World Academy of Science, Engineering and Technology
International Journal of Architectural and Environmental Engineering
Vol:8, No:3, 2014

TELUM Land Use Model: An Investigation of Data Requirements and Calibration Results for Chittenden County MPO, U.S.A.

Georgia Pozoukidou

Abstract—TELUM software is a land use model designed specifically to help metropolitan planning organizations (MPOs) prepare their transportation improvement programs and fulfill their numerous planning responsibilities. In this context obtaining, preparing, and validating socioeconomic forecasts are becoming fundamental tasks for an MPO in order to ensure that consistent population and employment data are provided to travel demand models. Chittenden County Metropolitan Planning Organization of Vermont State was used as a case study to test the applicability of TELUM land use model. The technical insights and lessons learned from the land use model application have transferable value for all MPOs faced with land use forecasting development and transportation modeling.

Keywords—Calibration data requirements, land use models, land use planning, Metropolitan Planning Organizations.

I. Introduction

It is a fact that for the last two decades Metropolitan Planning Organizations (MPOs) have played a significant role in shaping the future of metropolitan areas in the U.S. These agencies are responsible for major transportation and land use planning decisions and are required to assess the impact of their transportation and land use policies. Hence, these agencies are mandated to use sophisticated information management tools and complex land use modeling methods.

TELUM - Transportation Economic and Land Use Model, is a land use simulation model that can be used as a forecasting and policy analysis tool. Development of TELUM, which is part of a larger decision support system, was initiated and funded by Federal Highway Administration (U.S. Department of Transportation), with Rutgers University and North Jersey Planning Authority being responsible for designing and developing the system. Funding started in 1998 with $1 million per year, over six years period. In 2005 new funding was approved focusing on the implementation and widespread adoption of the system. Today the system is copyrighted and every MPO is eligible to use TELUM at no cost.

The purpose of this study is to examine how land use simulation models like TELUM can be used by middle-sized MPOs to address their planning responsibilities. The paper starts with the description of the employment and household allocation models emended in TELUM and the required data inputs for the forecasting procedure. It continues with the explanation of the calibration process, a constrained gradient search procedure, which is used to estimate the equations coefficients for TELUM's emended models. Finally the paper provides detail description and illustrations of applying TELUM in Chittenden County MPO South Burlington, VT, and discusses calibration results of the applied models.

Georgia Pozoukidou is with the Department of Planning and Development, Aristotle University of Thessaloniki, 54006 Thessaloniki, Greece (phone: +30 2310 995444; email: gpozoukid@plandevel.auth.gr).

II. Forecasting with TELUM

Basic parts of TELUM land use model are DRAM and EMPAL a residential and an employment location model emended with other auxiliary modules in one system. Following is a brief description of these two basic parts of TELUM, focusing mainly on the structure of the equations that models utilize. These equations are the final product of an extensive 40 year research of employment and residential location models that has been developed by Dr. S.H. Putman, Professor at the Department of City and Regional Planning, University of Pennsylvania [1].

A. The Employment Location Model – EMPAL

EMPAL is a modified version of the standard singly-constrained spatial interaction model. There are three modifications: 1) a multivariate, multiparametric attractiveness function is used, 2) a separate, weighted, lagged variable is included outside the spatial interaction formulation, and 3) a constraint procedure is included in the model, allowing zone and/or sector specific constraints [2].

EMPAL model normally uses for 4-8 employment sectors. Until EMPAL was released most of the work done, in the field of forecasting spatial distribution of employment, was splitting employment into two categories: Basic and non-Basic. The fact that EMPAL model uses four employment types was a significant modification. The parameters λ, α, β, a and b of the equation are estimated individually for each one of the employment types through the calibration process that will be discussed later. The equation structure used for EMPAL and for this project is as follows:

$$E^k_{j,t} = \lambda^k \sum_i P_{i,t-1} A^k_{i,t-1} W^k_{j,t-1} c^{\alpha^k}_{i,j,t} \exp(\beta^k c_{i,j,t}) + (1-\lambda^k) E^k_{j,t-1} \tag{1}$$

where

$$W^k_{j,t-1} = (E^k_{j,t-1})^{a^k} L^{b^k}_j \tag{2}$$

and

$$A^k_{i,t-1} = \left[\sum_\ell (E^k_{\ell,t-1})^{a^k} L^{b^k}_\ell c^{\alpha^k}_{i,\ell,t} \exp(\beta^k c_{i,\ell,t}) \right]^{-1} \tag{3}$$

where

$E^k_{j,t-1}$ = employment (place of work) of type k in zone j at time t-1

$E^k_{j,t}$ = employment (place of work) of type k in zone j at time t

L_j = total area of zone j

$c_{i,j,t}$ = impedance (travel time or cost) between z ones i and j and time t

$P_{i,t-1}$ = total number of households in zone I at time t-1

$\lambda^k, \alpha^k, \beta^k, a^k, b^k,$ = empirically derived parameters

B. The Residential Location Model - DRAM

DRAM is also a modified version of a singly – constrained spatial interaction model. There are two major modifications: 1) a multivariate, multiparametric attractiveness function is used, 2) a consistent balanced constraint procedure is included in the model, allowing zone and/or sector specific constraints. The model is normally used for 3-5 (the current maximum is 8) household categories whose parameters are individually estimated. A more detail description of model's structure is available in many texts written by Pr. Putman [2], [3]. The equation structure that we will be using is as follows:

$$N^n_i = \sum_j Q^n_j B^n_j W^n_i c^{\alpha^n}_{i,j} \quad (4)$$

where

$$Q^n_j = \sum_k a_{k,n} E^k_j \quad (5)$$

and

$$B^n_j = \left[\sum_j W^n_i c^{\alpha^n}_{i,j} \right]^{-1} \quad (6)$$

and

$$W^n_i = (L_i)^{q^n} (x_i)^{r^n} (L^r_i)^{s^n} \prod_{n'} \left(1 + \frac{N^{n'}_i}{\sum_n N^n_i} \right)^{b^n_{n'}} \quad (7)$$

where

E^k_j employment of type k (place of work) in zone j

N^n_i households of type n residing in zone i

L^v_i vacant developable land in zone i

x_i 1.0 plus the percentage of developable land already developed in zone i

L^r_i residential land in zone i

$a_{k,n}$ regional coefficient of type n households per type k employee

$c_{i,j}$ impedance (travel time or cost) between zones i and j

$\alpha^n, q^n, r^n, s^n, b^n_{n'}$ empirically derived parameters

DRAM is also capable of including additional attractiveness variables in the spatial potential term Wi, but there has been little use of this option in practice because such variables requires the subsequent development of a means for their updating in forecast runs of the model.

C. Required Data Inputs for the Forecasting Procedure

The forecasting procedure starts with EMPAL. The model usually uses four to eight employment types/sectors. For each one of employment type a parameter will be estimated. In order to proceed with the forecasting procedure EMPAL needs the following input data: employment by type in all zones, population by income in all zones, total area for all zones per zone, zone to zone travel cost or travel time between all zones.

After employment location forecasting is performed by EMPAL, residential location forecasts will be performed using DRAM module of TELUM. The model uses four to six household types, which represent different income groups i.e. high income, low income etc., the parameters of which will be individually estimated. To forecast the location of residents, DRAM needs the following input data: residents of all types in all zones at time t, land use for residential purposes in each zone at time t, the percentage of developable land that has already been developed in each zone at time t, the vacant developable land in each zone at time t, zone to zone travel cost, employment of all types in all zones at time t+1. The residential location forecasts produced by DRAM are then used as inputs to generate and distribute trips, split trips by mode and then assign vehicle trips to the transportation network.

D. Calibration

Calibration is the process of fitting DRAM and EMPAL models into the real world by estimating the parameters for each locator type (i.e. high income households, manufacturing etc), which will be used in models' equation. These parameters will be the ones that best fit in the general model structure of the dataset and will minimize the discrepancies between the model results and the real data. The calibration process used by CALIB module of TELLUM is based on the maximization of the likelihood function and employs a gradient search method [1].

Using CALIB we calculate partial derivatives (or estimate parameters) for each one of the locator types. Each locator type (government employment, low-income household etc) in EMPAL and DRAM models will have different "locating behavior" in a particular region. At the same time a particular locator type may also exhibit different "locating behavior" in different regions. Because of this, it is necessary to estimate the equation coefficients of the model equations separately for each locator type in each region. The process of estimating these equation coefficients is called model calibration. For each locator type CALIB runs are performed. It may take one or several CALIB runs for each locator type's full calibration.

One should examine the results of the calibration process in order to evaluate if the estimated partial derivative values are reasonable and acceptable as inputs in DRAM and EMPAL models. In order to do this one or more indicators of goodness of fit are used. The appropriate goodness of fit measure for DRAM and EMPAL calibration is the likelihood function, derived from the notion of maximum likelihood as used in econometrics [2]. Still, there are other measures of goodness of fit that someone can use to evaluate the calibration results such as:

R-Square: R-square value is an indicator of how well the model fits the data. The smaller the variability of the residual values around the regression line relatively to the overall variability the better is our prediction. The value of R-square can range from 0 to 1, where 1 indicates that we have accounted for almost all of the variability with the variables specified for the model. Despite the fact that R-square is a very common and reliable measure of goodness of fit the fact that DRAM and EMPAL incorporate non-linear equations limits its reliability. For that Best-Worst Likelihood Ratio is considered to be as most appropriate measure of fit.

Best-Worst likelihood Ratio: The "best fit" of a model is when the difference between the model's estimate of the depended variable and the observed values in the calibration dataset is as small as possible. The "perfect fit" would be if for each locator type and zone the estimated number of employees or households were the same as the observed. The "worst fit" would be when all values, in each one of the zones, of the depended variable are estimated by the mean of that variable.

From these two extreme likelihood variables a measure of relative goodness of fit is derived called likelihood ratio and it is the ratio of the difference of the computed likelihood minus the worst fit likelihood, divided by the difference of the best minus the worst likelihood. The value of this ratio range from 1 to 0, were 1 is the best/perfect fit and 0 is the worst fit. The B/W likelihood ratio takes the following equation form:

$$\varphi = \frac{L - L_w}{L_b - L_w} \tag{8}$$

Mean Absolute Percent of Error (MAPE): Of the several statistics, which are often used to test the results of forecasting models the Mean Absolute Percent of Error is one of the most appropriate measure of goodness of fit. MAPE examines the distribution of the residuals (or errors) between the observed data and model’s current best-fit estimates. More specifically it is the average of the absolute values of percent of error between the observed and the estimated by the model values.

When using MAPE as a measure of goodness of fit we should be aware that it does not take into account the size of the zones (population and employment wise). This can create distortions especially when we have large percentages of errors in small zones. For example a 10% percent error in a small zone that has population of 100 people has a different gravity in the total observations than a 10% error in a larger zone which has a population of 2000 people.

In order to avoid such misinterpretations it is wise to examine MAPE indicator for the biggest 25% and the smallest 25% of the zones and explore if we are likely to get mistakes because of zone sizes. If this is the case then another indicator is used for the same purpose, which is presented bellow.

MARMO: MARMO is very similar to MAPE. It also express the average of the absolute values of percent of errors between the observed set of data and the data estimated by the model (DRAM, EMPAL), but it is weighted by the size of the observation (actual count of population or employment). A 20% to 30% percent usually represents a good MARMO.

Regional Location Elasticities: Location elasticities measure the sensitivity of household and employment location to changes in the attractiveness variables of the DRAM and EMPAL models. Location elasticities are defined for each one of the employment and residential zones. For instance for a 1% increase in an attractiveness variable in a specific zone, the location elasticity measures the resulting percentage of change in the number of households and employees in that zone.

Location elasticities are static measures of model sensitivity, which means that location elasticity for a specific attractiveness variable is calculated assuming that the values of all other attractiveness variables remain the same or fixed. Because of that, location elasticities will change as the values of the DRAM and EMPAL attractiveness values change. In more detail the value of the location elasticity for a specific attractiveness variable and zone is a function of: the value of the calibrated parameter for the attractiveness variable, the number of the households or employees in the zone, the magnitude of the attractiveness variable and the relative attractiveness of the other zones in the region.

Location elasticities will be larger when the calibrated parameter for the attractiveness variable is large (in absolute value), the number of households or employees is small (in comparison to the rest of the zones in the region), or the value of the attractiveness variable is small (in comparison to the other zones in the region).

III. TELUM Application for Chittenden County: County Profile

Location and General Facts

Chittenden County is located in northeastern part of United States in Vermontstate. As of the 2010 census, the population was 156,545 and home to nearly a quarter of Vermont's total population. Chittenden is the most populous county in the state, with more than twice as many residents as Vermont's second-most populous county, Rutland. Chittenden County is part of the Area. The most significant cities within the study area are Burlington, South Burlington and Williston with Burlington being the biggest city in the area.

Chittenden County Regional Planning Commission (CCRPC) is the metropolitan planning organization and transportation planning agency for the greater Burlington, Vermont region. The CCRPC covers the area of Chittenden County and encompasses about 145,000 people in its 18 municipalities. The county is home to about 25 percent of the state's population [4].

Each year, the CCRPC oversees about $30 million in transportation investments. It evaluates and approves proposed transportation improvement projects and provides a forum for interagency cooperation and public input into funding decisions. It also sponsors and conducts studies, assists local municipalities with planning activities, and develops and updates County's long range transportation plan known as the Metropolitan Transportation Plan [4].

Demographic and Economic Facts

1. Population and Demographics

As of 2010 Chittenden County is home of 156,545 residents (Table I), which represents 25 percent of Vermont's population. In 2000, county's population totaled 146,571 indicating an increase of 6.8 percent in a 10-year period and a 0.7 percent annual increase.

TABLE I
CHITTENDEN COUNTY POPULATION AND PERCENT OF CHANGE

Area	2010	2000	Change ('10-'00)	Percent of Change ('00-'10)
Vermont	625,741	608,827	16,914	2,8%
Chittenden County	156,545	146,571	9,974	6,8%

Source: US Census

Of the 156,545 total county population in 2010, 23.8 percent were 18 year of age or younger (Table II). The median age for 2010 for Chittenden County was 36.2 years compared to 41.5 years for Vermont and 36.8 for US population.

TABLE II
CHITTENDEN COUNTY POPULATION AGE STRUCTURE 2010

		Percent of total Population/age group				
Geographic Area	Total Population	0- 19	20-44	45-64	65+	Median Age
Vermont	625,741	24	30.7	30.8	14.5	41.5
Chittenden County	156,545	23.8	36.4	27.5	12.3	36.2

Source: 2010 Census

The 2010 count also indicates that 12.3 percent of the population is above 65 years old. It is important to note that the age distribution indicates a "healthy" population distribution among the different age groups. This information is especially significant since it gives us an idea in regard to the size of future employment based on current population at the age group 0-19 years old.

TABLE III
POPULATION RACIAL COMPOSITION 2010

Geographic Area	Total Population	%White	%Other
Vermont	625,741	95.3	4.7
Chittenden County	156,545	92.5	7.5

Source: 2010 Census

In terms of racial composition the area is predominately white and represents 92.5 percent of the total population. The "other" category (Table III) includes Black or African American, Asian, Native or Hawaiian and other races. More detail information about the racial composition of the study area is provided by the General Demographic Profile Characteristics tables, US Census Bureau [5].

2. Housing Occupancy

In terms of housing occupancy rates it seems that Chittenden County has a high occupancy rate that is also higher than Vermont's (Table IV). This information is extremely valuable for this study since it indicates the available housing stock, which will partly determine if the city will further expand or if it is more likely to use existing housing stock.

TABLE IV
HOUSING OCCUPANCY 2010

Geographic Area	Total Housing Units	% Occupied housing units	%Vacant
Vermont	294,382	81.7	18.3
Chittenden County	65,722	94.1	5.9

Source: U.S. Census 2010

3. Income

Table V shows the median household income calculated as a mean for the time period of 2008-2012, for the study area, Vermot State and U.S. The numbers reveal that the state of Vermont has similar household income to U.S. average, when Chittenden County has a relative high median household income [6].

TABLE V
MEDIAN HOUSEHOLD INCOME 2008 -2012

Geographic Area	Median Household Income
National	$53,046
Vermont	$54,168
Chittenden County	$63,900

Source: US Department of Commerce

4. Employment

Chittenden County is the nucleus of the economic activity of the state. In 1960's, 1970's and 1980's the area experienced great economic development and attracted international companies from Switzerland, Germany and Toronto. Today (2010) there are over 93,000 Vermonters employed in the Burligthon labor market area. Chittenden County's employment base is largely within five private industry sectors (Table VI): Healthcare and social assistance; retail trade; manufacturing; accommodation and food service; and professional, scientific and technical services. It should be noted that Chittenden County is the "house" of the largest for-profit employer in the state –the major IBM complex.

Employment in the private sector declined between 2000 and 2010 (Table VI). Total non-farm employment in Chittenden County decreased from 95,354 to 93,231 between 2000 and 2010 – a loss of 2, 123 jobs, or - 2.2 percent. This was offset in part by an increase in public sector employment, but it was not sufficient to offset private sector losses (private sector: - 4,386 + public sector: 2,263 = net -2,123) [6], [7].

TABLE VI
PRIVATE SECTOR JOB CHANGE BY INDUSTRY IN CHITTENDEN COUNTY 2000-2010

NAICS	Industry	2000	2010	Change	Percent
11	Agriculture, forestry, fishing and hunting	176	97	-79	-45%
21	Mining	36	39	3	8%
22	Utilities	291	254	-37	-13%
23	Construction	5.305	4.205	-1100	-21%
31-33	Manufacturing	16.759	10.744	-6015	-36%
42	Wholesale trade	3.048	3.127	79	3%
44-45	Retail trade	12.139	12.536	397	3%
49	Transportation and warehousing	2.488	2.072	-416	-17%
51	Information	2.651	2.129	-522	-20%
52	Finance and insurance	3.546	3.126	-420	-12%
53	Real estate and rental and leasing	1.211	1.109	-102	-8%
54	Professional and technical services	6.014	6.734	720	12%
55	Management of companies and enterprises	(c)	318	n/a	n/a
56	Administrative and waste services	(s)	3.210	n/a	n/a
61	Educational services	2.296	2.052	-244	-11%
62	Health care and social assistance	11.031	14.059	3028	27%
71	Arts, entertainment, and recreation	1490	1.476	-14	-1%
72	Accommodation and food services	6.851	7.679	828	12%
81	Other services, except public administration	3150	2.658	-492	-16%
	TOTAL	78482	74096	-4386	-5.6%

Source: Vermont Department of Labor
Note: (c) indicates confidential data that is not available and (s) data that was suppressed to protect confidential information

Total non-farm employment in Chittenden County decreased from 95,354 to 93,231 between 2000 and 2010 – a loss of 2,123 jobs, or - 2.2 percent. This was offset in part by an increase in public sector employment, but it was not sufficient to offset private sector losses (private sector: - 4,386 + public sector: 2,263 = net -2,123) [6], [7].

IV. MODEL CALIBRATION

A. Data Availability

Following is a description of data used in order to perform the calibration procedure for TELUM DRAM and TELUM EMPAL.

1. Zones

The study area of Chittenden County is composed by 325 Transportation Analysis Zones (TAZ) which is the most commonly used spatial level to generate land use forecasts. The most significant cities within the study area are Burlington, South Burlington and Williston (Fig. 1).

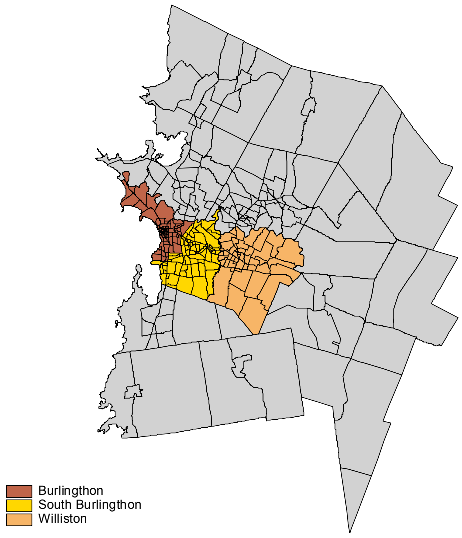

Fig 1 Chittenden county study area

2. Households

Available household data is limited to the total number of households per zone. Household spatial distribution and household density distribution (Fig. 2) indicates that Burlington, South Burlington and Williston are, as expected, the most densely populated parts of the county. For the five-year period of 2000-2005 there was an increase of households by 12 percent. The spatial distribution of the percent of households change for the five year period indicates sprawling trends in the metropolitan region.

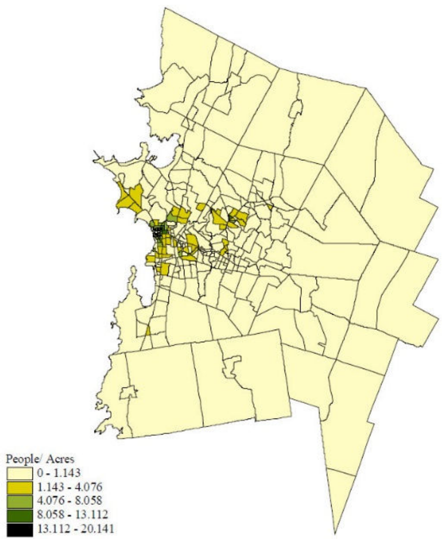

Fig. 2 Household density distribution 2000

It should be noted that the available household data is limited to the total number of households per zone. As a result necessary information to execute calibration runs is missing. In order to overcome this problem an assumption was made that the total number of households is equally divided between four types of household income groups. These are: High

Income Households, High Middle Income Households, Low Middle Income Households and Low Income Households.

3. Employment

Total employment and employment by type is available, for each zone and for both 2000 and 2005. Total employment number for 2000 is 95,354. This number is distributed among the following six employment types: Low, Medium Low, Medium High, High, School and Hotel. Each one of these categories was especially formulated for use in the relative travel demand model and represents certain SIC (Standard Industrial Classification) codes. Grouping the different SIC codes into the six employment types was made in accordance to the trips that each SIC code generates. Table VII shows the employment category by SIC code that each category entails (list is NOT exhaustive but indicative of what each code includes).

TABLE VII
TRIP GENERATION EMPLOYMENT CATEGORIES BY SIC

Employment type	Trip Generation Categories by SIC Code
Low	agriculture, mining, construction, manufacturing, utilities, transportation, warehousing, communication.
Medium Low	wholesale trade, banking and credit agencies, real estate, museums, government
Medium High	building materials, variety stores, market, specialty food stores, apparel stores, home stores, used merchandise, retail (stationary, jewelry, gift, novelty, luggage, sewing), hospitals, medical offices
High	Hardware stores, general merchandise, gas service stations, retail (music, electronics, hobbies, computrs, restaurants), Health services
School	Elementary school, college, university, daycare
Hotel	Hotel – motel

The diagram (Fig. 3) shows employment distribution for 2000, among the six sectors. Employment is mainly concentrated to the Medium High sector (28%) when an accountable part is distributed between High (20%) and Medium Low category (19%).

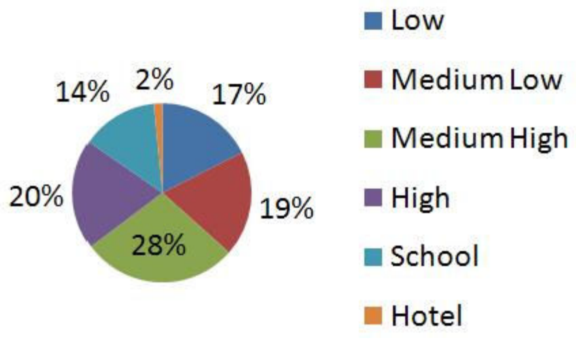

Fig. 3 Employment distribution by type 2000

A comparison between the total employment for 2000 and the projected employment for 2005 shows that there is a decrease of about 1% percent within the five year period. Spatial Distribution of total employment percent of change amongst the zones does not indicate specific patterns.

Distribution of employment by type changes in 2005. Low, Medium Low, and Medium High employment increase their share in contrast with High, School and Hotel which are losing employees. Medium High employment sector is still the one that occupies the highest amount of employees. Distribution of change for Medium High employment sector is shown below (Fig. 5).

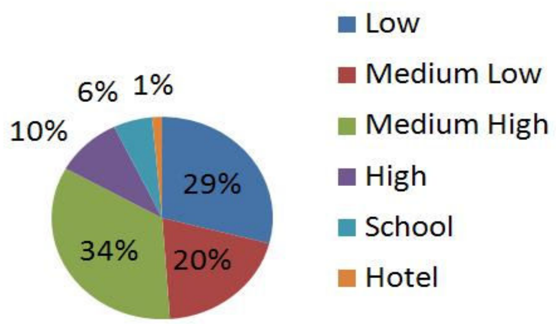

Fig. 4 Employment distribution by type 2005

It should be noted that a series of maps were created to study employment changes by type that each zone experienced but due to limited space a detail description it is not possible.

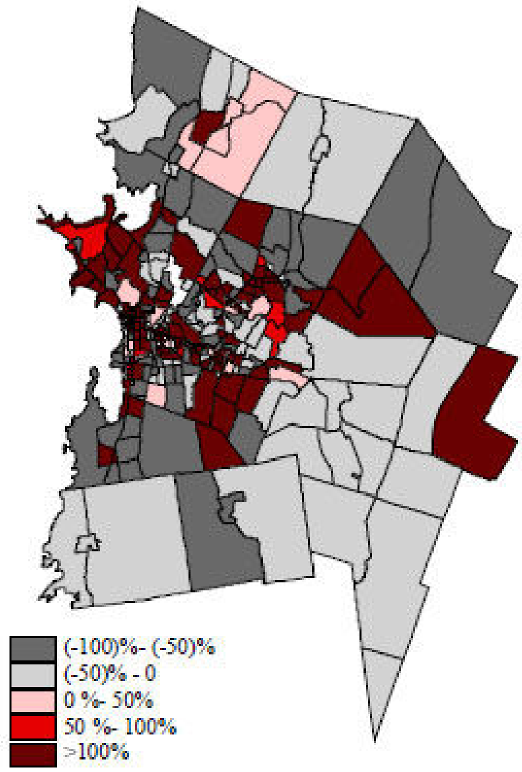

Fig. 5 Distribution of employment percent change for medium high employment

4. Land Use Data

Land Use data per TAZ was not available for the study area.

5. Zone to Zone Travel Time Cost

Travel time data consists of a travel time matrix that contains zone to zone travel time (in minutes). The data used here is the actual data used by the travel demand model that CCRPC utilizes.

Calibration Runs and Results

To perform calibration for TELUM-EMPALit is necessary to have employment data by employment type and by zone for two time points five years apart. In this case 2005 was considered to be the current time point and 2000 the lagged time point. Employment data is the only data required in two time points in the calibration process.

For TELUM-DRAM it is necessary to have employment data by employment type by zone, for the current time point. EMPAL current year should matches DRAM current year, which in this case is 2005. It is also necessary to have household data for one time point that will match current employment year. Household data was divided by type (4 to 8 categories of household types) and zones. Since we do not have households breakdown in different income categories but only total number in each zone, we divided the total number of households into four different categories by equally distributing households in each one of them. The four different household types used in the calibration runs were: High Income Households, High-Middle Income Households, Low-Middle Income Households, Low Income Households. The six different employment types used were: Low, Medium Low, Medium High, High, School and Hotel.

1. Employment Calibration Results

The measures of goodness of fit achieved from the calibration process were not satisfactory. Following is the table showing a summary of the calibration results. The four goodness-of-fit indicators are presented with their values.

TABLE VIII
GOODNESS OF FIT

	R-Square	B/W LR	MAPE	MARMO
Low	.0167	.1049	574.524%	99.897%
Med_Low	.0610	.1955	417.750%	92.087%
Med_High	.1010	.1707	476.062%	83.682%
High	.4100	.4291	153.495%	73.615%
School	.1570	.4376	96.904%	65.306%
Hotel	4264	.4684	56.717%	51.805%

R-Square and B/W LR are very low indicating that employment data poorly fits EMPAL model. At the same time MAPE values are very high indicating that the estimated parameters for DRAM an EMPAL models equation do not represent the best-fit model. This is because the percentage of error between observed and model's current best fit estimates for each one of the locator types are not within the acceptance range. Due to these results a more detailed examination of MAPE indicator was necessary. Following MAPE values are presented for the smallest and largest 25% for each one of the employment types.

Table IX shows that in most employment types MAPE has values within the acceptable range at least for the 25% smallest zones. Exception is the case of Medium High sector were MAPE for the 25% smallest zones is 3,670%. To avoid misinterpretations MARMO was used as a more secure indicator. MARMO values (Table X) are also very high indicating that the estimated parameters for DRAM and EMPAL models equation are not even close to the best fit model. The following table shows the estimated parameters for EMPAL equations.

TABLE IX
MEAN ABSOLUTE PERCENT OF ERROR

	Low	Med_Low	Med_High	High	School	Hotel-Motel
Min Observed Value	1	1	1	1	1	1
Max Observed Value	5625	3072	1354	1244	572	312
MAPE for smallest 25% of Zones (%)	.000	.000	3670	.000	.000	.000
% of smallest zones of region total	.31	.45	.28	.92	1.52	5.02
MAPE for largest 25% of Zones (%)	62.73	62.48	50.47	62.17	196.84	106.66
% of largest zones of region total	92.42	89.86	81.46	87.88	95.41	84.88

TABLE X
EMPAL CALIBRATION RESULTS

	α	β	EMP	LAND	λ	$1-\lambda$	R^2
Low	.0934	-.0541	.3040	.1778	1.0000	0	.0167
Med Low	-.9996	.0202	.3538	-.0799	.9373	0.062	.0610
Med High	-.2564	.0170	.1385	-.3214	.9740	0.02	.1010
High	.181839	-.210053	-.331482	-.251132	.2424	0.75	.4100
School	-.2712	.0164	.4925	.1333	1.0000	0	.1570
Hotel-Motel	-1.3166	-.0120	-.3855	-.2202	.3814	0.61	.4264

2. Household Calibration Results

Due to the peculiarity of the household data we run calibration for DRAM just for one income categories. The results are shown below.

TABLE XI
GOODNESS OF FIT

	Households
R-Square	.7497
B/W LR	.6990
MAPE	149.172%
MARMO	39.336%

R-square and B/W LR shows that household data fits almost perfectly DRAM equations. MAPE on the other hand has a relative high value that is why a further analysis was necessary. Analysis showed that MAPE value is inflated by the MAPE value of the 25 percent of the smallest zones, which is 614 percent as shown in Table XII.

TABLE XII
MEAN ABSOLUTE PERCENT OF ERROR

	Households
Min Observed Value	0
Max Observed Value	352
MAPE for Smallest 25% of Zones	614.583%
% of smallest zones of region total	.56%
MAPE for Largest 25% of Zones	30.792%
% of largest zones of region total	71.95%

3. Calibration Residuals

A more detail examination of residuals distribution will give us a better idea about the quality of the data and the validity of the results.

Also an examination of their locational distribution might show interesting spatial patterns. For example Fig. 6 shows the distribution of residuals for low employment and medium low employment sector. Red color indicates that the relative zones have been overestimated by the indicating percentage.

TABLE XIII
DRAM ESTIMATED PARAMETER VALUES

	Households
α	.9417
β	-1.2649
VACDEV	.6031
PERDEV	1.0075
RESLND	.3680
LIHH*	.3680
LAGHH	.2510

*the values are the same for all HH types

It is interesting to note that for all employment types and for households, all zones are overestimated and none of them is underestimated which indicates that in the calibration procedure there is a systematic pattern.

As explained earlier residuals indicate how accurate is the calibration and consequently how accurate is the forecasted population or employment values for each zone. The systematic pattern of overestimation that was detected both in employment and household values creates a lot of questions about the validity of our results. Also the fact that employment calibration results were really disappointing creates suspicions about the data used. For that a further examination of the household and employment distribution was necessary.

4. Validity of Household and Employment Projections

Following is a graph showing the relationship of Low Employment 2000 and Low employment 2005 (Fig. 7, Table XIV). In the five-year period 2000 to 2005 low employment sector experiences an increase of almost 55 percent. In many cases TAZs with zero low-employment in 2000 had very high numbers in 2005. This peculiarity creates the abnormal distribution that we see in the graph.

TABLE XIV
SUMMARY OF FIT FOR LOW EMP 2000 BY LOW EMP 2005
LINEAR FIT
LOWEMP00 = 42.010277 + 0.0142605 Z8_LOWEMP05
SUMMARY OF FIT

R Square	0.002487
RSquareAdj	-0.0006
Root Mean Square Error	104.8422
Mean of Response	43.14769
Observations (or Sum Wgts)	325

PARAMETER ESTIMATES

Term	Estimate	t Ratio	Prob>|t|
Intercept	42.010277	7.06	<.0001
LOWEMP05	0.0142605	0.90	0.3702

Similar to the low employment situation, is the situation with all six-employment sectors. A first evaluation might be that the projected 2005 employment numbers are not accurate. The accuracy of the number depends on the initial numbers (year 2000) and the projection method used.

In this case it is most possible that the number for each employment type was created just by breaking down the total employment number for 2005, which also explains the much better calibration results for total employment and households. For all these reasons we come to the conclusion that the data provided to us, was inadequate and cannot be used to get accurate projection results.

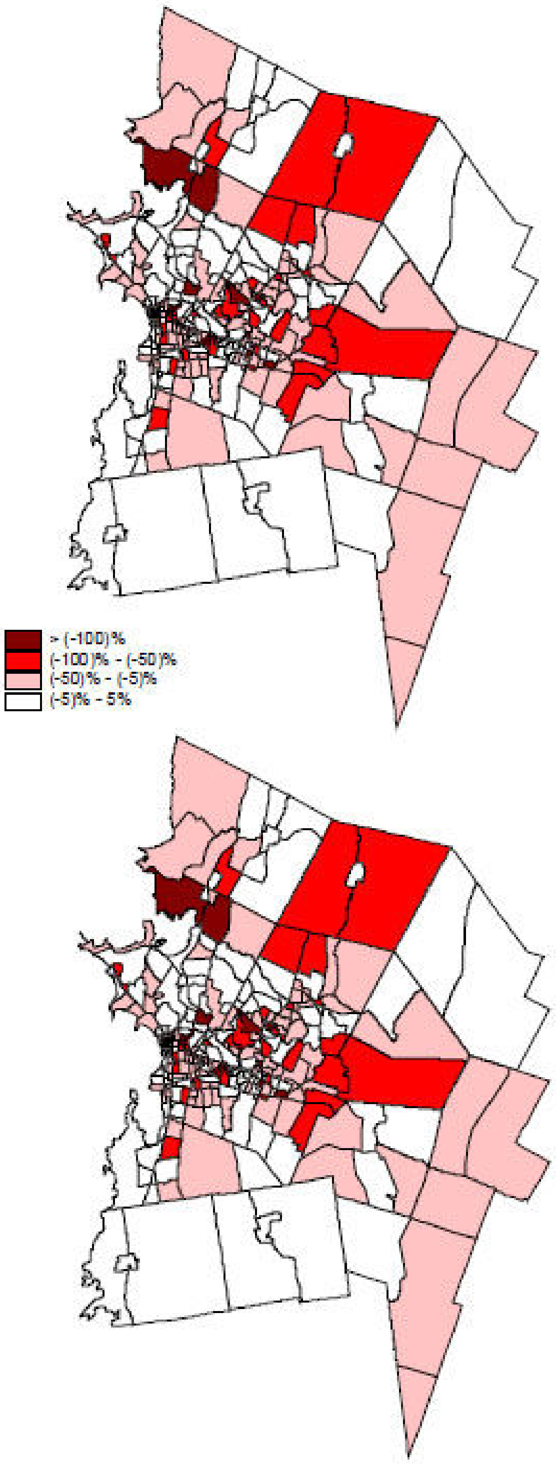

Fig. 6 Spatial distribution of residuals for low employment and medium low employment sector

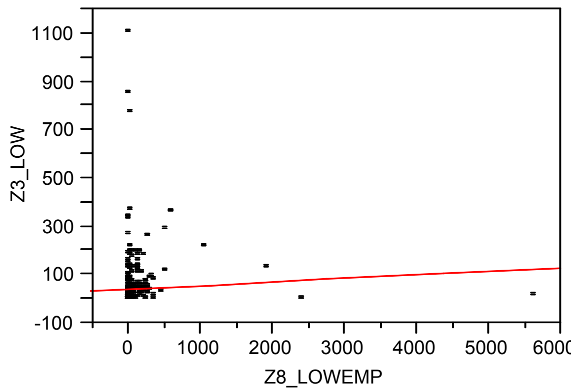

Fig. 7 Bivariate Fit of Low Emp 2000 by Low Emp 2005

V. Conclusions

The purpose of this research study was to examine how relative simple land use simulation models can be used by a middle-size MPO to address its planning responsibilities. In order to do that Chittenden County MPO South Burlington, VT was selected and TELUS Land Use Model was used.

The first and probably the most crucial part of forecasting future population and employment distribution is to fit the land use model into the real world by estimating the parameters for model's equation. These parameters are the ones that best fit the model and minimize the discrepancies between the model results and the real data. In this procedure (calibration process) the quality of input data (household and employment) is extremely significant for producing statistically valid land use inputs to travel models.

It is understandable that there is no universal data i/o interface and that the data used here was customized to interact with TELUS land use model. Calibration results showed that the quality of the data was insufficient to produce statistically reliable and replicable forecasts for agency use. As a result land use model outputs cannot be used as inputs in the transportation model.

Actually inadequate data and data availability in general is one of the reasons that land use models are underutilized in planning practice [8]. This study confirms that although we had a relatively simply land use modeling tool that was designed specifically for use by MPOs, its potential use the CCRPC agency would have been problematic. Therefore wide application of land use models as policy analysis tool depends as much as on data availability as does on the improvement of model formulation and predictability issues that scientific community is mainly focused on.

References

[1] Putman, Stephen H. "DRAM Residential Location and Land Use Model: 40 Years of Development and Application." *Residential Location Choice*. Springer Berlin Heidelberg, 2010, pp. 61-76.
[2] Putman S., *Integrated Urban Models 2*. Pion Limited, London, 1992.
[3] Putman S., *Integrated Urban Models: Policy Analysis of Transportation and Land Use*. Pion Limited, London, 1983.
[4] Chittenden County Regional Planning Commission, www.ccrpcvt.org [Accessed 15/01/2014].
[5] United States Census Bureau, http://factfinder2.census.gov [Accessed 15/01/2014].

[6] United States Department of Commerce - Bureau of Economic Analysis; Market Statistics, 2012.

[7] Department of Labor, State of Vermont, www.labor.vermont.gov [Accessed 13/01/2014].

[8] Pozoukidou G. (2010), Predicting the Outcome of Plans. Retrospect and Prospects of the Use of Models in Planning, in Brebbia C.A., HernandezS., Tiezzi E.(eds), Urban Regeneration and Sustainability, WIT Press, Boston, pp. 83-93.