetic space type that matched the recorded space type. Table 5 shows these two tables.

TABLE 5 Floor Space Synthesizer Output Analysis Worksheet

Output Score: 3924

Square Feet

Actual		Proportions		Assigned S	Assigned A	Assigned M	Assigned N	Assigned V	Assigned TOTAL
S	19,461,357	60.0%		5,770,657	833,894	740,745	804,150	0	8,149,445
A	4,798,279	14.8%		717,799	135,734	51,756	112,765	0	1,018,053
M	1,190,086	3.7%		315,000	68,500	57,233	23,100	0	463,833
N	6,976,814	21.5%		325,500	37,766	8,771	2,039,560	0	2,411,597
V	0	0.0%		4,569,229	1,427,126	504,069	744,255	0	7,244,678
TOTAL	32,426,536	100.0%		11,698,184	2,503,019	1,362,573	3,723,830	0	19,287,607
			Assigned Proportions	60.7%	13.0%	7.1%	19.3%	0.0%	100.0%

Parcels

Actual		Proportions		Assigned S	Assigned A	Assigned M	Assigned N	Assigned V	Assigned TOTAL
S	9,520	47.7%		1,579	1,668	830	5,361	82	9,520
A	2,455	12.3%		1,323	94	104	752	182	2,455
M	975	4.9%		630	137	54	154	0	975
N	878	4.4%		651	76	18	74	59	878
V	6,142	30.8%		36	40	24	22	6,020	6,142
TOTAL	19,970	100.0%		4,219	2,015	1,030	6,363	6,343	19,970
			Assigned Proportions	21.1%	10.1%	5.2%	31.9%	31.8%	100.0%

The focus was on finding values in the Match Coefficient table that lead to a good score, and hence a good assignment of space to parcels. But during analysis it was found that the mix of space types by TAZ in the floorspace inventory table was inappropriate for the parcel file; that in some TAZ's the parcel file indicated no parcels with certain types of housing. Thus, changes to the Floorspace Inventory table were made as well, between separate groups of runs, to move residential space between types.

Results and Analysis

Preference to Actual Built Parcels

In order to guide development toward parcels within urbanized Prattville, it was necessary to include an input in the Match Coefficient table causing the synthesizer to give preference to parcels with actual built space. Without this component, the synthesizer could not simulate the disparity in number of developed parcels that exists between urban and rural areas. The addition of this component ensured a realistic distribution of developed parcels during all the tests that followed, though space was assigned to fewer parcels after the addition of other attributes to the Match Coefficient table.

Giving preference to parcels with actual built space ensures that only a few actual vacant parcels are assigned space by the synthesizer.

Floor Space Inventory

The proportions of assigned square footage of each type are strongly influenced by the space type proportions used in the Floor Space Inventory.

The total amount of square feet listed in the Floor Space Inventory table was significantly less than the actual built space in the study area (as can be seen in Table 5). According to the assessor, 35.5 million square feet of built space exists in the study area. The floorspace inventory table contained only 20 million square feet. The system for generating TAZ space totals based on employment and population could be improved based on this discrepancy. Recall that for the other counties in the MPO, space data were not available, but increasing

Chunk Sizes

Several runs tested the effects of modifying the size of space chunks assigned by the synthesizer. Changing the residential chunk size from 500 to 50 square feet, and the non-residential chunk size from 150 to 50 square feet dramatically reduced the Output Score, meaning that the amount and type of space assigned to parcels decreased in similarity to development shown on the “official” map. As discussed in 5.3, reducing chunk sizes makes it easier for the synthesizer to stop assigning space to a parcel before exceeding the FAR Target, which spreads space to a larger number of parcels. This provided an indication that many of the FAR Targets were too low, and hence these were increased during the calibration exercise. In particular, every parcel without a matching attribute in the Match Coefficient table has a default FAR Target of zero. Therefore, *any* space assigned to these parcels exceeds their FAR Target, leading to an infinite penalty (equation 1). The Match Coefficients table was adjusted so that every parcel had some match that led to a non-zero FAR Target.

FAR Targets

Many tests involved the manipulation of FAR Targets in order to change the proportion of parcels assigned a particular type of space relative to another type. However, the ability of FAR Targets to control these proportions was limited, since the Floor Space Inventory also influences the proportions of space in TAZs assigned to a particular type.

In one test, FAR Targets for Mobile Home Dwellings were increased (from .5 to 1.5), in an effort to concentrate this type of development on fewer parcels. Though this adjustment should have allowed parcels assigned Mobile Home space to receive more square footage before the FAR penalty was activated, this was not the case. However, a test that reduced the FAR Target for Single-Family Dwellings, in order to distribute this type of development to more parcels, had the desired effect.

Match Values

Like decreasing chunk sizes and FAR Targets, reducing Match values also spread assigned space to a greater number of parcels, since the penalty calculated in equation 1 then became larger in proportion to the original Match values. Reducing Match values decreases the relative attractiveness of any particular parcel, distributing space more evenly among parcels in a TAZ. The synthesizer was much more responsive to changes in FAR Targets than Match values, although the two work together because Match values need to be reduced by the values calculated in equation 1 before a parcel will move from the highest scored position to a lower position in the scored list of parcels.

Lower FAR Targets allowed individual parcels to reach capacity quicker than parcels with higher FAR Targets, and the total inventory of space was assigned to a larger number of parcels.

Conclusions and Future Research

Data Availability

Throughout this study, the general availability of data used as synthesizer inputs was kept in mind. Since the most common use of the synthesizer is to compensate for data weaknesses, the synthesizer was calibrated using data that can be obtained by many users.

The Match Coefficient table must include an input to increase the number of parcels developed in urban areas relative to rural areas, within a TAZ. This study used assessor data indicating parcels containing built space. The synthesizer did not perform well without parcel data of this nature. Alternative attributes that could be used, if this attribute is not available, include an “urban vs. rural” attribute from census, or land classifications based on aerial photography. Jarosz [10] shows how aerial photography can be used, and the Atlanta Regional Commission has shown how aerial photos can be used to classify land use [4].

Users can apply their own knowledge of the regions where they live and work to improve synthesizer calibration. They may be able to create new datasets for input to the synthesizer. Knowledgeable users’ expectations of the results a synthesizer *should* guide them to modify inputs in a certain way. The floor space synthesizer enables users to apply information—from any source—about the type and quantity of built space in an area to individual parcels in a database, preparing a data set that can be utilized by a land-use and transportation model.

Additional Input Attributes

Numerous other attributes could be added to the Match Coefficient table in order to guide space assignment, varying according to the data that can be obtained for a particular region. Distance-to-roadway is an input that is widely available, and this data would cause the synthesizer to consider a particular parcel’s distance from major roads, better simulating actual space development patterns (e.g. encouraging the assignment of non-residential space very close to a highway while encouraging the assignment of residential space slightly further away). The lack of utility infrastructure on a particular parcel could be used to discourage space assignment. Development could be diverted from wetlands, floodplains, parks, schools, airports, and areas of excessive slope unless the synthesizer calculated a demand for space on such parcels strong enough to exceed assigned penalties, as in [11].

If quality floor space data is only available for a portion of a region, floor space could be pre-assigned to those parcels in this area, with the synthesizer distributing the remaining space.

Floor Space Inventory Improvements

The relationship between floorspace and population and employment is not consistent across the nation, and this study has shown that proportions borrowed from other modeling efforts are not appropriate for this county in Alabama. Any efforts to improve the floorspace use formulas are likely to be valuable. In the Montgomery AL region, further work could be done to determine the relationship between quantities of space by TAZ in the one county to the population and employment in the county. These improved relationships could then be used to determine space totals in the other counties extrapolating amongst counties in the same region.

Improvements to Floor Space Synthesizer Design

Changes could be made to the design of the synthesizer itself that may increase the validity of results and ease of calibration efforts.

Currently, the synthesizer is not responsive to numeric fields in the Parcel table, meaning it cannot give preference to parcels that have a higher (or lower) value for some numeric attribute. To use numerical attributes (e.g. distance-to-roadway and percentage of slope), users must first create a separate field in the Parcel table to identify parcels fitting into a specific range of the numeric attribute. Adding the ability to process numerical fields directly would make it easier for a wider variety of physical attributes to be used as synthesizer inputs.

In particular, the FAR Target could be modified by a parcel-level FAR calculated within a GIS system, so that the synthesizer's penalty for FAR in Equation 1 is guided by a more sophisticated consideration of parcel attributes than can be calculated using the Match Coefficients table.

A system could also be added to provide a score bonus based on parcel adjacency – so that similar types of space would cluster in the same way that similar types of space tend to cluster in reality to capture economies of scale.

A random perturbation to the score would break ties, and add entropy (randomness) to the resulting distribution. This would especially improve visualization, as the order of storage in the database currently breaks ties in scores, often leading to unrealistic clustering in the compass corners of TAZs.

In this study, much was learned about the behavior of the synthesizer, and many issues were raised that merit further investigation. The large number of tests completed, coupled with the limited number of inputs and relatively small study area, facilitated in-depth study of synthesizer results. The Output Score—developed during this study—provided a convenient, objective way to gauge the success of tests, and will help future users of the synthesizer. Many of the inconsistencies between actual and assigned space that were encountered during this study could be traced back to the zonal level Floor Space Inventory. This suggests that any effort to compare good space data, wherever it exists, with population and employment data, is likely to lead to improved inputs to the generation of a synthetic built form.

REFERENCES

1. Abraham, J.E., T. Weidner, J. Gliebe, C. Willison, and J.D. Hunt. "Three Methods for Synthesizing Base-Year Built Form for Integrated Land Use-Transport Models." *Transportation Research Record: Journal of the Transportation Research Board* 1902 (2005): 114-123.
2. Abraham, J. and JD Hunt "Documentation for SYNTHESIZER Program", February 2006, HBA Specto Incorporated, Calgary, AB (http://hbaspecto.com)
3. Alharthy, Abdullatif and James Bethel, Detailed Building Reconstruction from Airborne Laser Data, 20th ISPRS Conference, Istanbul, Turkey
4. Atlanta Regional Commission, Landpro GIS dataset, available in July 2008 from http://atlantaregional.com -> info centre -> gis data (2007 data is at http://www.atlantaregional.com/documents/gismeta_LandPro_07.zip)
5. Beckman, R. J., Keith A. B., and Michael D. M. Creating Synthetic Baseline Populations. Transportation Research A. Vol. 30. No. 6 (1996), 415-429.
6. Brenner, C. "Towards fully automatic generation of city models", International Archives of Photogrammetry and Remote Sensing, Volume 33, Part B3, Commission III, pp 85-92, Amsterdam, July 16-23, 2000. http://citeseer.ist.psu.edu/brenner00towards.html
7. Clay, M.J., K. Mullins, N. Woerner, K. Andersen, J.E. Abraham and J.D. Hunt, Developing an Integrated Land-Use/Transportation Model for Small to Medium-Sized Cities: Case Study of Montgomery, Alabama, presented at the 10th International Conference on Computers in Urban Planning and Urban Management, 11-13 July 2007 Iguassu Falls, PR, Brazil, and included on the conference CD-Rom
8. Haala, N., Brenner, C., and Andres, K-H. 1998. 3D urban GIS from Laser altimeter and 2D map data, ISPRS J. Photogrammetry &Remote Sensing, 32, Part 3/1, pp.339-346.
9. Hunt, JD, JE Abraham, D De Silva, M Zhong, J Bridges and J Mysko, Developing and Applying a Parcel-Level Simulation of Developer Actions in Baltimore. Presentated at the 87th Transportation Research Board Annual Meeting and possible publication on Transportation Research Record.
10. Jarosz, B., Using Assessor Parcel Data to Maintain Housing Unit Counts For Small Area Population Estimates, Chapter 5 in *Applied Demography for the 21st Century*, S.H. Murdock and D.A. Swanson, eds., Springer Science+Business Media, 2008, pp 89.101
11. Newkirk, Gregory A, Determining Vacant Buildable Lands Using Vector and Cadastral Data: Problems in Model Building, 20th Annual ESRI International User Conference, San Diego, California, (accessed at http://gis.esri.com/library/userconf/proc00/professional/papers/PAP263/P263.HTM)
12. Sester, Monika, Generalization Based on Least Squares Adjustment, Institute for Photogrammetry, Stuttgart University, Germany, International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B4. Amsterdam 2000, pp 931-938
13. H. Visser (editor), The MAP COMPARISON KIT: methods, software and applications, RIVM report 550002005/2004, Bilthoven, NL