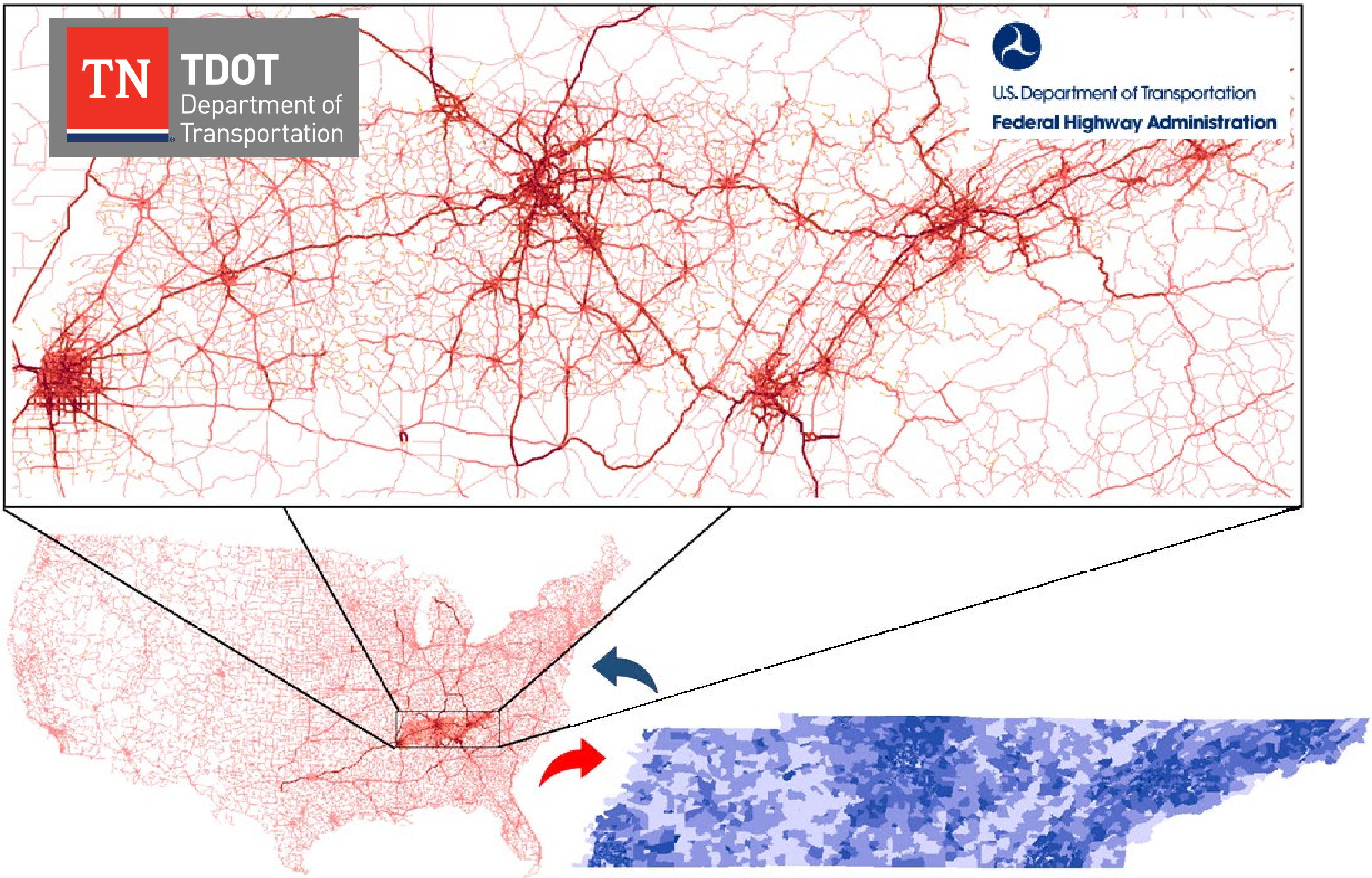

Developing Statewide Land Use Forecasting Model and Integrate with TDOT's Statewide Travel Demand Model

Research Final Report from University of Memphis | Sabyasachee (Sabya) Mishra, Mihalis M. Golias, Jerry Everett, and Ali Riahi Samani | December 31, 2021

Sponsored by Tennessee Department of Transportation Long Range Planning

Research Office & Federal Highway Administration

DISCLAIMER

This research was funded through the State Planning and Research (SPR) Program by the Tennessee Department of Transportation and the Federal Highway Administration under ***RES#: 2018-20 Research Project Title: Developing Statewide Land Use Forecasting Model and Integrate with TDOT's Statewide Travel Demand Model.***

This document is disseminated under the sponsorship of the Tennessee Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Tennessee and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the views of the author(s) who are solely responsible for the facts and accuracy of the material presented. The contents do not necessarily reflect the official views of the Tennessee Department of Transportation or the United States Department of Transportation.

Technical Report Documentation Page

1. Report No. RES2020-18	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle *Developing Statewide Land Use Forecasting Model and Integrate with TDOT's Statewide Travel Demand Model*		5. Report Date October 2021 6. Performing Organization Code
7. Author(s) Sabyasachee Mishra; Mihalis M. Golias; Jerry Everett; and Ali Riahi Samani		8. Performing Organization Report No.
9. Performing Organization Name and Address The University of Memphis Department of Civil Engineering Memphis, TN, 38152		10. Work Unit No. (TRAIS) 11. Contract or Grant No. RES2020-18
12. Sponsoring Agency Name and Address Tennessee Department of Transportation 505 Deaderick Street, Suite 900 Nashville, TN 37243		13. Type of Report and Period Covered Final Report August 2019 – December 2021 14. Sponsoring Agency Code
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.		

16. Abstract

In this research, a statewide forecasting land-use model is presented. The land-use model is called the Large-Scale Land Use Model (LS-LUM) and can be integrated with Tennessee Statewide Travel Demand Model (TSM v3). Massive data collection was followed in this project and households' information, employment information, housing information was collected at the Traffic Analysis Zones (TAZ) level. Moreover, parcel data for 94 counties were collected to be incorporated into the model. The model is developed for the base year 2010 and forecasts the demographic and socio-economic condition of the state of Tennessee from 2015 to 2050 with five years intervals. The developed model is validated using the backcasting approach and the goodness of fit and error measures are provided to show the accuracy of the model. To present the result of the model, an online dashboard is created providing forecasting results from 2015 to 2050 at TAZ and County levels Moreover, the online dashboard presents a brief statistical analysis. The developed land-use model is integrated with TSM v3 and results are provided. Since the model incorporates house conditions (total houses and vacant houses) and land use conditions in each TAZ (residential, commercial, industrial, agricultural, and developable lands), it provides a powerful tool for policy analysis. A software is developed for running the model using MATLAB Compiler Runtime, which enables sharing the model with other parties.

17. Key Words **Land Use Model, Travel Demand Model, Transportation Planning**		18. Distribution Statement No restriction. This document is available to the public from the sponsoring agency at the website http://www.tn.gov/.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 66	22. Price

Acknowledgement

The authors would like to acknowledge the assistance of many individuals at TDOT during this study. Particularly, the authors would like to express their gratitude to the project manager of this study, Dr. David Lee, for setting the direction for the study and for providing valuable suggestions in the development of the model. The authors would also like to thank Chris McPhilamy, Mohammad Molla, and their team in Long Range Planning Division for their assistance with the data collection and travel demand model. Besides, computational facilities at the University of Memphis are greatly acknowledged.

Executive Summary

What was the research need?

Tennessee Department of Transportation (TDOT) developed a new statewide travel demand model in 2015. Some land-use inputs such as household, population, household by income, household by size, employment, and employment by category were derived from regional Metropolitan Planning Organizations (MPOs) socio-economic forecasts along with the use of the National Household Travel Survey with Tennessee add-on data. While the current travel demand model uses 2010 and 2040 as the base and future years of analysis, no land-use model currently exists to provide additional future year data or the ability to do scenario planning. To strengthen scenario analysis and policy planning, the travel demand model will need adequate land use inputs. However, currently, there is no statewide land-use model in TN that can be used to generate inputs for the travel demand model. A need for the statewide land-use model is imperative to obtain (1) accuracy of future year land use forecast that represent long-range transportation improvements and planned zoning, (2) cumulative and indirect effects of transportation projects, (3) evaluation of economic effects of various state and regional policies, (4) land-use changes because of rapid changes in travel behavior owing to emerging technologies, (5) accurate choices of residential locations because of emerging greener and tech-savvy lifestyle choices, and finally (6) facilitation of the land-use model to be integrated with the travel demand model.

What were the research objectives?

The main objectives of this project were as follows:

- Developing a statewide land-use model
- Validating the model accuracy using backcasting and forecasting analysis
- Developing an online dashboard for presenting the output of the model
- Integrating the land-use model with the Tennessee Statewide Travel Demand model
- Developing an application software based on the land-use model

What was the research approach?

In this project, the main goal was to develop a statewide forecasting land-use that can be integrated with Tennessee Statewide Travel Demand Model (TSM v3). In this regard, a comprehensive data collection was followed and households' information, employment information, housing information were collected at the Traffic Analysis Zones (TAZ) level. Moreover, parcel data for 94 counties were collected to be incorporated into the model. An enhanced gravity approach was followed to develop the model using the data collected and considering the travel demand model. The proposed land use model was first tested on a small example. The model was developed for the base year 2010 and forecasts the demographic and socio-economic condition of the state of Tennessee from 2015 to 2050 with five years intervals. The developed model was validated using the backcasting approach and the goodness of fit and error measures are provided to show the accuracy of the model. To present the results of the model, an online dashboard has been created providing forecasting results from 2020 to 2050 at TAZ and County levels. The developed land-use model is integrated with TSM V3 and results are provided. Since the model incorporates house conditions (total houses and vacant houses) and land use conditions in each TAZ (residential, commercial, industrial, agricultural, and developable lands), it provides a powerful tool for policy analysis. A software application has been developed for running the model using MATLAB Compiler Runtime, which enables sharing the model with other parties.

What were the findings?

The key findings of the project are as follows:

- Collecting the parcel data and adding the land use consumption section increased the accuracy of the model significantly.
- The result of validating the model showed that the model has acceptable accuracy in forecasting demographic and socio-economic conditions of the state of Tennessee.
- The model performance under a disaggregated environment and at the TAZ level was acceptable while the run time was reasonable.
- The proposed land use model can retain its accuracy even after eight iterations and forecast the demographic and socioeconomic condition of the state of Tennessee in the year 2050.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction and Literature Review

The dynamic nature of urban systems involves the interaction of different agents such as infrastructure, facilities, administration, and individuals in an integrated environment. Transportation is crucial for the sustainability of an urban system. The significant increase in the private car uses had a major negative impact on the efficiency of transportation systems; the need for more research in the area of congestion management and travel demand modeling became crucial. This resulted in the development of the first generation of travel demand models in the 1950s in the U.S. (Southworth, 1995).

Researchers immediately realized the interdependence of transportation systems and land use patterns and land use models were developed that utilized economic theory and statistics to produce forecasts of future changes in land use, demographics, and socio-economic characteristics of a case study area (White, 2010). It was obvious that changes in transport systems could affect the patterns of urban development and location choices of households and employment. On the other hand, major changes in land-use patterns could affect the number of trips, their destinations, and modes. The interdependence of transportation and land use patterns resulted in the development of integrated land use and transportation models (ILUTM).

The first generation of land-use models was introduced around the 1960s and were aggregate models of spatial interaction and gravity models. Then, utility-based econometric and discrete choice models were developed. These two first classes of models mainly followed the top-down approach (Iacono et al., 2008). More advanced models were gradually developed since the late 1980s. These new models are mainly micro-simulation disaggregate models. Agent and rule-based models and Cellular Automata were also designed. Many of these models are considered to follow the bottom-up modeling approach. However, the classification of land-use models in separate categories can be misleading as many models from different categories can share common concepts and characteristics (White, 2010). Parallel to the evolution of land-use models, travel demand models also are evolving. The traditional four-step urban transportation planning systems (UTPS) were replaced by the more advanced activity-based models. The major concept behind the development of the activity-based models was that travel behavior and trip generation is determined upon individuals need to complete specific activities daily (Sivakumar, 2007)(Mishra et al., 2011).

The development of advanced micro-simulation land-use models and activity-based travel demand models created the need for a new generation of integrated land use-transport systems. New models such as ILUTE and ILUMASS were developed and existing models such as UrbanSim and MUSSA were updated to facilitate the needs for advanced research in the field of integrated land use-transport modeling.

Land-use models utilize economic theory and statistics to produce forecasts of future changes in land use, demographics, and socio-economic characteristics (White, 2010). The first land-use models were introduced around the 1960s and were aggregate models of spatial interaction and gravity models (Iacono et al., 2008). The model of Metropolis developed by Lowry in 1964 is the first operational land-use model. A new approach regarding the development of land-use models was introduced around the 1980s. This new approach suggested the development of econometric and discrete choice models that were based on utility theory. The first two categories include the spatial interaction and the econometric models are considered to follow the top-down modeling framework (Iacono et al., 2008). More advanced models were gradually developed since the late 1980s. These new models are mainly micro-simulation disaggregate models, including agent and rule-based systems and cellular automata. Many of these models are considered to follow the bottom-up modeling approach. Figure 1-1 provides an overview of the evolution of land-use models.

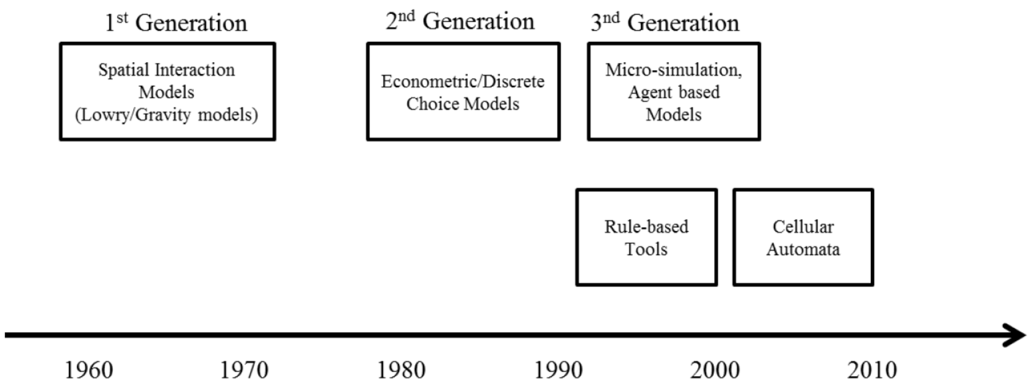

Figure 1-1. Land Use Model Evolution

1.1 Lowry/Lowry-Garin Models

Lowry (1964) developed a simple spatial interaction model, the “Metropolis Model,” which is widely used by many agencies in the U.S. primarily because of its simplicity and transparency. The model was designed to evaluate future changes in retail employment, residential population, and land use in the greater Pittsburgh Area. The Lowry model can also be a useful tool for evaluating future policies related to transportation planning and land use development. Activities are classified into three categories: basic sector (industrial, business, and administrative activities related to nonlocal customers), the retail sector (industrial, business, and administrative activities related to local customers), and household sector (focuses on the residential population). Employment is divided into Basic and Non-Basic (services). A singly constrained Lowry model spatially fixes the Basic employment. The Non-Basic employment and households are allocated to zones based on attractiveness coefficients derived from gravity model estimates until convergence occurs. Garin in 1966 suggested a significant revision of the model’s structure (Goldner, 1971). A vector and matrix version of the Lowry Model was introduced. Matrix operations were found to produce exact solutions, improving the performance of the model in total.

1.2 TOMM

TOMM (Time-Oriented Metropolitan Model) was initially introduced in 1964 as the first Lowry-derivative model (Crecine, 1968). TOMM is a spatial location model that shares many common characteristics with the Lowry/Metropolis model. The major difference between the two models is the more disaggregated nature of the TOMM model (e.g., classification of facilities, households, population, employment, etc.). The structure of the model includes the Exogenous employment sector (employment and activities outside the borders of an urban area), the

Endogenous commercial employment sector (employment and activities inside the borders of a metropolitan area), and the household or population sector (classification of households based on socio-economic characteristics).

1.3 PLUM

PLUM (Projective Land Use Model) was developed for the Bay Area Transportation Study Commission in 1968 (Crecine, 1968). It is a spatial model for activity and land use planning. PLUM has an incremental structure that utilizes land use data, market economics, and demographics to forecast future development. The inputs of the model include dwelling units, population, and employment data. The economic framework of the model includes the "population-serving" activities/employment (based on household and employee's spatial distribution) and the "basic" employment that describes the spatial allocation of endogenous industries.

1.4 TOPAZ/TOPMET

TOPAZ (a technique for the optimum placement of activities in zones) is an optimization model for identifying activity locations, designed at the Division of Building Research of the Commonwealth Scientific and Industrial Research Organization, Australia in the 1970s. TOPMAN, a planning version of TOPAZ also became available. TOPAZ was developed using the FORTRAN programming language. The required input data include among others: employment and population forecasts, transportation network characteristics, car ownership, and activity costs. The modeling process in TOPAZ starts with allocating employment and housing to the zonal level in such a way that minimizes infrastructure and transportation costs. Then trip patterns are created based on the assumption that demand can be predicted from spatial activity locations using entropy-maximization.

Then mode choice and travel costs are calculated. At the last part of the modeling process, data types are further aggregated at the urban or regional level. The outputs of the model include employment and population data, number of houses, and vacant land. The transport module of TOPAZ produces outputs such as trips per mode, emissions, and energy consumed. Economic results such as travel, activity costs, and accessibility measures can also be provided. TOPAZ has mainly been applied in different case studies in Australia.

1.5 IRUPD

The IRUPD model is a simulation land use and transport model, initially introduced in Germany in 1977 (Wegener, 1998). It's a spatial interaction and zone-based model and the major stock variables include population, employment, housing, and non-residential buildings. The model structure consists of six sub-models that are applied to identify employment and demand pattern changes. These sub-models involve a transport sub-model for estimating trips, sub-models for identifying the stock variables changes, and the results of public programs (e.g., infrastructure investments). Additional sub-models focus on identifying employment changes and residential/non-residential changes (e.g., new buildings, new houses, etc.). The outputs from the transport sub-model include the number of trips and different measures of performance such as trip time, trip costs, and emissions.

1.6 LILT

LILT (LEEDS Integrated Land Use-Transportation) model was introduced by Dr. Mackett in 1979 at the University of Leeds, England (Zhao & Chung, 2006). LILT is an entropy-maximization system for predicting future population, residential, and employment changes. It consists of three major components: a Lowry derivative location model, a four-step travel model, and a car ownership model. Depending on the level of accessibility, employment is defined as primary, secondary, and tertiary. LILT has been applied in different case studies in Germany, Japan, England, and Greece.

1.7 POLIS

The Projective Optimization Land Use Information System (POLIS) is a land-use transport model mainly applied in the greater San Francisco Bay Area (Prastacos, 1986). It was designed by the Association of Bay Area Governments and replaced previous models such as PLUM (Goldner, 1968). POLIS is a non-linear mathematical optimization model for producing forecasts of land use, employment, housing, population, and transportation changes. The optimal solution maximizes the profitability of employers and the utility related to travel choices for work and shopping. POLIS was different compared to the traditional Lowry derivative models as it integrated basic/non-basic employment and housing and it applied short of microeconomic theory. Housing choices are determined by travel-to-work behavior and housing availability. The location of retail facilities is affected by the proximity to areas with increased population. Additional parameters that are considered include the shopping centers' attractiveness, the accessibility to work location, and the economy of the study region.

1.8 HLFM Model

The Highway Land use Forecasting Model (HLFM) is a spatial interaction, Lowry derivative land-use model, introduced by Alan Horowitz (Dowling, 2005). HLFM utilizes land use, demographic and socioeconomic data to forecast future employment and population changes of a study area. HLFM is an equilibrium model since it focuses on the land use demand and supply equilibrium. Modeling in HLFM is an iterative process that starts with determining the number and the location of the basic industry employment in the study area. The allocation process follows, and the model estimates the conditional probabilities related to worker residential and service employment locations. The model then identifies the employment and the population of each district depending on the corresponding attractiveness of each choice. HLFM was designed to fully interact with travel demand models.

1.9 ITLUP/DRAM/EMPAL/METROPILUS

Putman (1983a) developed a derivative of the Lowry Model, the Integrated Transportation Land-Use Package (ITLUP). ITLUP includes two major models: Disaggregated Residential Allocation Model (DRAM) and Employment Allocation Model (EMPAL). DRAM works by allocating households based on the zonal attractiveness considering current residential development, the capacity derived from vacant and developable land, and other socio-economic characteristics of the zone. EMPAL also works in the same way by allocating employment based on the attractiveness of zones, considering an impedance cost matrix. DRAM/EMPAL has fewer data requirements compared to the original Lowry model.

DRAM/EMPAL is sensitive to changes in the basic sector employment and other investments that can potentially change the economic geography of the area. An important advantage of DRAM/EMPAL is its basis on generally available data such as population, households, and employment (Southworth, 1995). The market-clearing process is not modeled. A software-based version of ITLUP, called METROPILUS was later introduced.

1.10 FLUAM

The METROPLAN ORLANDO's Future Land Use Allocation Model (FLUAM) was introduced in 2006 (Data Transfer Solution, LLC, 2006). FLUAM is a geographic information system (GIS), a parcel-based tool for predicting population and employment changes and distributing forecasted data to TAZ. Forecasts are initially distributed at the parcel level based on historical land-use data and development trends and then are aggregated at the TAZ level. The aggregation at the TAZ level allows for developing transportation policies and plans to facilitate forecasted growth. A combination of the top-down and bottom-up approaches is applied for distributing and aggregating the forecasted data. The major inputs into the model include existing and future land use data, land use development factors/growth forecasts, etc.

1.11 TELUM

TELUM is an integrated land use and transport model developed to evaluate the effects of land use on transportation planning (Spasovic, 2013). It is part of the TELUS system, a computer-based system developed to assist transportation agencies in decision management. TELUM was introduced in 2006, focusing to assist small and medium-size MPOs to forecast the impact of future population and employment changes on land use. TELUM development was based on DRAM/EMPAL model, and it's integrated with GIS tools. The model structure consists of five modules: i) IDEU module for initial data entry, ii) DOPU for data organization and preparation, iii) TIPU for travel impedance data processing, iv) MCPU for model calibration, and v) MFCU module for model forecasting. The software outputs include employment and household density, land consumption, and density gradient. Different agencies such as the Missoula Area Council of Governments, the Des Moines (IA), and the Little Rock (AR) MPOs have used this model.

1.12 G-LUM

The Gravity-based Land Use Model (G-LUM) was developed by Professor Kara Kockelman and associates at the University of Texas at Austin (Valsaraj et al., 2007)(Paul & Zhou, 2009). G-LUM was used to validate the outputs of TELUM. The model structure is based on the formulation of the ITLUP package (Putnam, 1983) and includes three major sub-models for predicting changes in employment location, residential location, and land consumption. G-LUM was developed in MATLAB software and a graphical user interface (GUI) is also available. Model calibration is based on the comparison of lag with base year data. G-LUM is one the most recent model presented applying gravity theory. Five major data categories (employment data, household data, land use data, zone data, and inter-zonal travel times) are required for G-LUM implementation.

1.13 Land Use Allocation Model for Florida Turnpike

The Land Use Allocation Model (LUAM) was applied as part of the Turnpike State Model (TSM) integrated land use and transport model for the Florida Department of Transportation (FDOT) Turnpike Enterprise (Adler et al., 2007)(Lawe s., Lobb J., 2007). LUAM is a parcel-based growth model, developed in C++. Processing time usually requires 2-4 minutes. LUAM focuses on the allocation of the forecasted population and employment. Land allocation at the zonal level is based on four major parameters: household and employment density, developable land, and transportation accessibility. Land consumption is estimated using a logit model that produces the probability of land development in specific traffic analysis zones. Housing and employment developments (density) are determined using a linear model.

1.14 LUTSAM

LUTSAM (Land Use and Transportation Scenario Analysis and Microsimulation) was developed by two major partners; the Delaware Department of Transportation and the State Smart Transportation Initiative at the University of Wisconsin-Madison (Thompson-Graves S., DuRoss M., Subhani R., Holloway B., 2012). LUTSAM is an evaluation tool of land use and transportation alternative that integrates GIS, land use, and travel demand modeling and microsimulation. LUTSAM is a GIS and parcel-based model for evaluating smart growth policies, land use developments, and investments such as bicycle and pedestrian facilities. Modeling inputs include road networks, layer information, traffic analysis zones, and base maps. The scenario analysis and evaluation start with the identification of the study area and the location of new developments. The study area is divided into sub-regions and then the land use type and the density of each sub-area are determined. The road network and the sidewalks are also specified. Home location and the connectivity with the sidewalks and the roadway are then identified. The last part of the modeling process focuses on merging the new roadways and sidewalks with the existing networks. LUTSAM has been designed to operate in an integrated environment and the corresponding outputs can be utilized as inputs for travel demand models and simulation software.

1.15 PECAS

PECAS is a spatial input-output econometric model for allocating flows of exchanges such as goods, services, labor, and space from production to consumption points. Land use consumption due to the job and household growth can be simulated using Social Accounting Matrix (SAM). Nested Logit Models are applied to allocate flows based on exchange prices and market conditions. The exchange flows are then translated into transport demand for transportation networks. PECAS has been applied for developing land use-transport interaction models in different case studies around the U.S.

- PECAS model consists of two PECAS and two non-PECAS modules that operate in an integrated environment (Hunt et al., 2009). The PECAS modules include:
    - Space Development (SD) module: This module utilizes logit allocation models to identify the land and floor space changes due to developers' actions (new developments, demolitions, etc.).

    - Activity Allocation (AA) module: Logit models are also applied to allocate activities in space and model the interaction of activities through flows of commodities.
- The two non-PECAS modules include:
    - Transport Model (TR) module: An external transportation planning model is used to represent the transport network and the corresponding demands. The land-use model and the transport model are integrated through the translation of commodity flows into travel demand.
    - Economic Demographic Aggregate Forecasting Model (ED) module: ED module includes a set of different models to forecast household, population, and employment future changes.

The PECAS model has extensive data requirements including parcel boundaries, land prices, etc., that may not be available for the study region.

1.16 URBANSIM

UrbanSim was developed at the Center for Urban Simulation and Policy Analysis (CUSPA), University of Washington (Waddell, 2002)(Borning et al., 2008). UrbanSim can primarily evaluate the impact of alternative transportation, land use, and environmental policies. UrbanSim is an open-source tool that allows data analysis and processing at the grid, parcel, and zone levels. The option of integrating UrbanSim with travel demand models is available to users. UrbanSim is a microsimulation model, and its modular structure is based on utility theory. Household and employment location choices, real estate development, and prices can be modeled. A disaggregate classification of households is carried out, considering the number of individuals, workers, children, and the income of each household. Employment is also classified. The model can also be used to simulate disequilibrium conditions. UrbanSim is one of the most efficient integrated land use transport models for application in a regional case study. The model is now fully operational and has been implemented by different transportation agencies and organizations due to its advanced characteristics that are summarized below:

- Efficient geographical coverage at the regional level
- Spatial detail options that include grid, parcel, and zone versions of UrbanSim
- Integration with travel demand models (including both trip-based and activity-based)
- Consideration of multimodality
- Visualization capabilities for output representation that include tables, graphs, animation, and lately 3-D representation options

UrbanSim provides the option to develop extremely detailed models at the micro level that allow users to carry out complicated and efficient land use-transportation analysis and research. However, data requirements are quite extensive.

If the obstacle of collecting extensive and quality data has been overcome, UrbanSim can provide a set of different forecast outputs that are summarized as follows:

- Buildings by type, price, etc.
- Size of land, open space, etc.
- Households by income, size, etc.

- Employment by sector and building type
- Transportation accessibility, mode choice, Delay, etc.
- Greenhouse gas emission, energy use, etc.

Chapter 2 Development of Statewide Land-Use Model

In this chapter, the methodology applied for developing a statewide land-use model is discussed. However, before developing the land-use model, it is imperative to understand what the inputs for the statewide travel demand model are. Therefore, this chapter is divided into two sections. In the first section, Tennessee's statewide travel demand model and its data requirement are discussed. In the second section, the proposed land use model and its formulation are discussed.

2.1 Identification of Travel Demand Model's Data Requirement

The purpose of this research was to integrate the Tennessee Statewide Travel Demand Model (TSM V3) with a statewide land-use model, the proposed model should be able to generate the data requirement of statewide travel demand model input data needs and for forecasting years of analysis (base and future intermediate years). Moreover, treatment for internal and external zones should be the same.

The Tennessee statewide travel demand model consists of three different components: a short distance passenger model (trips less than 50 miles), a long-distance passenger model, and a freight model. The underlying geographic area of operation is at the TAZ level. The total number of TAZs in TSM is 3,687. Zonal attributes include the number of households, categorized by income, size, worker, presence of students, presence of seniors, and the number of vehicles; and the number of employments categorized by 20 sectors of North American Industry Classification System (NAICS) codes. The TSM V3 can be understood at a high level as comprised of input network and socioeconomic data together with some component demand models and a highway assignment model. The demand components can be gathered in three broad groups related to short-distance passenger demand, long-distance passenger demand, and freight and truck demand. The TST V3 uses TransCAD's implementation of the tri-conjugate Frank-Wolfe algorithm for multi-class user equilibrium traffic assignment (Bernardin Jr et al., 2017). The accessibility matrices which serve as input for the land-use model are obtained from TSM's assigned networks using the shortest path method.

Socioeconomic Data:

The socioeconomic data for all 3 phases (models) is identical to phase 2 (short and long trips) with the addition of transit and airport files and is based on the TAZ system. In the Tennessee statewide transportation model (TSM), TAZs are based on different sources:

- Urban areas:
  Aggregated MPO model zones for urban areas, including zones outside Tennessee for bistate MPOs
- Rural areas – Travel shed principle:
  Federal Highway Administration's (FHWA) Travel model improvement program
- Outside of Tennessee:
  Counties, counties combination, or an entire state

The total number of zones in TSM version 3 is 3687 while 35 of them are for states, although some states are split into multiple parts (Figure 2-1).

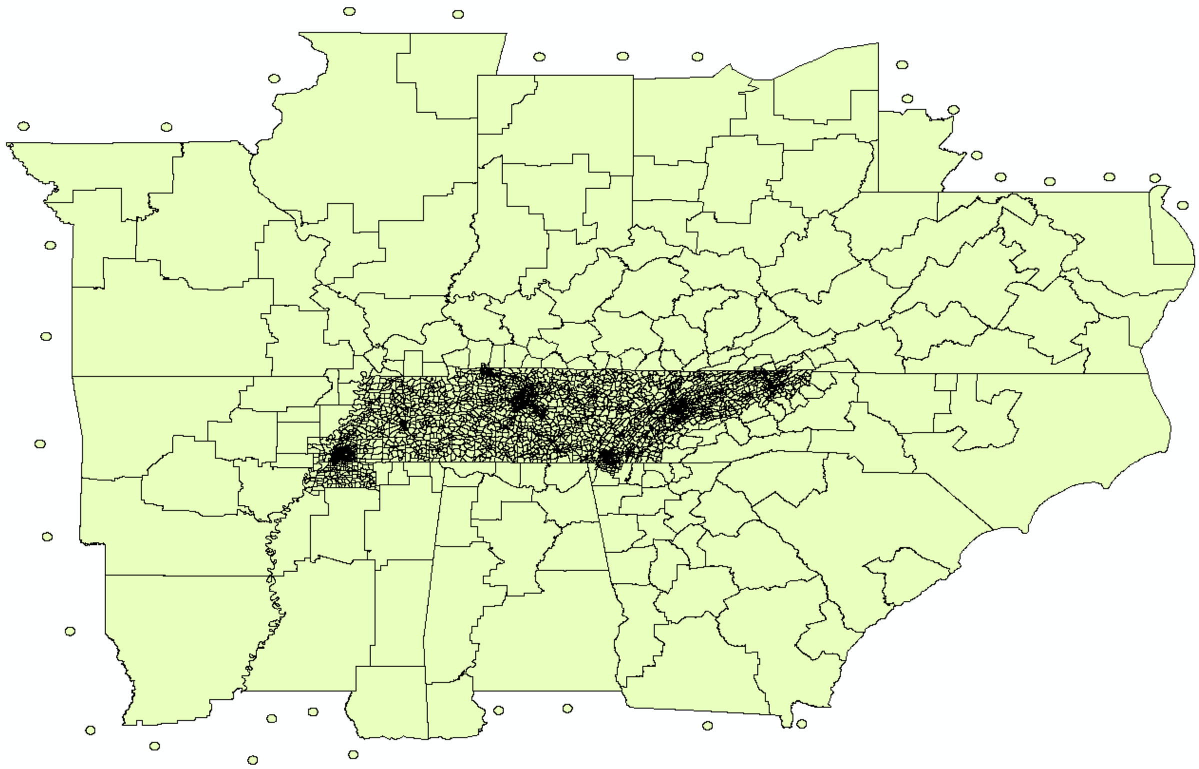

Figure 2-1. Tennessee and Halo Region Zone (Phase 2) Source: Tennessee statewide travel model (version 3) development and validation

Based on Tennessee statewide travel model (version 3) the required socio-economic data for developing a statewide travel model is tabulated in Table 2.1. The land used model should generate this data set at the TAZ level.

TABLE 2-1 TAZ SOCIO-ECONOMIC DATA FOR TN STATEWIDE TRAVEL DEMAND MODEL (SOURCE: TENNESSEE STATEWIDE TRAVEL MODEL (VERSION 3) DEVELOPMENT AND VALIDATION)

Filed	***Description***
TOTPOP	Total population
HHPOP	Household population
GQPOP	Group quarter population
HH	Total households
HH SIZE	Average household size
HHSIZE Average Household Size	HHSIZE Average Household Size
HHINC Average Household Income	HHINC Average Household Income
HHWRK Average Household Workers	HHWRK Average Household Workers
HHVEH Average Household Vehicles	HHVEH Average Household Vehicles
HHSTD Average Household Students	HHSTD Average Household Students
HHSIZE Average Household Size	HHSIZE Average Household Size
SENHH	Household seniors
TOTEMP Total Employment	TOTEMP Total Employment
EMP11	Employment in NAICS 11
EMP21	Employment in NAICS 21
EMP22	Employment in NAICS 22
EMP23	Employment in NAICS 23
EMP3133	Employment in NAICS 31-33
EMP42	Employment in NAICS 42
EMP4445	Employment in NAICS 44-45
EMP4849	Employment in NAICS 48-49
EMP51	Employment in NAICS 51
EMP52	Employment in NAICS 52
EMP53	Employment in NAICS 53
EMP54	Employment in NAICS 54
EMP55	Employment in NAICS 55
EMP56	Employment in NAICS 56
EMP61	Employment in NAICS 61
EMP62	Employment in NAICS 62
EMP71	Employment in NAICS 71
EMP72	Employment in NAICS 72
EMP81	Employment in NAICS 81
EMP92	Employment in NAICS 92

2.2 Development of Statewide Land Use Model

In this section, the proposed land use model is presented. Gravity-based models are between the first generation of the land-use models and drop between macro models. These models were first introduced by Lowry (1964) and since then several models have been provided based on this method. The macro model will utilize land use, demographic and socioeconomic data to forecast future employment and population changes of a study area. It will be an equilibrium model focusing on land use demand and supply equilibrium. Macro land-use modeling entails an iterative process that starts with determining the number and the location of the basic industry employment in the study area. The allocation process follows where the model estimates the conditional probabilities related to worker, residential, and service employment locations. The model then identifies the employment and the population of each district depending on the corresponding attractiveness of each choice. The gravity-based land-use models usually show a low accuracy in comparison with the micro-simulation model (URBANSIM, ILUTE, …), while the ease of implementation and short run-time makes this model a flexible tool for transportation planners to evaluate different policies. Moreover, the data requirement of this type of model is accessible. Gravity-based models are aggregated in nature, while in this report a disaggregated gravity model is provided that can provide the statewide travel demand model's inputs. The proposed model has similarities with the models introduced by (Putman, 1983b), Integrated transport Land Use Package (ITLUP), and Gravity-based Land Use Model (G-LUM) provided by (Paul & Zhou, 2009). The model presented in this research can be applied in smaller zone sizes like Traffic Analysis Zone (TAZ) or Blocks, while Putman's model has a restriction on the zone size and zones' population. Besides, the proposed model in this report has been developed based on the available data the research team could collect.

2.2.1 Model approach

In this section, the proposed land-use model's structure is provided. The model is called the Large-Scale Land Use Model (LS-LUM). LS-LUM contains two principal sections and two subsections. Principal sections estimate households and employments in different categories. These two models incorporate gravity theory for households and employments allocation. The first principal model is named HH-AL (Households Allocation) which is responsible for residential location choice. This model assigns households to each TAZ based on the total number of houses, the number of vacant houses, the amount of residential land (acres), total useable land, and the travel cost (time) between zones. The second principal model is called EMP-AL (Employments Allocation). This model allocates employment based on job opportunities, the amount of commercial, industrial, and agricultural land (acres), and travel costs between zones. Two subsections are responsible for updating house conditions and land-use consumption which are the components of principal sections. Two models provided in these subsections are called HC (House Condition) and LC (Land Consumption). HC models the number of total and vacant houses and LC Models the amount of land (acres) in five land-use classes (residential, commercial, industrial, agricultural, and developable or vacant). Multiple Linear Regression (MLR) is applied in these subsections. In the following sections, the formulations of these models are provided. Generally, in the following equations, i and j represent TAZs, $C_{i,j}$ represents the travel cost between zone i and zone j, and t represents the period (i.e., the year 2010). Moreover, n and k stand respectively for household categories (i.e., different household sizes) and employments categories (i.e., NAICS sectors categories).

I. **HH-AL (Households Allocation):**

$$N_{i,t}^{n} = \eta^{n} \sum_{j} a_{n}\, E_{j,t}^{T} \frac{W_{i,t-1}^{n}\, C_{i,j,t-1}^{\alpha^{n}}\ \exp\left(\beta^{n} C_{i,j,t-1}\right)}{\sum_{i} W_{i,t-1}^{n}\, C_{i,j,t-1}^{\alpha^{n}}\ \exp\left(\beta^{n} C_{i,j,t-1}\right)} + (1-\eta^{n}) N_{i,t-1}^{n} \tag{1}$$

Where,

$$W_{i,t-1}^{n} = (H_{i,t-1}^{T})^{o^{n}} (H_{i,t-1}^{V})^{p^{n}} (1 + \frac{L_{i,t-1}^{Res}}{L_{i,t-1}^{T}})^{q^{n}} \tag{2}$$

In Equation (1), $N_{i,t}^{n}$ is the number of households in category n in zone i in time t, a_n is the proportion of the population over employment in zone i, $E_{j,t}^{T}$ is the total number of employments, $W_{i,t-1}^{n}$ is the attractiveness function of zone i to which attract employment in zone j to live in zone i in year $t-1$. $W_{i,t-1}^{n}$ is a weighted multiplication of different components in a zone. In Equation (2), $H_{i,t-1}^{T}$ is the total number of houses, $H_{i,t-1}^{V}$is the number of vacant houses, $LP_{i,t-1}$ is the residential land value, $L_{i,t-1}^{Res}$ is the amount of residential land are in zone i in year $t-1$. Finally, η, α, β, o, p, and q are parameters estimated in the calibration procedure.

II. **EMP-AL (Employments Allocation):**

$$E_{j,t}^{k} = \lambda^{k} \sum_{i} N_{i,t-1}^{T} \frac{M_{j,t-1}^{k}\, C_{i,j,t-1}^{\alpha^{k}}\ \exp\left(\beta^{k} C_{i,j,t-1}\right)}{\sum_{i} M_{j,t-1}^{n}\, C_{i,j,t-1}^{\alpha^{k}}\ \exp\left(\beta^{k} C_{i,j,t-1}\right)} + (\,1-\lambda^{n}) E_{j,t-1}^{k} \tag{3}$$

Where,

$$M_{j,t-1}^{k} = (E_{j,t-1}^{k})^{g^{k}} (L_{j,t-1}^{Com} + L_{j,t-1}^{Ind} + L_{j,t-1}^{Agr})^{h^{k}} \tag{4}$$

In Equation (3), $E_{j,t}^{k}$ is the number of employments in category k, $N_{i,t-1}^{T}$ is the total number of households, and $M_{j,t-1}^{k}$ is the attractiveness function shows how much zone j is attractive for people living in zone i to find a job. $M_{j,t-1}^{k}$ is calculated based on, job opportunities in year $t-1$ ($E_{j,t-1}^{k}$), the amount of commercial ($L_{j,t-1}^{Com}$), industrial ($L_{j,t-1}^{Ind}$), and agricultural ($L_{j,t-1}^{Agr}$) land in zone j in year $t-1$. Finally, λ, α, β, g, and h are parameters estimated in the calibration procedure.

III. **HC (House Conditions):**

In this subsection, the total number of houses and the number of vacant houses in each TAZ are updated. First, the total number of houses in each TAZ is calculated by applying a Multiple Linear Regression. As Equation (5) shows, the total number of houses in zone i and year t ($H_{i,t}^{T}$) is the dependent variable; while the number of total houses in the previous year ($t-1$), the amount of vacant land ($L_{i,t-1}^{Vac}$), and the total number of households are the independent variables.

$$H_{i,t}^{T} = \theta_0 + \theta_1 (H_{i,t-1}^{T}) + \theta_2 \left(H_{i,t-1}^{V}\right) + \theta_3 \left(L_{i,t-1}^{Vac}\right) + \theta_4 \left(N_{i,t}^{T}\right) + \varepsilon \tag{5}$$

In Equation (5), $L_{i,t-1}^{Vac}$ is the amount of vacant or developable land in zone i and ε is the error associated with regression. In this equation, θ_0 is the intercept and θ_1 to θ_4 are coefficients estimated in calibration.

After calculating the total number of houses, the number of vacant houses in each TAZ can be estimated as follow:

$$H_{i,t}^{V} = H_{i,t}^{T} - \sum_{n} N_{i,t}^{n} \tag{6}$$

IV. LC (Land use Conditions):

Finally, in LC, the amount of land in different land-use classes is updated to feed the two principal models (HH-AL and EMP-AL) to forecast future years' demographic and socio-economic conditions.

$$L_{i,t}^{Res} = R_0 + R_1(L_{i,t-1}^{Vac}) + R_2(L_{i,t-1}^{Res}) + R_3(N_{i,t-1}^{T}) + R_4(N_{i,t}^{T}) + \varepsilon \tag{7}$$

$$L_{i,t}^{Com} = C_0 + C_1(L_{i,t-1}^{Vac}) + C_2(L_{i,t-1}^{Com}) + C_3(E_{i,t-1}^{Com}) + C_4(E_{i,t}^{Com}) + \varepsilon \tag{8}$$

$$L_{i,t}^{Ind} = I_0 + I_1(L_{i,t-1}^{Vac}) + I_2(L_{i,t-1}^{Ind}) + I_3(E_{i,t-1}^{Ind}) + I_4(E_{i,t}^{Ind}) + \varepsilon \tag{9}$$

$$L_{i,t}^{Agr} = A_0 + A_1\left(L_{i,t-1}^{Vac}\right) + A_2\left(L_{i,t-1}^{Agr}\right) + A_3\left(E_{i,t-1}^{Agr}\right) + A_4(E_{i,t}^{Agr}) + \varepsilon \tag{10}$$

$$L_{i,t}^{Vac} = L_{i,t-1}^{Vac} - \left(L_{i,t-1}^{Res} - L_{i,t}^{Res}\right) - \left(L_{i,t-1}^{Com} - L_{i,t}^{Com}\right) - \left(L_{i,t-1}^{Ind} - L_{i,t}^{Ind}\right) - \left(L_{i,t-1}^{Agr} - L_{i,t}^{Agr}\right) \tag{11}$$

In Equations (9) to (10), E^{Arg} refers to the number of employment in NAICS sector 11 (agriculture, forestry, fishing, and hunting), E^{Com} is the number of employment in NAICS sectors 44, 45, 51, 52, 53, and 72 (retail trade, finance and insurance, real estate and rental and leasing, accommodation and food services), and E^{Ind} is the number of employment in NAICS sectors 21, 31, 33, and 42 (mining, quarrying, oil and gas extraction, manufacturing, and wholesale trade).

LS-LUM is designed considering the capability of integrating with a travel demand model. This integrated modeling framework starts with forecasting employment in different categories and for each zone (see Figure 2-2). This section of the model gets the number of employments in the category k, the amount of agricultural, commercial, industrial lands, and travel cost in each zone and for the prior year. The output of this section is the forecasted employment (by different categories) in each TAZ. The output of the EMP-AL would serve as input for the HH-AL. The HH-AL incorporates the current total employment (from EMP-AL), the total number of houses, the number of vacant houses, and the proportion of residential to total land in each zone for the prior year; and the output is the number of households (by different categories, e.g., income). Then HC computation is processed by forecasting how many houses will be built in each TAZ. This section needs the total and the vacant number of houses, the amount of vacant land in the prior year, and the forecasted total number of households (from the HH-AL section). The output of HC feeds HH-AL by providing the number of total and vacant houses. Lastly, LC forecasts the amount of residential, commercial, industrial, agricultural, and vacant land. The output of LC directly affects other models' sections. By connecting LC to other sections, capturing the effect of land-use changes on the socio-economic character of each TAZ would be possible, and more accurate results can be obtained. The amount of commercial, industrial, and agricultural land modeled in this section will be added to EMP-Al and the amount of residential land will be added to HH-AL. Finally, the amount of vacant land is one of the components involved in forecasting the total number of houses in a zone. Moreover, the amount of vacant land in each zone works as a development restriction. Because in the model, if all the vacant lands had been allocated to other land-use classes, no more development will happen, and the model will stop adding a new area to other land-use classes (residential, commercial, industrial, and agricultural).

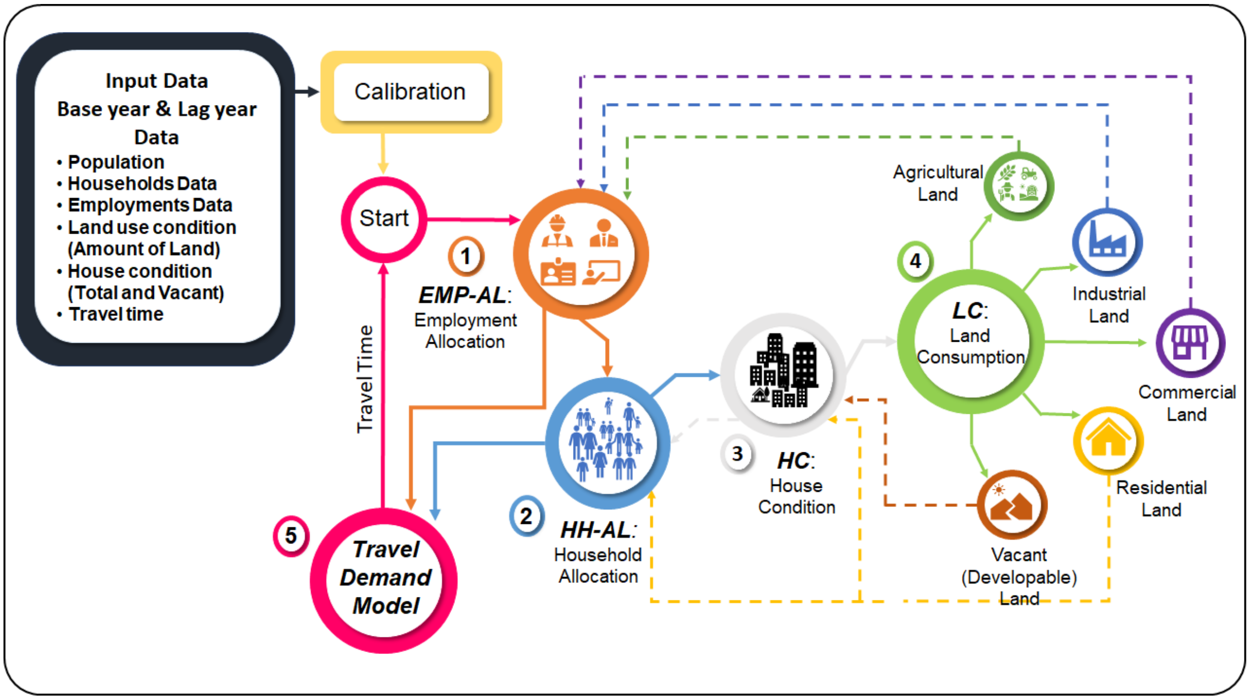

Figure 2-2. Proposed Integrated Land Use Transport Model's Flowchart (dashed lines represent one period (t-1) lagged feedback of information; each period is 5 years)

2.2.2 Model assumptions

The model is coded on MATLAB and follows two rules to control the population and employment attracted to each zone. First, households (population) and the number of employments in each TAZ were not allowed to fall or increase by or more than 5% in any (5-year) time interval. Second, growth in these counts was limited by the zone's population capacity. TAZs that violated the first rule were flagged, and then the corresponding number of households or jobs were taken from the unmarked TAZs, in proportion to their originally forecasted counts. Similarly, TAZs that violated the second rule were flagged, and then the "extra" households or jobs were reallocated to the unmarked TAZs, in proportion to their original counts. This reallocation process was run iteratively until all TAZs satisfied these three rules.

2.2.3 Calibration

The parameters of four models (HH-AL, EMP-AL, HC, and LC) need to be estimated through a calibration process. The calibration of the proposed models is categorized into two sections. The first section is dedicated to estimating the parameters of HC and LC. These two models are Multiple Linear Regression, and the intercept and coefficients are estimated using the least square method. The objective of the second section is to estimate the parameters of HH-AL and EMP-AL. The calibration is conducted through the maximum likelihood approach where the two following objective functions are defined. First, for HH-AL the objective function is as below:

$$Z_1 = Min \sum_i \sum_n \frac{\left(N^n_{i,t_{Obs}} - N^n_{i,t_{Est}}\right)^2}{\left(\sigma_{N^n_{i,t_{Obs}}}\right)^2} \tag{12}$$

Where, $N^n_{i,t_{Est}}$ is defined in Equations (1) and (2). In the objective function Z_1, $N^n_{i,t_{Obs}}$ and $N^n_{i,t_{Est}}$ are respectively, the number of observed and estimated households in category n and zone i and $\sigma_{N^n_{i,t_{Obs}}}$ is the standard deviation of observations. Where, the decision variables are η, α, β, o, p, and q (the calibration parameter mentioned in Equations (1) and (2). Moreover, a similar objective function is defined for EMP-AL as follow:

$$Z_2 = Min \sum_j \sum_k \frac{\left(E^k_{j,t_{Obs}} - E^k_{j,t_{Est}}\right)^2}{\left(\sigma_{E^k_{j,t_{Obs}}}\right)^2} \tag{13}$$

Where, $E^k_{j,t_{Est}}$ is defined in Equations (3) and (4). Similarly, in the objective function Z_2, $E^k_{j,t_{Obs}}$ and $E^k_{j,t_{Est}}$ are, the number of observed and estimated employment in category k and zone j and $\sigma_{E^k_{j,t_{Obs}}}$ is the standard deviation of observations respectively. Where, the decision variables are α, β, g, and h, defined in Equations (3) and (4).

Both objective functions Z_1 and Z_2 are non-linear, non-convex, and are not subject to any constraints. In the previous land-use models (TELUM and G-LUM), a gradient search method and the Nelder-Mead method with 12 different initial points were applied. Previous approaches add strict limitations to the solution approach. First, due to the non-convexity of objective functions, using the gradient search method would (Zhou et al., 2009) increase the chance of trapping in a local optimum solution. Second, the accuracy and final solution of the Nelder-Mead method with initial points are highly sensitive to the selection of initial points. Therefore, to eliminate these limitations, in this paper, an evolutionary algorithm is applied to solve the above-mentioned optimization problem. The following section discusses the proposed solution approach.

2.2.4 Data requirements

The proposed model (LS-LUM) needs six sets of input data which are households, employments, house conditions (total and vacant houses), amount of land in five land-use classes, and travel time. These data sets are needed for two periods of time with a time interval of five years. The household data, along with categories (total population, total households, household income, household size, household worker, household seniors, household students, quarter group) were collected from census data. This data set is available every 10 years. The employment data containing 20 categories of NAICS codes are available through Longitudinal Employment and Household Dynamics (LEHD). The house condition (the number of total and vacant houses) is collected through census data. The land-use condition in five land-use classes was collected from parcel data (*Tennessee Comptroller of the Treasury*, n.d.).

Lastly, travel time data is obtained from TSM V3. This section will be discussed in more detail in the next chapter (chapter 3 Data Collection).

TABLE 2-2 DATA REQUIREMENT FOR DEVELOPING LAND USE MODEL

Data set	***Source***	***Description***
Total population	Census data 2000 and 2010	Available at the block level
Total Households	Census data 2000 and 2010	Available at the block level
Households' income groups	Census data 2000 and 2010	Available at the block level
Households with different size	Census data 2000 and 2010	Available at the block level
Group quarters	Census data 2000 and 2010	Available at the block level
Households' seniors	Census data 2000 and 2010	Available at the block level
Total Employment	Longitude Employment Household Dynamics (LEHD)	Available form 2002
Employment of NAICS sections (20 sections)	Longitude employment - Household Dynamics (LEHD)	Available form 2002
House Situation (Vacant, Occupied, Total number of houses)	Census data 2000 and 2010	Available at the block level
The amount of land in five land use classes	Parcel data from the Tennessee Comptroller of the Treasury	Available for 2020
Geographic region shapefile (TAZ shapefile)	Tiger line shapefile data set	-
Total Available land	Tiger line shapefile data set	Shape area is used
Travel Time between each zone (Travel Cost)	TSM V3	-
TAZ's Land Restrictions	-	Not available

2.3 Three County Example

To test the applicability of the proposed model, the model is applied to a small area in the state of Tennessee. This section is provided to prove the concept of the proposed model. In this regard, the proposed land use model is applied to a region containing three counties of the state of Tennessee. The model applied to Shelby County, Tipton County, and Fayette County. By selecting these three counties the model performance in both high dense TAZ and TAZs with fewer population and employment will be tested. The study region has 387 TAZs and hosts 1,027,138 people in 2010. Figure 2-3 shows the geographic information of the small sample. In this figure, the state of Tennessee with its 95 counties is highlighted with red lines.

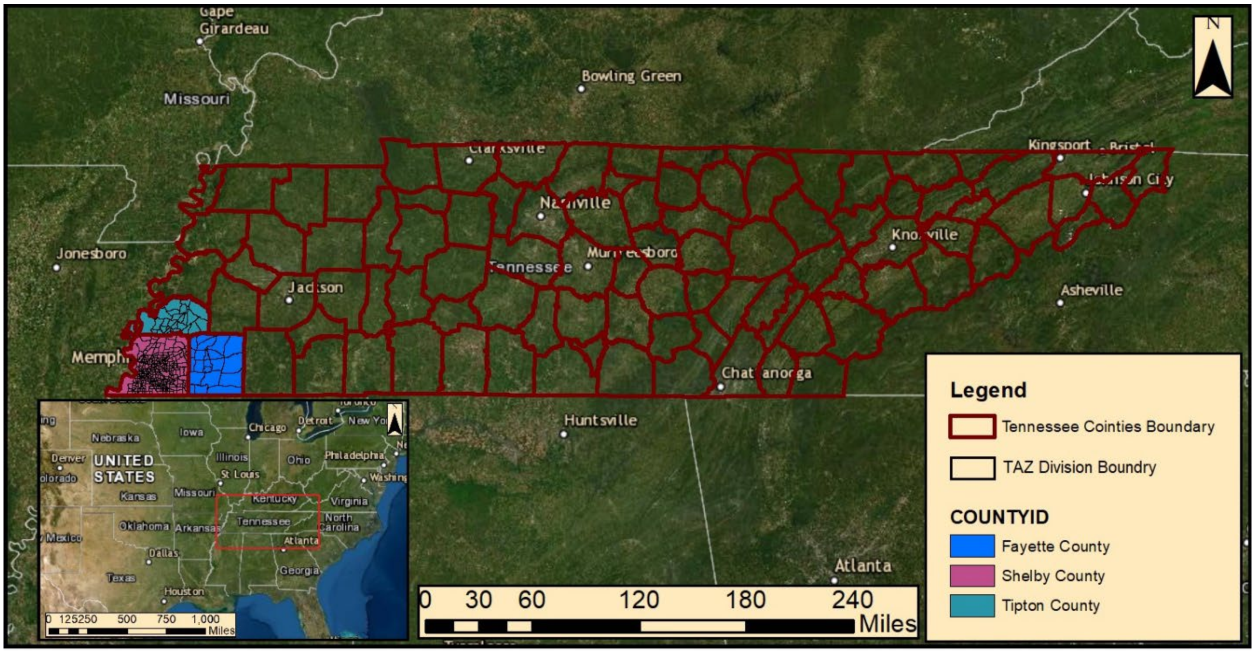

Figure 2-3. The Geographic area of the small sample problem (three-counties sample)

2.3.1 Developing the model for the base year 2010 (3-counties example)

To show the model performance and its application, the data requirement for running the land-use model is collected for Shelby, Tipton, and Fayette counties for two periods of time, where the year 2010 is the base year and the year 2005 is the lag year. In the following, first, the result of developing the model for the year 2010 is provided to show how the model is fitted. In this regard, the goodness-of-fit measure (R^2) and the error, Mean Absolute Percentage Error (*MAPE*) are provided. To illustrate the model performance in highly disaggregate conditions, the model was developed for 9 categories of households (total population, total households, households with 1 to 6 persons, and households with 7 or more persons) and 21 categories of employments (total employment and 20 NACIS section codes). In addition to households and employment, the proposed model can estimate land-use conditions in the study area. As Table 2-3 shows, the R^2 and *MAPE* values of the above-mentioned land use fields are provided for these three counties. As this table shows the proposed model fitted very well to the sample since the R^2 of all land use fields is greater than 0.66. Based on Chin's study (Chin, 1998) which proposed a rule of thumb for acceptable R^2, where R^2 greater than 0.66 is substantial, between 0.33 and 0.66 is moderate, and less than 0.33 is weak. Therefore, it is concluded that the model accuracy is acceptable. The interesting fact about the proposed model (LS-LUM) is that the goodness of the fit (R^2) and the error value (*MAPE*) for the land use conditions (residential, commercial, industrial, agricultural, and vacant land) demonstrate the high accuracy of the proposed model. The high accuracy of these land-use fields is important from another point of view. Since the output of LC (the amount of land in five land use classes) affects the accuracy of households and employments estimation indirectly, it is crucial to estimate land-use classes with high accuracy.

In addition to presenting the accuracy of the model in numbers, Figure 2-4 is provided to present the correlation plot of the estimated and observed 8 land use categories. Since the purpose of this section is to show the model applicability, only the correlation plots of 8 land use fields are provided in this chapter. Correlation plots are one the most common methods in showing the model validity. In this kind of plot, one of the axes represents the observations and another axis presents the estimated value. The more the plot shows a straight 45-degree line, the estimated values are closer to the observations. As Figure 2-4 shows, the correlation plots of estimated values through the proposed model (LS-LUM) are close to observations.

TABLE 2-3 THE R^2 AND THE MAPE VALUE OF DEVELOPING THE MODEL FOR THE YEAR 2010 (3-COUNTIES EXAMPLE)

Land use Filed	R^2	*MAPE (%)*
Total Population	0.878	14.13
Total Households	0.911	13.18
Households with 1 Person	0.956	16.58
Households with 2 Persons	0.921	13.87
Households with 3 Persons	0.868	15.67
Households with 4 Persons	0.867	18.32
Households with 5 Persons	0.834	20.46
Households with 6 Persons	0.848	22.19
Households with 7 or more Persons	0.899	23.91
Total Employment	0.978	35.57
Employment in NAICS 11	0.904	156.26
Employment in NAICS 21	0.679	326.79
Employment in NAICS 22	0.789	103.07
Employment in NAICS 23	0.918	213.85
Employment in NAICS 3133	0.940	97.76
Employment in NAICS 42	0.931	205.3
Employment in NAICS 4445	0.919	65.31
Employment in NAICS 4849	0.988	83.77
Employment in NAICS 51	0.716	271.27
Employment in NAICS 52	0.977	85.58
Employment in NAICS 53	0.852	115.88
Employment in NAICS 54	0.936	85.93
Employment in NAICS 55	0.839	302.9
Employment in NAICS 56	0.821	175.75
Employment in NAICS 61	0.977	160.6
Employment in NAICS 62	0.918	130.18
Employment in NAICS 71	0.885	133.82
Employment in NAICS 72	0.914	222.09
Employment in NAICS 81	0.934	62.68
Employment in NAICS 92	0.996	87.33
Residential Land	0.998	6.32
Commercial Land	0.961	15.99
Industrial Land	0.923	53.81
Agricultural Land	0.999	12.34
Vacant Land	0.926	25.64

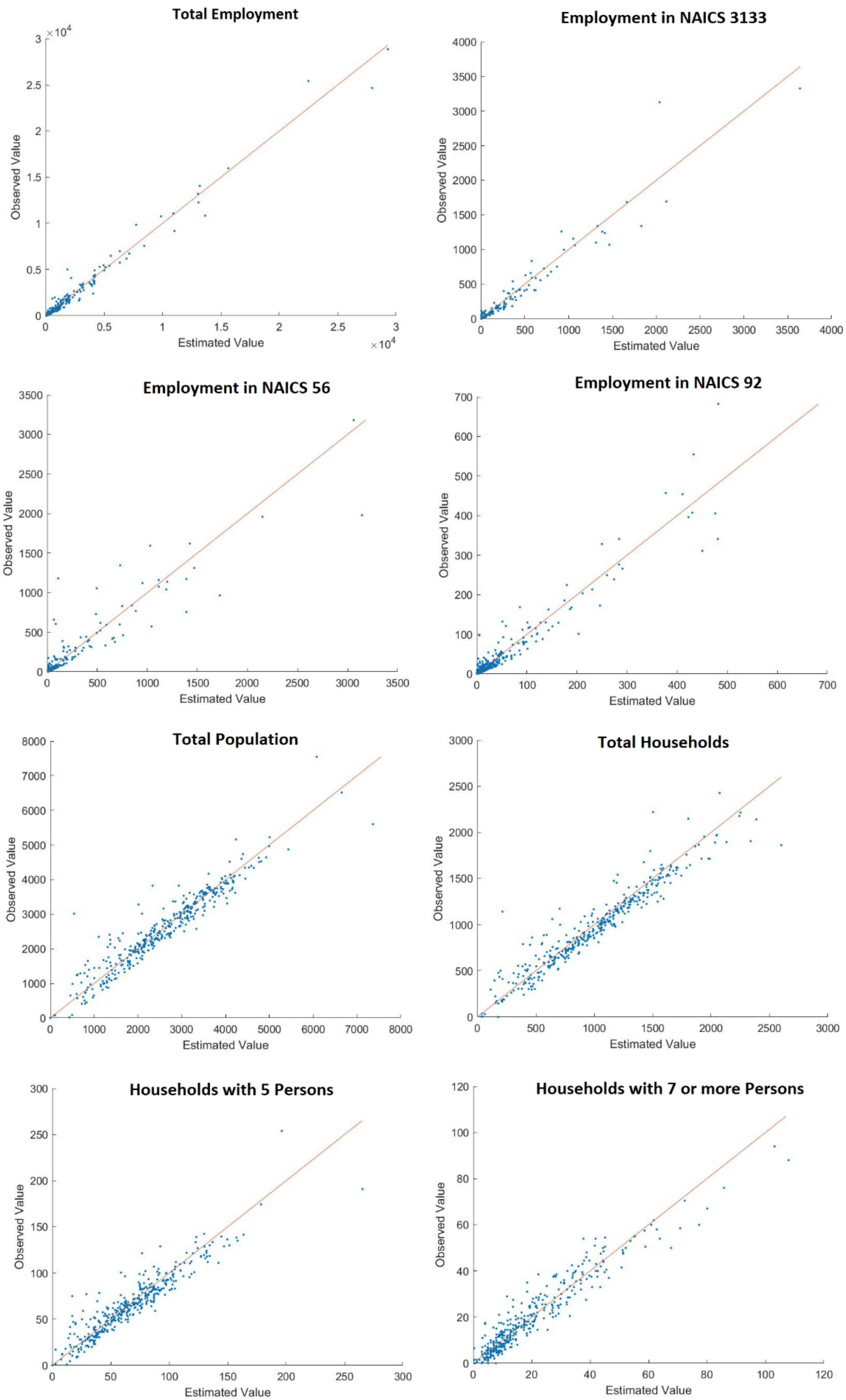

Figure 2-4. The correlation plot for 8 land use fields for the year 2010 (3-Counties example)

2.3.2 Backcasting validation for the 3-counties example

In addition to presenting the accuracy of developing the proposed model for the base year (2010), in this section, the model backcasting accuracy is provided for the 3-counties example. Similar to the previous section, the measure of the goodness of the fit (R^2) and the error measure (MAPE) are presented to show the model's accuracy in backcasting the demographic, socio-economic, the land use conditions in 3-counties example. Table 2.4 demonstrates the R^2 and the MAPE value of 9 household categories, 21 employment categories, and the land use condition in five categories. Results show the model backcasted the land use fields with acceptable accuracy (R^2 greater than 0.66). Moreover, similar to presenting the correlation plot for the developing model for 2010, in Figure 2-5, the correlation plots of 8 land use fields are provided. As this figure shows, backcasted results by the proposed model (LS-LUM) are close to the observed value for the year 2005 and show a close to 45-degree shape.

TABLE 2-4 THE R^2 AND THE MAPE VALUE OF BACKCASTING THE MODEL FOR THE YEAR 2005 (3-COUNTIES EXAMPLE)

Land use Filed	*R^2*	*MAPE (%)*
Total Population	0.878	14.13
Total Households	0.911	13.18
Households with 1 Person	0.956	16.58
Households with 2 Persons	0.921	13.87
Households with 3 Persons	0.868	15.67
Households with 4 Persons	0.867	18.32
Households with 5 Persons	0.834	20.46
Households with 6 Persons	0.848	22.19
Households with 7 or more Persons	0.899	23.91
Total Employment	0.978	35.57
Employment in NAICS 11	0.904	156.26
Employment in NAICS 21	0.679	326.79
Employment in NAICS 22	0.789	103.07
Employment in NAICS 23	0.918	213.85
Employment in NAICS 3133	0.940	97.76
Employment in NAICS 42	0.931	205.3
Employment in NAICS 4445	0.919	65.31
Employment in NAICS 4849	0.988	83.77
Employment in NAICS 51	0.716	271.27
Employment in NAICS 52	0.977	85.58
Employment in NAICS 53	0.852	115.88
Employment in NAICS 54	0.936	85.93
Employment in NAICS 55	0.839	302.9
Employment in NAICS 56	0.821	175.75
Employment in NAICS 61	0.977	160.6
Employment in NAICS 62	0.918	130.18
Employment in NAICS 71	0.885	133.82
Employment in NAICS 72	0.914	222.09
Employment in NAICS 81	0.934	62.68
Employment in NAICS 92	0.996	87.33
Residential Land	0.998	6.32
Commercial Land	0.961	15.99
Industrial Land	0.923	53.81
Agricultural Land	0.999	12.34
Vacant Land	0.926	25.64

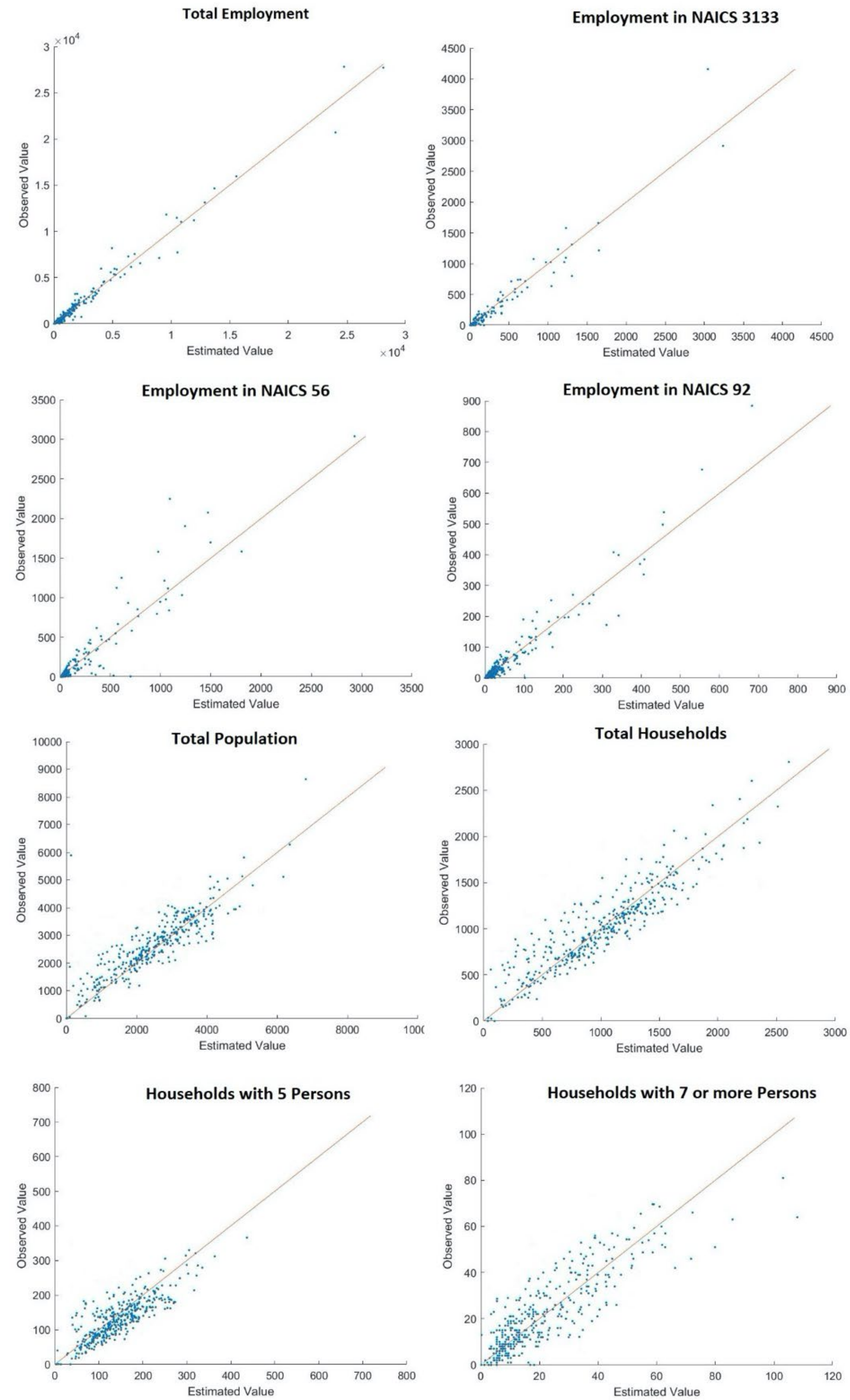

Figure 2-5. The correlation plot for 8 land use fields for the year 2005 (3-Counties sample)

Chapter 3 Data collection and reconciliation

Chapter 3 is dedicated to data collection. The data collection involves gathering various data sets such as census, Longitudinal Employment and Household Dynamics (LEHD), statewide parcel data, zoning plans and restriction for the base and future years, TN statewide travel demand model (TSM V3) along with spatial data such as traffic analysis zones, networks, skims, and the model set up. Running the proposed land use model necessitates the presence of data set for two periods of time (base year and lag year). For this project, total population, total households, the total number of houses, the number of occupied and vacant houses, total employment, and employment in 20 sections of NAICs for the years 2000 and 2010 are collected. In addition, the research team dedicated huge efforts to collecting parcel data for the 95 counties of the state of Tennessee. Generally, in this project data sets are collected from three sources:

I. Census data:
   Census data for both years 2000 and 2010 were collected. These data sets contain demographic information at different levels. The smallest geographic level in these data sets is the block level. Total population, total households, number of houses, house condition (vacant and occupied).
   Household income information is available at the block group level.
   Collection data from census 2000 need a specific procedure which is explained in Appendix A.
   Since the geographic division of the blocks and block group varies in 2010 and 2000, to use the data collected from these two data sources, the blocks and block group shapefile (boundaries) should be downloaded from Tiger-line's sources which are available using the following link: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html.
II. Longitudinal Employment and Household Dynamics (LEHD):
   Information regarding the number of employments can be collected from 2002 to 2018 for all 20 categories of NAICS. Because the data for the year 2000 was needed, the data for this section was generated applying interpolation.
III. Tennessee Comptroller of Treasury:
   The parcel data for 84 counties of the state of Tennessee is available through the Tennessee Comptroller of Treasury website. The parcel data is available for the year 2018-2020. The amount of land in each TAZ in five land use classes, residential, commercial, industrial, agricultural, and vacant (developable) land, extracted from this source. This data set was collected for all counties and aggregated to the TAZ level.

3.1 Data Collected

In this section, a part of demographic and socio-economic data collected for both years 2000 and 2010 are presented in Figures 3-1 to 3-9. These data sets are presented at the TAZ level.

Total Population

Figure 3-1. Total Population in the state of Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Total Households

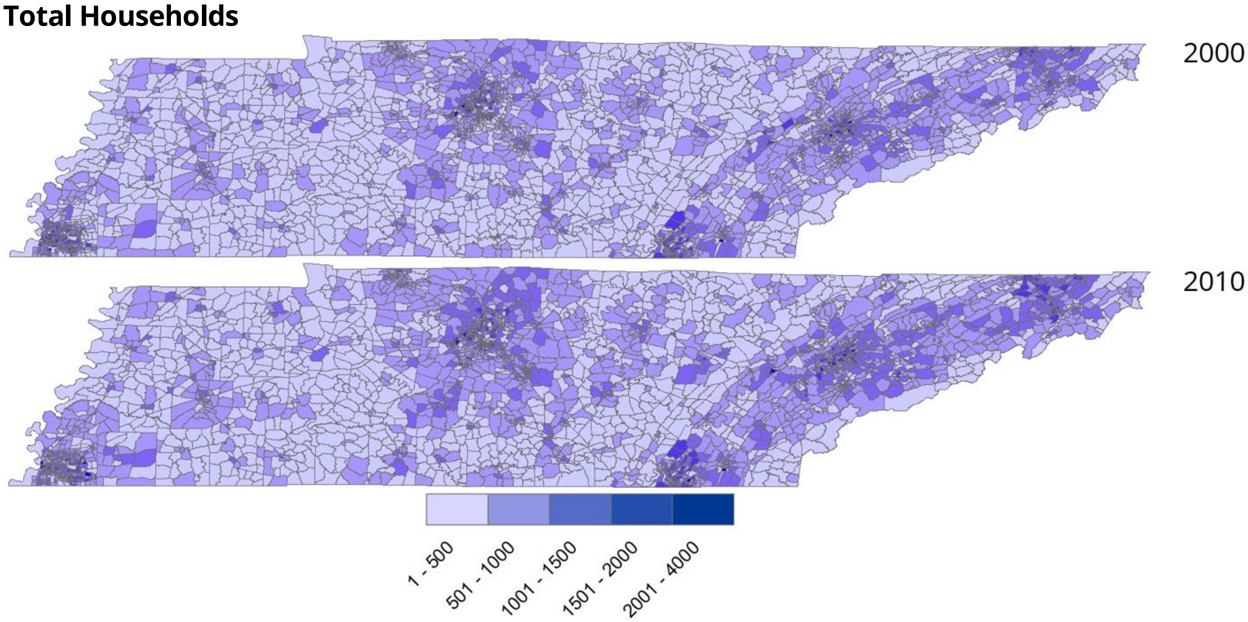

Figure 3-2. Total Households in the state of Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Total Number of Houses

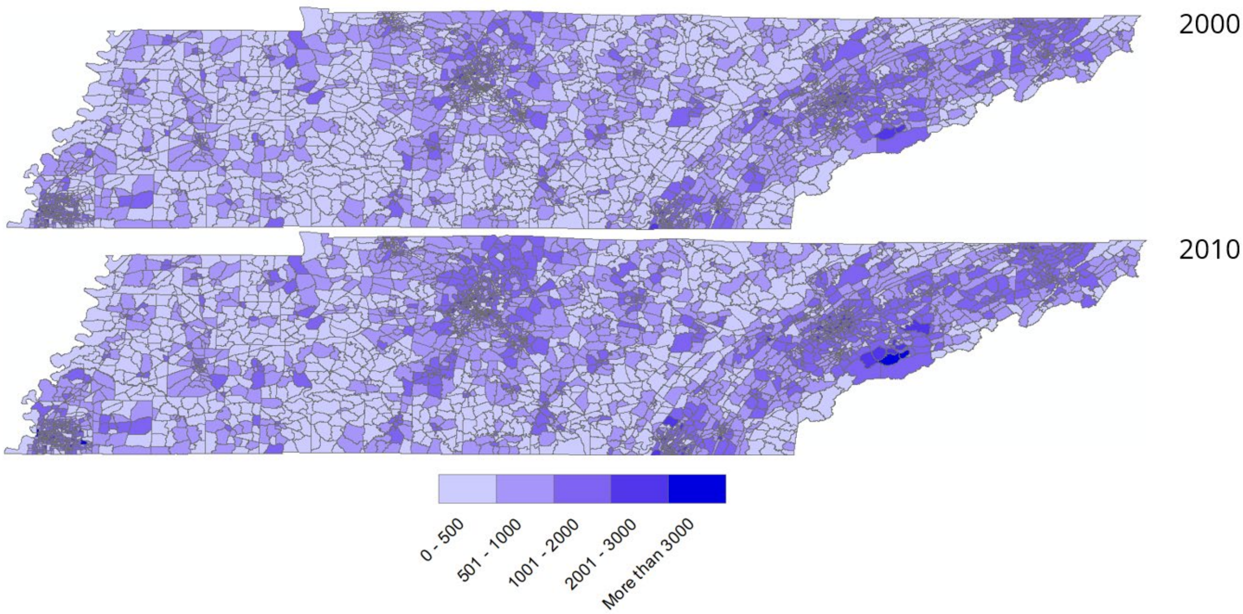

Figure 3-3. Total Houses in the state of Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Number of Occupied Houses

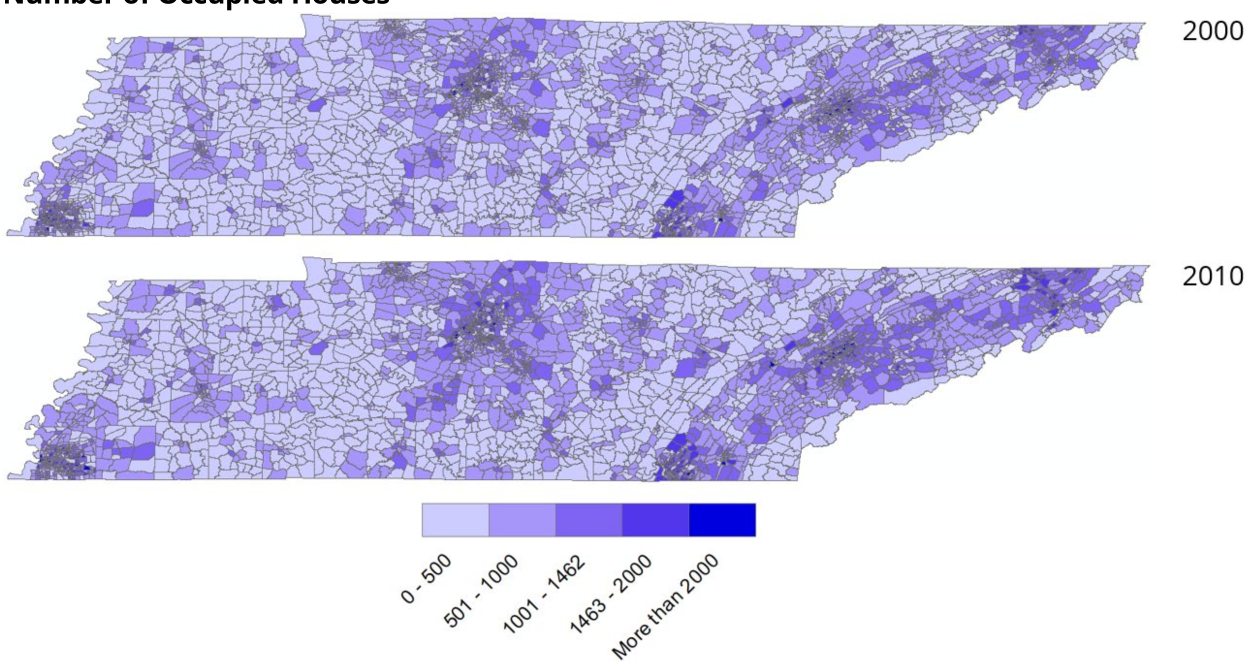

Figure 3-4. Number of Occupied Houses in the state of Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Number of Vacant Houses

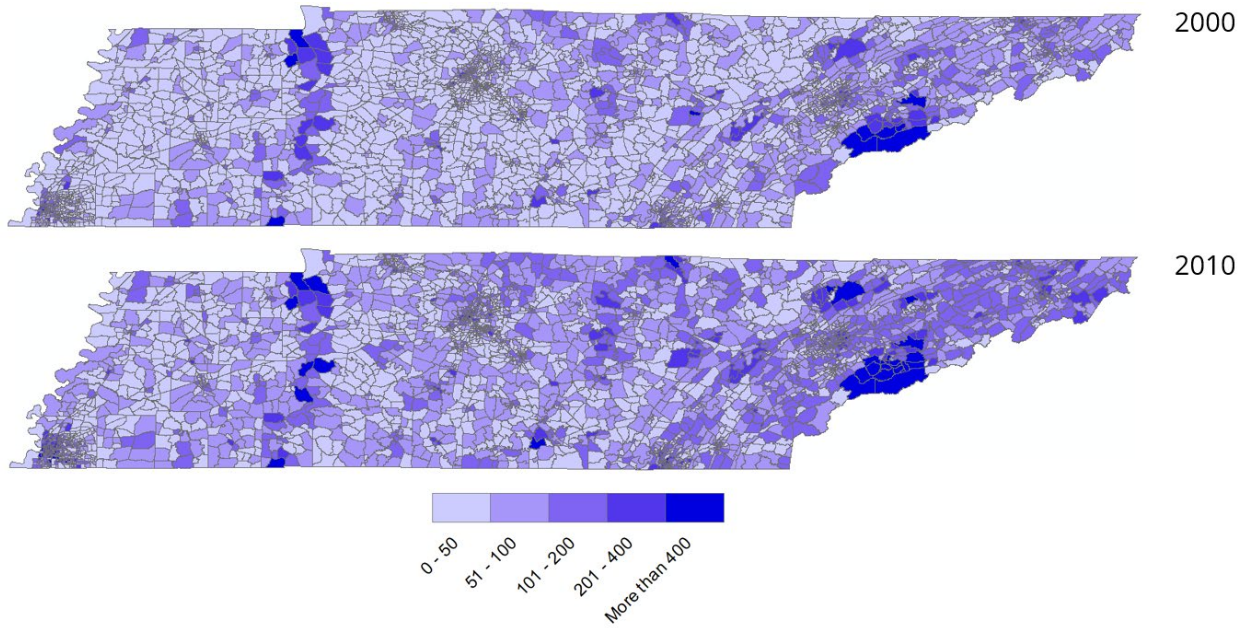

Figure 3-5. Number of Vacant Houses in Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Total Employments

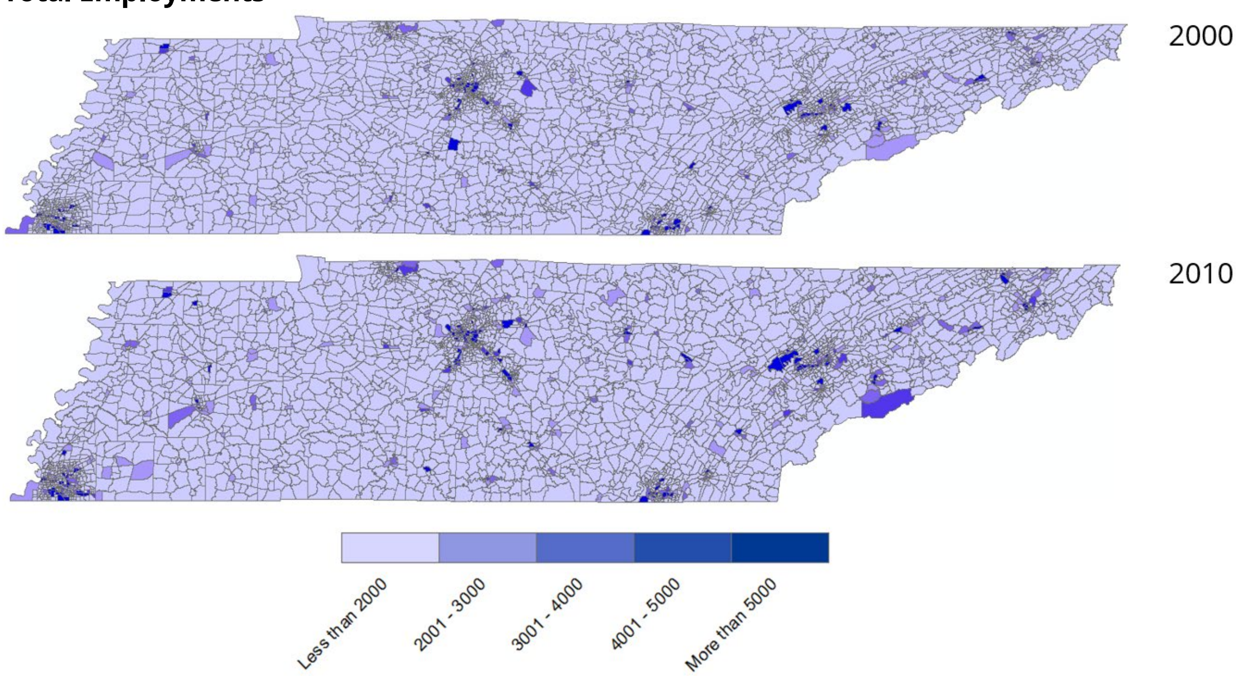

Figure 3-6. Total Employment in Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Employment in NAICS 48-49

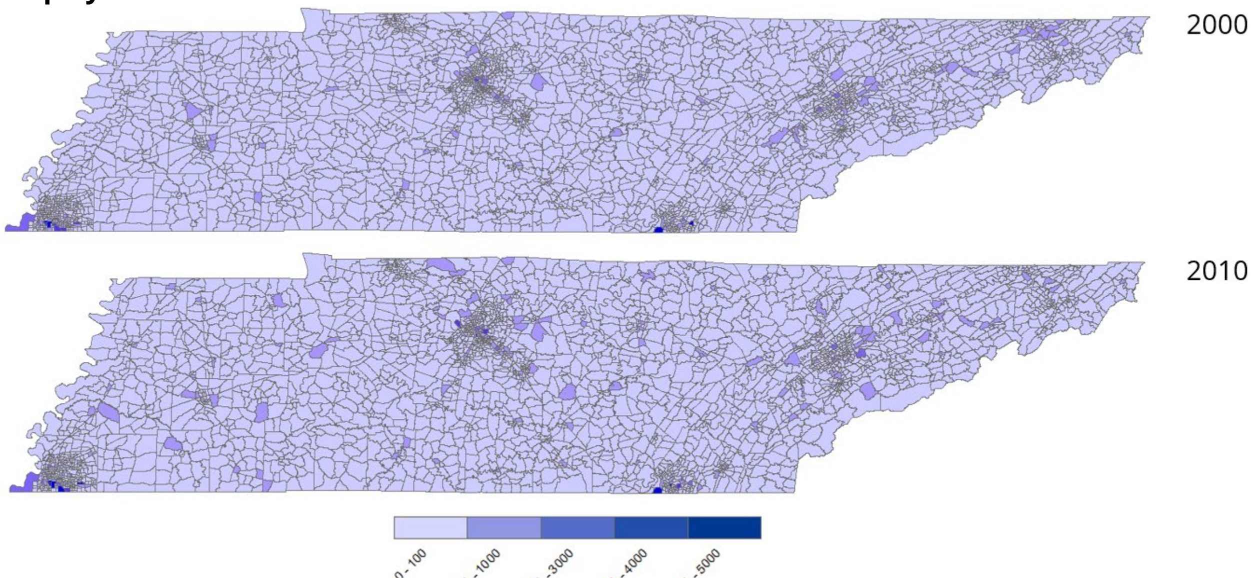

Figure 3-7. Employment in NAICS 48-49 in the state of Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Employment in NAICS 62

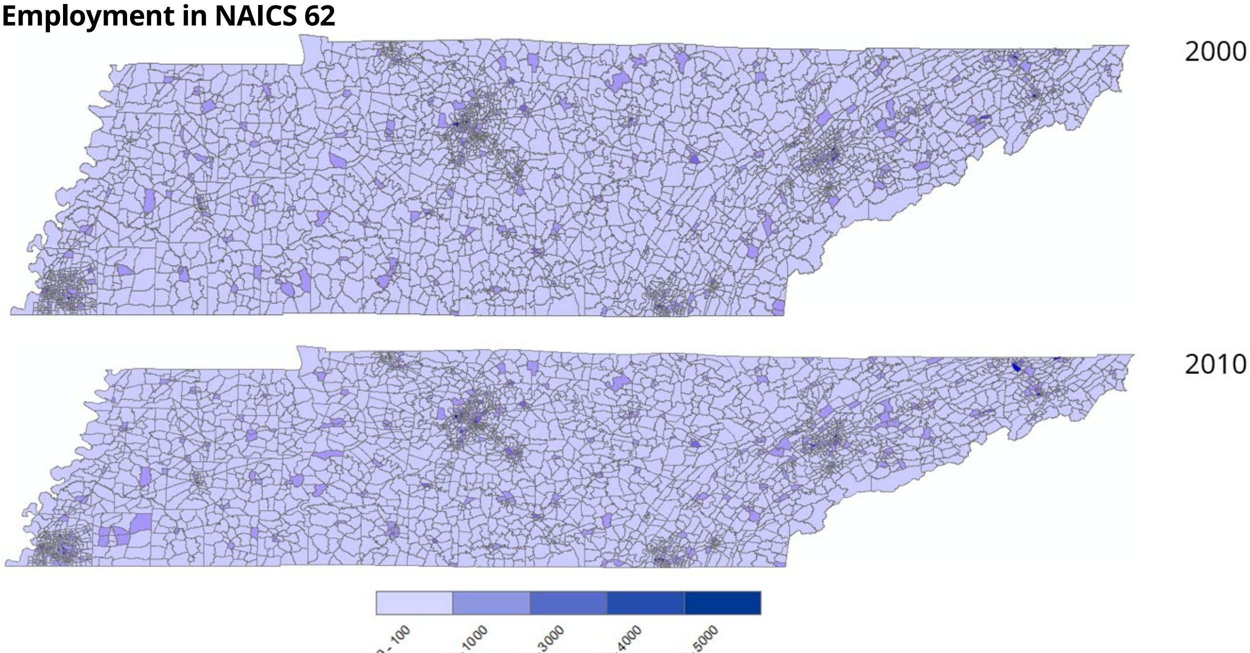

Figure 3-8. Employment in NAICS 62 in Tennessee in TAZ level for Base and Lag year (2010 and 2000)

Employment in NAICS 92

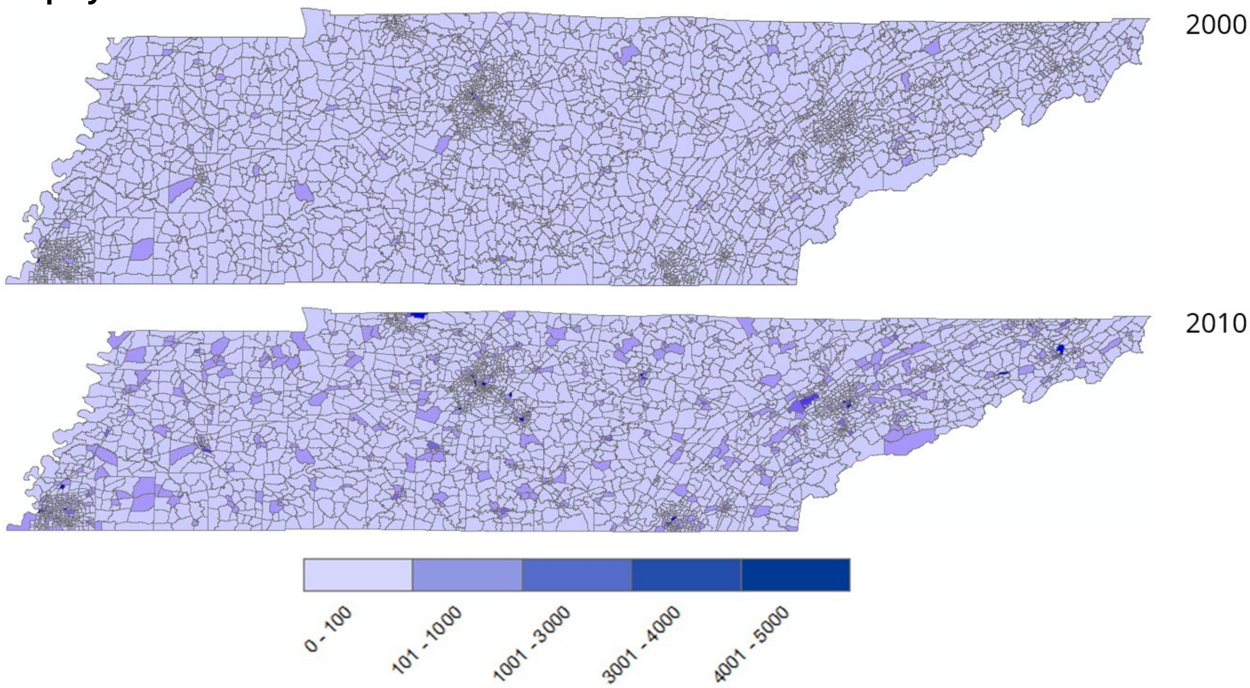

Figure 3-9. Employment in NAICS 92 in Tennessee in TAZ level for Base and Lag year (2010 and 2000)

3.2 Parcel Data

Land parcel databases, which are also known as cadasters, describe the rights, interests, and value of the property. The legal boundaries of land parcels are defined in the deed to a property. A surveyor confirms these measurements anytime the property is subdivided or platted, or in boundary disputes. Ownership of land parcels is an important part of the financial, legal, and real estate systems of a society. Real estate tax parcels are typically graphic representations of land ownership to support property taxing functions. These maps are often used as parcel maps for a jurisdiction. The aggregate set of land parcels represents the distribution of the real property assets of a community and its ownership, forms the basis for all land use and zoning decisions, and represents the location of residences, businesses, and public lands. In other words, almost every aspect of government and business can be associated with a land parcel (Council, 2007). Availability of parcel data gives the chance of reviewing different land-use models and can increase the accuracy of the model presented in chapter 2.

Although many land parcel data exist in the United States, they are not entirely in digital form, they are not in a common format, and they are certainly not consistently available across the nation. In the state of Tennessee, parcel data is available for 84 counties. Data for 11 counties (Bradley, Chester, Davidson, Hamilton, Hickman, Knox, Montgomery, Rutherford, Shelby, Sumner, and Williamson) was not available through the Tennessee Comptroller website. The research team collected the parcel data for these 11 counties by contacting each county. An example of parcel data is presented in Figures 3-10. The first task was to aggregate the parcel data to the TAZ level for five land use classes. This task was conducted by using the land-use codes attached to each parcel which indicate the land use of each parcel (e.g., residential, commercial, public, utility). Moreover, parcel data was available for the year 2018-2020 (varies between counties); while the data was needed for the years 2010 and 2005. Therefore, the parcel data extended for different years by using the built year of each parcel.

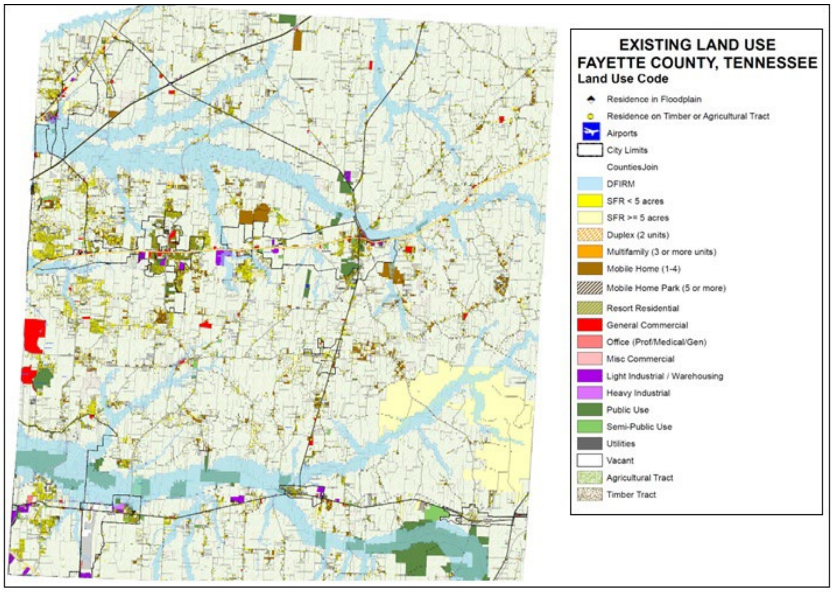

Figure 3-10. Fayette County Parcel data

3.3 Transportation Network

The transportation network was collected from the Tennessee Statewide Travel Demand model to calculate the travel time (travel cost) between each paired TAZ (see Fig 3-11).

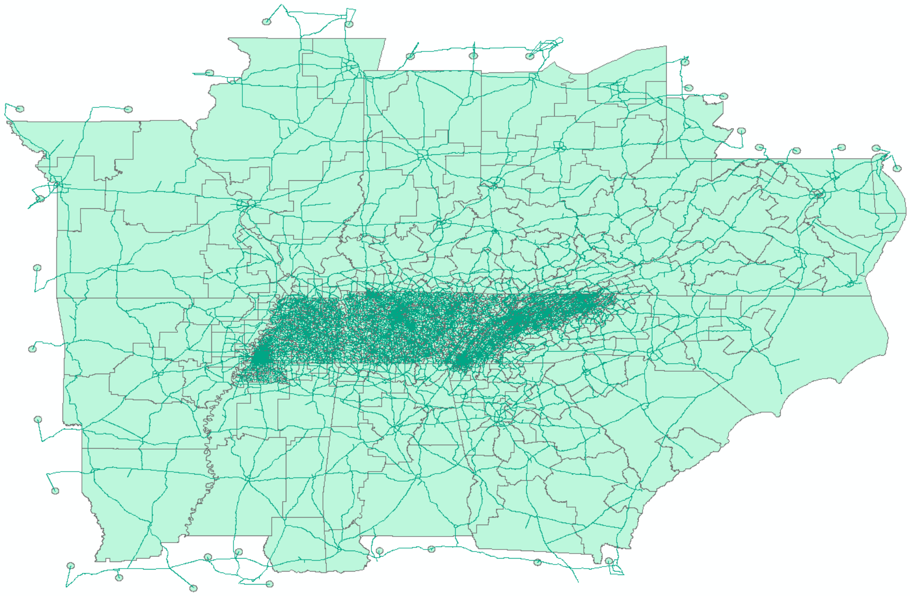

Figure 3-11. Tennessee statewide Transportation network

Chapter 4 Development of Statewide Land Use Model

In this chapter of this report, the procedure and result of applying the proposed model on 95 counties of the state of Tennessee are provided. The study region contains 3293 TAZs (Figure 4-1). Three metropolitans are located in this state (Memphis in the west, Nashville in the center, and Knoxville in the east). The state of Tennessee hosts a population of 4,781,279 and total employments of 1,924,1238 in 2010. The proposed model needs the presence of data sets for 2 periods of time. We suggest considering the time interval of five years between each period. The model is developed for the year 2010 (considering the lag year 2005).

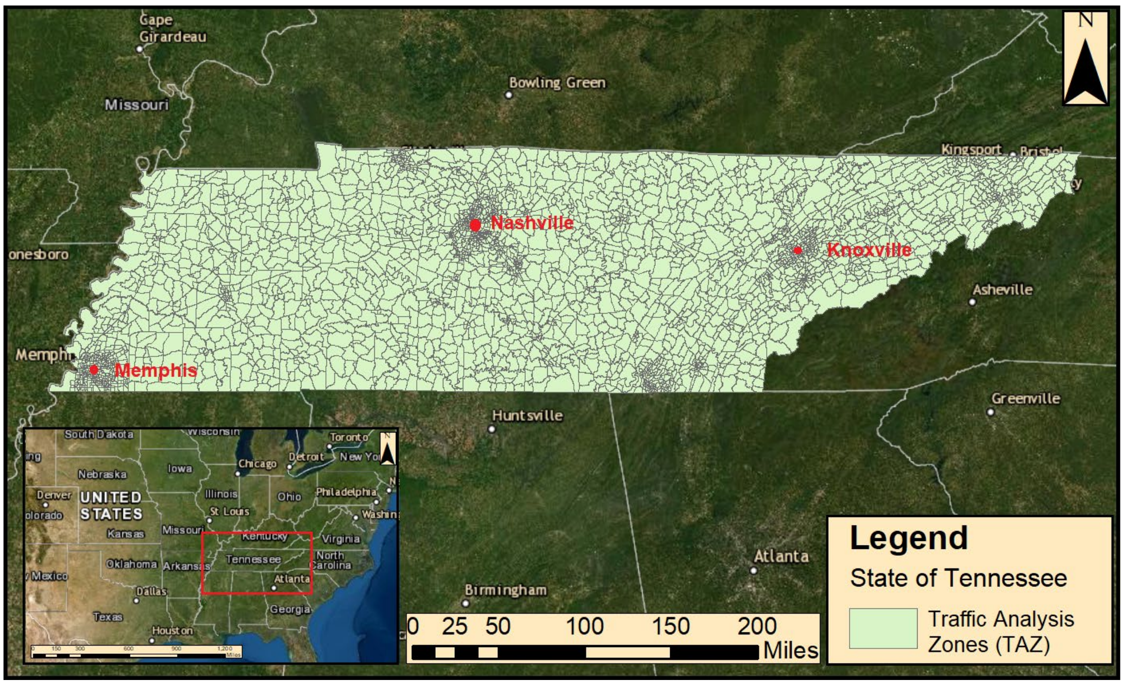

Figure 4-1. The State of Tennessee with 3293 TAZs

To develop the model for the base year 2010, the data for the year 2005 is required. Data set related to employment were collected from LEHD sources. However, because census data is released every 10 years, population, households, and house condition data sets were generated by interpolating between data collected for 2010 and 2000. In the following sections of this chapter, the result of developing the model on 95 counties of the state of Tennessee is provided. For doing so, the goodness of the fit (R^2), the Percentage of Good Prediction (PGP), and the value of Mean Absolute Error Percentage (MAPE) are provided to show how the proposed land use model (LS-LUM) is fitted to the study area. Since MAPE has limits, especially in the case of too many zeros, PGP is added to show the error and accuracy of the proposed model. PGP is an index to measure the goodness of prediction, and is calculated using the following formulation:

$$PGP = 1 - \frac{1}{2}\frac{\sum_{k=1}^{p}|O_k - E_k|}{\sum_{k=1}^{p} O_k} \tag{14}$$

PGP varies from 0, when no agreement is found, to 1, when the observation and estimation are identical. Similar to R^2, when PGP greater than 0.66 shows an acceptable prediction.

Table 4-1 shows the R^2, PGP, and the MAPE value of 9 household categories (total population, total households, households with 1 to 6 persons, and households with 7 or more persons), employments in 21 categories (total employments, employment in 20 NAICS sections), and land use condition for five land use classes. As this table shows, the proposed land use model is fitted very well to the study area. The model shows high accuracy in estimating the households, where the R^2 of all categories is greater than 0.85. Moreover, the error measure (MAPE) for the household section shows a very small value, which means that the results are completely reliable. In addition to households, the model shows high accuracy in estimated land-use conditions. The accuracy of the proposed model is acceptable in the employment section since the R^2 is greater than 0.66. The error measure in this section shows that the accuracy of the model in this section is lower than the households and land use condition section. In addition to R^2 and MAPE, the PGP for each category show an appropriate value. PGP is greater than 0.66 in all sections and categories.

TABLE 4-1 THE R2, PGP, AND THE MAPE (%) VALUE OF DEVELOPING THE LAND USE MODEL FOR THE YEAR 2010 (95 COUNTIES)

Land use Filed	***R^2***	***MAPE (%)***	***PGP***
Total Population	0.957	12.98	0.967
Total Households	0.965	12.52	0.967
Households with 1 Person	0.968	14.82	0.958
Households with 2 Persons	0.955	12.32	0.96
Households with 3 Persons	0.947	12.59	0.961
Households with 4 Persons	0.933	16.22	0.951
Households with 5 Persons	0.915	21.15	0.936
Households with 6 Persons	0.893	30.18	0.901
Households with 7 or more Persons	0.903	38.42	0.857
Total Employment	0.939	67.69	0.901
Employment in NAICS 11	0.925	106.52	0.802
Employment in NAICS 21	0.844	111.89	0.743
Employment in NAICS 22	0.885	130.6	0.871
Employment in NAICS 23	0.879	112.65	0.841
Employment in NAICS 3133	0.767	222.35	0.763
Employment in NAICS 42	0.898	162.24	0.824
Employment in NAICS 4445	0.907	172.12	0.872
Employment in NAICS 4849	0.967	51.78	0.762
Employment in NAICS 51	0.827	94.52	0.725
Employment in NAICS 52	0.967	87.84	0.862
Employment in NAICS 53	0.868	90.08	0.821
Employment in NAICS 54	0.830	82.3	0.818
Employment in NAICS 55	0.664	363.57	0.679
Employment in NAICS 56	0.794	58.76	0.736
Employment in NAICS 61	0.908	280.83	0.872
Employment in NAICS 62	0.831	70.24	0.848
Employment in NAICS 71	0.938	93.95	0.814
Employment in NAICS 72	0.935	271.6	0.867
Employment in NAICS 81	0.866	78.24	0.749
Employment in NAICS 92	0.973	162.35	0.688
Residential Land	0.997	3.32	0.993
Commercial Land	0.969	11.17	0.957
Industrial Land	0.818	66.12	0.926
Agricultural Land	0.999	19.28	0.994
Vacant Land	0.906	25.64	0.896

R^2 and PGP greater than 0.66 and MAPE less than 200 are considered as acceptable accuracy.

In addition to Table 7, a combination of histograms and correlation plots are provided for presenting the backcasting and forecasting validation of the proposed model (LS-LUM). Correlation plots are a common method in presenting validation results in land-use studies. However, in the case of large-scale problems, due to a large number of zones, this method cannot give detailed information regarding the strengths and weaknesses of the models. In this study, correlation plots are combined with the histogram of data distribution, both for observation and estimated values. This approach provides a better insight into the model's accuracy by showing the differences between observed and estimated values at different intervals (buckets).

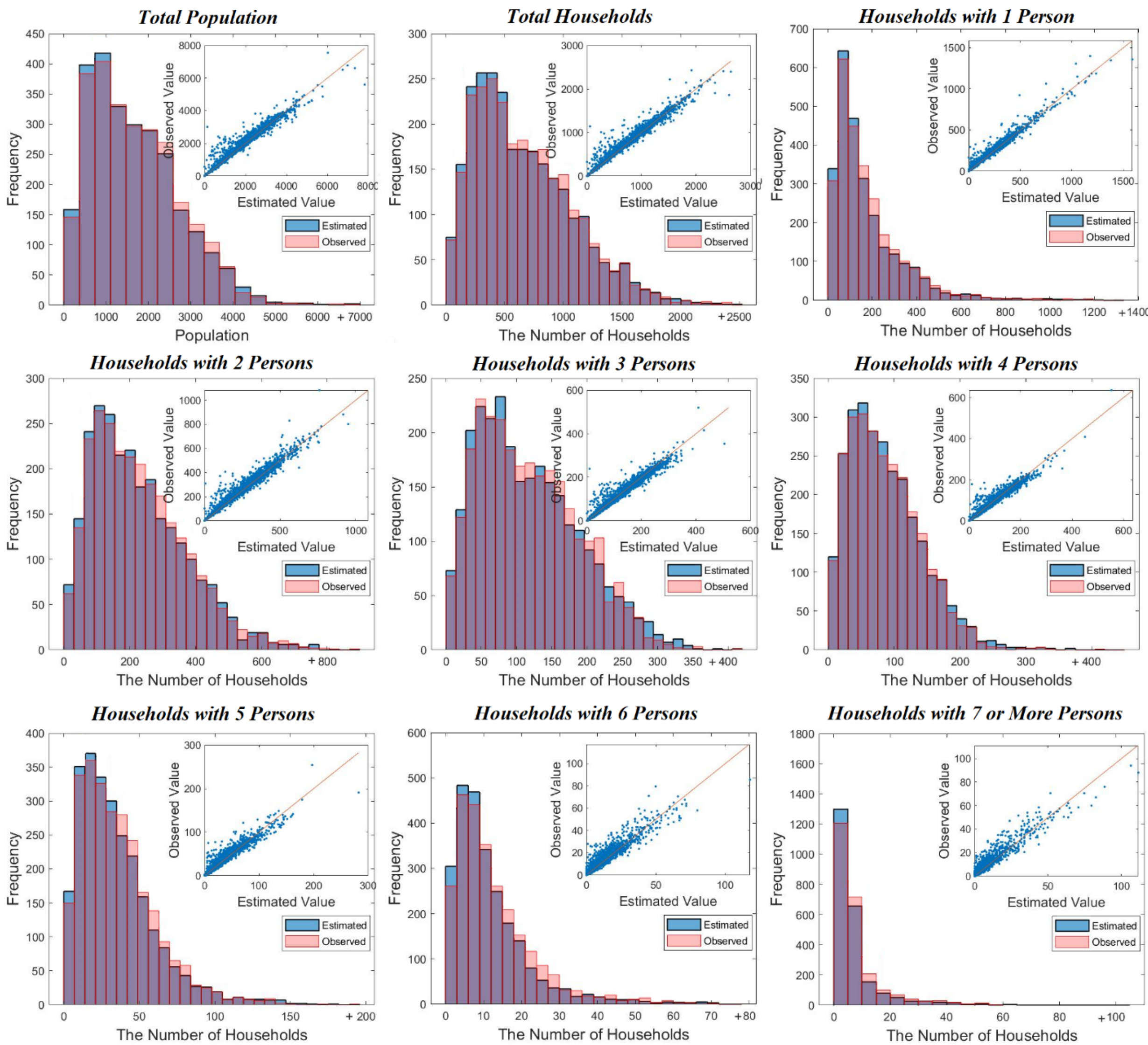

Figure 4-2. Histogram of estimated and observed households by LS-LUM in the year 2010. The inside plot is the correlation plot of observed and estimated households

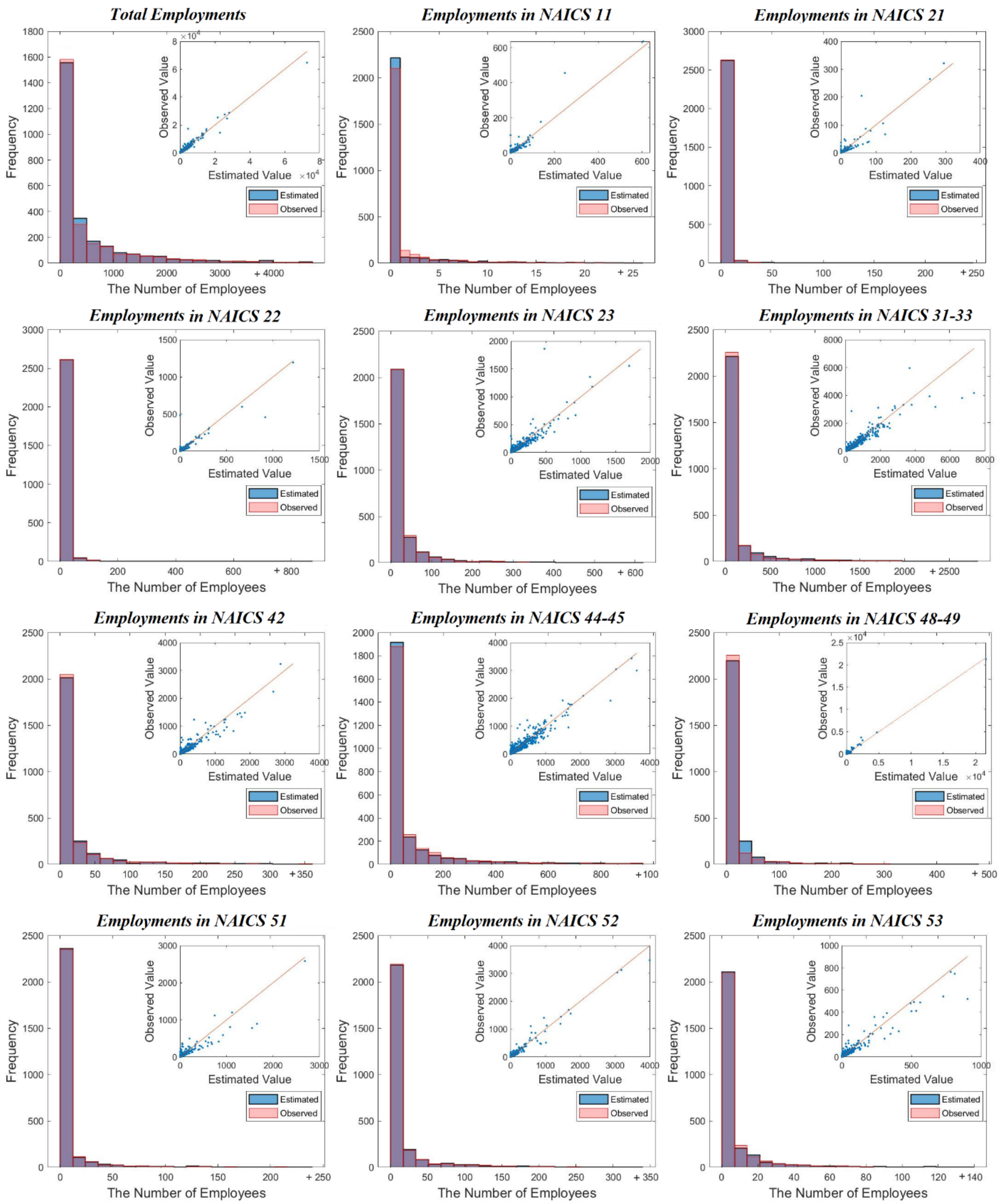

Figure 4-3. Histogram of estimated and observed households by LS-LUM in the year 2005. The inside plot is the correlation plot of observed and estimated employment (Part 1)

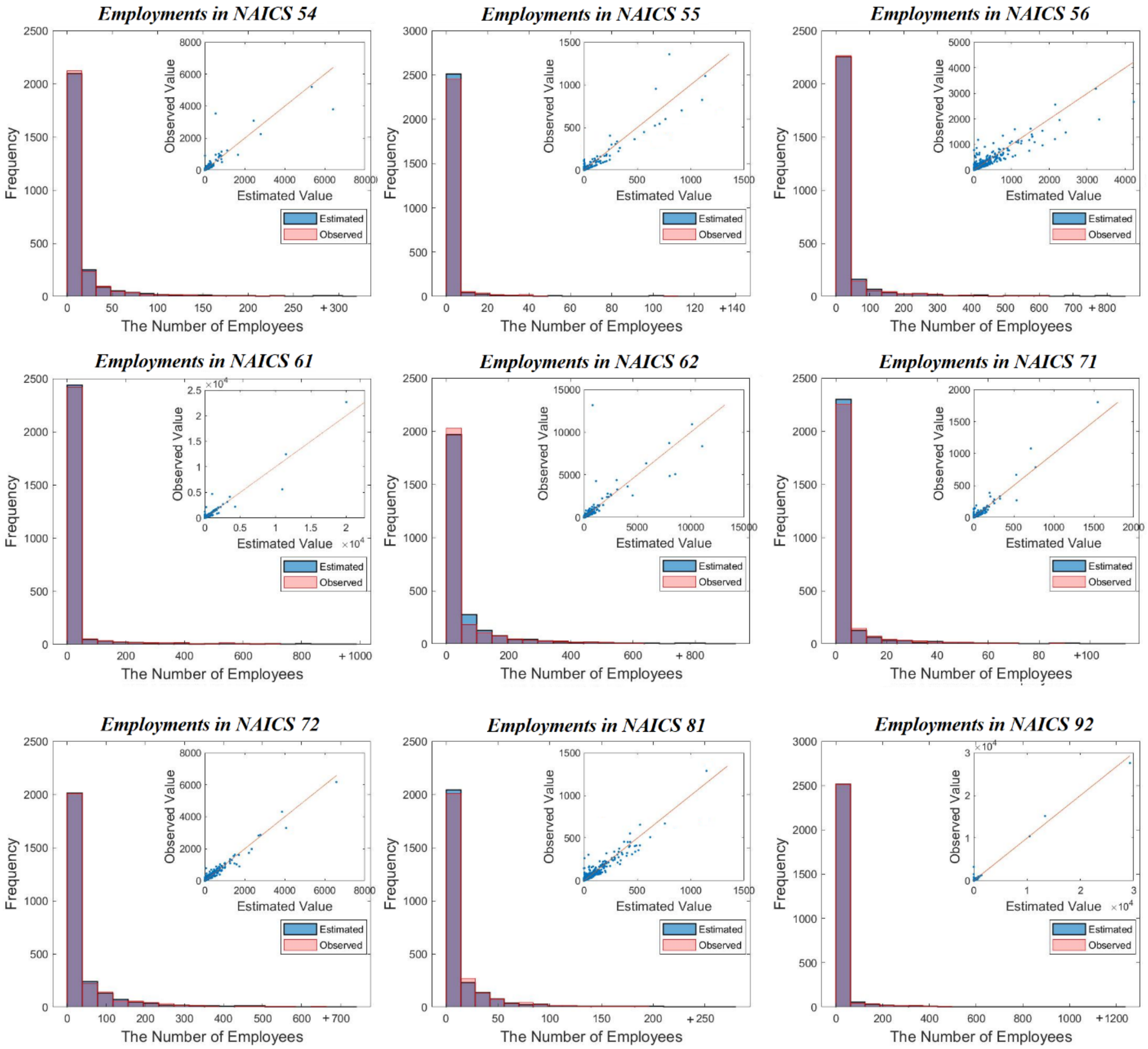

Figure 4-4. Histogram of estimated and observed employment by LS-LUM in the year 2010. The inside plot is the correlation plot of observed and estimated employment (Part 2)

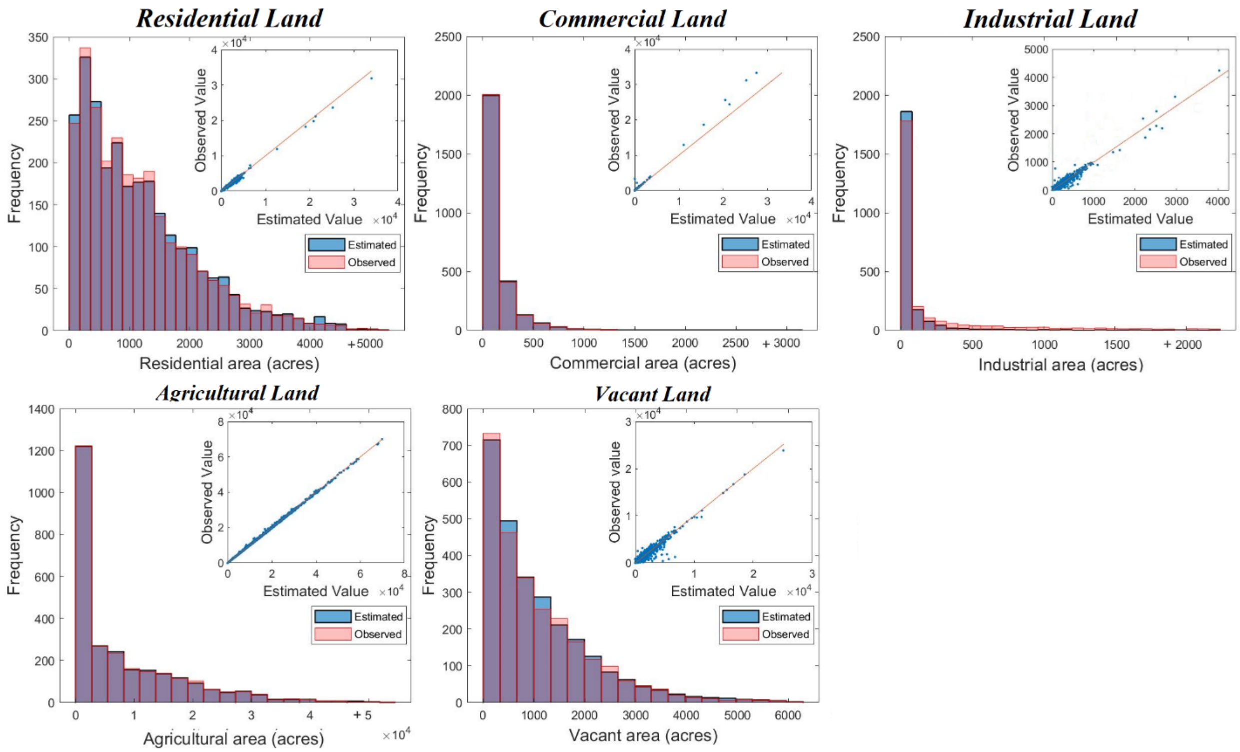

Figure 4-5. Histogram of estimated and observed land-use condition by LS-LUM in five land-use classes in the year 2010. The inside plot is the correlation plot of observed and estimated land-use conditions

Chapter 5 Land Use Model Validation

In this chapter, the validation of the proposed land use model (LS-LUM) is provided. The land-use model validation includes assessing the accuracy of the output produced when compared with observed data. This section of this report discusses the result of the model's validation and accuracy. To illustrate the model applicability and validity, the proposed land use model is implemented in the state of Tennessee, the United States, and the backcasting forecasting results are presented. In this regard, the proposed model is developed (calibrated) for the base year (2010), then backcasted for the year 2005 and forecasted the demographic and socioeconomic for the year 2015. At each step, the goodness-of-fit measure (R^2), Percentage of Good Prediction (PGP), and Mean Absolute Percentage Error (MAPE) are provided. Moreover, in each section, visual results are provided to illustrate the difference between the model and observed data. Generally, a model with a higher R^2 and PGP and smaller MAPE is a better model. Based on Chin's study (Chin, 1998) which proposed a rule of thumb for acceptable R^2, where R^2 greater than 0.66 is substantial, between 0.33 and 0.66 is moderate, and less than 0.33 is weak. In this study, R^2 greater than 0.66 is considered acceptable. Similar to R^2, the same criterion is applied for PGP. Therefore, in this report, PGP greater than 0.66 is considered substantial. Correlation plots are a common method in presenting validation results in land-use studies. In this method of presenting the accuracy of the results, one of the axes shows the observed data and the other axes show the estimated value. The extent to which the points illustrated in this 2-dimensional plot shape a 45-degree line, the model accuracy is higher. However, in the case of large-scale problems, due to the large number of zones, this method cannot give detailed information regarding the strengths and weaknesses of the models. In this research, correlation plots are combined with the histogram of data distribution, both for observation and estimated values. This approach provides a better insight into the model's accuracy by showing the differences between observed and estimated values at different intervals (buckets).

This chapter is divided into two sections. In the first section, the backcasting validation results of the proposed model are provided. In the second section, the forecasting validation results of the proposed model are provided.

5.1 Backcasting Validation

After developing the models for the base year 2010, the proposed model backcasted the households, employment, and land-use condition in the year 2005. Table 5.1 shows backcasting validation results for all categories. As this table shows the proposed model could estimate the households and employment categories with acceptable accuracy. The model accuracy in the household section is very well. All the households are estimated where R^2 for all categories is greater than 0.80 and the MAPE error shows a very small percentage in all categories. The model accuracy in estimating the land use classes is very high. The goodness of fit of these sections is greater than 0.90, which showed that the model can estimate these sections very well. In addition to Table 5-1, to present the model validity, correlation plots of the estimated and observed data are illustrated for the backcasting results in Figures 5-1 to 5-4. As mentioned before, in these figures the histogram of the estimated and observed data are provided to give a better insight into the accuracy of the model in estimating related categories at different intervals.

TABLE 5-1 THE R^2, MAPE, AND PERCENTAGE OF GOOD PREDICTION (PGP) OF BACKCASTING FOR THE YEAR 2005

Land use Filed	R^2	***MAPE (%)***	***PGP***
Total Population	0.95	7.87	0.969
Total Households	0.956	9.26	0.97
Households with 1 Person	0.967	14.17	0.955
Households with 2 Persons	0.951	14.45	0.962
Households with 3 Persons	0.938	13.88	0.963
Households with 4 Persons	0.927	11.23	0.955
Households with 5 Persons	0.905	17.55	0.934
Households with 6 Persons	0.832	27.21	0.913
Households with 7 or more Persons	0.905	29.42	0.863
Total Employment	0.956	199.9	0.904
Employment in NAICS 11	0.893	72.73	0.771
Employment in NAICS 21	0.848	77.01	0.709
Employment in NAICS 22	0.839	83.36	0.862
Employment in NAICS 23	0.831	54.52	0.844
Employment in NAICS 3133	0.834	180.47	0.827
Employment in NAICS 42	0.893	92.18	0.827
Employment in NAICS 4445	0.915	118.88	0.866
Employment in NAICS 4849	0.987	654.8	0.76
Employment in NAICS 51	0.816	49.95	0.741
Employment in NAICS 52	0.959	100.21	0.866
Employment in NAICS 53	0.878	66.6	0.825
Employment in NAICS 54	0.77	82.02	0.821
Employment in NAICS 55	0.644	50.28	0.681
Employment in NAICS 56	0.799	198.47	0.737
Employment in NAICS 61	0.910	74.95	0.869
Employment in NAICS 62	0.899	112.78	0.796
Employment in NAICS 71	0.921	50.64	0.798
Employment in NAICS 72	0.953	206.88	0.856
Employment in NAICS 81	0.971	58.62	0.844
Employment in NAICS 92	0.842	80.51	0.641
Residential Land	0.997	3.32	0.974
Commercial Land	0.974	13.53	0.962
Industrial Land	0.910	75.28	0.83
Agricultural Land	0.998	21.27	0.999
Vacant Land	0.933	29.31	0.926

The R^2 and PGP greater than 0.66 and MAPE less than 200 are considered as acceptable accuracy.

In Figure 5-1, the histogram and the correlation plot of the estimated and observed households for the year 2005 are presented. The correlation plots in all household categories show a close to the 45-degree line, which means that the model is estimating the households in 2005 accurately. Moreover, the histogram provided in each figure can show that the differences between estimated and observed data at different intervals are acceptable.

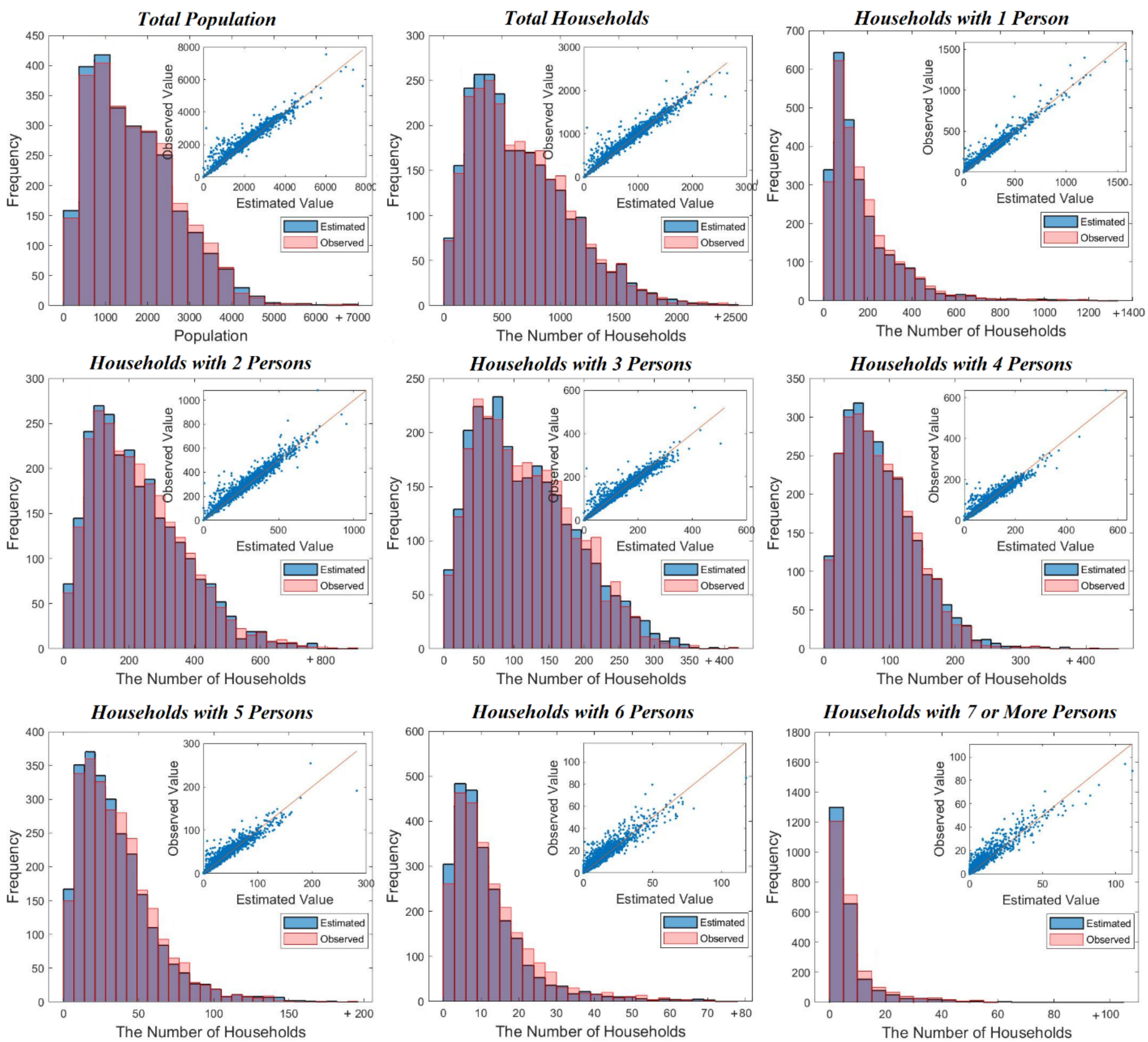

Figure 5-1. Histogram of estimated and observed households by the proposed model in the year 2005. The inside plot is the correlation plot of observed and estimated households.

Similar to the household section, the histogram and correlation results are provided for employment categories in 2005, which are presented in Figures 5-2 and 5-3 (to show the results higher resolution, the plot related to the employment section are provided in two separate figures). As these figures show, in the majority of the employment sections (NAICS) the correlation plot shows a line close to a 45-degree line. Besides, the histogram plots show that the distribution of the employment between TAZs is close to zero and the majority of the TAZ has few employments and the difference between observation and estimated data is more significant in the first and second intervals in the employment section.

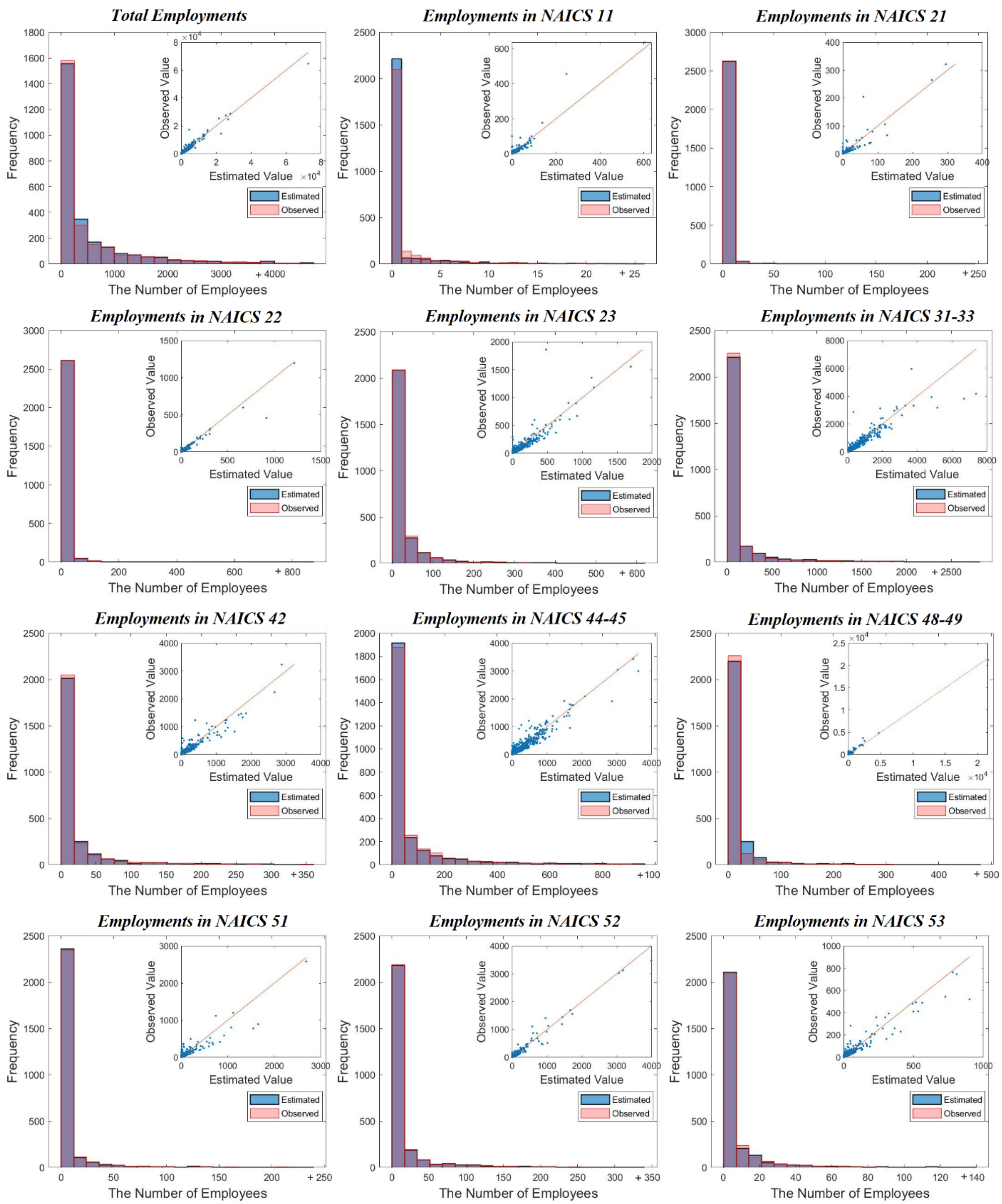

Figure 5-2. Histogram of estimated and observed households by the proposed model in the year 2005. The inside plot is the correlation plot of observed and estimated employment (Part 1).

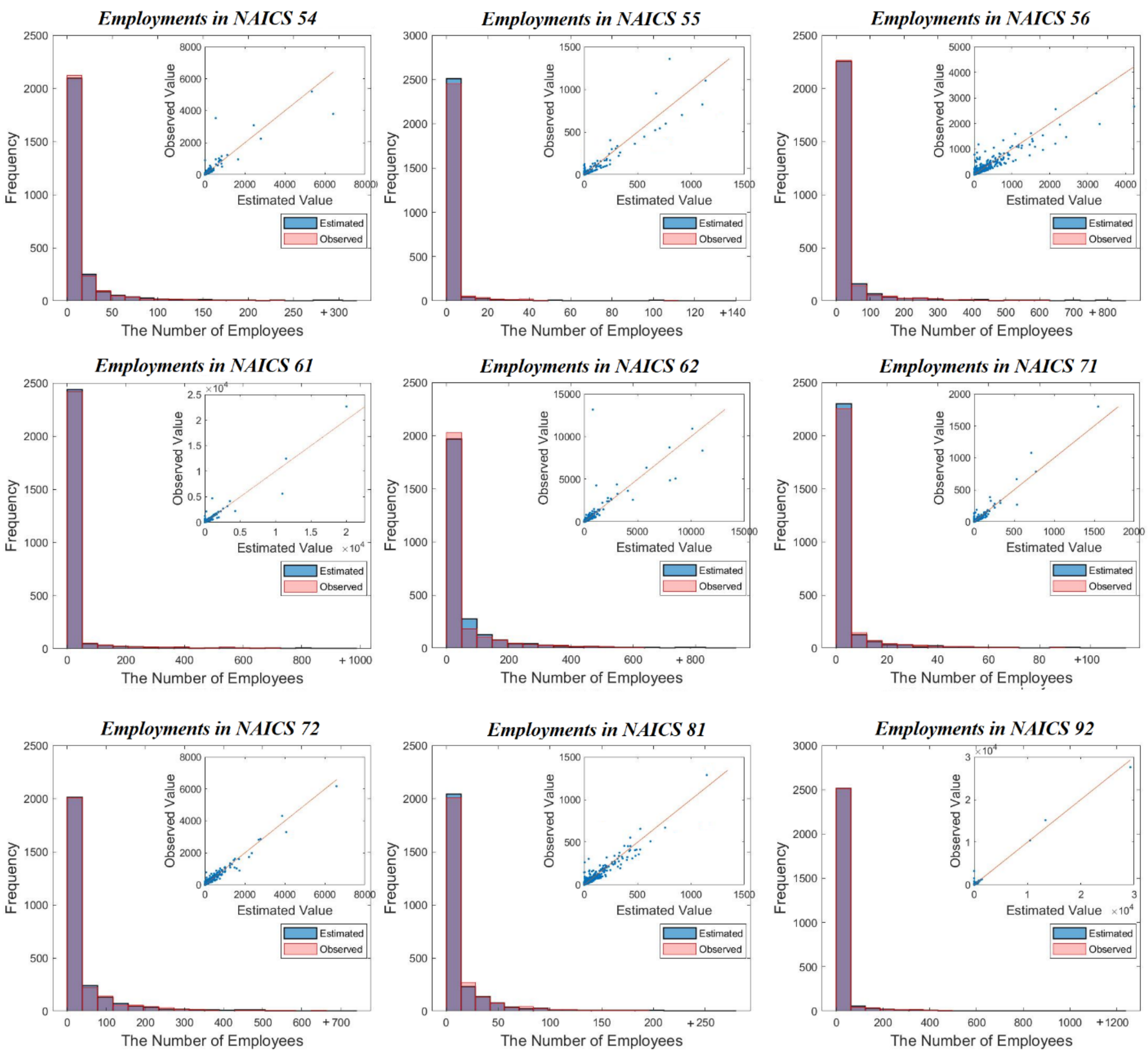

Figure 5-3. Histogram of estimated and observed employment by the proposed model in the year 2005. The inside plot is the correlation plot of observed and estimated employment (Part 2).

Finally, for the land use classes, the histogram and the correlation plots for the year 2005 are provided in Figure 5-4. This figure can easily show that proposed the model can estimate the land use classes with very good accuracy. In three land-use classes, residential, commercial, and agricultural areas, the correlation plot is very close to the 45-degree line (in agriculture the plot is a line actually); and the model could estimate the industrial and vacant land with acceptable accuracy.

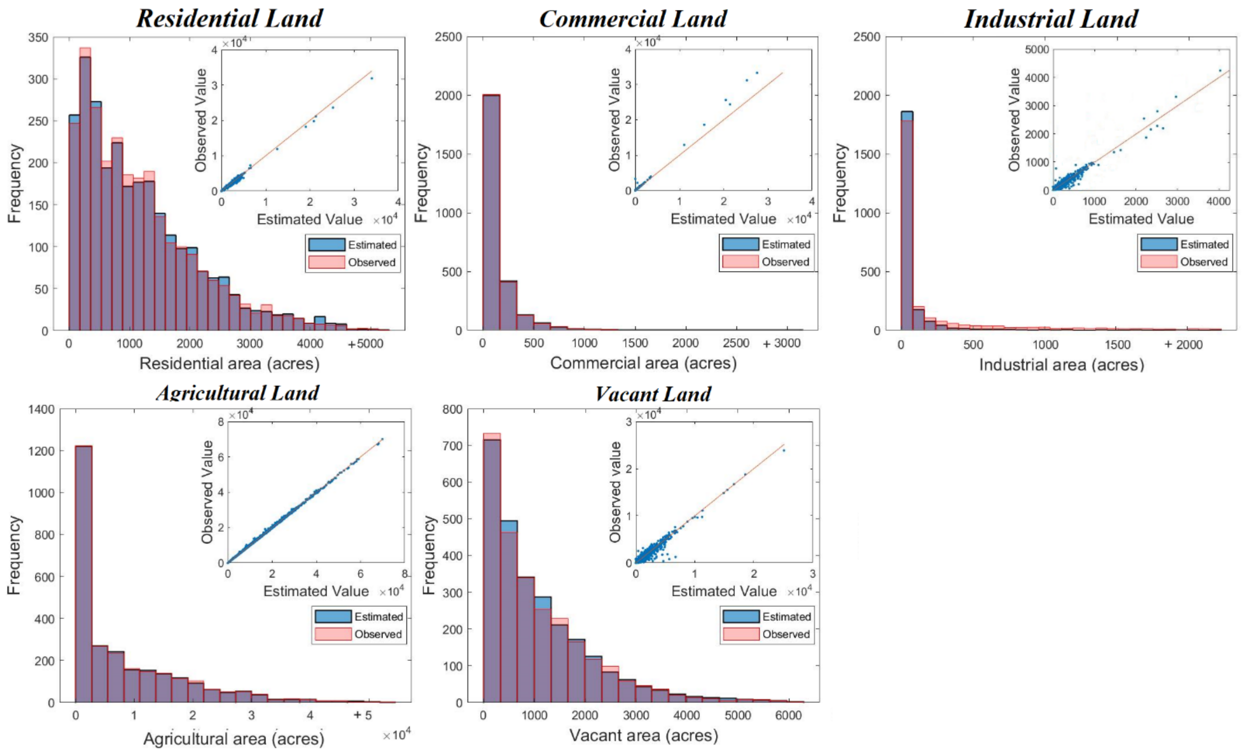

Figure 5-4. Histogram of estimated and observed land-use condition by the proposed model in five land-use classes in the year 2005. The inside plot is the correlation plot of observed and estimated land-use conditions.

5.2 Forecasting Validation

Forecasting validation had limits due to the latest available household data being from the year 2010; the model forecasting accuracy is provided for employment and the land use condition. Table 5-2 shows the R^2 and *MAPE* of the estimated employment in the year 2015 and land use condition in the year 2015. In the employment sections, the proposed model has deficits in forecasting NACIS 11, 21, and 53. Generally, in the employment section, the accuracy of forecasting is reduced in comparison to backcasting validation. However, the model still shows high accuracy in estimating the land use condition. The R^2 of all land use classes is more than 0.85, and the value of *MAPE* is acceptable. In addition to Table 5-2, similar to backcasting validation, the histogram and correlation plots are provided for the forecasting validation. Figures 5-5 to 5-7 show the validation results for the employment and land use classes in 2015. First, like backcasting, to show the results with higher resolution, the plot related to the employment section is provided in two separate figures. As Figures 5-5 and 5-6 show, in most of the employment sections (NAICS), the correlation plot shows a line close to a 45-degree line. Besides, the histogram plots show that the distribution of the employment between TAZs is close to zero and the majority of the TAZ has few employments and the difference between observation and estimated data is more significant in the first and second intervals in the employment section.

Figure 5-7 shows the proposed model can estimate the land use classes with very good accuracy. In three land-use classes, residential, commercial, and agricultural areas, the correlation plot is very close to the 45-degree line (in agriculture the plot is a line actually); and the model could estimate the industrial and vacant land with acceptable accuracy.

TABLE 5-2 THE R^2, MAPE, AND PERCENTAGE OF GOOD PREDICTION (PGP) OF BACKCASTING FOR THE YEAR 2015

Land use Filed	***R^2***	***MAPE (%)***	***PGP***
Total Employment	0.943	249.73	0.873
Employment in NAICS 11	0.584	121.69	0.662
Employment in NAICS 21	0.54	111.98	0.518
Employment in NAICS 22	0.926	99.97	0.787
Employment in NAICS 23	0.699	161.44	0.762
Employment in NAICS 3133	0.693	791.83	0.722
Employment in NAICS 42	0.794	337.51	0.764
Employment in NAICS 4445	0.785	286.61	0.829
Employment in NAICS 4849	0.923	693.08	0.535
Employment in NAICS 51	0.835	107.09	0.730
Employment in NAICS 52	0.664	145.09	0.730
Employment in NAICS 53	0.556	98.95	0.717
Employment in NAICS 54	0.718	218.74	0.745
Employment in NAICS 55	0.634	121.9	0.672
Employment in NAICS 56	0.67	324.2	0.688
Employment in NAICS 61	0.721	247.17	0.795
Employment in NAICS 62	0.956	239.83	0.768
Employment in NAICS 71	0.855	121.67	0.752
Employment in NAICS 72	0.832	328.71	0.821
Employment in NAICS 81	0.663	104.05	0.762
Employment in NAICS 92	0.694	173.06	0.599
Residential Land	0.994	3.64	0.983
Commercial Land	0.967	11.28	0.938
Industrial Land	0.865	73.29	0.806
Agricultural Land	0.999	20.96	0.993
Vacant Land	0.904	39.26	0.915

The R^2 and PGP greater than 0.66 and MAPE less than 200 are considered as acceptable accuracy.

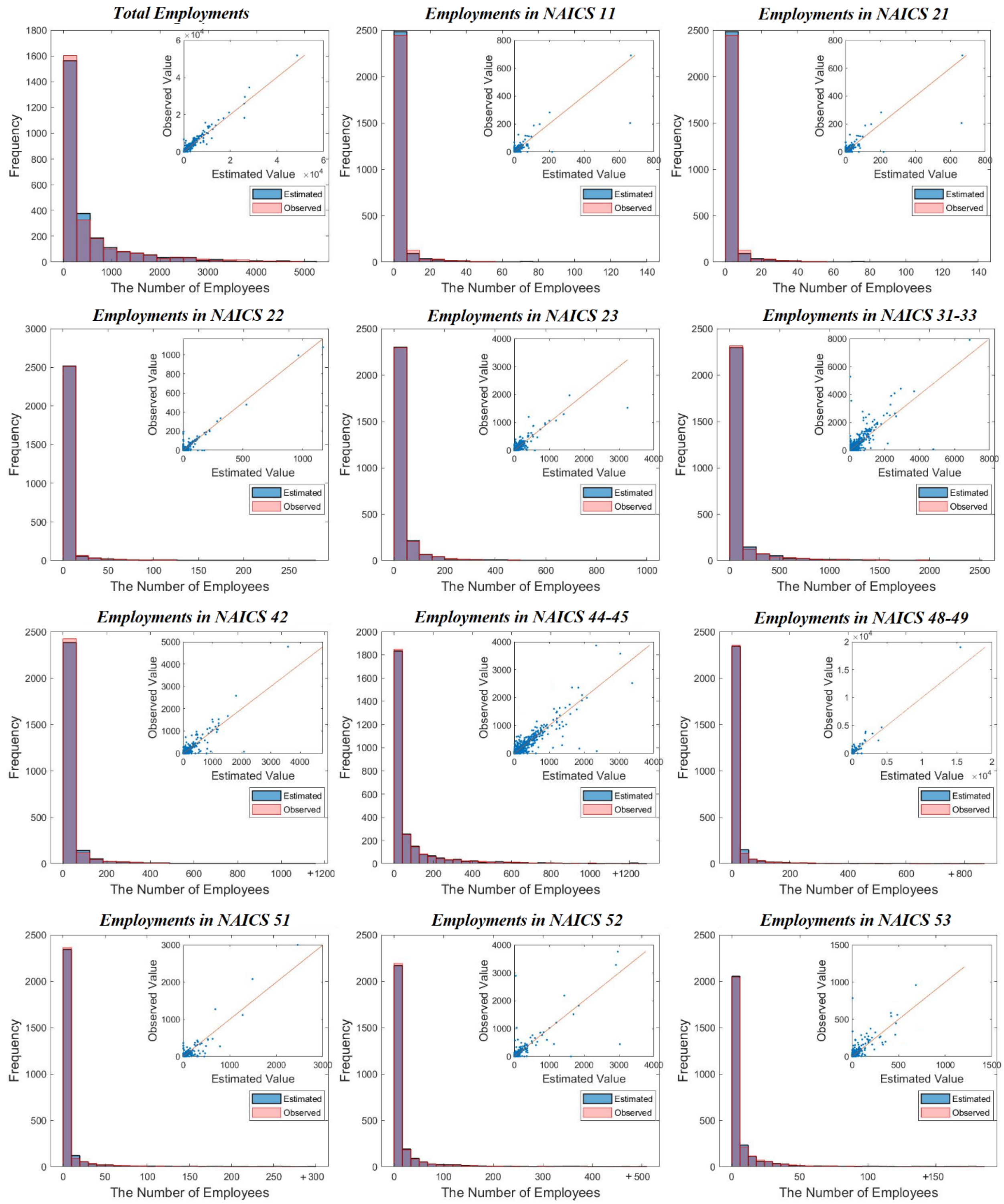

Figure 5-5. Histogram of estimated and observed employment by the proposed model in the year 2015. The inside plot is the correlation plot of observed and estimated employment (Part 1).

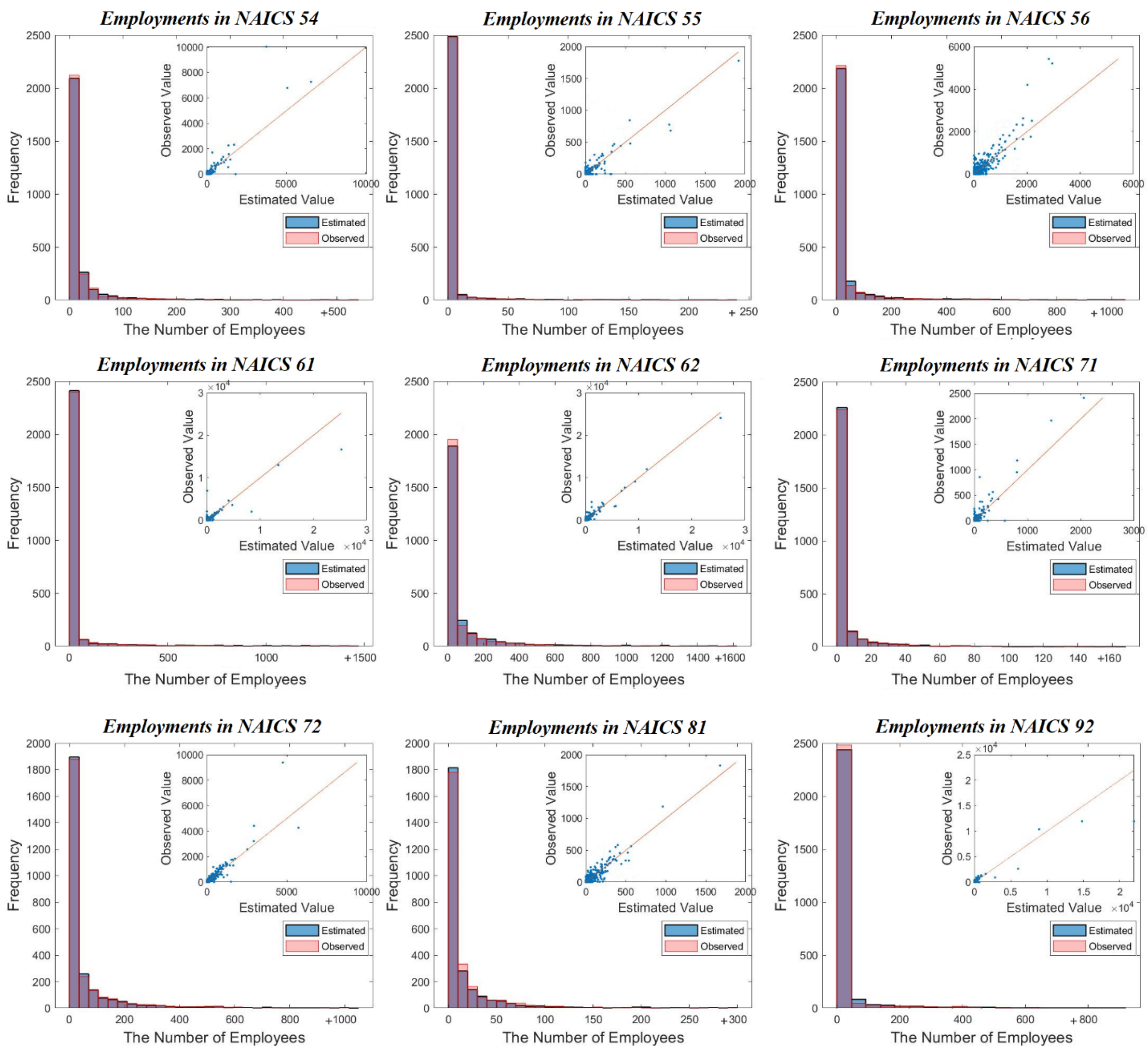

Figure 5-6. Histogram of estimated and observed employment by the proposed model in the year 2015. The inside plot is the correlation plot of observed and estimated employment (Part 2).

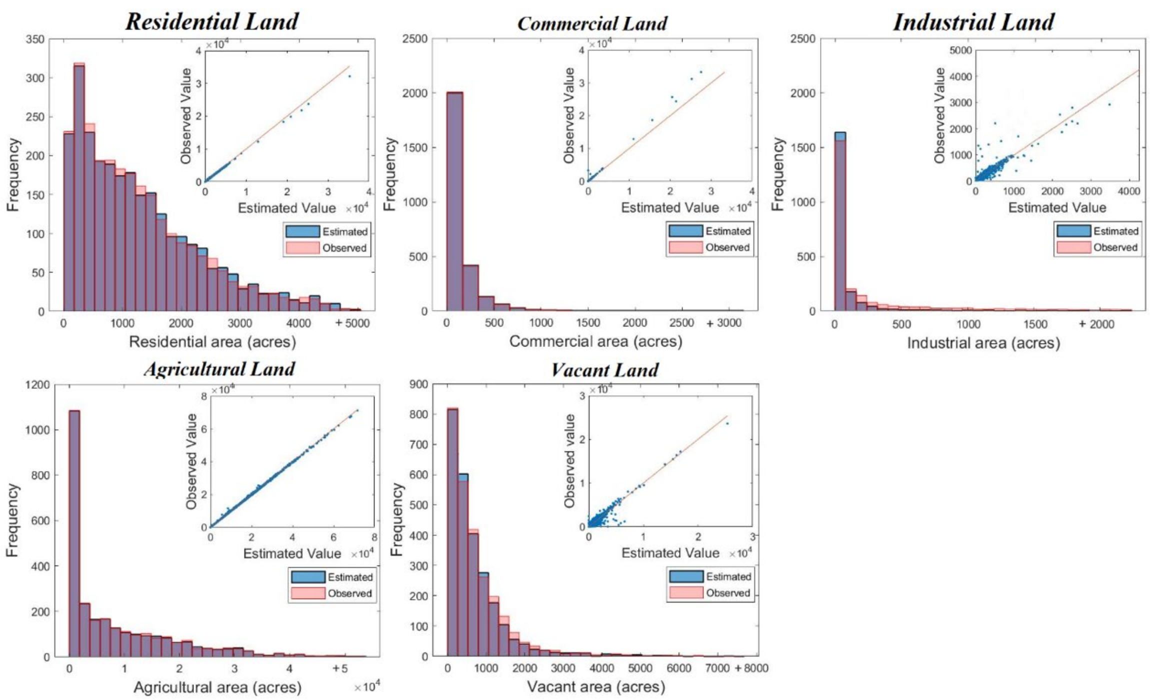

Figure 5-7. Histogram of estimated and observed land-use condition by the proposed model in five land use classes in the year 2015. The inside plot is the correlation plot of observed and estimated land-use conditions.

Chapter 6 Land Use Forecasting

Land-use models should be able to forecast socio-economic and demographic data not only for the base and future years but also for interim years such as five years in the interval. Because the current statewide travel demand model serves 2010 and 2040 as the base and future years, the proposed land use model (LS-LUM) is designed as such the output of the land-use model will be saved for each of the interim years until the future year. The purpose of this chapter is twofold: A) providing the LS-LUM's forecasting accuracy for horizon years 2015 to 2050 (considering 5-year intervals), B) presenting the state of Tennessee's socio-economic and demographic condition from 2015 to 2050.

It is crucially important to evaluate if the land-use model will retain its accuracy after forecasting for a long-range of horizon years. In this regard, the forecasting results of LS-LUM are evaluated for the horizon years 2015, 2020, 2025, 2030, 2040, 2045, and 2050. Data was collected from the Woods & Pool data set to measure the accuracy of LS-LUM. Since the Woods & Pool data sets are at the county level, the results of LS-LUM are aggregated from TAZ level to county level. Tables 6-1 and 6-2 present the goodness of fit measure (R^2), the error measure (MAPE), and Percentage of Good Prediction (PGP) for Total population, Total households, Total employment, and employment in 20 sectors of NAICS categories from 2015 to 2050. As tables 6-1 and 6-2 show, the LS-LUM could retain its accuracy after forecasting in a sequence of eight prediction intervals, but some employments categories that show moderate accuracy in forecasting the year 2040 and higher.

In addition to presenting LS-LUM's accuracy in the forecasting of horizon years, to present all forecasted demographic and socio-economic conditions of the state of Tennessee, an online dashboard was developed to help the user with observing, downloading, and using the forecasted values. To access the online dashboard, please visit https://arcg.is/0fTO0H. A screenshot of the dashboard is presented in Figure 6-1. In addition to a visual demonstration of the forecasted demographic and socio-economic condition, this dashboard provides a brief statistical analysis.

To illustrate the output of the LS-LUM, as an example, in this chapter the results of forecasting total population and total employment in the state of Tennessee are provided visually, to show how the demographic and socio-economic condition of the state will change from 2015 to 2050. Figures 6-2 and 6-3 present the changes in the total population from 2010 to 2050 where the state of Tennessee will witness a significant population growth. Moreover, Figures 6-4 and 6-5 present the changes in total employment from 2010 to 2050, where the number of total employment will increase gradually.

TABLE 6-1 THE R^2, MAPE (%), AND PGP OF FORECASTING HORIZON YEARS 2015 TO 2030 IN COUNTY LEVEL

Year	***2015***			***2020***			***2025***			***2030***		
Land use Filed	**R^2**	**MAPE (%)**	**PGP**	**R^2**	**MAPE (%)**	**PGP**	**R^2**	**MAPE (%)**	**PGP**	**R^2**	**MAPE (%)**	**PGP**
Total Population	0.984	23.7	0.95	0.968	33.2	0.925	0.943	44.5	0.901	0.895	60.2	0.869
Total Households	0.986	21.7	0.955	0.982	24.9	0.944	0.977	27.7	0.934	0.97	31.4	0.924
Total Employment	0.895	78.6	0.825	0.855	92.3	0.804	0.819	103.2	0.783	0.8	108.4	0.779
Employment in NAICS 11	0.576	104.5	0.556	0.554	102.7	0.554	0.586	102.5	0.556	0.556	102.3	0.556
Employment in NAICS 21	0.645	158	0.605	0.595	164.2	0.595	0.599	160.6	0.593	0.589	156.7	0.589
Employment in NAICS 22	0.511	172.7	0.66	0.48	179.1	0.653	0.439	189.2	0.64	0.409	197.2	0.639
Employment in NAICS 23	0.755	105.5	0.757	0.68	120.5	0.729	0.68	120	0.729	0.687	118.3	0.733
Employment in NAICS 3133	0.915	45.1	0.878	0.885	50.7	0.856	0.874	52.6	0.852	0.864	54.6	0.85
Employment in NAICS 42	0.971	51.4	0.881	0.97	54.3	0.88	0.97	53.8	0.876	0.974	48.6	0.883
Employment in NAICS 4445	0.957	45.4	0.878	0.947	49.9	0.867	0.947	48.6	0.868	0.949	47	0.874
Employment in NAICS 4849	0.902	112.3	0.812	0.846	134.5	0.778	0.818	141	0.765	0.788	146.6	0.754
Employment in NAICS 51	0.96	66.3	0.886	0.918	97.4	0.855	0.907	103.9	0.852	0.891	112	0.844
Employment in NAICS 52	0.847	106.8	0.802	0.783	126.6	0.767	0.731	139.7	0.741	0.736	137.8	0.74
Employment in NAICS 53	0.437	198.9	0.625	0.398	205	0.617	0.364	208.4	0.608	0.352	208.1	0.606
Employment in NAICS 54	0.769	146.3	0.756	0.726	161.8	0.743	0.67	180.5	0.722	0.658	186.9	0.727
Employment in NAICS 55	0.892	109.4	0.824	0.828	141.7	0.79	0.702	192.4	0.728	0.734	187.3	0.763
Employment in NAICS 56	0.837	123.8	0.789	0.819	126.9	0.779	0.775	140.6	0.753	0.744	149	0.75
Employment in NAICS 61	0.33	209.8	0.601	0.311	210.5	0.597	0.29	211.6	0.592	0.272	214	0.588
Employment in NAICS 62	0.967	52.1	0.895	0.942	70.1	0.875	0.89	97.1	0.838	0.841	117.6	0.815
Employment in NAICS 71	0.491	245.4	0.697	0.446	255.6	0.68	0.4	268.4	0.658	0.372	277.1	0.656
Employment in NAICS 72	0.961	49.6	0.886	0.932	65.3	0.859	0.912	72.4	0.838	0.898	76	0.832
Employment in NAICS 81	0.511	172.7	0.66	0.48	179.1	0.653	0.439	189.2	0.64	0.409	197.2	0.639
Employment in NAICS 92	0.443	152.3	0.659	0.418	153.8	0.653	0.391	153.5	0.644	0.373	152.1	0.646

TABLE 6-2 THE R2, MAPE (%), AND PGP OF FORECASTING HORIZON YEARS 2035 TO 2050 IN COUNTY LEVEL

Year	*2035*			*2040*			*2045*			*2050*		
Land use Filed	R^2	MAPE (%)	PGP	R^2	MAPE (%)	PGP	R^2	MAPE (%)	PGP	R^2	MAPE (%)	PGP
Total Population	0.942	44.6	0.898	0.744	93.1	0.804	0.684	109.2	0.774	0.644	125.8	0.742
Total Households	0.963	35.2	0.917	0.953	38.9	0.9	0.945	42.2	0.892	0.937	45.1	0.886
Total Employment	0.773	115.5	0.767	0.765	116.7	0.765	0.755	118.5	0.76	0.75	120.3	0.76
Employment in NAICS 11	0.556	102.5	0.556	0.558	104.1	0.558	0.555	127.8	0.555	0.563	102.9	0.563
Employment in NAICS 21	0.586	152.5	0.586	0.583	146.7	0.583	0.581	144.3	0.581	0.58	140.7	0.58
Employment in NAICS 22	0.38	214.6	0.653	0.354	211.6	0.627	0.332	217.1	0.622	0.314	223.1	0.619
Employment in NAICS 23	0.686	118	0.732	0.711	112.6	0.744	0.719	110.5	0.747	0.724	109.1	0.751
Employment in NAICS 3133	0.868	53.5	0.853	0.829	59.8	0.836	0.811	62.5	0.825	0.791	66.4	0.813
Employment in NAICS 42	0.973	49.1	0.879	0.975	45.7	0.891	0.968	50.2	0.885	0.951	61.4	0.864
Employment in NAICS 4445	0.942	49.6	0.868	0.945	47.9	0.879	0.935	51.4	0.876	0.918	58.4	0.87
Employment in NAICS 4849	0.764	148.9	0.741	0.718	154.6	0.728	0.685	156.7	0.723	0.654	161.8	0.72
Employment in NAICS 51	0.867	123.9	0.832	0.845	133.6	0.823	0.812	146.4	0.812	0.773	162.7	0.803
Employment in NAICS 52	0.727	139.6	0.734	0.724	140.2	0.731	0.723	140.9	0.728	0.721	142.8	0.728
Employment in NAICS 53	0.33	209.8	0.601	0.311	210.5	0.597	0.29	211.6	0.592	0.272	214	0.588
Employment in NAICS 54	0.605	204.3	0.707	0.595	209.8	0.711	0.565	220.8	0.703	0.537	231.4	0.695
Employment in NAICS 55	0.69	208.1	0.754	0.65	227	0.735	0.612	245.6	0.716	0.575	265	0.698
Employment in NAICS 56	0.69	162.6	0.73	0.692	160.5	0.731	0.675	164	0.724	0.662	166.8	0.719
Employment in NAICS 61	0.33	209.8	0.601	0.311	210.5	0.597	0.29	211.6	0.592	0.272	214	0.588
Employment in NAICS 62	0.762	145	0.775	0.769	143.4	0.779	0.753	148.8	0.773	0.746	152.4	0.771
Employment in NAICS 71	0.332	288.4	0.642	0.316	293.3	0.638	0.294	299.4	0.63	0.277	307.2	0.625
Employment in NAICS 72	0.857	87.7	0.807	0.859	84.6	0.816	0.825	92.2	0.805	0.784	101	0.796
Employment in NAICS 81	0.5	204.9	0.633	0.354	211.6	0.627	0.332	217.1	0.622	0.314	223.1	0.619
Employment in NAICS 92	0.381	147.8	0.649	0.338	149.3	0.645	0.328	147	0.646	0.317	146	0.647

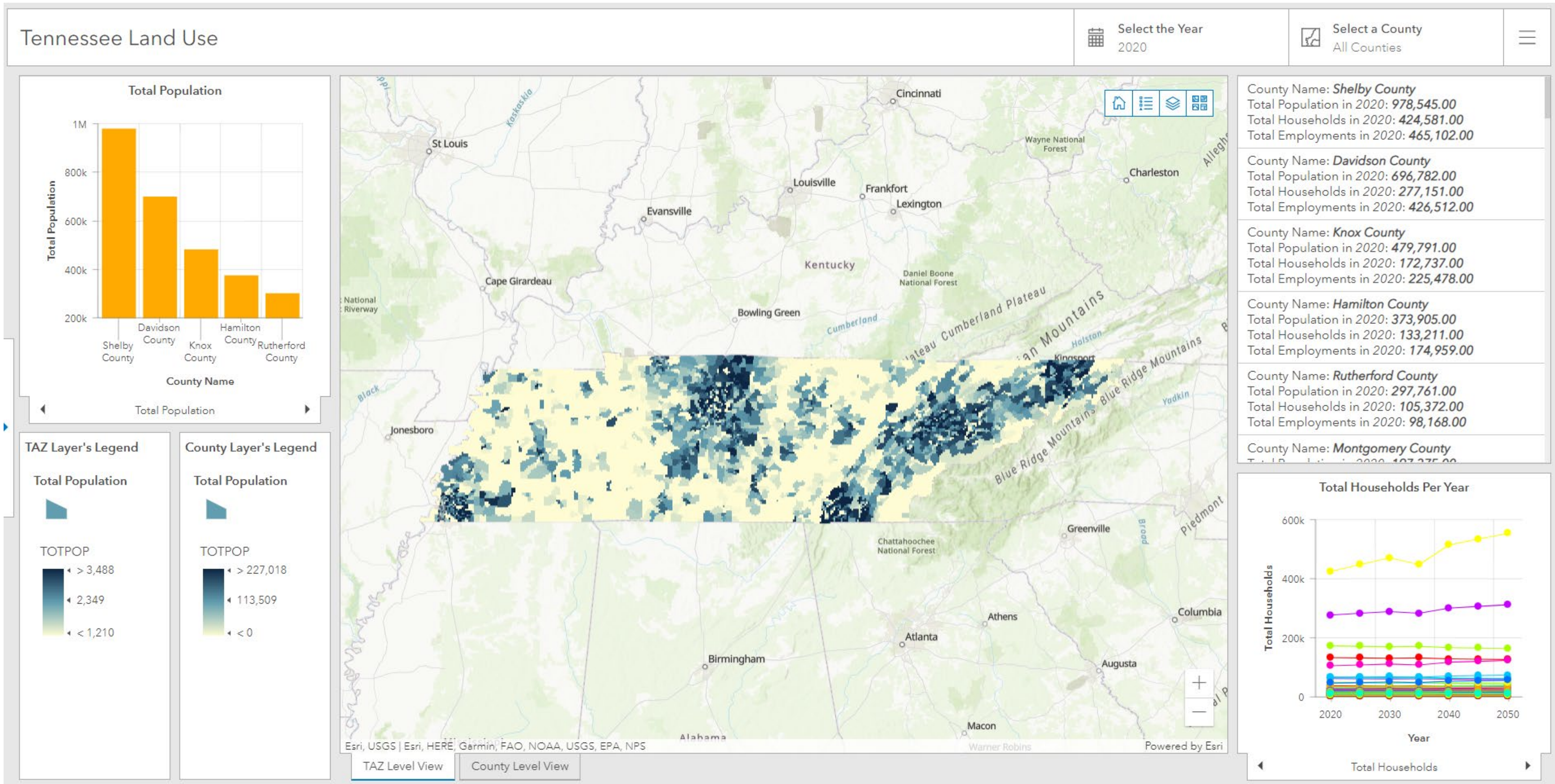

Figure 6-1 The developed online dashboard for presenting the results and some online statical analysis

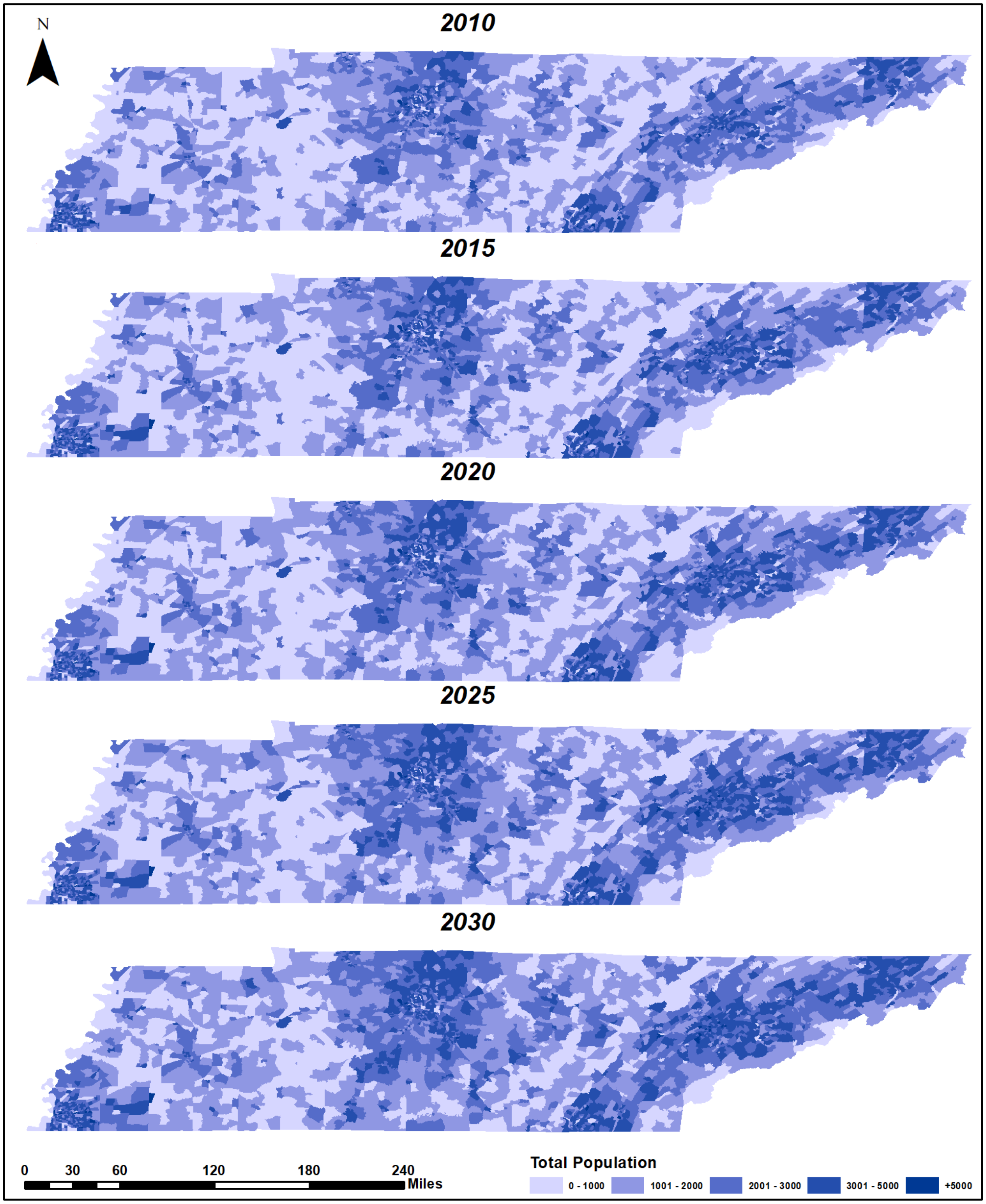

Figure 6-2. The forecasted total population for the state of Tennessee from 2010 to 2030

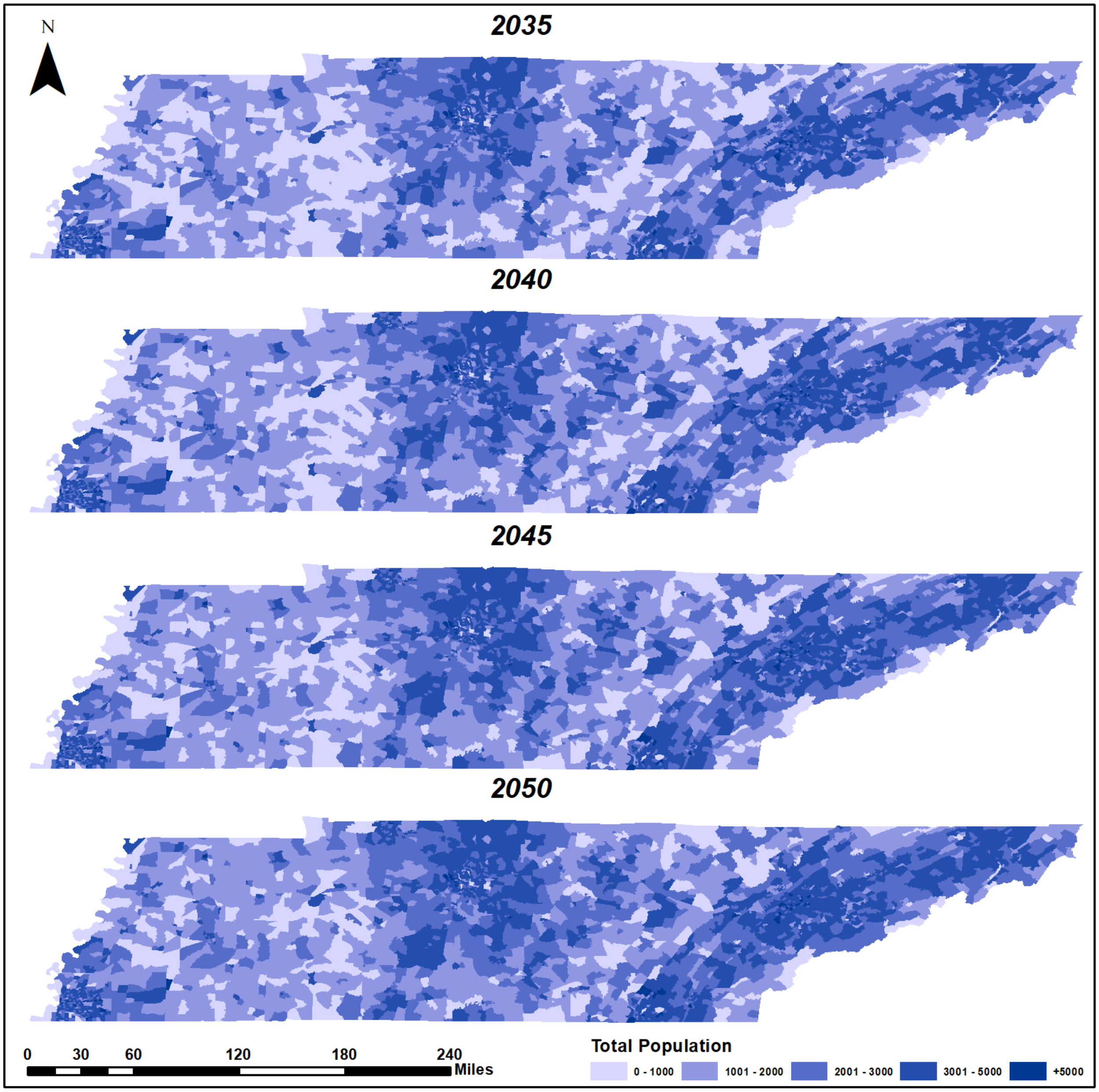

Figure 6-3. The forecasted total population for the state of Tennessee from 2010 to 2030

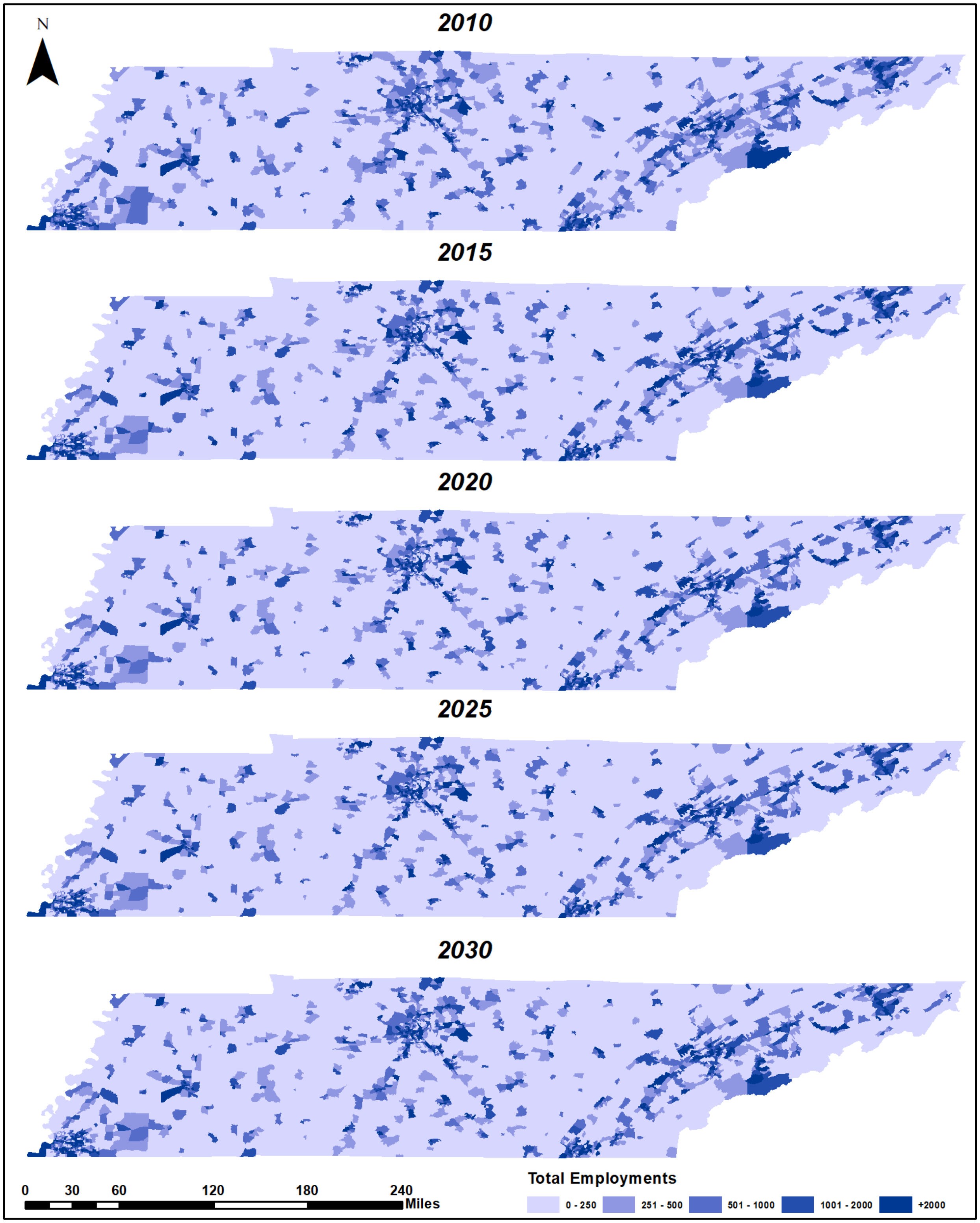

Figure 6-4. The forecasted total employment for the state of Tennessee from 2010 to 2030

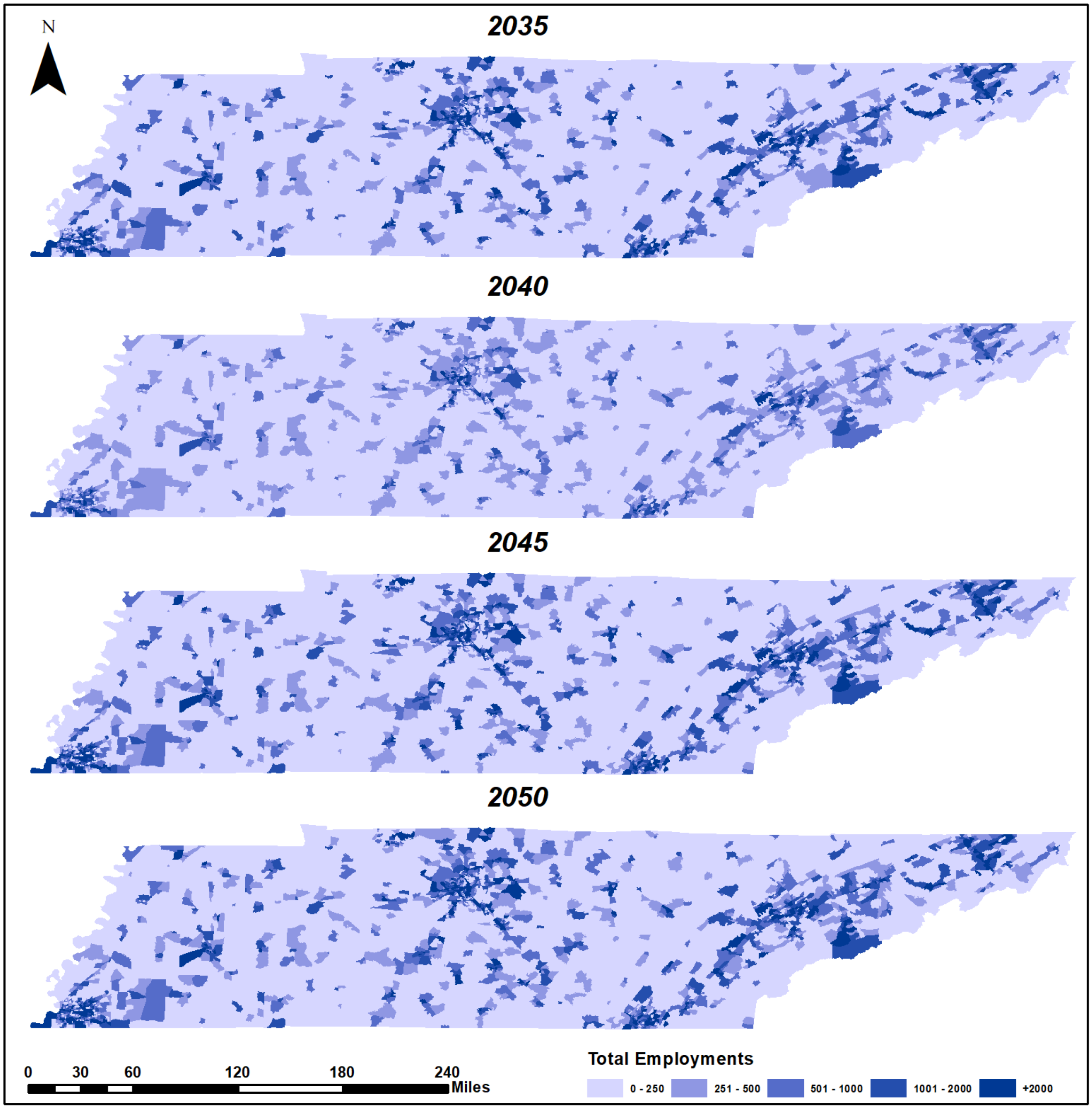

Figure 6-5. The forecasted total employment for the state of Tennessee from 2010 to 2030

Chapter 7 Integration with Statewide Travel Demand Model

A land-use model when integrated with the travel demand model provides a complete package to analyze interdependencies between land use and transportation. In this chapter, the proposed land use model in this report (LS-LUM) is integrated with Tennessee Statewide Travel Demand Model, version3 (TSM V3). The integration is such that the software or set of scripts developed can run the land-use model, generate visualizations of the land-use model output, and various performance measures, and provide output that can be used as input to the travel demand model. In this chapter, first, brief information about the TSM V3 is provided and then the integrated land-use transportation model and its functionality are discussed.

7.1 Travel Demand Model

The Travel Demand Model, which is integrated with the land-use model and the travel time derived form, is the Tennessee Statewide Travel Model (TSM) version 3 (RSG, 2014). This version of TSM is a traditional four-step travel demand model consisting of three different components: a short distance passenger model (trips less than 50 miles), a long-distance passenger model, and a freight model. The underlying geographic area of operation is at the TAZ level. The total number of TAZs in TSM V3 is 3,687. Zonal attributes include the number of households, categorized by income, size, worker, presence of students, presence of seniors, and the number of vehicles; and the number of employments categorized by 20 sectors of NAICS codes. The TSM V3 can be understood at a high level as comprised of input network and socioeconomic data together with some component demand models and a highway assignment model. The demand components can be gathered in three broad groups related to short-distance passenger demand, long-distance passenger demand, and freight and truck demand. The TSM V3 uses TransCAD's implementation of the tri-conjugate Frank-Wolfe algorithm for multi-class user equilibrium traffic assignment (Bernardin Jr et al., 2017). The accessibility matrices which serve as input for the land-use model are obtained from TSM's assigned networks using the shortest path method. In the following, the overview of TSM V3 is provided, this overview is presented from Tennessee Statewide Travel Model (version 3) Development and Validation Technical Report.

The TSM V3 development focused on the incorporation of advanced functionality and sensitivity particularly for freight and long-distance travel which were not able to be included in the version 2 model due to the schedule of Phase 2. Phase 3 also built on the success of the TSM V2's use of American Transportation Research Institute (ATRI) data with the incorporation of additional new big data from AirSage. In particular, the goals of version 3 can be broken down into four areas or topics: freight modeling/forecasting, highway forecasting, accuracy and sensitivity, and long-distance travel. Significant accomplishments were made in each of these areas. The TSM V3offers improved freight forecasting and analysis with a freight model based on and driven by commodity flows with functionality to allow the user to analyze the commodities carried by trucks on a particular facility or test various mode share scenarios including testing new intermodal facilities with user assumed diversion, test new rail lines or extensions, or test new port facilities. The TSM V3 offers enhanced highway forecasting with AM & PM peak period volumes, rough demand for high occupancy vehicle (HOV) lanes, and more realistic networks with truck prohibitions due to vertical clearance and refined capacities considering grades, freeway weave sections, etc. The TSM V3 offers improved accuracy and sensitivity by incorporating the effects of population aging, built environment effects and walkability, and psychological boundaries. The TSM V3's integration with FHWA's new national long-distance passenger model offers consistency with FHWA forecasting, a model based on recent long-distance travel data from AirSage, and the ability to test long-distance modal alternatives (new air service, new intercity bus service, new intercity passenger rail). Several of these accomplishments went beyond the actual scope requirements for Phase 3, particularly concerning the ability to test different modal scenarios (concerning freight, short and long-distance passenger demand).

The advanced components of the TSM V3 are applied within an overall framework of a data-driven model. The demand model components, using information from the network and zones, predict changes in demand which are applied to actual data on the existing travel patterns. This approach combines the sensitivity of advanced models with the accuracy of real base year data. Such data-driven approaches are used in other statewide models that have also incorporated big OD data such as Indiana and Florida's models. The TSM V3 can be understood at a high level as comprised of input network and socioeconomic data together with several component demand models and a highway assignment model. The demand components can be gathered in three broad groups related to short-distance passenger demand, long-distance passenger demand, and freight & truck demand. See Figure 7-1.

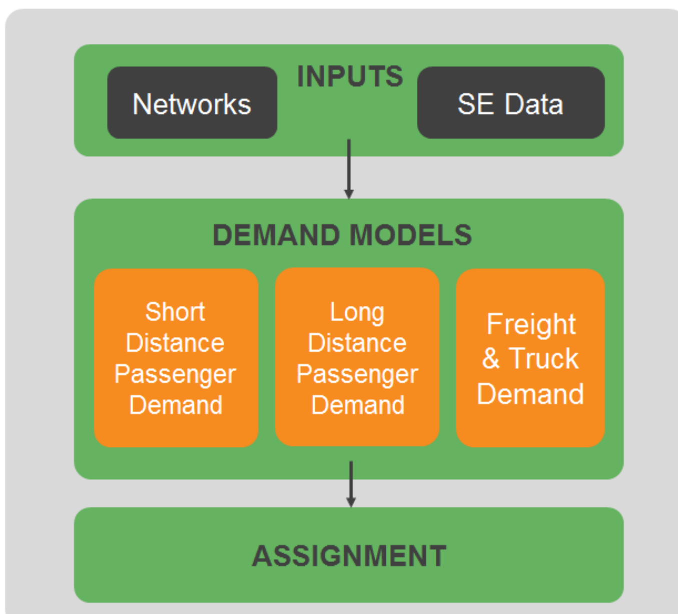

Figure 7-1. Tennessee statewide travel demand model general framework

As with its predecessor, the TSM V2, the TSM V2 achieves a high level of validation, with particularly good highway assignment error statistics. Moreover, it achieves this validation and provides all its new functionality efficiently, maintaining reasonable runtimes. The full model run with daily assignment takes roughly three hours or five hours for a run including AM and PM peak assignments (based on a machine with 12 physical cores and 32 GB RAM). The TSM V3 combination of functionality, validation, and efficiency make it an example of best practice in statewide modeling; while a few other statewide models may be comparable 20 or even slightly better in one of these dimensions, no other statewide model compares favorably in all these dimensions.

7.2 Integrating LS-LUM with Tennessee Statewide Travel Demand Model

The main purpose of this research was to develop a statewide land-use model which can be integrated with Tennessee Statewide Travel Demand Model. The previous chapter showed that the developed land-use model (LS-LUM) can forecast and generate demographic and socioeconomic conditions of the study area at the TAZ level and with acceptable accuracy. Integrated land use transport framework provides a strong tool to evaluate the interdependence effect of land use and transportation system. Where the land-use model forecasts the location of population and employment based on travel time (accessibility) and the travel demand model forecasts the traffic flow based on population and employment locations and transportation network. In this regard, the LS-LUM output was imported as the input data to the TSM V3 for the years 2015 to 2050 (every five years), and the total traffic flow, Vehicle High Frequency (VHF), and Vehicle Hours Traveled are calculated. For instance, Figure 7-2 presents the result of TSM V3 in terms of total flow.

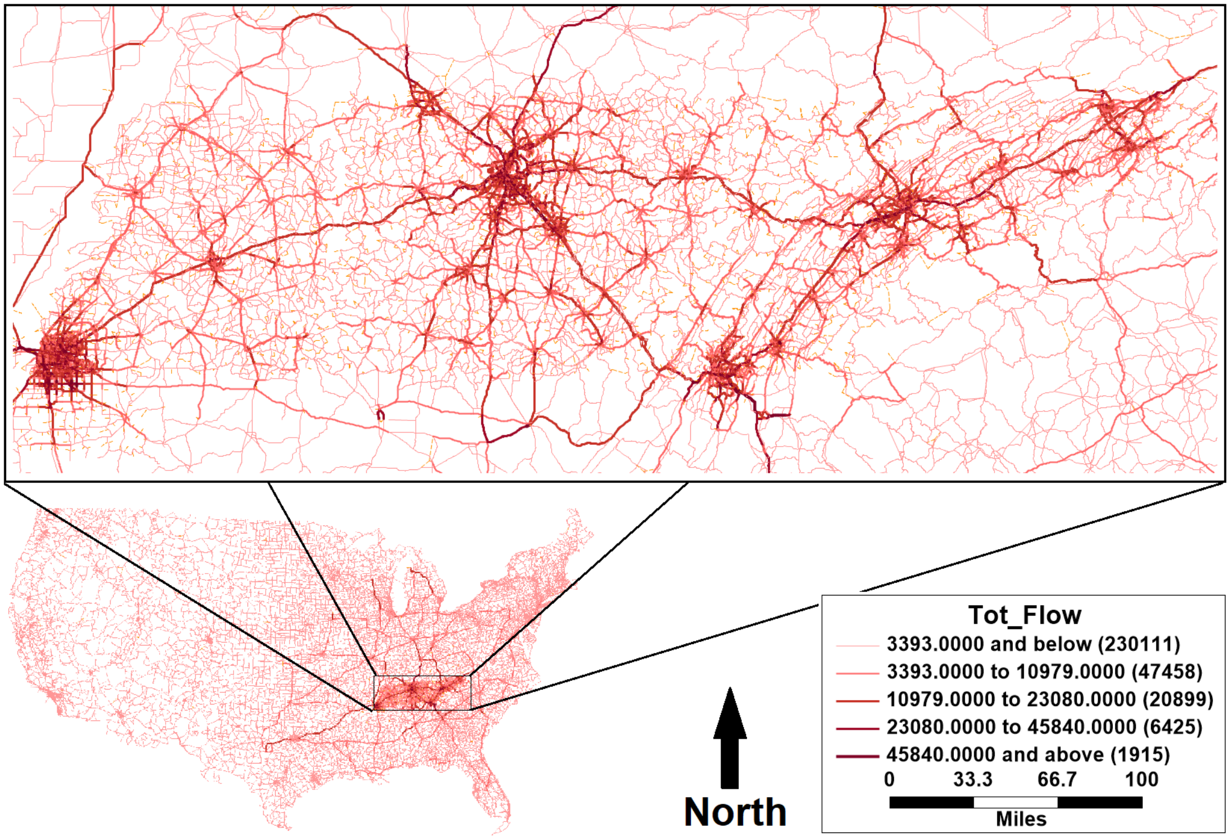

Figure 7-2. Total Traffic Flow for the year 2030

Chapter 8 Policy Analysis

Policy analysis is an integral part of assessing the model's responsiveness to changes in the input files representing a specific scenario. In this chapter, a hypothetical example for policy analysis is provided to test the strength of the developed land use model in policy testing.

The developed land-use model (LS-LUM) provides powerful tools for applying policy analysis on a statewide scale. Transportation planners can manipulate the land use condition (i.e., the amount of residential, commercial, industrial, and agricultural areas), the housing condition (i.e., the number of total houses and vacant houses in a TAZ), the population density, and the number of job opportunities in an area, to assess the changes in the demographic and socio-economic condition of future years. Moreover, the integrated land-use transport system gives this opportunity to apply transportation network-related scenarios and evaluate the effect of any improvement or deterioration of transportation networks in an area on population and employment locations.

In this regard, a hypothetical problem is defined to show and test the LS-LUM capability in assessing different policies. It is assumed that a local transportation agency is interested in evaluating how turning an agricultural area in a TAZ into a residential area would affect the demographic condition of that TAZ and its surrounding TAZs. Also, this transition is planned to be applied in the year 2030. Table 8-1 shows the study area's total population, households, total employment, and land use condition for the year 2030. The agency is planning to turn 2,000 acres of agricultural areas in TAZ #1 into residential areas. Therefore, first, the model is executed to forecast the condition of the year 2035 considering the current condition. The results of the forecasting year 2035 considering the current condition are provided in Table 8-2.

Then the land use condition is changed for the year 2030 (the new policy is applied). In this condition, the residential areas in TAZ #1 would be 3487 and the agricultural areas turn into 12552. Having this new input data, the forecasting model is executed to forecast the year 2035 with new land-use conditions. The results are tabulated in Table 8-3.

TABLE 8-1 THE CONDITION OF THE EXAMPLE PROBLEM IN THE YEAR 2030 BEFORE APPLYING THE NEW POLICY

TAZ ID	*Tot. Pop.*	*Tot. HH.*	*Tot. Emp.*	*Res. area*	*Com. area*	*Ind. area*	*Agr. area*	*Vac. area*
1	1579	529	0	1487	26	0	14552	89
2	1796	588	15	2103	2	0	10509	78
3	2381	816	2535	401	187	12	0	63
4	2280	813	2615	2328	395	51	7996	128
5	3008	1109	898	2038	256	12	6135	155
6	3333	996	4292	726	114	77	1132	329
7	917	361	79	1050	7	0	22861	56

TABLE 8-2 THE CONDITION OF THE EXAMPLE PROBLEM IN THE YEAR 2035 BEFORE APPLYING THE NEW POLICY

TAZ ID	Tot. Pop.	HH	Tot. Emp.	Res. area	Com. area	Ind. area	Agr. area	Vac. area
1	1685	561	0	1515	26	0	14519	94
2	1900	620	17	2130	2	0	10477	83
3	2483	846	2564	409	190	10	0	54
4	2382	843	2648	2372	395	40	7977	114
5	3106	1136	934	2081	259	10	6123	123
6	3429	1024	4227	758	122	59	1147	291
7	1009	395	87	1067	8	0	22699	200

As table 8-3 shows, changing the land use condition in TAZ #1 causes changes in total population, total households, and total employment in the study area.

TABLE 8-3 THE CONDITION OF THE EXAMPLE PROBLEM IN THE YEAR 2035 AFTER APPLYING THE NEW POLICY

TAZ ID	Tot. Pop.	HH	Tot. Emp.	Res. area	Com. area	Ind. area	Agr. area	Vac. area
1	1764	572	0	2516	26	0	12510	95
2	1889	618	18	2130	2	0	10477	83
3	2481	825	2564	409	190	10	0	54
4	2378	843	2647	2372	395	40	7977	114
5	3106	1136	934	2081	259	10	6123	123
6	3429	1024	4229	758	122	59	1147	291
7	999	394	86	1067	8	0	22699	200

There are a few points that transportation planners should have considered while they are using the LS-LUM for conducting policy analysis. First, if you are changing the land use condition, always consider the vacant area and total lands. The sum of the land areas must not exceed the total land in a TAZ. Moreover, if you are increasing a land-use section (e.g., residential area) the added area should come from either the vacant areas or other land use sections. Second, when you are adding new houses to a TAZ, the total number of houses and the number of vacant houses in that TAZ should change correspondingly.

Chapter 9 Conclusion and Recommendations

The purpose of this research was to develop a statewide forecasting land-use model that can be integrated with Tennessee Statewide Travel Demand Model (TSM v3), with reasonable computational time and acceptable accuracy. A complete data collection approach was followed in this research and a comprehensive data set from publicly available data was collected. Demographic information of the state of Tennessee, containing total population, total households, and households' size, income, and seniors were collected at the TAZ level. Moreover, total employment and employment in 20 NAICS sectors were collected for all TAZs. This study incorporated parcel data for collecting land-use conditions. The content, currency, structure, and coverage of parcel data sets vary significantly across jurisdictions and regions. These differences create a challenge to develop standardized data. Assembling and standardizing parcel-level data from individual states and counties is more complicated than simply contacting each state or county and arranging for a data transfer. Some of the challenges include (1) an understanding of data availability and completeness, (2) the willingness of local governments to provide data, and (3) the varying content, format, and structure of data among counties. For this research project, standardized parcel data were collected using the Tennessee Comptroller of Treasury website.

The proposed model, Large-Scale Land-Use Model (LS-LUM), incorporates the gravity theory approach for allocating population and employment locations while using an enhanced formulation and strong solution approach. The improved model formulation consists of new variables addition in the form of total and vacant houses. LS-LUM involves the amount of commercial, industrial, and agricultural land in predicting the number of employments. A new evolutionary computation-based solution approach is presented to enhance accuracy and optimality. The overall results showed that LS-LUM provides acceptable accuracy in forecasting demographic and socio-economic conditions of the state of Tennessee and can retain its accuracy after running for eight-time intervals (until the year 2050). The LS-LUM forecasts the required data for the TSM V3 with 5 years interval from 2015 to 2050.

The LS-LUM incorporates three approaches to improve accuracy. The first approach was to include the components of land-use conditions in the employment section directly. LS-LUM involves the amount of land (acres) in commercial, industrial, and agricultural classes to estimate the number of employments. The second approach was involving house conditions in forecasting households. Results showed that when the HC section is added to the model, the model accuracy and stability increase. The results of household allocation in LS-LUM showed acceptable R2, PGP, and MAPE in both developing and backcasting. Finally, LS-LUM compared to other land-use models at this level can forecast land use conditions with a very good accuracy, which not only increases the accuracy of forecasting but also, provides a powerful tool for policy analysis. Five types of land use conditions are forecasted by the developed land-use model, residential, commercial, industrial, agricultural, and developable land. In addition, the LS-LUM can forecast the housing condition (total number of houses and the number of vacant houses) at the TAZ level.

In this project, an online dashboard using ArcGIS online was developed to present the forecasting results of the project. This online dashboard provides a brief statistical analysis of the forecasted values and illustrates the output of the model at both TAZ and county levels. Having this online dashboard, the output of the model can be shared and downloaded with the permission of the TDOT.

Finally, an interface was developed especially for the state of Tennessee. This software is developed using MATLAB Compiler Runtime (MCR) and can be installed and shared with users who do not have an installed MATLAB. The interface of the model is presented in Figure 9-1. A manual is provided separately for users on how they can use the software.

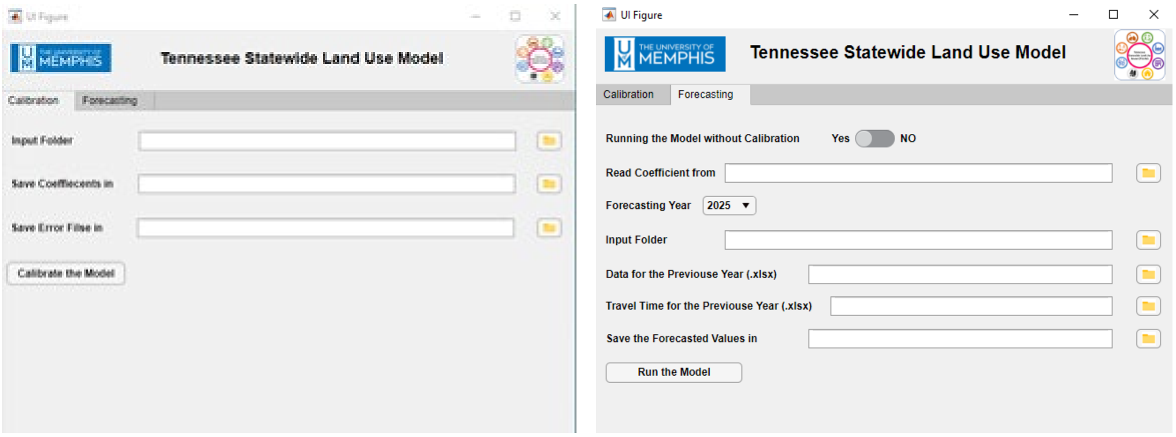

Figure 9-1. The developed software for the Tennessee Statewide Land Use Model using the LS-LUM

9.1 Recommendations for Future Studies

This project developed the first statewide land-use model for the state of Tennessee. The validation results showed acceptable accuracy in forecasting the demographic and socio-economic variables, making it suitable for integration with TSM V3. An integrated land use and transportation model was delivered as a part of this project which can be used to test a number of scenarios and policies in the future. While the current model is practice ready, there are number of ways the model can be improved in the future.

The first direction for future projects can start with improving the model by adding new components such as land price and households' salaries. Adding new components would increase the accuracy of the model and provide a better tool for policy analysis. Moreover, incorporating other modeling approaches like Meso and Micro models can forecasting accuracy. This is a long-term potential enhancement of the model. Conducting policy and scenario analysis is another avenue of future research. LS-LUM and the provided software offer a powerful tool for policy analysis, therefore many new policies can be tested at the level of the state. Here are some of the policies proposed for applying and comparing the results in the state of Tennessee:

- Changes in gas prices and their effect on land use and transportation
- Clustered zoning/densification versus sprawl
- Incentivized tax policies on the household, vehicle, and establishment ownership

- Technology, and built environment effect on land-use
- Shared-ride, transit, and other transportation initiatives on land-use planning.

In addition, the land use input data and forecasts can be reviewed by the local metropolitan planning organizations, transportation planning organizations, regional planning organizations, and other agencies in their respective jurisdictions. The quality of the land-use model depends on the input data and reasonableness of the forecasts. This project presents the first version of the statewide land-use model for the state of Tennessee that can be used for various planning projects, corridor studies, project impact analysis, input to travel demand models, and changes in land-use patterns along with socio-economic data over time (say from 2020 to 2050). The land-use model needs to be updated to incorporate changes in zone structure, and other facets of socio-economic data in the state. Figure 9-2 demonstrates the proposed road map for future studies on Tennessee Statewide Land Use Model.

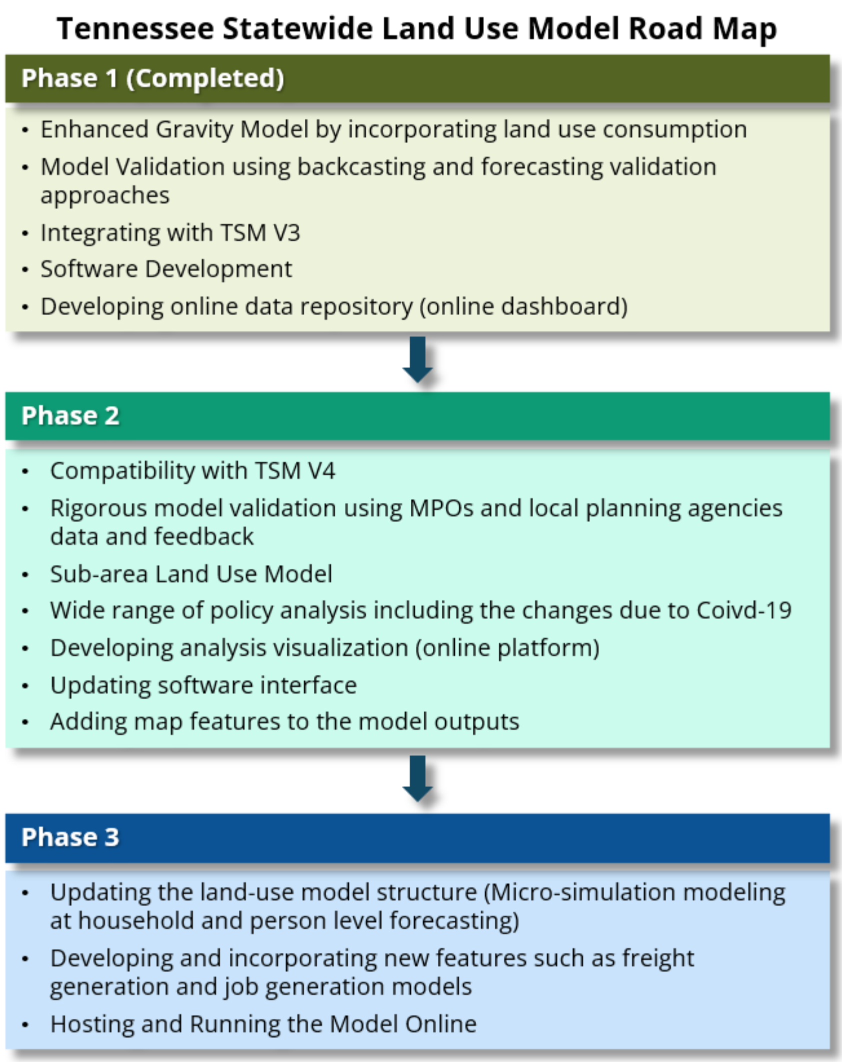

Figure 9-2. Tennessee land use model road map

References

Adler, T. J., Dehghani, Y., Doherty, M., & Olsen, W. (2007). *Florida's Turnpike State Model: Development and Validation of an Integrated Land Use and Travel Forecasting Model*.

Bernardin Jr, V. L., Ferdous, N., Sadrsadat, H., Trevino, S., & Chen, C.-C. (2017). Integration of National Long-Distance Passenger Travel Demand Model with Tennessee Statewide Model and Calibration to Big Data. *Transportation Research Record*, *2653*(1), 75–81.

Borning, A., Waddell, P., & Förster, R. (2008). UrbanSim: Using simulation to inform public deliberation and decision-making. In *Digital government* (pp. 439–464). Springer.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295–336.

Council, N. R. (2007). *National Land Parcel Data: A Vision for the Future*.

Crecine, J. P. (1968). *A dynamic model of urban structure*. RAND CORP SANTA MONICA CA.

Dowling, R. G. (2005). *Predicting air quality effects of traffic-flow improvements: Final report and user's guide* (Vol. 535). Transportation Research Board.

Goldner, W. (1968). *Projective land-use model (PLUM): A model for the spatial allocation of activities and land uses in a metropolitan region* (Vol. 219). Bay Area Transportation Study Commission.

Goldner, W. (1971). The Lowry model heritage. *Journal of the American Institute of Planners*, *37*(2), 100–110.

Hunt, J. D., Abraham, J. E., & Silva, D. D. (2009). PECAS—for Spatial Economic Modeling. *Calgary, Alberta: HBA Specto Incorporated*.

Iacono, M., Levinson, D., & El-Geneidy, A. (2008). Models of transportation and land-use change: A guide to the territory. *Journal of Planning Literature*, *22*(4), 323–340.

Lawe s., Lobb J., H. K. (2007). Statewide Land-Use Allocation Model for Florida, *11th TRB National Transportation Planning Application Conference, Daytona Beach, Florida*.

Lowry, I. S. (1964). *A model of metropolis*. RAND CORP SANTA MONICA CALIF.

Mishra, S., Ye, X., Ducca, F., & Knaap, G.-J. (2011). A functional integrated land use-transportation model for analyzing transportation impacts in the Maryland-Washington, DC Region. *Sustainability: Science, Practice and Policy*, *7*(2), 60–69.

Paul, V. V., & Zhou, B. (2009). Documenting for use of the Gravity-Based Land-use Model (G-LUM): With and Without In-Built Travel Demand Model Assumption.

Prastacos, P. (1986). An Integrated Land-Use—Transportation Model for the San Francisco Region: 1. Design and Mathematical Structure. *Environment and Planning A*, *18*(3), 307–322.

Putman, S. H. (1983a). Integrated Urban Models: Policy Analysis of Transportation and land use. *London, UK*.

Putman, S. H. (1983b). Integrated Urban Models: Policy Analysis of Transportation and land use. Pion Limited. *London, UK*.

RSG. (2014). *Tennessee Statewide Travel Model (Version 3) Development and Validation Report*.

Sivakumar, A. (2007). Modeling transport: A synthesis of transport modeling methodologies. *Imperial College of London*.

Southworth, F. (1995). *A Technical Review of Urban Land Use—Transportation Models as Tools for Evaluating Vehicle Travel Reduction Strategies*.

Spasovic, L. N. (2013). *TELUM-interactive Software for Integrated Land Use and Transportation Modeling. New Jersey Institute of Technology*.

Tennessee Comptroller of the Treasury. (n.d.). https://comptroller.tn.gov/office-functions/pa/gisredistricting/land-use-data.html

Thompson-Graves S., DuRoss M., Subhani R., Holloway B., and S. E. (2012). *"Development of the State Smart Transportation Initiative's DelDOT 3-D Micro Model Process—A scenario planning tool to evaluate urban form, land use, and multimodal investment impacts on mobility." State Smart Transportation Initiative. Retrieved 02/28/20*.

Valsaraj, V., Kockelman, K., Duthie, J., & Zhou, B. (2007). *Forecasting Employment & Population in Texas: An Investigation on TELUM Requirements, Assumptions, and Results, Including a Study of Zone Size Effects, for Austin and Waco Regions*. Austin: http://www. ce. utexas. edu/prof/kockelman/G-LUM_Website ....

Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, *68*(3), 297–314.

Wegener, M. (1998). *The IRPUD model: Overview*.

White, W. l. (2010). *FSUTMS/Land Use Model Integration. Status Update. Presented to the Florida Model Task Force*.

Zhao, F., & Chung, S. (2006). *A study of alternative land use forecasting models*.

Zhou, B., Kockelman, K. M., & Lemp, J. D. (2009). Applications of integrated transport and gravity-based land-use models for policy analysis. *Transportation Research Record*, *2133*(1), 123–132.