

DOCUMENTATION FOR USE OF THE GRAVITY-BASED LAND USE MODEL (G-LUM): WITH AND WITHOUT IN-BUILT TRAVEL DEMAND MODEL ASSUMPTIONS

Prepared by Siva Kakaraparthi, Sumala Tirumalachetty, Dr. Kara Kockelman, Jen Duthie, Binny Paul, Varun Valsaraj and Brenda Zhou

Updated: September 2009

1. INTRODUCTION

Long-term traffic forecasts require long-term land use forecasts, which may be affected by mid and near-term transportation infrastructure decisions, as well as traveler choices and associated changes in travel times and the relative access of competing locations. A long-standing methodology (Steve Putman's ITLUP ®) for land use forecasting relies on a system of gravity-based equations, with proprietary code managed by consultants. To this end, the Gravity-based Land Use Model (G-LUM) has been coded and made accessible via freeware to facilitate greater use of land use models, under local control, with more transparency in the forecasting process.

G-LUM consists of three sub-models: EMPLOC, RESLOC, and LUDENSITY. The equations involved in each of three sub-models are presented and discussed in Appendix A of this report. Calibration provides estimates of parameter values that generate the "best fit" of forecasts for the base year based on lag year data. G-LUM's parameter estimation process is discussed in detail in the next section. Forecasts are generated in specified time increments. A time step begins with the execution of the EMPLOC model, followed by RESLOC, and then LUDENSITY. Figure 1's flowchart indicates the steps involved in G-LUM's estimation and application. The user has an option to run the program separately, or along with a travel demand model. In the later case, travel times in G-LUM are updated using the results of the travel demand model after each time step.

The GUI developed for this code enables users to run the MATLAB-based G-LUM code without a MATLAB license (which costs about \$2,000). However, it requires installing the freely available compiler software, which is 100MB in size. Full version of MATLAB can be purchased at <http://www.mathworks.com/store/productIndexLink.do>.

MATLAB's compiler (version 4.3) was used to develop this GUI and can be obtained from a full version of MATLAB 7.0¹. The compiler software can also be obtained via Dr. Kara Kockelman (kkockelm@mail.utexas.edu).

¹ MATLAB's latest version was not used, due to the absence of a full license but also to reduce the size of the compiler program required for loading the G-LUM GUI.

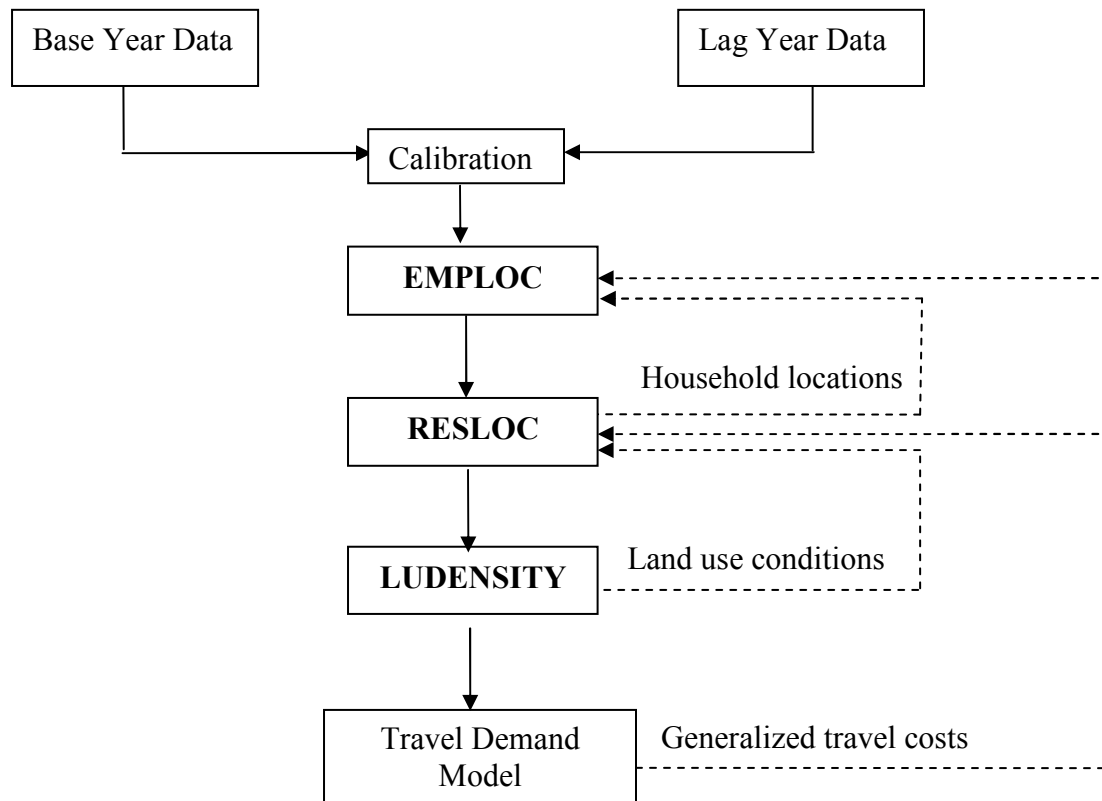


Figure 1. Flowchart for G-LUM Modules, in Concert with (external) Travel Demand Model

Note: Dashed lines represent one-period ($t-1$) lagged feedback of information.

For more information on the G-LUM model, including applications to Texas’s Austin and Waco regions, please see the “ForecastingPopulationAndEmploymentInTexas.pdf” report located at http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/homepage.htm.

2. DATA PREPARATION

Fourteen input files are required to estimate parameters and predict future year household and job distributions throughout a system of zones using this model. The required data for Texas regions of Austin (base year = 2005; lag year = 2000; # zones = 1074; # districts = 109) and Waco (base year = 2005; lag year = 2000; # zones = 283; # districts = 44) are provided in the Code + Data section of the http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/homepage.htm website at both the district (large-zone) and TAZ (neighborhood-zone) levels with data listed by ascending FID number. Below is a description of each required file. It is imperative that the zone-based data be listed in the same order in each file (e.g., if employment data is listed in order of ascending zone number, then all other data must be listed in ascending zone order).

- 1) **EMP_lag.txt** contains employment by sector in the lag year for each zone. The number of industry sectors (or the number of columns in the EMP_lag.txt file) is determined by users and can vary across regions due to data availability (between 3 and 6). A common employment classification will emphasize three categories, as follows: (1) basic jobs (often including agriculture, forestry, fishing, mining, construction, manufacturing, services of transportation, communications, electric, gas, and sanitary, and wholesale trade), (2) retail jobs (i.e., retail trade), and (3) service jobs (including finance, insurance, real estate, services, and public administration). See the Austin and Waco applications on line at:
http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/InvestigationOfTELUM.pdf. The first column should correspond to basic employment, the second to retail and the third to service. Sector data should be separated by tabs or spaces, and the data for each zone (e.g., TAZ) should be on a new line and listed in the same order for all sectors.
- 2) **EMP_base.txt** contains employment by sector (equal to the number of sectors in the EMP_lag.txt file and in the same order) in the base year for each zone. As with the EMP_lag.txt file, sector data should be separated by tabs or spaces, and the data for each zone (e.g., each TAZ) should be on a new line. Also needs to have the same number and type of job categories as EMP_lag.txt.
- 3) **HH_lag.txt** contains the number of households by type in the lag year for each zone. The number of household types (or the number of columns in the HH_lag.txt file) is determined by users and can vary across regions (between 1 and 6). Household income is commonly used to classify households (e.g., low-income, medium-income, and high-income households). Household data should be separated by tabs or spaces, and the data for each zone (e.g., TAZ) should be on a new line and listed in the same order as in all the other household types.
- 4) **HH_base.txt** contains the number of households by type in the base year for each zone. As with the HH_lag.txt file, household data for each household type should be separated by tabs or spaces, and the data for each zone (e.g., TAZ) should be on a new line and listed in the same order as in the HH_lag.txt file. Also, the number of zones and types of household should match those in the HH_lag.txt file.
- 5) **ZoneSize.txt** contains the total size (area) of each zone. The data for each zone should be separated by tabs or spaces, and values for distinct zones should be on distinct lines. All land area across all files need to have consistent units (e.g., square miles, acres, etc).
- 6) **landb.txt** contains the amount of land dedicated to basic employment in the base year in each zone. The data for each zone should be separated by tabs, spaces, and data for distinct zones should be on distinct lines. As described above, basic employment generally includes jobs in the agricultural, forestry, fishing, mining, construction, manufacturing, and other sectors.

- 7) **landc.txt** contains the amount of land dedicated to non-basic (or “commercial”) employment in the base year in each zone. The data for each zone should be separated by tabs or spaces, and data for distinct zones should be on distinct lines.
- 8) **landd.txt** contains the amount of undeveloped – but developable – land in the base year in each zone. (Area for streets and other undevelopable lands should not be included here.) The data for each zone should be separated by tabs or spaces, and data for distinct zones should be on distinct lines.
- 9) **landu.txt** contains the amount of unusable land (i.e., land preserved from any development, such as lakes or state parks) in the base year in each zone. The data for each zone should be separated by tabs or spaces, and data for distinct zones should be on distinct lines.
- 10) **lands.txt** contains the amount of land dedicated to streets and highways in the base year in each zone. The data for each zone should be separated by tabs or spaces, and data for distinct zones should be on distinct lines.
- 11) **landr.txt** contains the amount of land dedicated to residential use in the base year in each zone. The data for each zone should be separated by tabs or spaces, and data for distinct zones should be on distinct lines. As a final note, all the files on areas of land by type need to use consistent units (e.g., square miles, acres, etc.).
- 12) **TT.txt** contains the travel time (or travel cost) between each pair of zones in the base year. The data should be input as a matrix with one row for each origin zone and one column for each destination zone, with tabs or spaces between data entries in each row.
- 13) **EmpProj.txt** contains the control totals for employment over the prediction time period. The data should be input as a matrix with one row for each time period and one column for each employment type, with tabs or spaces between data entries in each row. These should be consistent with the difference in base year and lag year used while estimating the parameters.
- 14) **HHProj.txt** contains the control totals for employment over the prediction time period. The data should be input as a matrix with one row for each time period and one column for each employment type, with tabs or spaces between data entries in each row. These should be consistent with the difference in base year and lag year used while estimating the parameters.

Note: EmpProj.txt and HHProj.txt are not required for calibration of parameters but are necessary for future year projections.

3. INSTALLING THE GUI

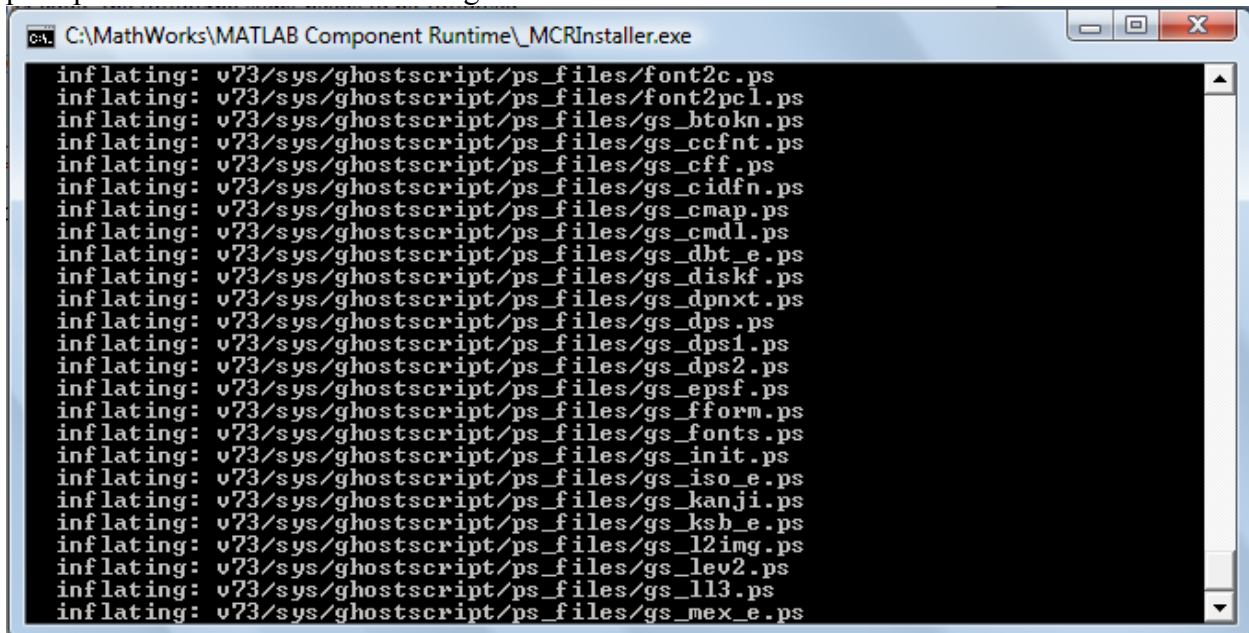
3.1 Installing the GUI

To run the G-LUM code, the following steps need to be followed:

Step 1: Install MCRInstaller.exe present in the unzipped folder.

Step 2: In *Windows XP*, ignore this step 2 and proceed to step 3.

In *Windows Vista*: Navigate to the folder C:\Programfiles\ and copy the Math Works folder into C:\ (i.e., onto the C drive of your hard disk). Now, run the file _MCRInstaller.exe present in the folder C:\Math Works\ Matlab Component Runtime. You will see a command prompt window which looks like Figure 2.



```
C:\MathWorks\MATLAB Component Runtime\_MCRInstaller.exe
inflating: v73/sys/ghostscript/ps_files/font2c.ps
inflating: v73/sys/ghostscript/ps_files/font2pcl.ps
inflating: v73/sys/ghostscript/ps_files/gs_btokn.ps
inflating: v73/sys/ghostscript/ps_files/gs_ccfnt.ps
inflating: v73/sys/ghostscript/ps_files/gs_cff.ps
inflating: v73/sys/ghostscript/ps_files/gs_cidfn.ps
inflating: v73/sys/ghostscript/ps_files/gs_cmap.ps
inflating: v73/sys/ghostscript/ps_files/gs_cmd1.ps
inflating: v73/sys/ghostscript/ps_files/gs_dbt_e.ps
inflating: v73/sys/ghostscript/ps_files/gs_diskf.ps
inflating: v73/sys/ghostscript/ps_files/gs_dpnext.ps
inflating: v73/sys/ghostscript/ps_files/gs_dps.ps
inflating: v73/sys/ghostscript/ps_files/gs_dps1.ps
inflating: v73/sys/ghostscript/ps_files/gs_dps2.ps
inflating: v73/sys/ghostscript/ps_files/gs_epsf.ps
inflating: v73/sys/ghostscript/ps_files/gs_fform.ps
inflating: v73/sys/ghostscript/ps_files/gs_fonts.ps
inflating: v73/sys/ghostscript/ps_files/gs_init.ps
inflating: v73/sys/ghostscript/ps_files/gs_iso_e.ps
inflating: v73/sys/ghostscript/ps_files/gs_kanji.ps
inflating: v73/sys/ghostscript/ps_files/gs_ksh_e.ps
inflating: v73/sys/ghostscript/ps_files/gs_l2img.ps
inflating: v73/sys/ghostscript/ps_files/gs_lev2.ps
inflating: v73/sys/ghostscript/ps_files/gs_l13.ps
inflating: v73/sys/ghostscript/ps_files/gs_mex_e.ps
```

Figure 2: Installing the GUI

Step 3:

- Click on the Start button in the task bar
- Right click on 'My Computer' and click on properties
- In *Windows XP*, click on the advanced settings tab.

In *Windows Vista*, click on 'Advanced System Settings' present on the left side of the window.

- Click on the Environmental Variables button.
- In system variables, go to the 'path' present in the list of system variables.
- Edit the path by appending the “;C:\Program files\MathWorks\MATLAB Component Runtime\v73\bin\win32”.

Note: Analysts must use a semicolon (;) before adding the path to the C:\...\win32 subdirectory, and be careful in handling this step as it can alter the system properties.

Step 4: Copy the GUI_new.exe and GUI_new.ctf present in the 'GUI code for G-LUM' folder of the website to the folder C:\MathWorks\MATLAB Component Runtime\v73\runtime\win32.

Step 5: Double Click on the file GUI_new.exe. You will see the command prompt window, as shown below, in Figure 3.

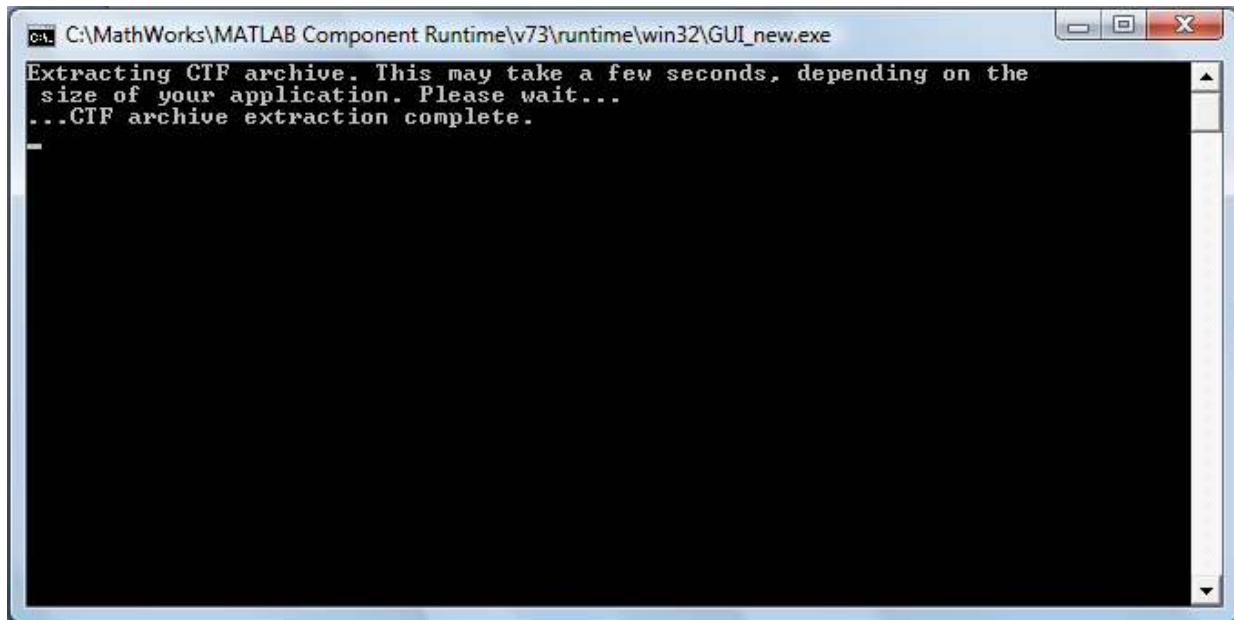


Figure 3: Running the GUI

Wait for a minute to get the GUI window.

Step 6: Unzip the G-LUM Finale folder present in the 'GUI code for G-LUM' folder of the website into the C- drive. This is sample Austin data. If you are planning to use your own data, then mimic the dataset for Austin. Data inputs needs to be provided in the input folders of calibration and prediction folders. The first twelve (12) input files listed in the data preparation section will be present in the *calibration input folder*, while the 13th and 14th input file will be present in the *prediction input folder*.

3.2 Running G-LUM

G-LUM can be run either using the GUI (as specified in 3.2.1) or using MATLAB (as specified in 3.2.2). A GUI was developed at The University of Texas at Austin to allow users to run G-LUM without owning a MATLAB license.

3.2.1 Running the GUI

Step 7: The GUI has three push buttons for parameter estimation and one push buttons for model predictions.

- 1) **RESLOC Calibration** is used to estimate the household allocation models' parameters. The model parameters can be estimated by just clicking the push button labeled 'RESLOC Calibration' and the command prompt window shows a series of ZZ values popping up. The ZZ values report a goodness of fit for the parameter estimates, the lower the ZZ value the better the estimate. Note the final ZZ value of the run. Change the initial values in the *initial_values.txt* and aim for a lower ZZ value. Reasonable ranges are as follows: $\eta = [0,1]$; $\alpha = [0,3]$; $\beta = [-.01,-.1]$; $qq = [.01;.1]$; $rr = [0,1]$; $ss = [0,1]$; $bnn[i] = [-5,5]$. (Note: these are simply suggested ranges and may change depending on

the units used or the problem). Try different values of the parameters until it does not seem possible to achieve a lower ZZ value.

The columns of initial_values.txt file correspond to eta, alpha, beta, qq, rr, ss, bnn[1], bnn[2], bnn[3], bnn[4] in the first four rows for RESLOC.

- 2) **EMPLOC Calibration** is used to estimate parameters in the jobs allocation model. The procedure to run this is the same as that used for RESLOC Calibration. Reasonable ranges of initial values are lambda = [0,1]; omega = [1,5]; rho = [-0.001,-0.010]; aa = [0.01,1]; bb= [0.01,1]. Try different initial values until it does not seem possible to achieve a lower ZZ value.

The columns of initial_values.txt file correspond to lambda, omega, rho, aa, and bb in the rows 5 to 10 for EMPLOC.

- 3) **LU DENSITY Calibration** is used to estimate the parameters of the model's land use component. The procedure to run this is the same as shown for RESLOC. Reasonable ranges of initial values for kk, gg, and pp are [0, 1]. Try different initial values until it does not seem possible to achieve a lower ZZ value. The last row in the initial_values.txt file is for LU DENSITY and corresponds to kk, gg and pp initial values.

After calibrating the three components, copy the contents of the output folder (which is present in the calibration folder) to the input folder (which is present in the prediction folder). In other words, all files shown under C:\G-LUM\Finale\Calibration\Output should be saved in C:\G-LUM\Finale\Prediction\Input.

- 4) **PREDICTION:** To predict, the user should enter the total number of future time periods to be predicted from the base year. For example, if the base year is 2005 and one seeks to forecast land use patterns through 2015, then one should enter the number of time periods that are required to reach 2015 in the "Edit" box. The size of each time period depends upon the control totals specified (For example, if control totals are provided at intervals of 5 years, then the size of each time period is five years and two time periods are needed to reach 2015 from 2005). The user should then click on the push button labeled 'Prediction with TDM'. You can see a statement showing the success of your prediction in the command prompt window (i.e., the words "Finished Estimation for "XX" time periods" signal successful completion).

G-LUM enables users to run their own TDM externally, based on land use forecasts in each time step (or every several time steps, for example) and update travel times/travel costs as needed. This module also runs in accordance to the number of time periods entered by the user. Results are outputted to the 'output' subfolder in the prediction folder

Note: The edit box will accept numbered values greater than zero and no letters. The exogenous control totals entered for different time periods should be consistent with the difference in base year and lag year used while estimating the parameters (as discussed above).

3.2.2 Running the MATLAB Code (G-LUM)

- 1) Open the files RESLOC.m, EMPLOC.m, LUDENSITY.m, NLLS_RESLOC_SIM_1.m, NLLS_RESLOC_SIM_2.m, NLLS_RESLOC_SIM_3.m, NLLS_RESLOC_SIM_4.m, NLLS_EMPLOC_SIM.m, NLLS_LUDENSITY_G_SIM.m, NLLS_LUDENSITY_K_SIM.m, NLLS_LUDENSITY_P_SIM present in the calibration folder and prediction.m, and fourstep.m in the prediction folders.
- 2) Change the input and output paths to appropriate folders and then run RESLOC.m, EMPLOC.m, and LUDENSITY.m.
- 3) Copy the files in the output folder present in the calibration folder to the input folder in the prediction folder.
- 4) Comment out the lines 590 and 591 in the prediction.m file by adding a “%” in front, as follows:

```
%c1=fourstep(Lrproject,Lbproject,Lcproject,HHproject,EMPproject,c,X,Y,Z,count);  
%c=c1;
```

Run the file prediction.m. Results are outputted to the ‘output’ subfolder in the prediction folder

4. VIEWING RESULTS

Output data can be viewed in ArcGIS or any GIS mapping software by mapping the output data with a spatial layer. To map within MATLAB, visit <http://www.spatial-econometrics.com/> and navigate to download the MATLAB Arc Map Toolbox file (courtesy of James P. Le Sage of Texas State University – San Marcos). Also, several open-source GIS options are available for download at <http://opensourcegis.org/>.

REFERENCES

Caliper. (2000) Travel Demand Modeling with TransCAD4.7. *TransCAD Manual*, p. 172

Kockelman, K. (2008) Gravity-based Land Use Model Website. University of Texas at Austin. Available at http://www.ce.utexas.edu/prof/kockelman/G-LUM_Website/homepage.htm.

Lemp, J.D. (2007) Travel Demand Forecasting Models: Development, Application, and Comparison of Aggregate and Activity Based Approaches for the Austin, Texas Region. Master’s Thesis, The University of Texas at Austin.

Sheffi, Y. (1985) *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Martin W.A., and N.A., McGuckin. NCHRP Report 365 (1998) *Travel Estimation Techniques for Urban Planning*. National Cooperative Highway Research Program, Transportation Research Board. Washington, D.C.

APPENDIX

Model Description

ITG-LUM consists of a land use model (G-LUM) with three sub-components, RESLOC, EMPLOC, and LUDENSITY and a four-step travel demand model. A description of the sub-models of G-LUM can be found in Kockelman (2008). The only change to G-LUM is that LUDENSITY has been replaced by simple rules for how much land each type of use requires. The rules selected for the test application will be presented along with the section on ‘Sample Network and Input Data.’

RESLOC

The equations obtained from calibrating RESLOC, the residential land use allocation module, are shown below for each household type n . Here, i and j are used as zone indices, k is an index for employment type, n is an index for household type, and t is an index for time period. The number of households of type n residing in zone i at time t $N_{i,t}^n$ is given by

$$N_{i,t}^n = \eta^n \sum_j Q_{j,t}^n \frac{W_{i,t-1}^n c_{i,j,t-1}^{\alpha^n} \exp(\beta^n c_{i,j,t-1})}{\sum_i W_{i,t-1}^n c_{i,j,t-1}^{\alpha^n} \exp(\beta^n c_{i,j,t-1})} + (1 - \eta^n) N_{i,t-1}^n \quad (1)$$

where

$$Q_{j,t}^n = \sum_k a_{k,n} \frac{E_{j,t}^k}{1 - u_k} \quad (2)$$

$$W_{i,t-1}^n = (L_{i,t-1}^v)^{q^n} (1 + x_{i,t-1})^{r^n} (L_{i,t-1}^r)^{s^n} \prod_{n'} \left[\left(1 + \frac{N_{i,t-1}^{n'}}{\sum_n N_{i,t-1}^n} \right)^{b_n^n} \right] \quad (3)$$

and $c_{i,j,t-1}$ is impedance (travel time and/or cost) between zones i and j at time $t-1$; $a_{k,n}$ is the number of type n households per type k employee in the study region; $E_{j,t}^k$ is employment (number of jobs) of type k in zone j at time t ; u_k is the unemployment rate for job type k ; $L_{i,t-1}^v$ is vacant developable land in zone i at time $t-1$; $x_{i,t-1}$ is the proportion of developable land already developed in zone i at time $t-1$; $L_{i,t-1}^r$ is residential land in zone i at time $t-1$; and $\eta^n, \alpha^n, \beta^n, q^n, r^n, s^n$ and b_n^n are parameters estimated in model calibration. In addition, $Q_{j,t}^n$ converts employment to households, and $W_{i,t-1}^n$ represents the attractiveness of zone i for household type n at time $t-1$.

EMPLOC

The equations obtained from calibrating EMPLOC, the employment land use allocation module, are shown below for employment type k (basic, retail, service).

$$E_{j,t}^k = \lambda^k \sum_i N_{T,i,t-1} \frac{M_{j,t-1}^k c_{i,j,t-1}^{\omega^k} \exp(\rho^k c_{i,j,t-1})}{\sum_j M_{j,t-1}^k c_{i,j,t-1}^{\omega^k} \exp(\rho^k c_{i,j,t-1})} + (1 - \lambda^k) E_{j,t-1}^k \quad (4)$$

where

$$M_{j,t-1}^k = (E_{j,t-1}^k)^{a^k} (L_j)^{b^k} \quad (5)$$

and $N_{T,i,t-1}$ is total households in zone i at time $t-1$; L_j is the total area of zone j ;

λ^k , ω^k , ρ^k , a^k and b^k are estimated in model calibration; and $M_{j,t-1}^k$ represents the attractiveness of zone j for employment type k at time $t-1$.

Of the three employment types, service employment is least dependent on existing service employment, and retail employment is most dependent on existing retail employment.

LUDENSITY

The equations obtained from calibrating LUDENSITY, the land allocation module, are shown below for each land type. Other notation needed is L_b – amount of land devoted to basic employment, L_c – amount of land devoted to commercial (retail + service) employment, L_d – amount of developable land, L_d – amount of developed land.

Residential land use:

$$\frac{L_{r,i,t}}{N_{T,i,t}} = k_0 L_{D,i,t-1}^{k_1} \left(\frac{L_{d,i,t-1}}{L_{D,i,t-1}} \right)^{k_2} \left(\frac{L_{b,i,t-1}}{L_{D,i,t-1}} \right)^{k_3} \left(\frac{L_{c,i,t-1}}{L_{D,i,t-1}} \right)^{k_4} \times \left(\frac{N_{1,i,t}}{N_{T,i,t}} \right)^{k_5} \left(\frac{N_{2,i,t}}{N_{T,i,t}} \right)^{k_6} \left(\frac{N_{3,i,t}}{N_{T,i,t}} \right)^{k_7} \left(\frac{N_{4,i,t}}{N_{T,i,t}} \right)^{k_8} \left(\frac{N_{5,i,t}}{N_{T,i,t}} \right)^{k_9} \left(\frac{N_{6,i,t}}{N_{T,i,t}} \right)^{k_{10}} \quad (6)$$

Basic Land Use:

$$\frac{L_{b,i,t}}{E_{b,i,t}} = g_0 L_{D,i,t-1}^{g_1} \left(\frac{L_{d,i,t-1}}{L_{D,i,t-1}} \right)^{g_2} \left(\frac{E_{b,i,t}}{E_{T,i,t}} \right)^{g_3} \left(\frac{L_{b,i,t-1}}{L_{D,i,t-1}} \right)^{g_4} \left(\frac{L_{r,i,t-1}}{L_{D,i,t-1}} \right)^{g_5} \quad (7)$$

Commercial Land Use:

$$\frac{L_{c,i,t}}{E_{c,i,t}} = p_0 L_{D,i,t}^{p_1} \left(\frac{L_{d,i,t}}{L_{D,i,t}} \right)^{p_2} \left(\frac{E_{c,i,t}}{E_{T,i,t}} \right)^{p_3} \left(\frac{L_{c,i,t}}{L_{D,i,t}} \right)^{p_4} \left(\frac{L_{r,i,t}}{L_{D,i,t}} \right)^{p_5} \quad (8)$$

E stands for employment by type, as defined above (b = basic, c = commercial, including both retail and service jobs); N represents households by type; and the k 's, g 's and p 's are estimated parameters.

To summarize: This documentation explains the data requirements and procedure for installing and running the ITG-LUM and G-LUM using both GUI and MATLAB. The document's appendix provides further details of the land use model and the travel demand model. The purpose of this software is to provide a friendly user interface while allowing various users to run the model (which was coded in MATLAB) without specialized software installed.

Common Errors in Using GUI

While installing the GUI, care should be taken in choosing correct paths. For example, C:\Program files\MathWorks\MATLAB Component Runtime\v73\bin\win32 is different from C:\Program files\MathWorks\MATLAB Component Runtime\v73\runtime\win32.