

2050 Land Use and Socio-Economic Data Forecast for the French Broad River Metropolitan Planning Organization

Executive Summary

The 2050 Land Use and Socio-Economic Forecast for the Greater Asheville Region, prepared for the French Broad River Metropolitan Planning Organization (FBRMPO), presents a comprehensive projection of population, employment, and land use change across a five-county area encompassing Buncombe, Haywood, Henderson, Madison, and Transylvania counties. This forecast supports the development of the region's long-range Metropolitan Transportation Plan (MTP) by providing scenario-based insights into future growth patterns and transportation needs.

Current projections assume steady regional growth, with the number of households increasing from 223,100 in 2020 to 297,091 by 2050, and employment rising from 248,972 to 317,553 over the same period. Four alternative land use scenarios were analyzed, representing different ways that growth might be spatially distributed across the region:

1. **Baseline** – A “business as usual” projection based on recent trends and pre-existing policy frameworks.
2. **Consolidated** – A scenario favoring compact, walkable development in transit-accessible areas, aimed at reducing vehicle miles traveled (VMT) and supporting environmental goals.
3. **Dispersed** – A low-density, more rural growth pattern that assumes stricter development limits and encourages greenfield expansion.
4. **Accelerated** – A high-growth scenario that doubles the household growth rate, modeling a more aggressive housing and labor retention policy.

Analysis shows that the **Consolidated** scenario performs best in terms of transportation efficiency, reducing regional VMT by 9% compared to the Baseline by 2050. This reduction implies substantial benefits for air quality, public health, and infrastructure efficiency. Conversely, the Dispersed scenario yields slightly higher VMT than the Baseline due to longer average travel distances, while the Accelerated scenario results in significantly higher VMT due to the increased number of households, despite a reduction in external commuting. As a result of this analysis, the FBRMPO Board adopted the **Consolidated** scenario as the preferred regional growth strategy, prioritizing investment in infrastructure and development policies that support concentrated growth in urban centers and walkable neighborhoods.

Project Overview

Background

The French Broad River Metropolitan Planning Organization (FBRMPO) is the Federally designated long-range transportation planning agency for the Greater Asheville region. Local governments in the MPO planning area include: Buncombe, Haywood, Henderson, and Madison Counties, and the municipalities of Asheville, Biltmore Forest, Black Mountain, Canton, Clyde, Flat Rock, Fletcher, Hendersonville, Laurel Park, Maggie Valley, Mars Hill, Mills River, Montreat, Waynesville, Weaverville, and Woodfin. The FBRMPO is staffed by the Land of Sky Regional Council and includes the Land of Sky Rural Planning Organization, which also serves Transylvania Council, parts of Madison County, and the municipalities of Brevard, Rosman, Marshall, and Hot Springs.

The North Carolina DOT maintains a travel forecasting model for the Greater Asheville Region, including Buncombe, Haywood, Henderson, Madison *and* Transylvania counties. As such the study area of this travel model encompasses both the FBRMPO and Land of Sky planning areas. Since, as described in following sections, the purpose of the 2050 socio-economic land use forecast study is in part to support traffic modeling activities, the same study area is used for both land use forecasting and travel modeling.

Goals & Objectives

The FBRMPO serves the important role of prioritizing transportation improvement projects within its planning area, and developing a long-range Metropolitan Transportation Plan (MTP) every five years. The MTP helps establish a framework for coordination with local (County and City) as well as State government (e.g. NCDOT) partners. The FBRMPO develops its MTP based upon deficiencies and needs identified using the region's travel model, assuming a specific growth scenario, which is selected from among several alternatives.

The objective of the 2050 Socio-Economic Land Use Forecast Study is to develop four alternative land use scenarios and refine these quantitatively to a sufficient level of detail for evaluation using the region's travel demand model, so that high-level performance comparisons can be made. The FBRMPO board then adopts one future land use scenario as its preferred growth pattern, which is used as the basis for project prioritization and needs assessment through the MTP process.

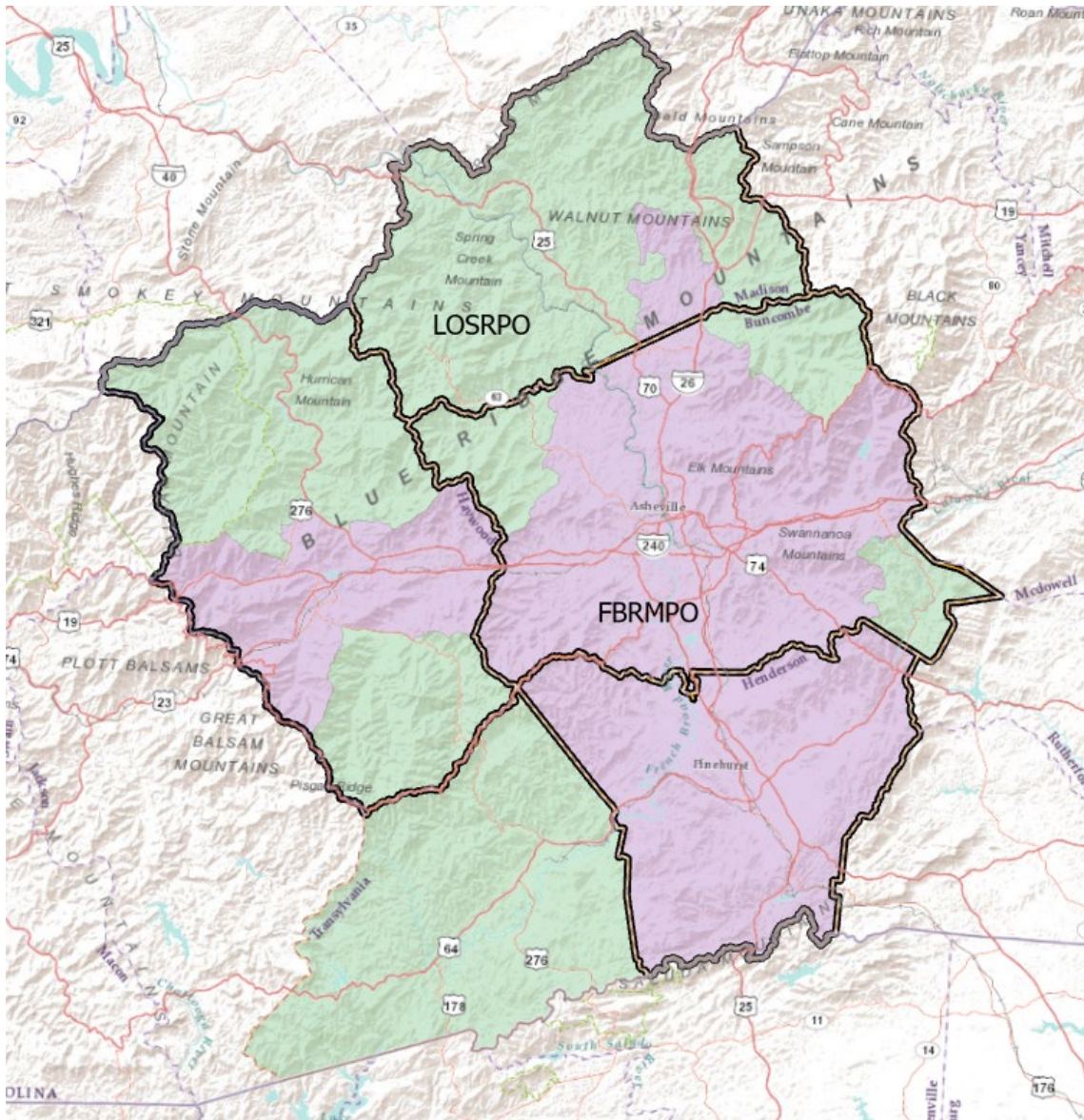


Figure 1. Relationship between the forecasting model study area, FBRMPO planning area, and Land of Sky Regional Planning Organization (LOSRPO) areas.

Unique Challenges

During the conduct of this study, flooding and avalanches caused by Hurricane Helene decimated parts of the FBRMPO and Land of Sky planning areas, leaving many residents without water and power. The storm caused extensive property damage; entire neighborhoods, such as Asheville's River City Arts District, faced an existential threat. We owe a debt of sincere gratitude to the local staff who kept working with us despite intense disruption in their personal lives, and did their best to get us access to information we could use to make educated guesses about what the land use pattern might begin look like as the region recovered from the storm. Though not in our original scope of work, the consulting team dedicated extra effort to gathering a basic understanding of the storm and its impacts, in order to include some representation of effects in scenario forecasts.

In addition to the above, it is worth noting that this forecast is the first of its kind since the COVID-19 pandemic, which greatly disrupted normal economic activity and travel patterns in 2020. This complicates normal operating procedures: whereas a decennial Census year would typically offer the best possible opportunity to re-calibrate travel and land use forecasting models, the 2020 Asheville region travel demand model is in practice based more upon late 2019 travel patterns, not mid-pandemic conditions. Similarly, the timeframe between 2014 and 2019 provided cleaner comparisons for gauging the sensitivity of development location choices to travel conditions. Since our land use forecast model was developed in 2024, and both Census and TAZ geographies changed in 2020, making direct comparisons between 2019 and 2024 difficult, we did not calibrate our land use forecasting models to post-COVID-19 data. However, we did compare the models' output to available socio-economic data for 2024 and make adjustments to our scenario forecasts based on the results of these comparisons.

Input Data & Assumptions

Population Forecast

Our study utilized population projections provided by the North Carolina State Demographer as a starting point for developing forecasts of households and employment in the region.

At the direction of FBRMPO staff and the FBRMPO Prioritization Sub-Committee, (serving as a technical advisory committee for this study) the consulting team adopted the State Demographer's population projections for all but one scenario, an "Accelerated" growth condition. Consistent with a similar scenario analyzed in the previous EPA-funded resiliency study, the population growth rate was doubled under this scenario, while the employment growth rate remained the same as in the other

scenarios. This was intended to capture a hypothetical case in which the region is able to aggressively recruit and retain workers (perhaps through housing policy or quality-of-life improvements), achieving a more even local balance between jobs and labor supply.

The state projections do not provide information about employment or households. Therefore, we purchased other economic forecast data: the Comprehensive Economic and Demographic Data Source (CEDDS) offered by Woods and Poole, and population, household, and employment data provided by Moody's Analytics (Economy.com). In our previous study, Moody's projections were found to be the most appropriate source for the FBRMPO land use and travel model study area. Both private forecast sources show lower population growth than the State Demographer's forecast; our employment forecast was therefore adjusted to compensate for the differences.

Existing Land Use

FBRMPO staff maintain a parcel-level GIS layer of existing land use conditions, named "ELUSE", which aggregates information from local partners (i.e. counties and cities), applying a simplified land use classification scheme which has remained highly consistent over more than a decade. A table describing this land use classification scheme is shown in Figure 4.

In addition to classification of current land use for each parcel, fields in the layer indicate when and if parcels have been split. The 2020 ELUSE data provided to Manhan by FBRMPO include data for 2015 and 2010 land use designations as well.

U.S. Census Data

Many publicly available Census datasets provide additional understanding of current and recent socio-economic land use patterns in the five-county study area.

Census Tabulation Blocks

The U.S. Census Bureau publishes housing and population counts at a Tabulation Block level for each decennial Census year (e.g. 2010, 2020). These provide a helpful point of comparison to the parcel-level 2020 existing land use data, though the ELUSE parcel boundaries do not precisely align to 2020 Census Block boundaries. The 2020 Census Block boundaries also differ from the 2010-vintage Census Blocks, which form the basis of block group and higher tabulations used for understanding historical trends between 2010 and 2019.

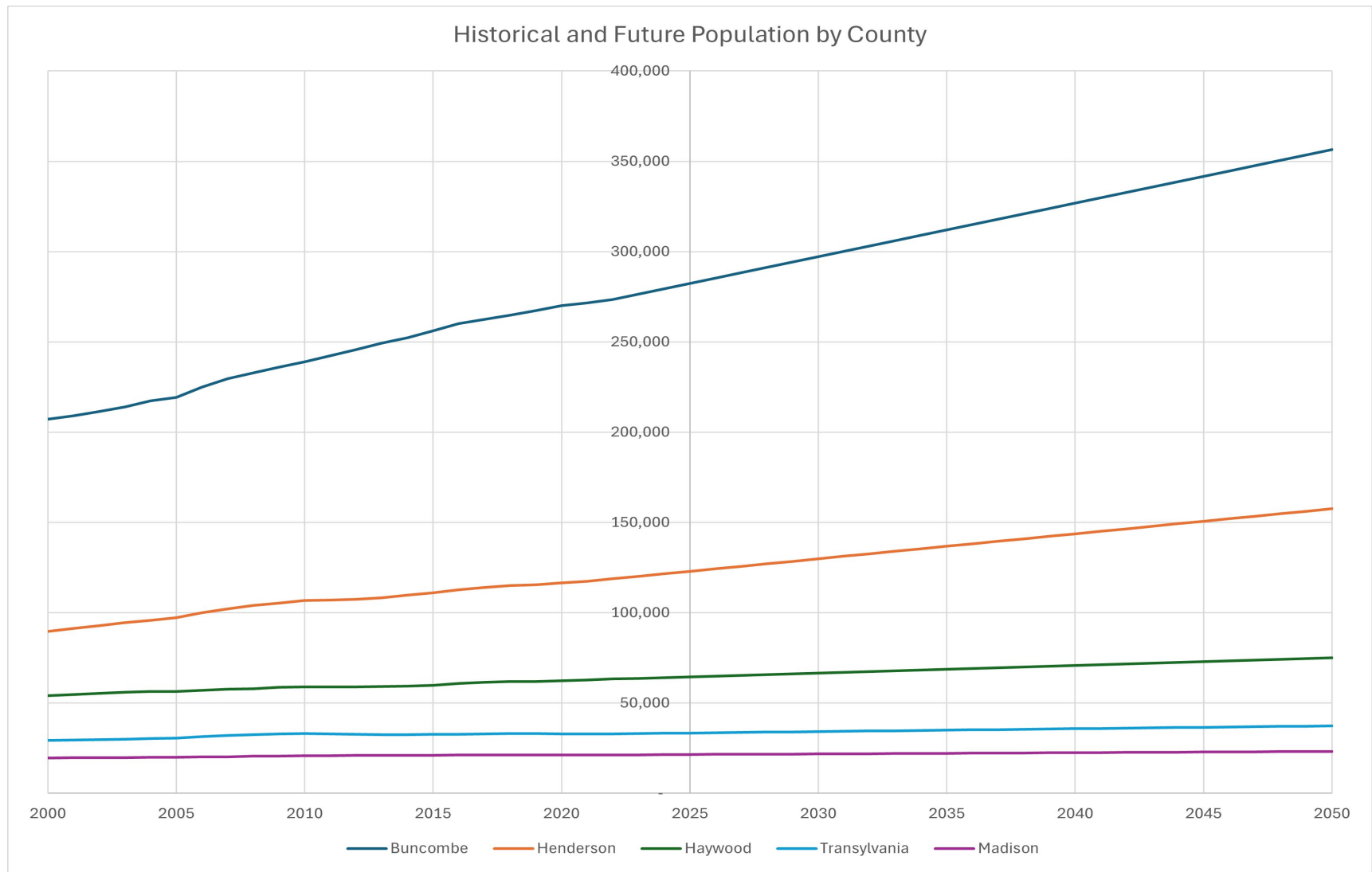


Figure 2. North Carolina State Demographer population projections for the Study Area

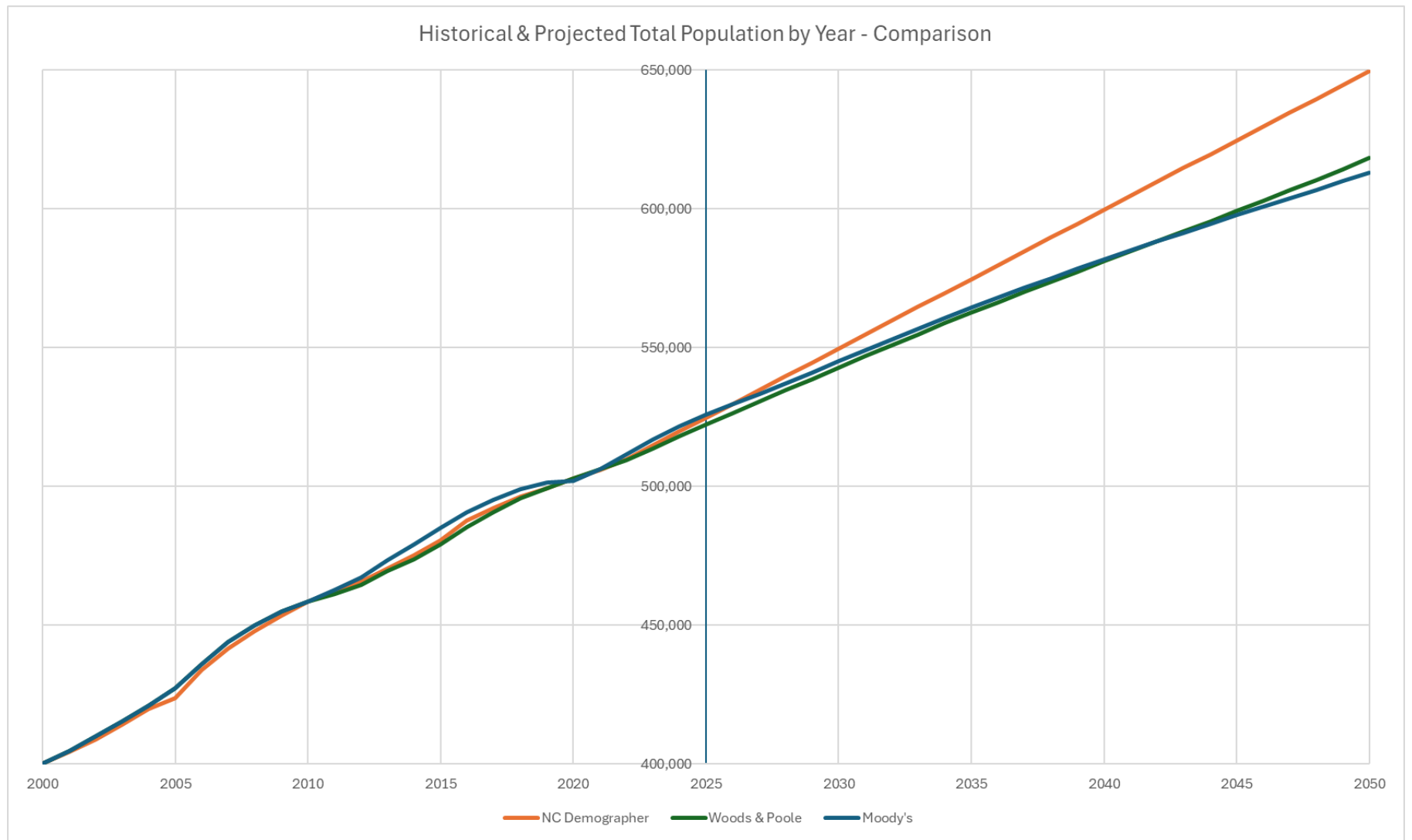


Figure 3. Comparison between state and private population projections for the study area.

Aggregate Type	RLUC	Description
Undeveloped	VACANT	Other land without structures
	UTILITY	Well land, water/sewer right-of-way, powerline and other easements; intakes
	FARM	Farmland and Forestland (horticultural and forestry activities)
Undevelopable	WATER	Surface Water (lakes, rivers, ponds)
	NATARA	Protected natural area, mostly undeveloped parks and protected natural areas, excluding lake surface area
	GRNSPC	Developed parks, golf courses, cemeteries
	RDROW	Road right-of-way (not parceled)
Residential	GROUP	Group Quarters (see census definitions)
	SENIOR	Senior Living to see text for description of facilities
	RR	Rural residential (0.50 du/ac)
	VL	Very low residential (0.51 to 1.00 du/ac)
	L	Low residential (1.01 to 3.00 du/ac)
	ML	Low-medium residential (3.01 to 5.00 du/ac)
	M	Medium residential (5.01 to 8.00 du/ac)
	MH	Medium-high residential (8.01 to 12.00 du/ac)
	H	High residential (12.01 to 16.00 du/ac)
	VH	Very high residential (16.01 to 20.00 du/ac)
Industrial	UH	Urban high residential (Over 20.00 du/ac)
	IND	Industrial (Factories, warehouses, storage units,)
Commercial	SERVICE	Service (Restaurant, Bar, Club, Cafeteria, fast-food)
	HWYRET	Highway Retail (Retail adjacent to highway to typically Gas stations, mini-marts, fast-food).
	RETAIL	Retail (Department stores, boutiques, auto shops/dealers, banks)
	MALL	Regional Mall (High Density Retail)
	ENT	Entertainment (Bowling alleys, gymnasiums, theatres)
Lodging	COMSTP	Commercial Strip (typical strip center with a combination of retail, highway retail and service jobs).
	LODGING	Hotels/Motels/Inns
Office	OFFICE	Offices
	HIOFFC	Highrise Office (>3 floors or exceptionally high FAR)
Mixed	MU	Mixed use parcels
Institutional	HOSP TL	Hospital or Medical Complex
	CIVIC	Civic Buildings (churches, Camps, community centers etc.)
	GOV	Community Infrastructure or building that is typically owned by a government entity not a school or handled in the special classification.
	SCHOOL	School (public and private schools serving grades K-12, post-secondary schools and universities)
Special	SPECIAL	Special Class (airport, landfill, quarry, power plant, wastewater treatment plant, parking deck or other requiring special consideration)

Figure 4. ELUSE land use classification scheme.

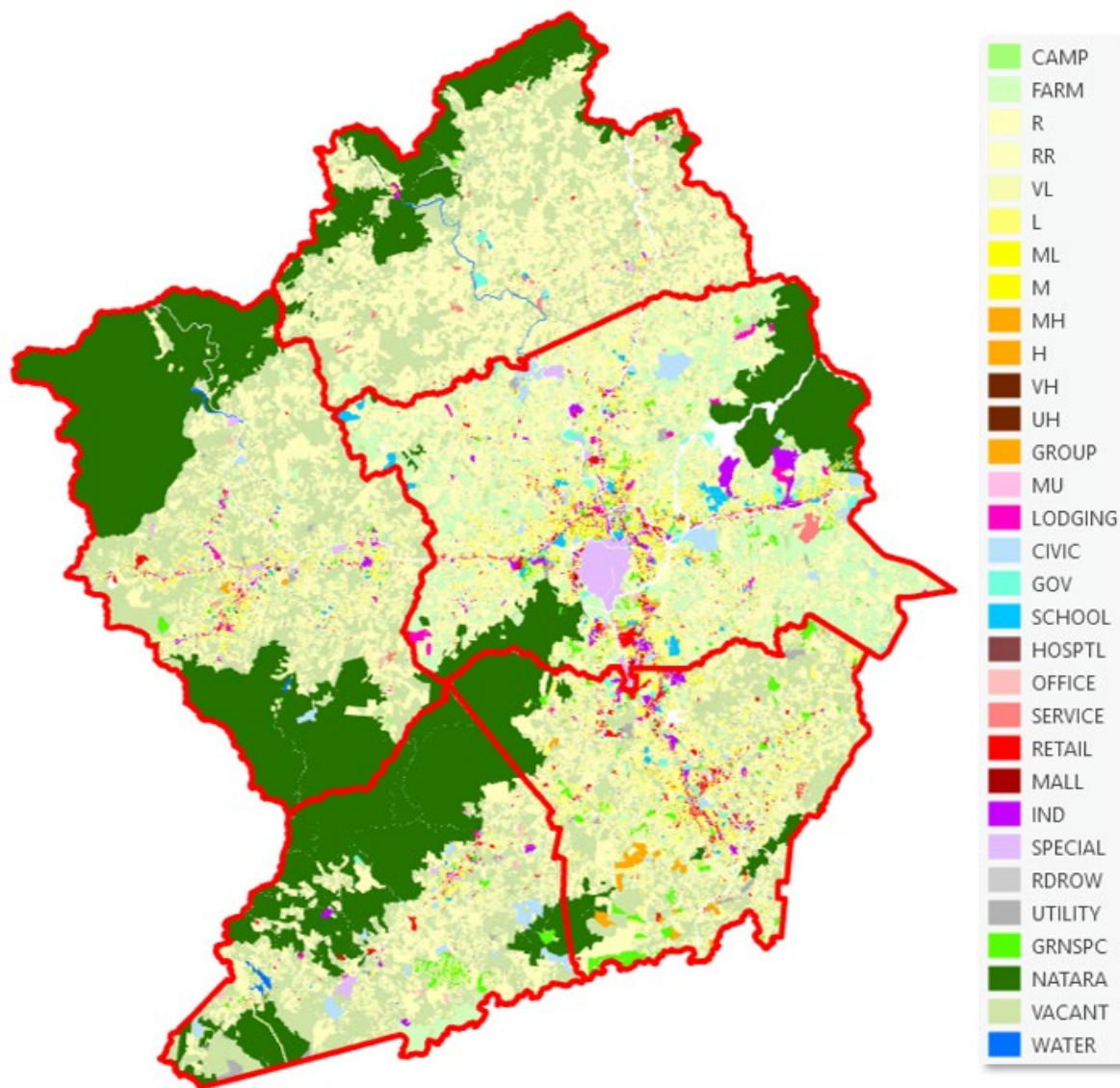


Figure 5. Parcel-level ELUSE map for the five-county study area.

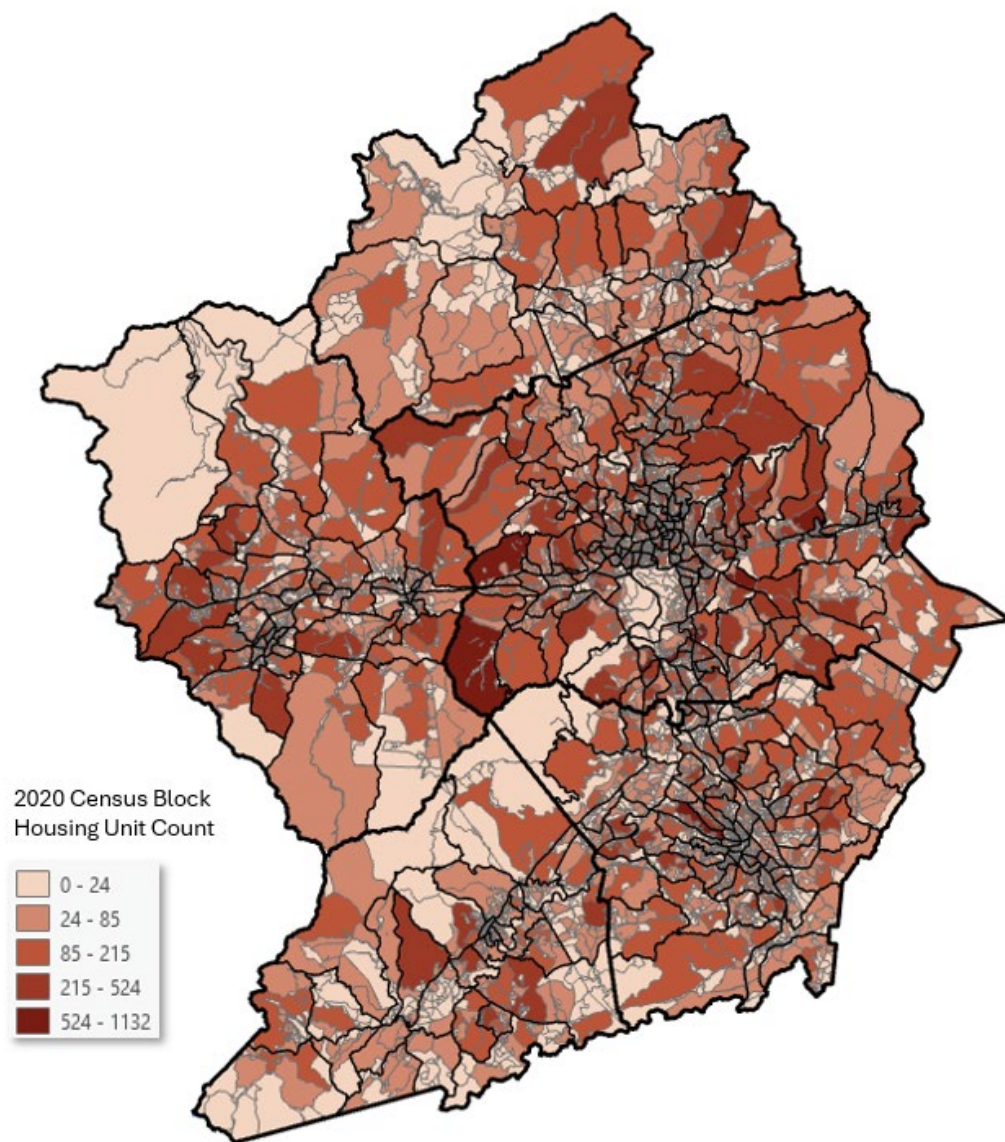


Figure 6. U.S. Census 2020 Block Groups and Blocks with Housing Unit Counts

U.S. Census American Community Survey (ACS) Block Group Tabulations

Many useful statistics, such as housing units by number of units in structure, or household counts by size and income category, are available at a block group level from the U.S. Census Bureau. One-year and five-year rolling estimates are published annually for most data series. ACS block group data downloaded and used by the consulting team include:

- Housing units by number of units in structure
- Households by income group
- Median household income

U.S. Census ACS Public Use Microdata Statistics (PUMS)

In addition to block group summary statistics, anonymized individual records from the ACS are available for Public Use Microdata Areas (PUMAs), providing access to customized cross-tabulations and statistics not available through standard block group tabulations. For example, these data allow deeper investigation of vehicle ownership rates, or number of school-age children, by household size and income.

Environmental Protection Agency (EPA) Smart Location Database

The Smart Location Database (or SLD) is a special product developed by consultants to the EPA which leverages U.S. Census ACS block group tabulations and other data sources (such as General Transit Feed Specification, or GTFS data, as well as routable highway networks from the navigation firm HERE) to calculate many metrics which have been found useful by planners. The most recent available edition of this dataset is 2019.

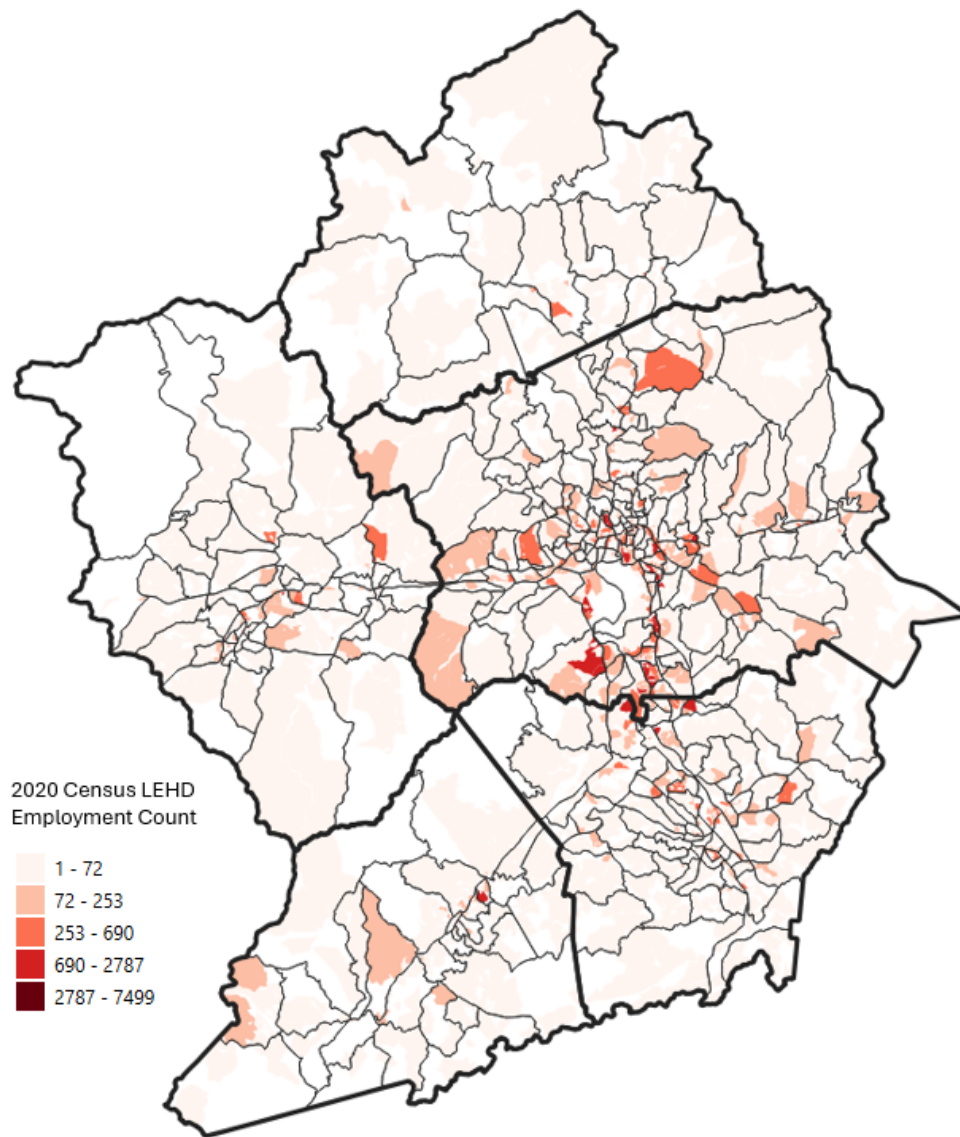


Figure 7. Block-level 2020 LEHD employment statistics with block group and county boundaries

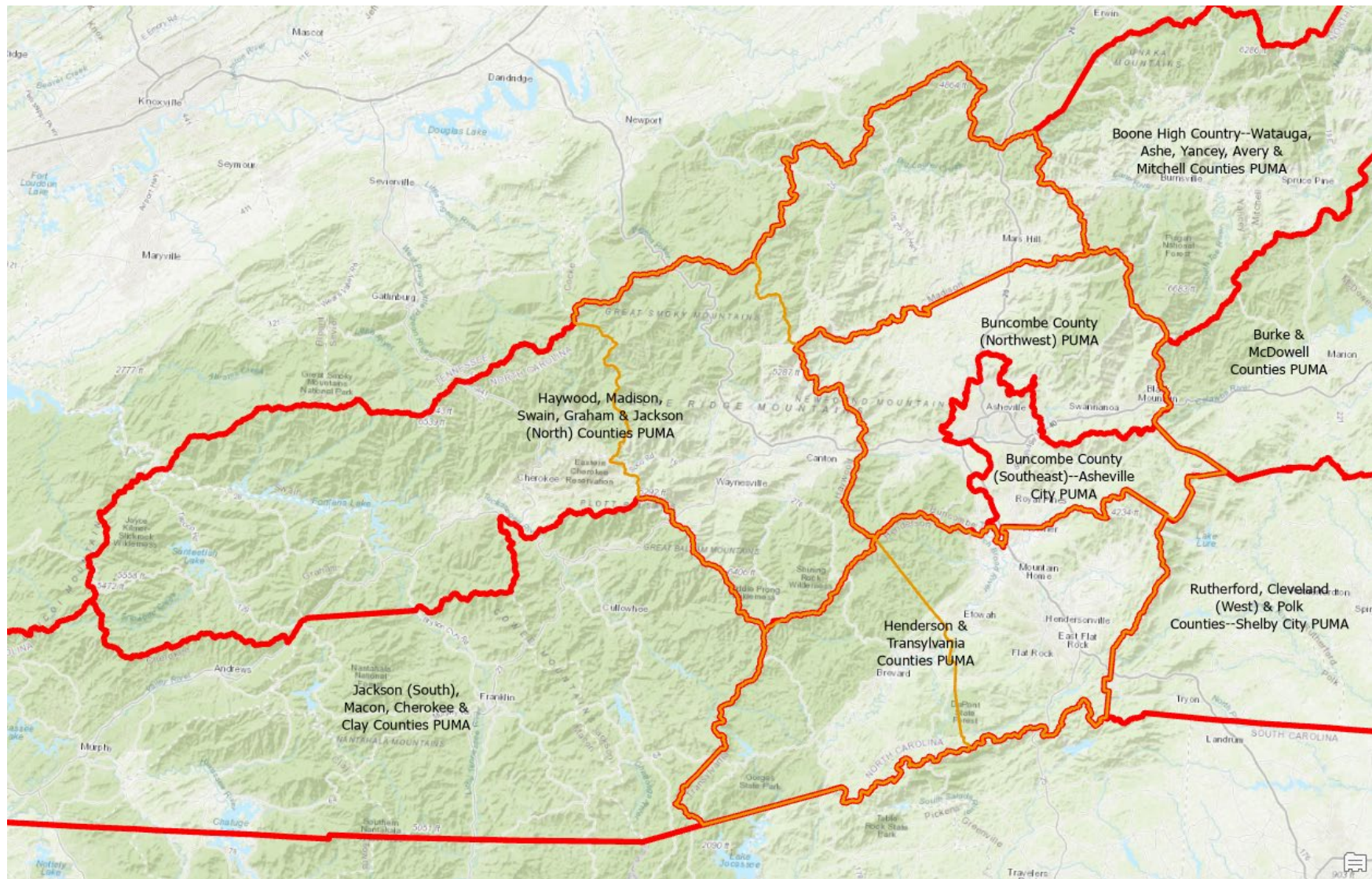


Figure 8. U.S. Census 2010 Public Use Microdata Areas for the FBRMPO region

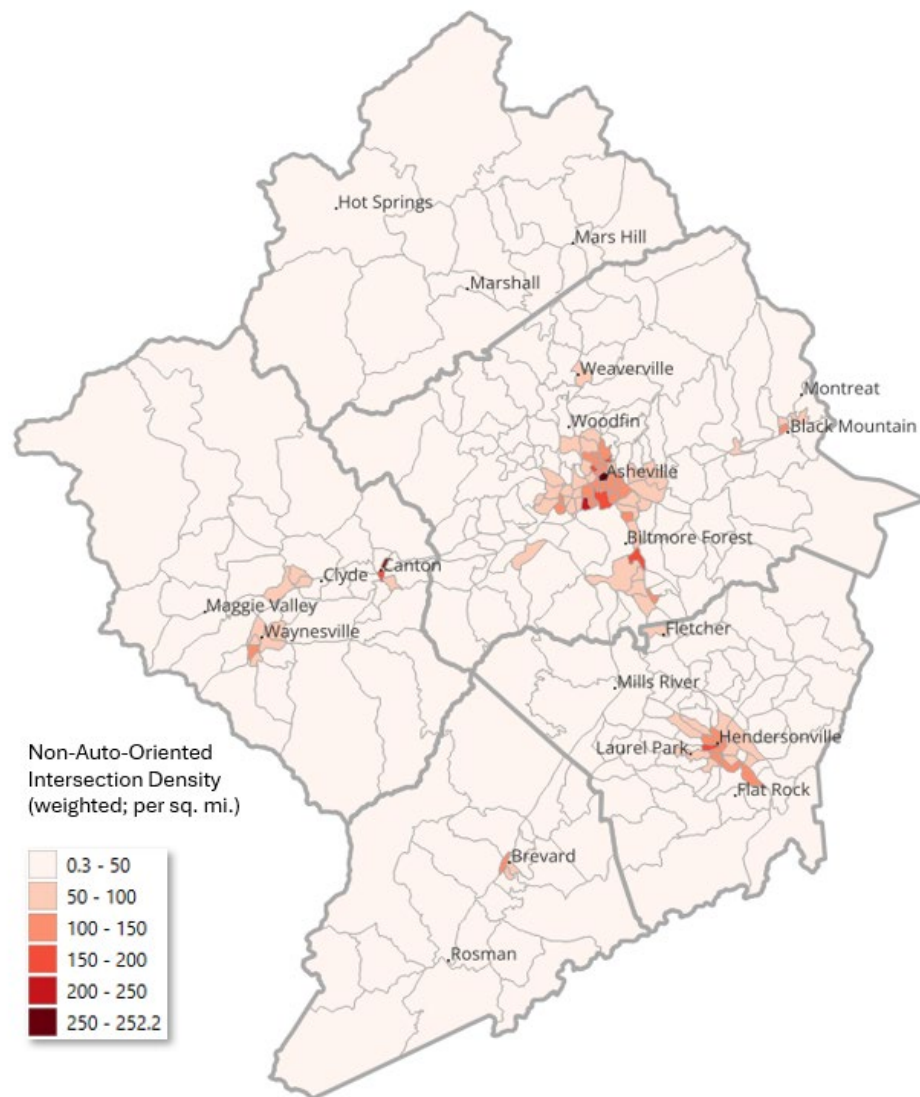


Figure 9. EPA SLD intersection density data at block group level.

InfoUSA Business Establishment Point Data

The region's travel demand forecasting model requires as input employment by transportation analysis zone categorized according to super-sectors that are defined as aggregations of three-digit NAICS sectors. This requirement cannot be met by the Census LEHD/LODES employment data noted above, so NCDOT purchased, processed, and made available to the consulting team a privately-developed dataset of business establishment points with five-digit NAICS classification and job counts by site. This dataset formed the basis for compiling base year (2020) employment data for input to the region's travel demand model

Travel Model Base Year

NCDOT and FBRMPO cooperatively develop and maintain a regional Transportation Analysis Zone (TAZ) layer and associated socio-economic data for input to the regional travel demand model for the five-county study area. Prior to the consulting team's notice to proceed, the boundaries of this layer were re-aligned with 2020 Census Tabulation Blocks, for easier comparison between public and private data sources.

It is worth noting that the 2020 travel model base year is closer to 2019 conditions in reality, due to the exceptional and anomalous nature of employment and travel patterns during the COVID-19 pandemic, whose effect was most strongly felt during 2020. The travel model base year condition might be thought of as a counterfactual case in which the COVID-19 pandemic never happened. Though non-essential employment levels did respond very quickly to shutdowns and furloughs triggered in response to the pandemic, and some households made location shifts to facilitate remote work or outdoor recreation during this time, those shifts are probably not captured in the base year data.

The previous (2015) generation travel model was also used to some extent, despite being based on a different TAZ system. Specifically, we derived measures of household access to jobs and labor market access using the calibrated 2015 congested origin-destination highway skim tables and 2015 socio-economic input data. This allowed us to create a predictive model with five-year time lags, such that 2020 transportation network conditions influence 2025 land use, 2025 affects 2030 land use, and so on to the 2050 planning horizon. Standard attribute transfer techniques were applied using GIS software as needed to relate 2015 accessibility measures to 2020 TAZs.

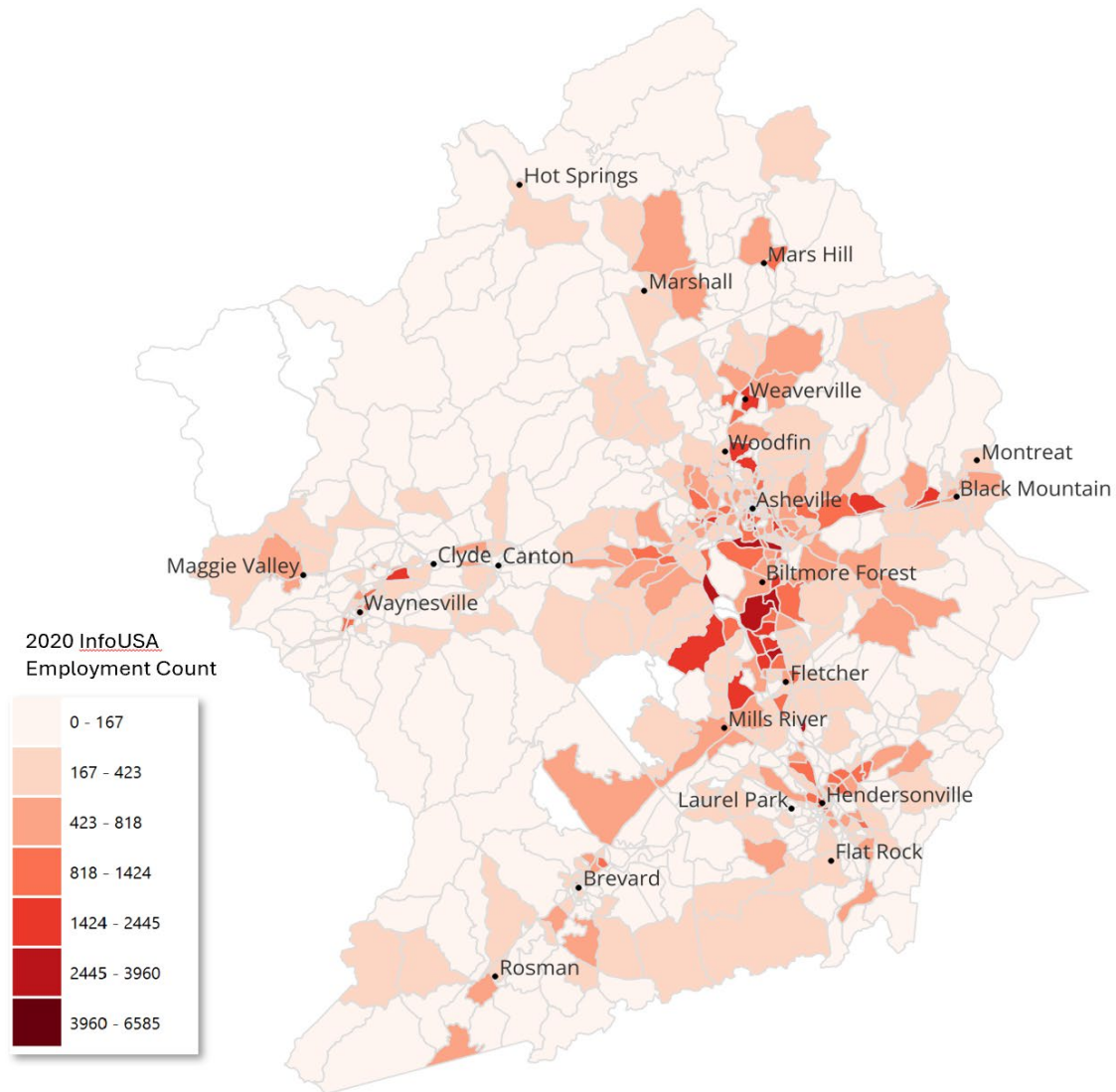


Figure 10. 2020
Transportation Analysis
Zones with InfoUSA job
counts

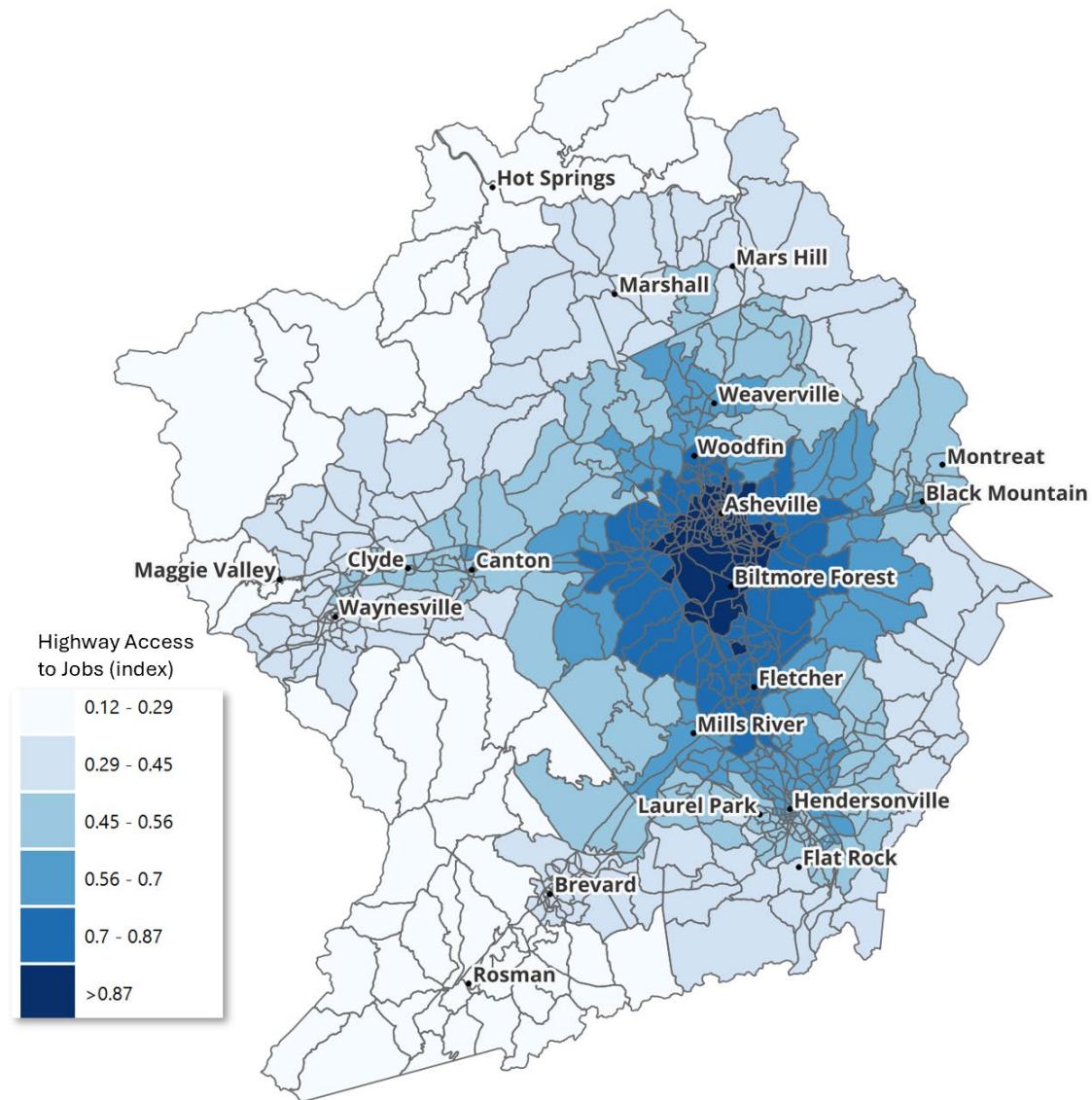


Figure 11. Year 2015
highway access to jobs,
transferred to 2020 TAZ
system

Helene Impacts

In late 2024, Hurricane Helene triggered extensive flooding and landslides throughout the study area, affecting the region's land use patterns. In the immediate aftermath, and as of this writing, the impacts of this natural disaster on land use remain incompletely understood. None of the publicly available data sources mentioned previously (i.e. the 2020 ELUSE data, U.S. Census Data, and 2020 base year travel model input data) capture the impact Helene may have had or will continue to have on land use patterns, moving forward. To provide the consulting team with a better understanding of these impacts, FBRMPO staff arranged to grant limited access to a GIS data source known as ICEYE, which allows for detailed mapping of inundation and flood extent during the hurricane event. We combined these data with unemployment statistics and insights gathered from the FINMAP resources available to the public to build a simulation of hurricane-induced land use impacts, akin to a “digital twin” in that it represents a model of a process in motion, rather than a future or historical condition.

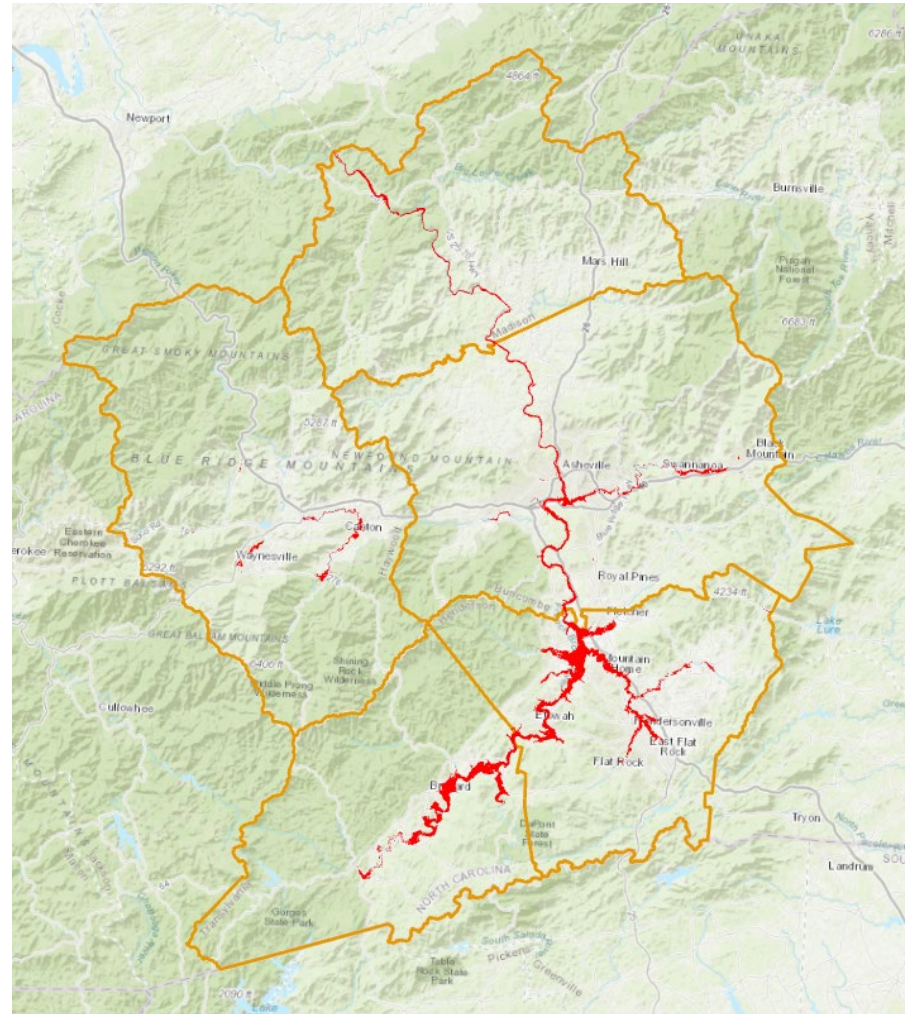


Figure 12. Helene flood extent (in red, Source: ICEYE)

Land Use Policies & Regulations

GIS data were provided by FBRMPO to allow mapping of applicable zoning codes where available throughout the study region. Rather than attempt to translate and hard-code these restrictions as explicit rules in our allocation model, we used them as input to the land use change model, which is described in a later section. Zoning data were obtained for the following jurisdictions:

- Asheville
- Biltmore Forest
- Black Mountain
- Brevard
- Buncombe County
- Flat Rock
- Fletcher
- Haywood County
- Henderson County
- Hendersonville
- Hot Springs
- Laurel Park
- Madison County
- Mars Hill
- Mills River
- Montreat
- Weaverville
- Woodfin

The consulting team, with help from FBRMPO staff, also obtained and reviewed comprehensive plans and future land use maps for much of the study area. We found that most such documents offered only qualitative insights regarding aspirational goals held by cities and counties, not hard indications of likely growth patterns. Thus these plans were not directly input to our forecast, but rather relied upon to develop qualitative understanding of the region.

Future Real Estate Development Projects

In most regions, local planners can enumerate a list of real estate development projects that are highly likely to be built within a short-term time range (e.g. 5-10 years). These include permitted projects as well as those that may not have a permit but are nonetheless widely considered to be likely to move forward through the planning and permitting process. The Manhan team working on the 2045 socio-economic land use forecast for the FBRMPO region previously gathered a list of such projects, some of which had expected build dates falling within the post-2020 timeframe of the current study.

For the current study, we planned to use a web-based development data crowdsourcing tool to QA/QC and validate our previously gathered development project list. Unfortunately, our release of this tool to local partners coincided with the Helene event and aftermath, making it almost impossible to solicit local input. Instead, we performed some in-house filtering of our project list and included it as an “index” of recent trends which would influence, but not dictate, short-term allocations. Later, by January 2025, local partners were in fact able to provide some information regarding likely near-term real estate development projects; while the non-residential component of these were too qualitative to include explicitly we did incorporate housing projects from this list in our interim future household allocation process.

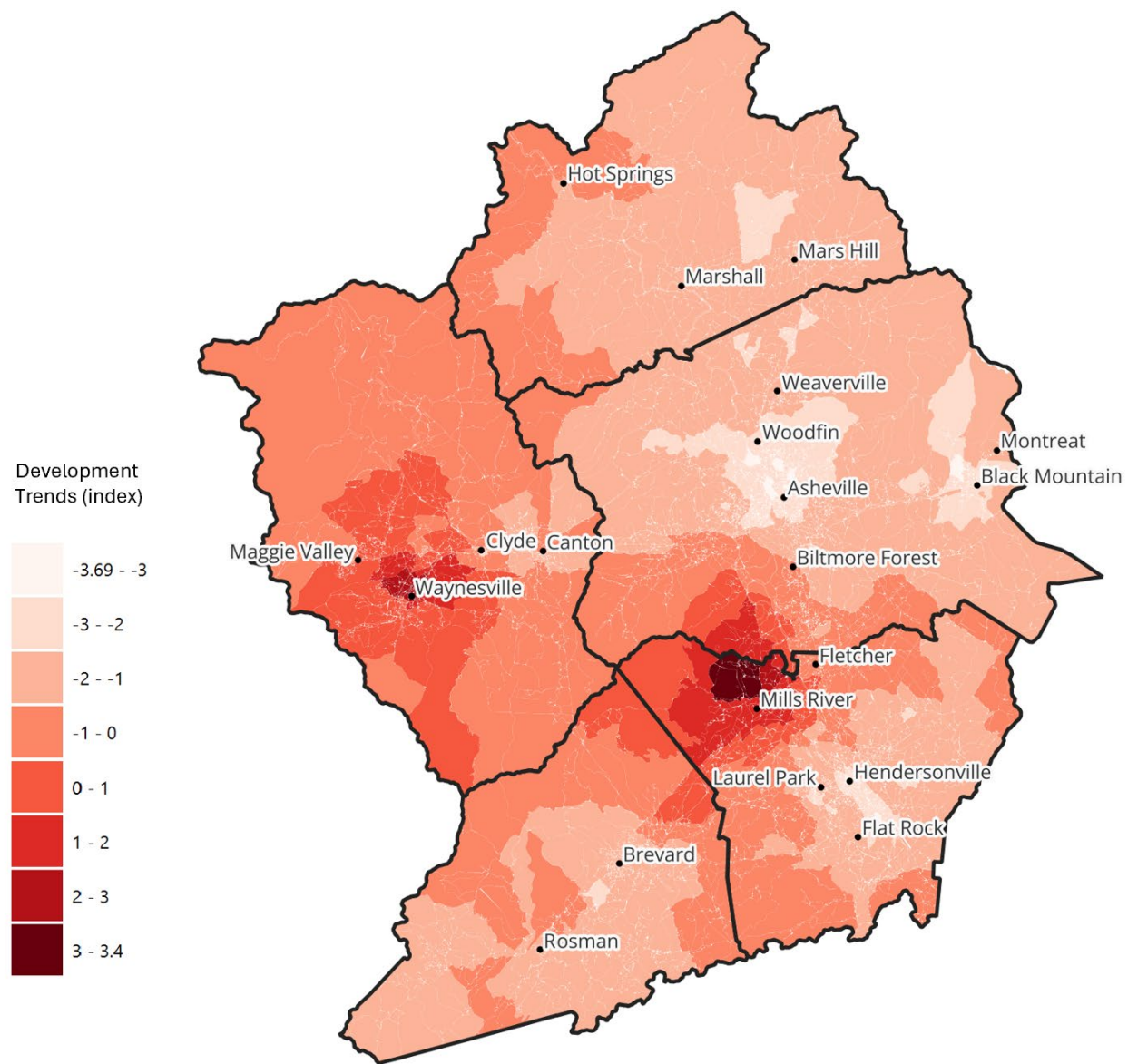


Figure 13. Development trends index, based upon previously collected inventory of upcoming real estate projects.

Forecast Methodology

Manhan developed a data-driven approach to preparing socio-economic land use forecasts for the five-county Greater Asheville study area. In so doing, we created a system of interrelated sub-models, which we shall refer to as the Asheville Region Land Use Model (ARLUM). The diagram in Figure 14 provides a high-level overview of the ARLUM system. The following sections provide more technical details about each of the sub-models as well as input and output data flows.

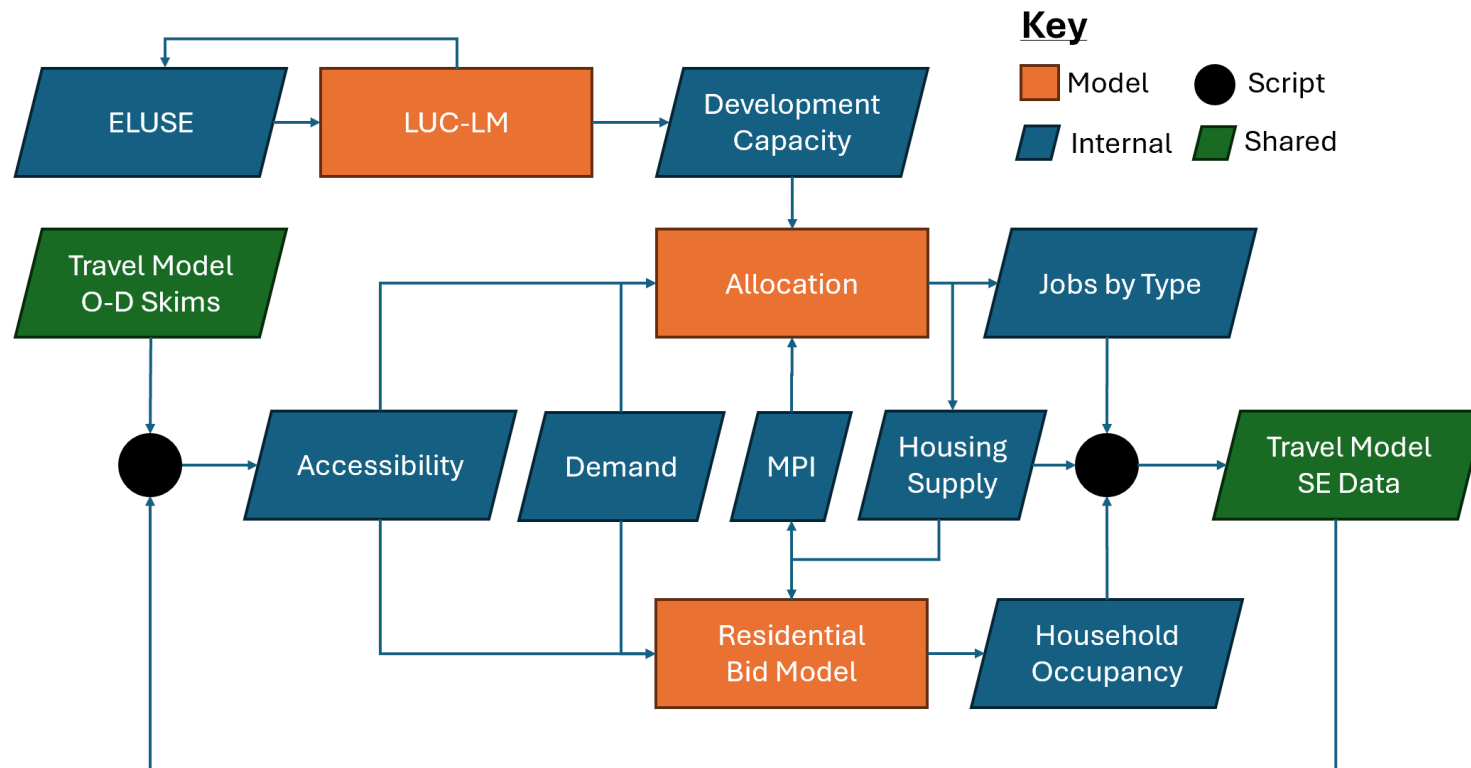


Figure 14. High-level ARLUM process schematic.

Employment & Household Forecasts

As noted previously the North Carolina State Demographer does not provide projections of households, nor jobs. Meanwhile, the privately-developed employment and household forecasts that we obtained were based upon population projections that differ from the state's assumptions adopted as the basis for our study. Therefore, we applied the following methodology to adjust the Moody's household and employment forecasts for greater consistency with the State's projections:

1. An adjusted household forecast was computed by multiplying the State population forecast by the average household size (persons per household) given by Moody's.
2. The adjusted employment forecast was then computed by multiplying the adjusted households forecast by the jobs per household ratio found given by Moody's data.

Woods and Poole data were mainly used to further disaggregate the adjusted Moody data. For example, the Woods and Poole CEDDS provides employment projections for 20 industry sectors, as well as breaking out its household forecast by income group.

Prior to allocation, the control totals described above are converted into five-year increments of change, or "deltas", for allocation. This allowed for an allocation process that is additive, rather than multiplicative, starting with the base year spatial distribution of socio-economic land use gathered by NCDOT and FBRMPO, and adding or subtracting increments of change, rather than applying growth rates or shifts in shares of growth over time. Increments of change may be negative under either of two conditions:

5. In the housing market, there is an underlying "churn" of housing being abandoned at the same time as new housing is built, which results in "hidden" declines that are masked by an overall increase in the supply of housing stock across the region. By quantifying this churn and allocating some negative as well as positive growth increments, we are able to capture localized declines as well as redevelopment and housing stock turnover in more dynamic neighborhoods.
6. Certain employment sectors are projected to have declines during the forecast period; in some cases only temporarily, during specific five-year time increments, or in other cases throughout the entire forecast series. We allocate negative increments in these cases; however we do not attempt to quantify and allocate churn in general for the non-residential real estate market.

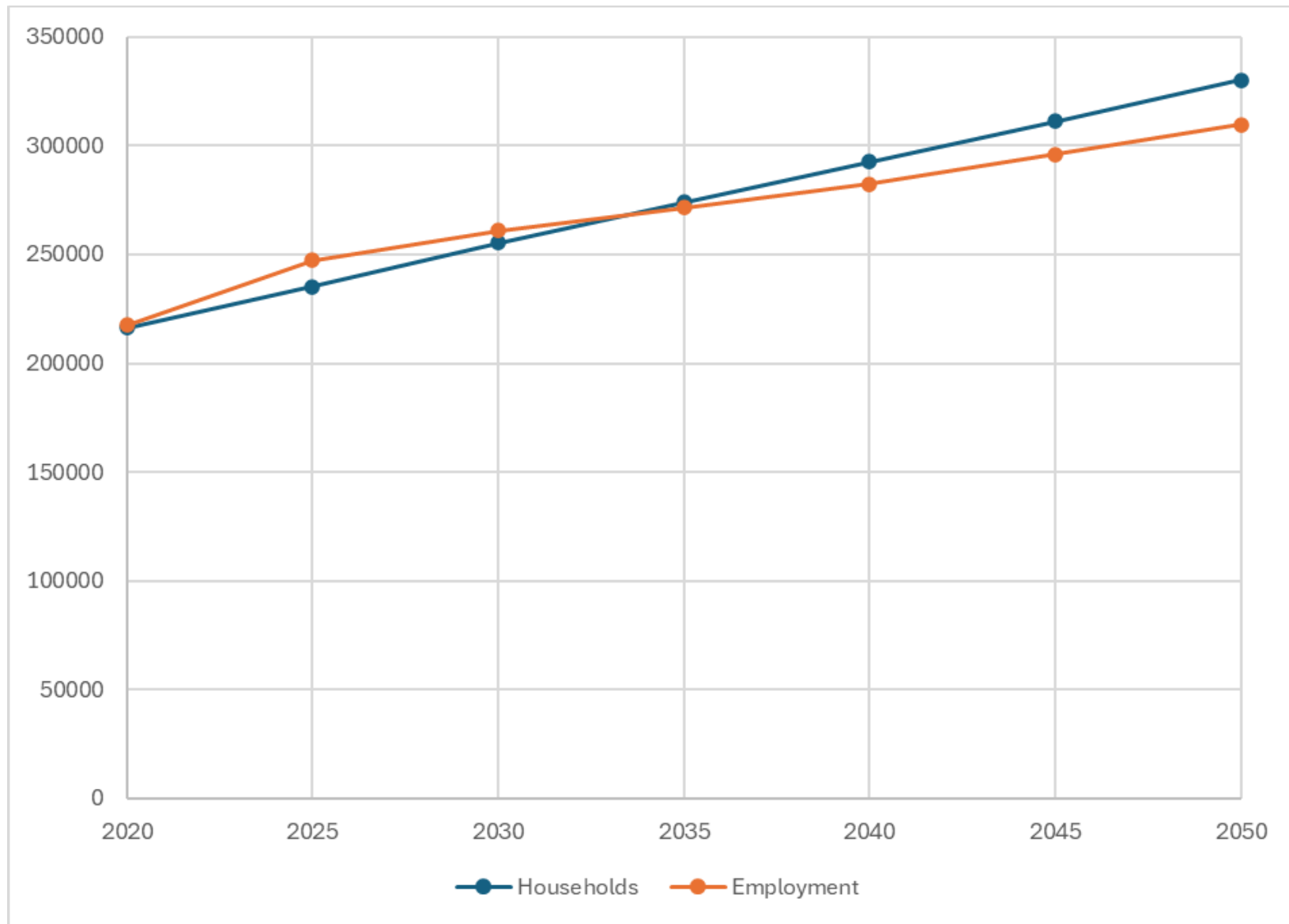


Figure 15. Total jobs and households, 2020-2050, five-county study area

Residential Bid-Auction Model

Bid-rent theory is a well-established core principle of urban economics which holds that land, and property built on that land, tends to be occupied to the highest-paying use in a competitive urban real estate market with diverse buyers who value property and location characteristics differently. This principle is embedded in the “highest and best use” criterion used by appraisers to determine how to value property.

The bid-*auction* approach to land use modeling is derived from bid-rent theory, providing a set of mathematical equations for predicting the outcomes of a property auction. In effect, it models the seller’s choice of tenant or the landlord’s choice of buyer. In either case a key factor determining the outcome of such an auction (which may be implied, rather than a real “auction” as for livestock or on eBay) is the buyer’s willingness and ability to pay, and theoretically, the price paid by the buyer or renter is the “expected maximum bid” among all those that might be placed by potential occupants of a property. A complete bid-rent system is created if the output of a bid-auction model is input to a price prediction model.

By themselves, bid-auction models are convenient tools for predicting the occupancy characteristics of real estate units. This is especially helpful in residential real estate markets, where developers and land use planners typically quantify housing in terms of single-family and multi-family unit counts, yet travel models need to know the average size and income of the households living in those units, in order to predict their travel behavior. Bid-auction models translate between housing units and households in an intelligent manner, taking into account the prevailing household size and income distribution as well as the characteristics of housing units and the locations in which they are found. This includes accessibility measures, such as access to jobs from the place of residence. Bid models are especially good at capturing sensitivity to such variables because they allow the response to vary depending upon the specific type of household—larger households generally demand more space, for example, and are willing to trade it for a compensating reduction in access to jobs. See Appendix B for more details regarding bid model parameters; explanatory variables considered in the ARLUM residential bid model include:

- Bidding household size
- Bidding household income
- Zonal household density
- Zonal job density
- Drive access to jobs

The bid model outputs are used to predict two zonal inputs required by the travel model: average household size and median income. Average household size is calculated by taking the joint size and income distribution of households predicted by the bid model, multiplying by persons per household for each size-income category (tabulated from 2019 Census ACS PUMS data), and dividing by the total number of households. The median income is predicted using a regression model based on household income distributions, which was estimated from 2019 Census ACS Block Group tabulations data. Both are corrected to compensate for deviations between model output and base year zonal data.

The ARLUM bid model does not produce “rent”, or property value metrics, but rather a Market Pressure Index (MPI), which represents the net effect of competition on the value of a given location and product type in the real estate market. The MPI of the ARLUM bid model is an input to the housing Allocation model, such that more housing of a certain product type will be allocated to areas with greater MPI values for that product type.

Allocation Model

In ARLUM, land uses are spatially allocated using the Open Land Allocation Framework (OLAF), an open-source codebase which was developed by Manhan as a replacement for the CommunityViz Allocator 5 model in the course of a previous study. We had proposed to use CommunityViz, but found that it did not serve our clients’ needs due to the following:

- Lack of stable ArcGIS Pro support;
- Inefficient operation at the parcel level of analysis;
- Atypical implementation of location choice models;
- Challenges for integration with travel models; and
- Barriers to model sharing and access due to licensing.

CommunityViz was used by FBRMPO consultants for 2040 long-range land use forecasts made ten years ago, but during the 2045 study, Manhan found that this CommunityViz model was not available in complete operational form. At that time we built a more open-source land use forecasting model based upon the R programming language and the open-source mu-Land bid-rent module created with funding from the Lincoln Institute of Land Policy. The current phase of development continues in that vein, yet the Allocation model is based upon Python, rather than R, to facilitate integration with ArcGIS Pro geoprocessing.

The algorithm implemented coded in the Allocation Python script is as follows:

1. The input dataset and control file are read, including global parameters as well as geographic attributes.
2. Variables are initialized by evaluating a set of expressions coded in the control file on all of the geographies in the input dataset. Hard-coded development (such as scheduled development projects) may be added at this time.
3. A development “queue” comprising all of the real estate units defined in the control file is enumerated and shuffled (so that no one unit type has priority over another). Note that allocation units may not equate to housing units or jobs; one unit may represent five jobs, for example.
4. For each unit in the queue:
 - 4.1. A sample of possible geographic location options is drawn from the input dataset. This is performed by first applying a query that can be used to exclude ineligible areas (e.g. areas with no development capacity), then sampling from the remaining records.
 - 4.2. A “value” expression, unique to the current real estate unit type, is evaluated for each viable location option. This is used to compute choice probabilities for each of the sampled locations, using a multinomial logit formula. More detail regarding the value expressions is provided in Appendix A; factors considered include:
 - Residential market pressure index
 - Intersection density
 - Sewer and water service
 - Average elevation and slope (or presence of steep slopes)
 - Proximity to minor arterials and major collectors
 - Proximity to freeway ramps
 - Labor and consumer market access via auto (lagged)
 - Transit service
 - Parcel land use type

- 4.3. A location option is selected randomly, with probability determined by the choice model described above. A field defined in the control file for the real estate unit type is incremented by the unit size. (Note: there can be unit types which represent decrease in supply, instead of increase.)
5. After all units are allocated, a set of expressions coded in the control file are evaluated for all geographic areas in the dataset. This is generally used to re-compute total units by type, incrementing and decrementing as needed.

The Allocation control file allows for extensive user configuration in order to handle special cases. For example:

- In the 2020-2025 increment, a “development trend” index, derived from the previously collected list of real estate projects, was incorporated into the value functions, to encourage development to follow actually observed trends.
- In 2025, a specially-configured Allocation run simulates the impacts of flooding caused by Hurricane Helene, removing jobs at affected businesses, and relocating affected households.
- In the 2025-2030 increment, committed real estate development projects are pre-loaded as a fixed offset of growth with known spatial distribution or location, and deducted from the housing increments to be algorithmically allocated.
- Development location preferences were modified slightly between scenarios to create alternative spatial distributions of growth. For example, observed preferences for areas with more walkable street design (measured using intersection density as a proxy) were boosted in the Consolidated growth scenario, as was business preference for locations served by transit.
- The filter function applied before sampling potential candidate locations is implemented as a capacity constraint, with build-out capacity rates derived from analysis of base year ELUSE data. For the Consolidated growth scenario, the capacity of zones previously designated by FBRMPO staff as “Walkable Urban Places” (or WalkUPs) was boosted, whereas in the Dispersed growth scenario, lower build-out densities were assumed across the entire region.
- A secondary, TAZ-level Allocation run is performed as a post-process to the main model, in order to assign growth in lodging to zones within the study area.

Build-Out and Land Use Change Analysis

As noted above, the build-out capacity of a block for each land use type is an important input to the Allocation process, since it strictly constrains the amount of development that can occur there. This input was initially calculated based upon the 2020 ELUSE data, applying a set of maximum expected density rates to each parcel. For undevelopable and natural protected lands, this maximum density would be set to zero, prohibiting any land use of any kind. The residential land use types in the

ELUSE classification scheme are density-based, so the definition of each includes a maximum density value, usable for calculating build-out capacities. For non-residential land uses, the maximum expected density values were estimated in a more analytical manner:

1. InfoUSA business establishment points were combined with ELUSE parcels using GIS to tabulate jobs by land use type. Illogical combinations of land use type and industry code were excluded.
2. The distribution of estimated job density (jobs per acre) by land use type was summarized and the 80th percentile density for each land use type was selected as the assumed build-out rate.

There are multiple reasons for selecting 80th percentile values, rather than the maximum observed values, or 80% of the maximum observed values, such as:

- Though detailed, the InfoUSA and ELUSE data can be unreliable, especially for small businesses and niche land uses. The potential for spatial mismatch between job sites and parcels is also significant. Some extreme outlier density values may be observed due to data quality issues.
- The build-out density rates affect only future allocated growth, not existing. Thus choosing a lower value does not cause the model to eliminate existing development, only control and reduce the potential for over-allocation to already-dense areas.
- The default build-out density rates are intended to represent the normal, “business as usual” baseline case. For certain scenarios, these are selectively increased or decreased; thus choosing an extremely high capacity assumption would leave no room for differentiation between the scenarios based upon changes to policies affecting build-out in certain areas.

The non-residential build-out density rates assumed for each scenario are shown in Figure 16.

	Density (jobs per acre)	
Type	Dispersed Scenario	Other Scenarios
Industrial	16.4	5.2

Commercial	41.6	13.1
Lodging	11.1	4
Office	57.9	19.2
Mixed-Use	30.8	6.1
Institutional	14	3.1
Special	25	1.1

Figure 16. Build-out (maximum) densities by land use type

Land Use Change Language Model

Historically, land use change modeling has been implemented using grid-based datasets, where each cell has a land use classification that is evaluated in the context of its neighbors' status. Land use changes are predicted for each gridcell based upon simple rules or using neural networks trained using machine learning techniques. The ARLUM land use change model has some similarities to such models, but leverages modern generative AI technology to add desired capabilities.

Thanks to the continued maintenance of historical land use classification in the ELUSE dataset, it is possible to construct land use histories as sequences of text tokens representing the state of the land use at a given time. The land use change model predicts the next token in each sequence. This next-token prediction is a task that language models have recently been shown to perform well, especially language models based upon the neural network architecture known as the “transformer.” Transformer-based language models, even relatively small ones, have also been shown to quickly pick up on examples of the kind of sequence completion desired, and mimic these in their output; so, putting land use transition sequences of neighbor parcels in the string provided to a transformer-based land use sequence prediction model, we can expect those to influence the output, like how neighboring grid cells influence the subject of analysis in a traditional land use change model.

Like the existing land use data, the zoning data consist of a set of textual values in character fields of GIS layers that can be related spatially to parcels. Other potentially relevant information which was not originally stored as text can be converted easily, as shown in Figure 17.

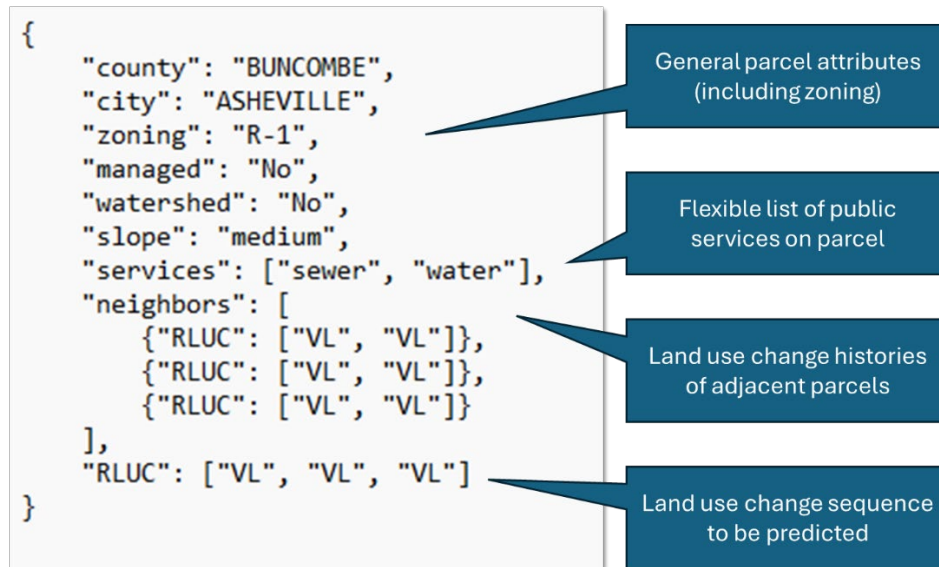


Figure 17. Excerpt from the LUC-LM training dataset.

We decided to base our model on a pre-trained language model from the StarCoder series trained by the BigCode consortium, for a variety of reasons:

- fine-tuning a model is far more efficient than training one from scratch, generally speaking;
- StarCoder is trained only on code repositories, so it has seen many examples of structured data stored as text, in formats similar to that of our corpus, with relatively less irrelevant general English language training material;
- furthermore the StarCoder models pose little or no risk of exposing either Manhan or FBRMPO to claims of copyright infringement, since their developers put significant effort into including only “permissively licensed” (e.g. open-source) code repositories in their training dataset. This should be especially un-concerning given the limited and highly technical scope of the task being performed by the language model in this case.

Note that this application of generative AI is different from the way commercially available large language models (LLMs) are commonly used today: we did not seek to “chat” with an AI about how land use patterns would evolve in the Greater Asheville region over the next thirty years. In particular, the use case described here does not require nor in fact benefit from the kind of “one-shot” or “zero-shot” behavior for which language models with billions of parameters are praised (in which they appear to perform arbitrary tasks with few or no examples provided by the user). Accordingly, we selected the smallest possible model in the StarCoder series (one with 117 million parameters) so as to minimize run times and energy costs.

In our validation tests, the Land Use Change Language Model (or LUC-LM) achieved 92.2% accuracy predicting non-residential development outcomes and over 80% accuracy for residential cases. This out-performed a gridcell-based model we previously trained using a more conventional neural network. However, even with optimizations applied, LUC-LM was too slow to apply by evaluating every parcel in the region at every five-year interval. This would also be unnecessary and unrealistic, because many parcels cannot change (due to protected or undevelopable status) or are very unlikely to change (due to mature or stable land use patterns). Note also that a parcel may increase its intensity of land use (i.e. adding jobs or households) without changing its basic land use type according to the ELUSE scheme.

To recognize these realities and limit the number of parcels evaluated by the land use change model, a screening model was developed to predict the likelihood that a parcel would change land use type. Job and household densities are input variables to this model, as well as the parcel’s current land use, such that when there is increasing household density surrounding a parcel which is non-residential, that potentially creates pressure on the non-residential parcel to switch to residential. This creates a kind of feedback loop from the allocation outputs to the land use change model, and makes LUC-LM more sensitive to demand.

The parcel-level outputs of LUC-LM for each five-year time step are aggregated to the Census Block level used by the Allocation process, and the same maximum density assumptions described previously are applied in order to re-calculate build-out density. What the addition of LUC-LM to the system allows is the “unlocking” of extra capacity in later forecast years due to developers and governments making decisions that result in land use changes. Note that these are sensitive to zoning codes where those apply, but the model also works where few or no zoning laws exist, representing normative developer and local government behavior.

Tourism-Related Activity

The Greater Asheville region is a popular tourist destination thanks to its beautiful mountains and other natural scenery as well as a thriving downtown and attractions such as the Biltmore mansion. The COVID-19 pandemic temporarily suppressed visitor activity, yet it surged back and spurred a wave of hotel construction in the region.

A post-process allocates lodgings by type to TAZs within the region, based upon a measure of highway access to special attractions derived from the travel model, and the amount of service employment in each respective zone. While not 100% equivalent to hotel and accommodation lodging, the Service category does contain the NAICS sector(s) which correspond to hotels and related employment. Recognizing this relationship prevents the model from allocating hotels to zones with little or no service employment.

Alternative Scenarios

Initially, seven candidate scenarios were discussed with FBRMPO staff and the Prioritization Subcommittee. These were narrowed down to focus on the following four scenarios:

- **Baseline:** a “business as usual” growth pattern based upon continuation of trends.
- **Consolidated:** a growth pattern that seeks to place more development in low VMT areas.
- **Dispersed:** a growth pattern characterized by strict density limits and more rural development.
- **Accelerated:** a hypothetical condition assuming a higher than expected population growth rate.

All of the above scenarios, except the Accelerated scenario, utilize household and employment forecasts based upon the state demographer population projections presented earlier. All scenarios also start from the same base data and do not differ in outputs for the historical years from 2020 to 2025.

Baseline Scenario

There are certain stages at which assumptions regarding “business as usual” development change significantly. These include:

- From 2020 to ~2024 (which is actually a pre-Helene 2025 run), a “recent development trends” index, representing shares of committed development identified in the previous 2045 forecast study, is included. Most of those projects were found to have likely been completed between 2020 (the model base year) and the present year. The index serves

to improve the quality of the forecast by allowing the simulation during this first timestep to be informed by what we know about actual development during that period.

- In addition, a “correction factor” was applied to compensate for initial deviations in 2024 output from observed spatial distributions (sourced from the ESRI Current Demographics dataset). These deviations likely stem from changes in behavior due to the impact of the COVID-19 pandemic, which was otherwise not captured in any of our parameter estimation datasets.
- For 2025, a special Helene-only iteration of the land use model was run. This run removes about 10,000 workspaces in areas affected by flooding, without re-allocation. An index representing likelihood and severity of flooding, derived from ICEYE and FINMAP data, drives the selection of jobs for removal. A similar index also drives relocation of households; however these are re-allocated to other locations within the region based upon the same site suitability index used in normal allocation processes. This is based upon research we reviewed finding that most households affected by natural disasters that choose to move do not relocate very far from where they lived before the disaster, and are likelier to move if they believe that they can upgrade their housing situation in the process.
- High-probability housing projects submitted to FBRMPO by local governments are pre-loaded to the blocks where they are located for the 2030 allocation time period. Most of this was approved or planned prior to Helene and the project information was gathered after Helene, implying that it probably was not canceled due to economic uncertainty following the storm.
- From 2025 to 2035, development capacity that falls within “floodway” zones is removed from consideration by the Allocation model, based upon the assumption that growth will not be allowed by cities in areas that were flooded by Helene or at high risk of future flooding. The restriction is relaxed from 2035 to 2050, based on the hypothesis that some communities may choose to allow such development again as the memory of the Helene disaster fades.

From 2035 on, the second half (or last 15 years) of the 30-year forecast is model-driven, meaning that the output is generated based upon interactions between the travel model and the land use change, allocation, and bid-auction sub-models of the land use forecasting system. To the extent that these models have parameters calibrated to match observed recent historical behavior (prior to 2020 in most cases), the output reflects a “business as usual” approach to real estate development and land use management. It may also be considered a “positive” forecast describing a likely future if no major changes were to occur, rather than a “normative” forecast that seeks to achieve a desired outcome.

Consolidated Scenario

In contrast to the Baseline scenario, the Consolidated scenario is defined with a specific policy goal in mind of guiding growth in such a manner as to reduce vehicle-miles traveled and improve resident quality of life by concentrating development in more walkable neighborhoods as well as other areas which offer alternatives to private auto mobility. More specifically, it differs from the Baseline scenario in the following assumptions:

- The general residential developer preference for neighborhoods with walkable street design (measured by intersection density, a proxy metric commonly adopted by transportation planners and available in the EPA Smart Location Database for 2019) was boosted under this scenario. Note that we did not assert a preference where one did not previously exist, nor turn a negative term into a positive one; we simply increased an observed preference. Most members of the FBRMPO Prioritization Subcommittee commented that this is likely representative of actual shifts in behavior, a plausible shift given the nationwide surge in outdoor recreation during the COVID-19 pandemic, not captured in our pre-2020 calibration data. However, if the preference shifts did not materialize in the future, the boost reflected in our forecast could be taken to represent any of a variety of possible economic incentive or subsidies that could be offered to households and housing developers in reward for locating in walkable communities.
- Service employment, the largest growth sector in the Asheville region, was found in our analysis to have some preference for workplace locations in areas served by public transit; we boosted this preference. Some other sectors were found to be slightly aversive to public transit; we zeroed out those parameters so that those jobs would at least not avoid areas well-served by the bus system. Again, if this change in preferences should not materialize organically, there are mechanisms for offering developers and businesses financial incentives to locate in such areas, working through Transportation Management Associations (TMAs) or local economic development organizations.
- FBRMPO staff had previously identified a set of geographic areas called “Walkable Urban Places”, or “WalkUPs”, which were key to the formulation of the growth scenario that was adopted in the 2045 MTP. For this round, we boosted the development capacity of blocks located within WalkUP areas. This may require action on the part of local planning and zoning officials to allow up-zoning or increase density limits within walkable neighborhoods.
- In the land use change model, we added a penalty to deter development of hitherto undeveloped land. This should generally prevent the creation of large amounts of capacity on the exurban outskirts of the region, limiting supply to areas that are more well-served by existing infrastructure. The implementation of this measure would largely depend upon the favorability of local government planning, zoning, and permitting actions towards greenfield development;

regional coordination would be required to communicate the shift in policy and build consensus around conservation and preservation of open and green space as a community goal.

- The Baseline and other scenarios reflect continuation of the observed trend towards an increasing share of single-family housing in the available supply stock. By contrast, in the Accelerated scenario a 50/50 balance between new single-family and multi-family unit types is asserted for new housing. There is some basis for this in the fact that the “scheduled” development, or known near-future housing projects submitted by local partners, is heavily weighted in type towards multi-family housing. Additionally, the continued existence of a housing affordability crisis generally indicates pent-up latent demand for more housing that can be most economically provided as housing units in multi-family buildings (or auxiliary units, which are not really distinguishable from multi-family housing in our analysis due to limitations in how residential land uses are classified in the ELUSE dataset).

Dispersed Scenario

This scenario examines the hypothesis that density causes congestion and other urban problems; as such, it is diametrically opposite the Consolidated scenario in many ways. Key differences in assumptions from the Baseline include:

- Maximum allowed density constraints were lowered across the entire region for all land use types. Since these constraints are only checked when new units are allocated to a block, the modified limits do not remove or demolish existing high-density urban places; however, these would not be allowed to grow and no new development would be allowed to reach similar levels of density.
- Conversion of undeveloped, unprotected land is given a boost in the land use change model. This means that while development is limited in already-dense areas, creation of new capacity is encouraged via greenfield development.

No changes to behavioral parameters or preferences were assumed in this scenario. The split of single-family and multi-family housing development is the same as in the Baseline scenario.

Accelerated Scenario

As discussed previously, both of the private-sector population forecasts we reviewed showed lower population growth than the adopted forecast developed by the North Carolina State Demographer. Investigation of the reasons for this revealed that the Woods & Poole projections show lower working-age population than the official projections. This may drive some of the difference between the public and private forecast data sources.

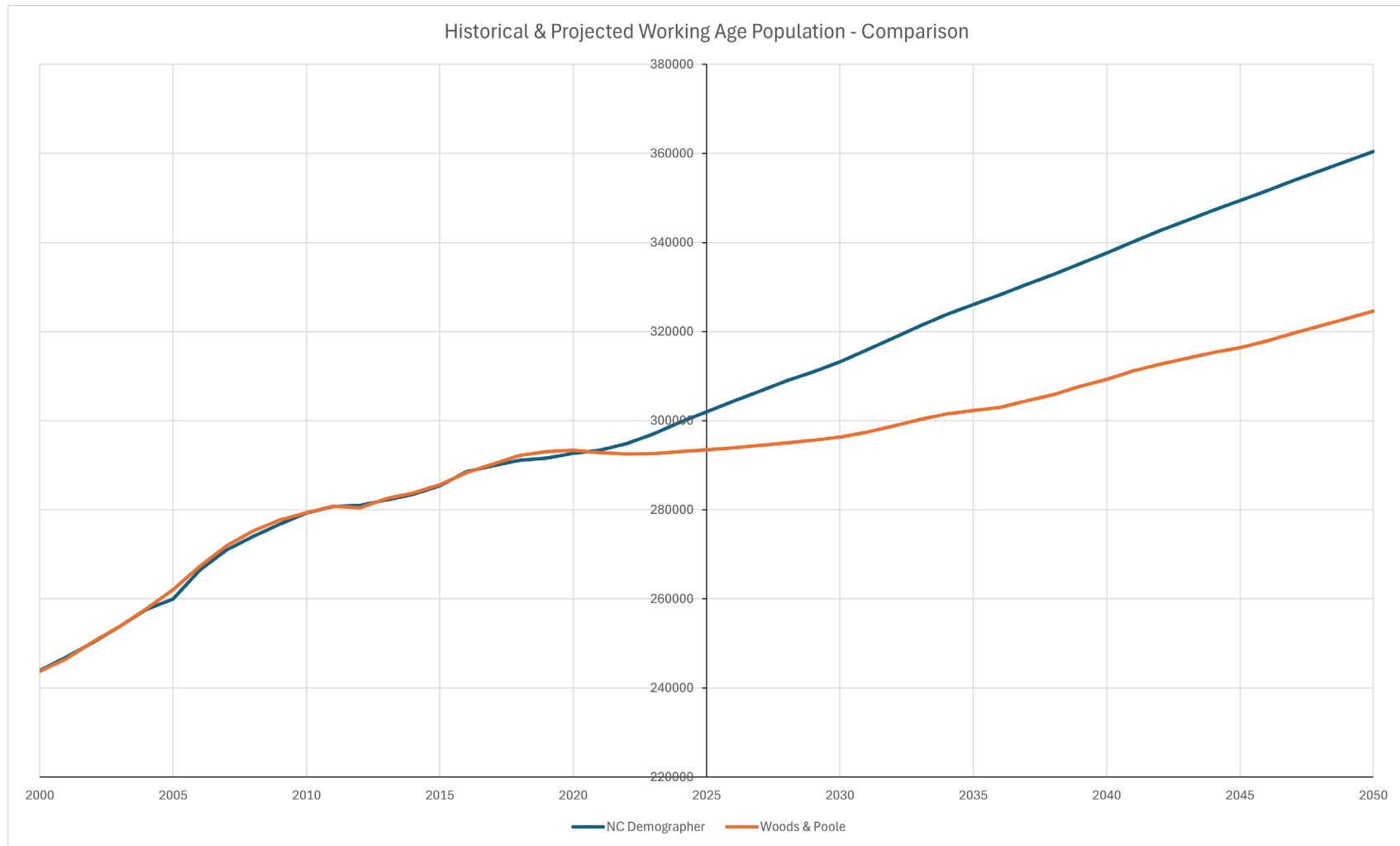


Figure 18. Comparison between state and private working-age population projections for the study area.

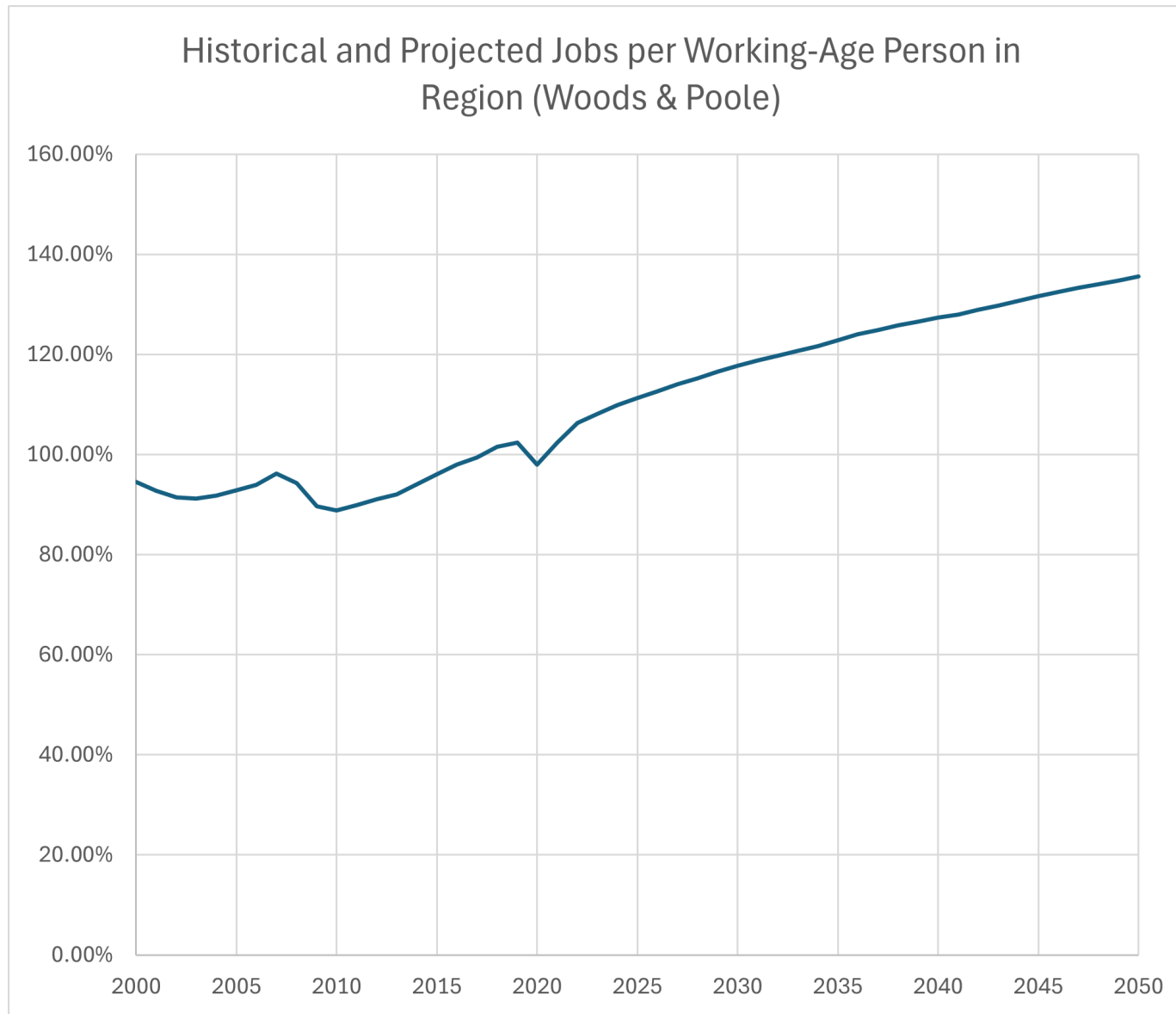


Figure 19. Historical and projected jobs per working-age resident in the study area, according to Woods & Poole CEDDS.

As shown in Figure 19, Woods & Poole also appears to be extrapolating a trend of increasing jobs per working-age resident in the region, which began in 2010 and passed the 100% mark just prior to 2020. This means that the region is likely currently importing workers from outside the five-county area, e.g. via supercommuting, or telecommuting, or migration of workplaces to exurban locations on the fringe of the region.

According to 2016-2020 5-year ACS Commuting Flows data, over 11,000 workers commute into the study area from counties within a 10-mile straight-line buffer of the border. This is actually more than the 7,000-person gap between area jobs and working-age residents reported by Woods and Poole for 2019, so the hypothesis that some workers may be leaving the region for more affordable housing is plausible. High ratios of jobs per working-age resident seems to be a relatively recent trend, however, and given the state demographer's methodology it is not surprising that this trend wouldn't play a major role in their future projections.

The Accelerated growth scenario asks what would happen if the trend of importing workers were reversed, rather than extrapolated as in the Woods & Poole forecast. As such it contains the following major differences in assumptions from the Baseline:

- The household growth rate is doubled, starting in 2030. This assumption was taken from the “higher demographic growth rates” scenario tested in a 2022 Environmental Protection Agency (EPA) study examining “Resilience Under Scenarios of Global Change”.
- External-internal trip volumes were reduced in the travel model, reflecting the idea that, with greater housing capacity within the region, fewer workers would be living outside the region and commuting into it every day.

The Accelerated scenario is also a useful sensitivity test of the model system as a whole. By pushing beyond the limits of observed growth trends this scenario may provide insights into how congestion—which is driven by household travel activity—shapes the spatial distribution of growth in a land-use/transportation interaction (LUTI) model like the one developed for FBRMPO.

Summary Findings

Our most decisive metric for distinguishing the transportation performance impact of these different land use scenarios is vehicle-miles traveled, or VMT. VMT is strongly correlated with many highway system externalities, such as air pollution and CO2 emissions, traffic safety impacts, and overall dependence on automotive transportation, which implies less active transportation and concomitant community health impacts.

Facility Type	Baseline VMT	Share	VMT Difference From Baseline Scenario					
			Accelerated		Consolidated		Dispersed	
Centroid Connectors	2,115,519	9%	679,230	32%	-318,758	-15%	72,140	3%
Minor Thoroughfares	4,314,834	19%	2,030,186	47%	-727,453	-17%	158,636	4%
Boulevards	626,299	3%	141,164	23%	-18,890	-3%	22,696	4%
Expressways	409,089	2%	56,442	14%	-55,633	-14%	460	0%
Freeways	9,318,213	41%	706,615	8%	-228,810	-2%	59,643	1%
Other Major Thoroughfares	5,928,666	26%	1,425,369	24%	-645,109	-11%	118,935	2%
All Types	22,712,619		5,039,006	22%	-1,994,654	-9%	432,511	2%

Figure 20. Baseline scenario VMT by highway system element as well as differences between scenarios.

As shown in Figure 20, the Consolidated scenario achieves a 9% reduction from Baseline 2050 levels. The Dispersed scenario shows slightly higher VMT than the Baseline 2050 level, due to the generally sprawling pattern associated with lower density levels and greater propensity for greenfield development, creating longer average trip lengths. The Accelerated scenario has higher VMT than all other scenarios due to the much greater number of households generating trips.

Another key metric of system performance is the share of regional trips made by auto, bus transit, and non-motorized modes of travel such as walking or biking. Utilization of non-auto modes is generally associated with better public health outcomes. The Asheville regional travel demand model predicts that the Consolidated land use scenario would yield the highest walk/bike and transit mode shares of the four scenarios tested, though cars would still be used for the vast majority of trips.

Scenario	Auto		Walk/Bike		Transit	
	Trips	Share	Trips	Share	Trips	Share
Baseline	1,917,718	94.2%	107,785	5.3%	11,011	0.5%
Consolidated	1,726,887	93.2%	113,127	6.1%	12,800	0.7%
Dispersed	1,935,389	94.5%	102,162	5.0%	9,844	0.5%
Accelerated	2,528,672	94.9%	123,871	4.6%	11,750	0.4%

Figure 21. Year 2050 trips by mode and mode shares for the four scenarios (daily total, all trip purposes).

The FBRMPO board approved the adoption of the Consolidated scenario as the preferred growth pattern for the region, per the recommendation of FBRMPO staff, the MPO's Prioritization Subcommittee, and the Technical Coordination Committee (TCC). Figure 21 provides a high-level summary of the growth characteristics of this scenario in tabular format, showing households and employment for 2020, 2030, 2040 and 2050 broken down by county as well as the largest city in each county. In general, the assumptions input to the Consolidated scenario result in a pattern where, in each county, the share of households in the largest city increases over time, i.e. households are concentrated in cities. For example, in 2020, Asheville had a roughly 37% share of Buncombe County households; in the 2050 Consolidated scenario, that share rises to 45%. Hendersonville's share of Henderson County households also increases, from 45% to 50%. The picture for employment, however, is more complex: Hendersonville, Mars Hill, and Brevard increase their shares of Henderson and Madison County jobs, yet Asheville and Waynesville decrease in share of their respective counties' employment bases (despite increasing overall).

	2020		2030			2040			2050		
	#	%	Δ	#	%	Δ	#	%	Δ	#	%
Buncombe	129,142		16,333	145,475		16,778	162,253		15,932	178,185	
Asheville city	47,686	36.9	9,289	56,975	39.2	12,338	69,313	42.7	10,486	79,799	44.8
Other Buncombe towns	81,456	63.1	7,044	88,500	60.8	4,440	92,940	57.3	5,446	98,386	55.2
Haywood	34,887		6,200	41,087		4,893	45,980		4,375	50,355	
Waynesville township	11,515	33.0	2,120	13,635	33.2	2,707	16,342	35.5	2,084	18,426	36.6
Other Haywood towns	23,372	67.0	4,080	27,452	66.8	2,186	29,638	64.5	2,290	31,929	63.4
Henderson	56,439		9,374	65,813		4,735	70,548		5,001	75,550	
Hendersonville township	25,117	44.5	5,347	30,464	46.3	3,916	34,381	48.7	3,037	37,418	49.5
Other Henderson towns	31,322	55.5	4,027	35,349	53.7	819	36,168	51.3	1,964	38,132	50.5
Madison	11,044		489	11,533		(569)	10,964		44	11,008	
Mars Hill township	2,032	18.4	133	2,165	18.8	(4)	2,161	19.7	84	2,244	20.4
Other Madison towns	9,012	81.6	356	9,368	81.2	(565)	8,803	80.3	(40)	8,763	79.6
Transylvania	19,020		549	19,570		(110)	19,460		535	19,995	
Brevard township	5,703	30.0	366	6,069	31.0	344	6,413	33.0	442	6,855	34.3
Other Transylvania towns	13,317	70.0	184	13,501	69.0	(454)	13,047	67.0	93	13,140	65.7
Grand Total	250,532		32,945	283,477		25,728	309,205		25,887	335,092	

Figure 22. Households by County and largest cities/towns, 2020-2050, Consolidated scenario. Note: U.S. Census County subdivisions used for tabulations by city and/or township. Key: “#” denotes the number or count of households within the designated geography, “Δ” denotes change between years, and “%” denotes the share of county totals for a given year.

	2020		2030			2040			2050		
	#	%	Δ	#	%	Δ	#	%	Δ	#	%
Buncombe	176,494		16,848	193,343		11,710	205,052		13,869	218,921	
Asheville city	118,964	67.4	7,879	126,843	65.6	7,432	134,275	65.5	8,175	142,450	65.1
Other Buncombe towns	57,530	32.6	8,969	66,499	34.4	4,277	70,777	34.5	5,694	76,471	34.9
Haywood	26,623		3,313	29,936		2,059	31,995		2,941	34,936	
Waynesville township	14,256	53.5	863	15,119	50.5	854	15,972	49.9	1,164	17,137	49.1
Other Haywood towns	12,367	46.5	2,450	14,817	49.5	1,205	16,022	50.1	1,777	17,799	50.9
Henderson	58,549		8,796	67,345		5,419	72,763		6,521	79,284	
Hendersonville township	32,064	54.8	5,420	37,484	55.7	3,588	41,072	56.4	4,078	45,150	56.9
Other Henderson towns	26,485	45.2	3,376	29,861	44.3	1,831	31,692	43.6	2,443	34,135	43.1
Madison	5,978		1,397	7,375		516	7,892		896	8,788	
Mars Hill township	1,509	25.2	492	2,001	27.1	207	2,208	28.0	325	2,532	28.8
Other Madison towns	4,469	74.8	906	5,375	72.9	309	5,684	72.0	572	6,256	71.2
Transylvania	15,833		2,947	18,780		1,698	20,477		2,480	22,957	
Brevard township	9,248	58.4	1,712	10,960	58.4	1,138	12,098	59.1	1,616	13,714	59.7
Other Transylvania towns	6,585	41.6	1,235	7,820	41.6	559	8,379	40.9	864	9,243	40.3
Grand Total	283,477		33,301	316,778		21,401	338,179		26,707	364,886	

Figure 23. Employment by County and largest cities/towns, 2020-2050, Consolidated scenario. Note: U.S. Census County subdivisions used for tabulations by city and/or township. Key: “#” denotes the number or count of jobs within the designated geography, “Δ” denotes change between years, and “%” denotes the share of county totals for a given year.

Caveats and Next Steps

Though current general plans and future land use maps were reviewed by the consulting team during the preparation of this forecast, they did not directly inform any of the allocation scenarios, due to the lack of any specific quantitative information found in these documents of sufficient detail for use as input to the forecast model system. The consultant recommends that general plans be reviewed in relationship to block-level and TAZ-level allocation outputs for the Consolidated for consistency in the future, as part of ongoing dialogue and coordination between local and regional government agencies.

By delivering a user-configurable land use forecasting system, rather than simply a static forecast, the consultant hopes to allow modifications to the Consolidated scenario as needed in the future, within the broad set of community preferences that it reflects. This may especially be necessary as information regarding Helene impacts becomes clearer. Future real estate projects and transportation infrastructure improvements can be added and evaluated to determine their real estate impact.

Figure 24. Change in households,
Census County Subdivisions,
2020-2050, Consolidated Scenario

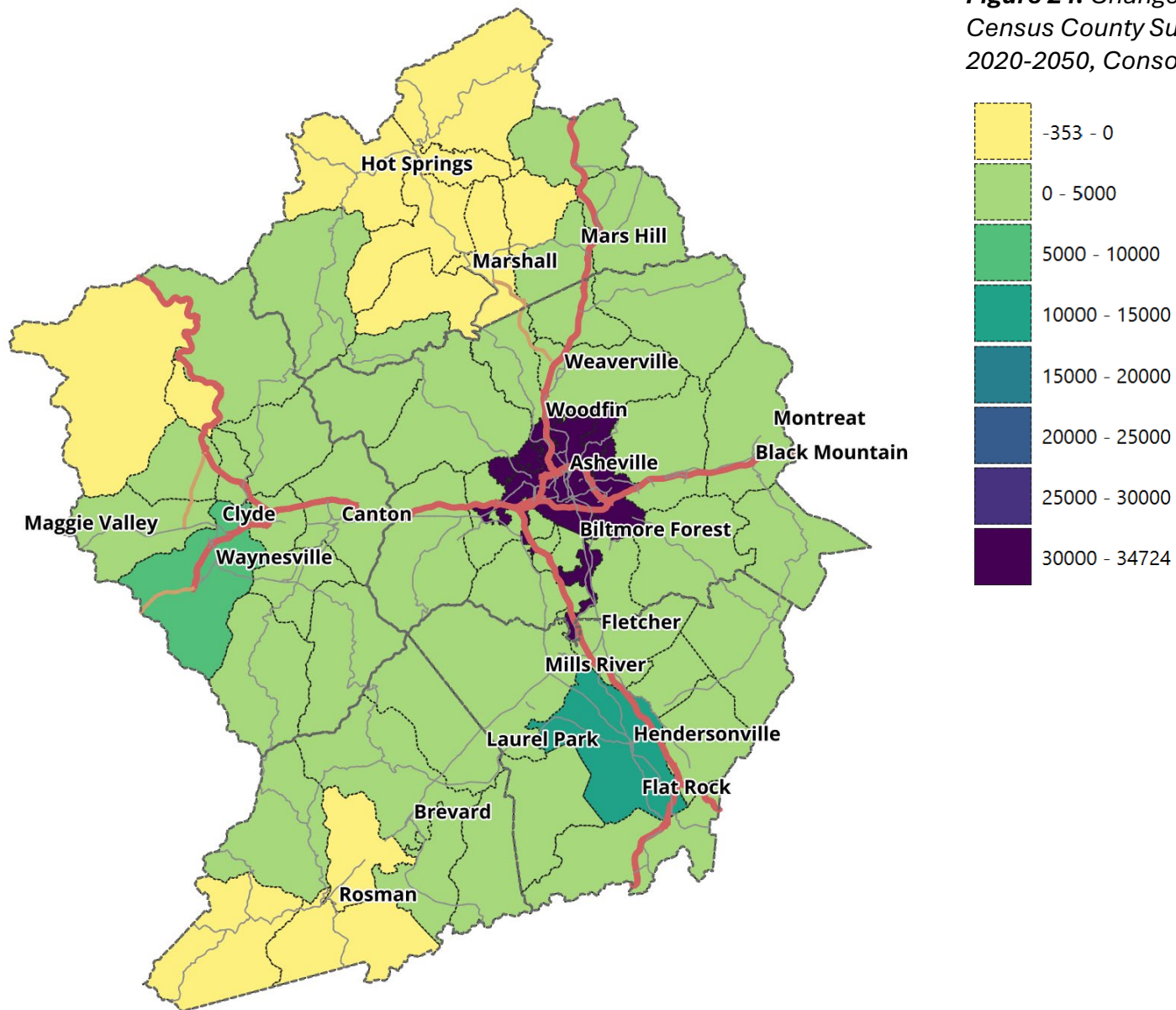
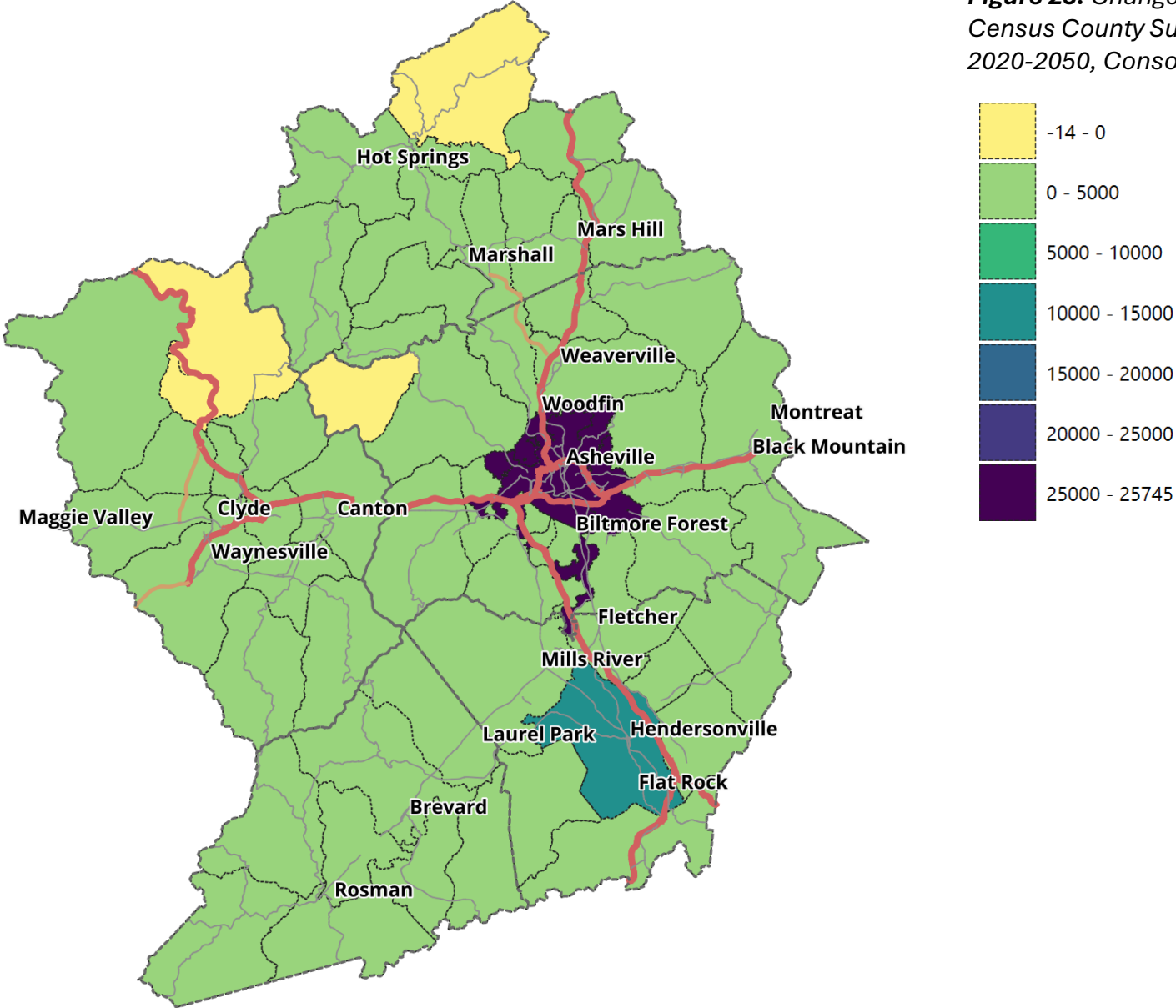


Figure 25. Change in employment, Census County Subdivisions, 2020-2050, Consolidated Scenario



Appendix A. Allocation Model Parameters

At the core of the allocation model is a mathematical representation of how developers choose where to locate housing and workspace in the region. Developers are assumed to specialize in a particular real estate product type, such as single-family housing or retail floorspace, and evaluate potential locations based on their suitability for that type of development. The equation predicting how much developers of a certain real estate product type value a given location is sometimes called a “site suitability function”.

While some land use models use site suitability functions asserted based on expert knowledge or local understanding, the consultant for this study applied a more data-driven approach, as follows:

1. Data regarding the spatial distribution of housing and non-residential development were gathered for 2014 and 2019, in order to calculate changes, or shifts, in the shares of development by product type over a five-year period. For the housing site suitability function, 5-year U.S. Census American Community Survey block group tabulations were used as the source for these data, because it included structure type details not provided in the FBRMPO travel model input files (i.e. counts of housing units by single-family, multi-family and other types). For non-residential development, we used the socio-economic zonal input data for the travel model.
2. For housing, the build-out potential of each block group or zone for each development type was also calculated. This is an important input because, all else being equal, development is likelier to go where there is capacity. In effect, the “excess” capacity, or available capacity after deducting existing 2014 development, was used to pre-weight the location options in the estimation process, taking into account their higher likelihood of selection before other factors are considered.
3. Explanatory variables were gathered, describing the characteristics of each location option (block group or TAZ). For non-residential development, these include the total area by non-residential land use type in the 2015 ELUSE dataset, playing a similar role to the build-out potential defined for housing. Other sources of data included FBRMPO GIS data, travel model outputs, and the housing market pressure index (MPI) calculated using the bid model described in Appendix B.
4. A multinomial logit model was fitted in R to estimate sensitivity of shifts in shares to location characteristics.

The estimated GLM model coefficients for the housing market site suitability function are presented below, in Figure A-1.

Variable	Meaning	Source	Estimate	Std. Error	z value	Pr(> z)	
logsum	Market pressure index	Bid model	0.187583	0.001617	116.035	< 2e-16	***
intdens.N	Intersection density (normalized)	EPA SLD v.3	1.540372	0.016445	93.669	< 2e-16	***
pct_sewer	Percent of CBG with sewer service	FBRMPO GIS	0.019317	0.008533	2.264	0.02358	*
is_MF	1 if multi-family (MF), 0 otherwise		0.657856	0.061943	10.62	< 2e-16	***
is_OR	1 if not MF nor single-family, else 0		3.248328	0.041661	77.97	< 2e-16	***
elevation	Average CBG elevation	USGS DEM	4.947418	0.04183	118.275	< 2e-16	***
avg_slope	Average CBG elevation	USGS DEM	-1.27185	0.067136	-18.944	< 2e-16	***
pct_water	Percent of CBG with water service	FBRMPO GIS	0.326802	0.010966	29.801	< 2e-16	***
elevation:pct_steep	Elevation * percent steep slopes	USGS DEM	-0.01628	0.000376	-43.274	< 2e-16	***
elevation:is_MF	Elevation (MF-specific effect)	USGS DEM	2.552743	0.113463	22.498	< 2e-16	***
is_MF:avg_slope	Slope (MF-specific effect)	USGS DEM	-2.09495	0.108083	-19.383	< 2e-16	***
is_MF:pct_water	Water service (MF-specific effect)	FBRMPO GIS	0.197213	0.025849	7.629	2.36E-14	***
elevation:is_OR	Elevation (other housing units)	USGS DEM	-1.82757	0.089468	-20.427	< 2e-16	***
is_OR:avg_slope	Slope (other housing units)	USGS DEM	0.274398	0.097456	2.816	0.00487	**
is_OR:pct_water	Water service (other housing)	FBRMPO GIS	-0.13997	0.018835	-7.431	1.08E-13	***

Figure A-1. Estimated coefficients, housing site suitability model.

All coefficients are highly significant. While there is no equivalent to the R-squared measure of fit for multinomial logit models, the McFadden Rho-squared for this model is extremely high (95%). Note that:

- The “logsum” variable, or market pressure index, is derived from the bid model, and as such includes the net effect of variables which are valued differently by different household types. These variables and their effects are listed in Appendix B.
- The coefficients representing effects specific to multi-family and other non-single-family housing unit types must be added to the baseline effect to get the true effect. For example, while the other-specific effect of water service is negative, the net effect is still positive ($0.326802 - 0.13997 = 0.186832$).

The non-residential model parameters are presented below, in Figure A-2. Due to the high number of variable interactions, the table is formatted differently, with the different non-residential land use types arrayed across columns. Though not shown, all estimated coefficients are highly significant ($p < 2e-16$, ***).

Row Labels	HTRET	IND	OFF	RET	SER	Meaning
Constant	-7.343613	-8.540735	-7.549109	-7.098898		Bias constant for each workspace type (except Service)
ma_mc_shr	0.322564	-0.059508	0.901185	-0.18525	-0.402948	Share of zone area close to minor arterials and major collectors
mkt_access_lag	0.937357	2.792779	1.260383	1.43287		Lagged labor and consumer (household) market access via auto
pct_steep	-0.407564	-0.485591	-0.536263	-0.927859	-8.838052	Percent of zone area that has steep slopes (30% or more)
ramp_shr	0.396426	-0.485176	-0.107142	0.664613	-0.372877	Share of zone area close to freeway/highway ramps (1 mile)
sewer_pct	0.38011	0.253665	0.61074	0.344077	0.518	Percent of zone area with sewer service
shr_Commercial	1.214567	1.873853	-0.643416	1.600973	-11.87021	Share of zone area classified as “Commercial” in ELUSE
shr_Industrial	0.489804	2.910368	-1.884893	-0.744662	-7.829837	Share of zone area classified as “Industrial” in ELUSE
shr_Institutional	-0.557212	0.573033	0.666865	-0.846898	-5.685225	Share of zone area classified as “Institutional” in ELUSE
shr_Lodging	0.004509	-1.688257	0.213098	-2.360163	-5.494106	Share of zone area classified as “Lodging” in ELUSE
shr_Mixed	-5.170282	6.040875	0.764729	-2.463273	-9.439943	Share of zone area classified as “Mixed-use” in ELUSE
shr_Office	-0.772683	1.862934	0.697896	-0.237846	-4.562582	Share of zone area classified as “Office” in ELUSE
shr_Residential	-0.208622	-0.05977	-1.117999	-1.111405	-9.097148	Share of zone area classified as “Commercial” in ELUSE
shr_Special	0.455217	0.029488	-0.299634	-1.202912	-4.971938	Share of zone area classified as “Commercial” in ELUSE
transit_pct	-0.124981	-1.213921	-0.088398	-0.507519	1.369714	Percent of zone within transit walk-shed

Figure A-2. Estimated coefficients, non-residential site suitability model.

Again, the McFadden Rho-squared for this model is extremely high (99%). Note that:

- There is no non-residential bid model, since none was needed, and hence no “logsum” nor “market pressure” term.
- Whereas in the residential market, highway jobs access is indirectly considered via the bid model, here auto access to the workplace by households (playing the role of either workers or consumers) is considered directly.
- Many of the explanatory variables represent different shares, proportions or percentages of total zone area; this may also be interpreted as the probability that a given parcel in that zone has the characteristic in question (e.g. steep slopes or sewer service)

Two “special case” Allocation runs are performed: one around the 2024-2025 timestep in the forecast series, to simulate damage and relocations caused by Hurricane Helene, and another TAZ-level allocation run in every year to allocate lodging-related growth. These are described in the sub-sections that follow.

Flood Damage Sub-Model

After extensive research and consideration, we determined that the same algorithm which was used to allocate normal development and turnover of real estate supply stock could also simulate the likely changes brought by Hurricane Helene. It is worth noting that at the time of performance of this study, situational awareness remained tentative; it was not possible to determine exactly which parcels suffered catastrophic damage nor who would ultimately leave and under what terms (i.e. a FEMA buyout versus simply selling or abandoning property). A more simulation-based approach provided a way to represent likely disaster impacts, without hard-coding these using only limited knowledge.

The first stage of our simulation is to evaluate the likelihood of severe and catastrophic damage to property in a given Census block. To do this, we obtained flood impact simulation data from the North Carolina Flood Inundation Mapping and Alert Network (FINMAN). Though themselves simulated, these data are appealing because the service provides estimates of dollar value damage which can be related to property values at the building level, for different hypothetical flood levels. We assume that as the dollar value of property damage approaches the building value, the likelihood of the property being abandoned or sold rather than repaired increases. We furthermore found that we could use a binomial logit (logistic regression) function to model this ratio as it is affected by development type and flood depth, as presented in Figure A-3.

Coefficients	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.93979	0.21219	-9.142	< 2e-16	***
Flood_Depth	0.11669	0.01938	6.021	1.73E-09	***
is_residential	0.88505	0.24347	3.635	0.000278	***

Figure A-3. *Estimated catastrophic flood damage prediction model.*

All coefficients in the damage model are highly significant and the McFadden Rho is acceptable, at 41.8%. Note that only properties within the inundated area were included in the estimation sample. Thus the probability of catastrophic damage is calculated as the probability of being within the inundation area, times the probability of catastrophic damage, predicted using the above function of property type and flood depth. We were grateful to receive permission to use ICEYE inundation and flood depth data for this exercise.

The mere likelihood of catastrophic damage does not necessarily predict how households or businesses will respond to that circumstance. Based upon our literature review, we found evidence that in other disasters of a similar nature, households that moved did not tend to go very far nor leave the region; rather they would seek to find somewhere close to their original neighborhood, ideally with somewhat preferable characteristics compared to their affected home. Thus we did not deduct households in properties that we believed would be severely affected from the regional control totals, but simply removed them from the affected area and gave them the opportunity to relocate elsewhere, following their usual site suitability preferences. The assumed total relocation potential by product type was determined in consultation with FBRMPO staff.

To understand the magnitude of impact on workplaces, we reviewed monthly employment statistics for the region, and found a gap of around 10,000 jobs between pre-Helene and post-Helene conditions. This matches well with our best estimate of jobs within the inundated area (determined by overlaying ICEYE and InfoUSA business establishment point data). The drop in reported employment seems to imply that jobs within the impacted areas were simply lost, not moved to other worksites. This is supported by anecdotal evidence.

The Allocation model is used to both pseudo-randomly select households and jobs to remove from within the impacted area, with a likelihood driven by the catastrophic damage function described above, as well as relocating households outside that area.

Lodging Post-Process Allocation

A TAZ-level application of the Allocation algorithm was used to spatially assign growth in hotel/resort rooms, short-term rental units, cabins, campgrounds, and bed and breakfast lodgings. Key to this model is the development of an index of auto access to tourist attractions. This index is calculated similarly to other accessibility variables in the model system, such as access to jobs, except that in place of the employment total at the destination zone, the index sums the YearlyViz (yearly visitors) field, weighted by a value decaying from 1 to 0 depending on travel time to the destination. In effect this metric represents the net accessibility to tourist attractions of a given potential lodging site. We also noted that, since accommodations employment overlaps with the combination of NAICS three-digit sectors in the travel model's employment classification scheme, it would make sense for hotel and resort rooms to be located in the same zones as Service employment. Finally, given the character of the Greater Asheville region, we considered elevation as a potential component in the lodging site suitability function.

After gathering zonal statistics on average elevation, we were able to directly use the travel model input zonal socio-economic data as an estimation dataset for a simple multinomial logit location choice model. Resulting coefficients are presented in Figure A-4, below.

Coefficients	Meaning	Estimate	Std. Error	t value	Pr(> t)	
special_attr_access_n.diff	Auto access to special attractions	1.340459	0.242788	5.521	5.06E-08	***
ln_service.diff	Service employment in zone	0.329729	0.075977	4.34	1.68E-05	***
mean_elevation.diff:is_RV_Camp	Elevation effect, RV campgrounds	0.001918	0.000279	6.879	1.54E-11	***
mean_elevation.diff:is_BnB	Elevation effect, bed and breakfasts	0.000816	0.000365	2.239	0.025558	*
mean_elevation.diff:is_STR	Elevation effect, short-term rentals	0.000255	0.000167	1.528	0.12716	
ln_service.diff:is_STR	Service jobs effect discount, STR	-0.26541	0.077035	-3.445	0.000611	***
ln_service.diff:is_RV_Camp	Service jobs effect discount, RV camps	-0.23388	0.081083	-2.884	0.004065	**
ln_service.diff:is_BnB	Service jobs effect discount, BnBs	-0.38114	0.157482	-2.42	0.015814	*
ln_service.diff:is_Cabins	Service jobs effect discount, Cabins	-0.2658	0.075834	-3.505	0.000491	***

Figure A-4. Lodging site suitability function(s).

Most of the coefficients in this model are statistically significant. We included one term which is not significant because it was logical in comparison with other parameters: the net effect of elevation on site suitability for short-term rentals is less than that for bed and breakfasts or RV campgrounds, which makes sense because AirBnB is as much an urban phenomenon as a clearinghouse for getaway lodgings. A much stronger association with Service employment was found for hotel and resort rooms than other lodgings, which makes sense given the number and variety of staff that such facilities employ.

Appendix B. Bid Model Parameters

A “bid model” is a multinomial logit model choice model representing a property seller’s choice of buyer (or landlord’s choice of tenant), taking into account real estate market segmentation. The residential bid model serves two roles within ARLUM:

- It disaggregates households by size and income category, supporting the calculation of zonal population and median income estimates required for input to the travel model;
- It generates a Market Pressure Index (MPI) representing the net desirability of each housing type and location given the differential preferences and prevalence of the various household market segments, which is an input to the housing allocation model.

The residential bid model that was previously implemented for FBRMPO by Manhan in the 2045 Socio-Economic Land Use Forecast Study was an instance of a “standard” set of functions estimated using nationwide Census data. We developed a new residential bid model for the 2050 Socio-Economic Land Use Forecast Study in order to make the tool more localized, as well as to test inclusion of highway access to jobs in the model specification. This was not possible using the previous set of bid functions, because they had been developed using individual-level Census Public Use Microdata Samples (PUMS), which lack geographic specificity; sensitivity to highway accessibility metrics had been added at the level of overall spatial allocation without differential sensitivity by household type.

The bid model estimation dataset used in the current study was also derived from PUMS; however these data were aggregated and re-weighted using an iterative proportional fitting (IPF) process to match the marginal distribution of households by size and income category at the Census Block Group level. This allowed linking a variety of location-specific variables such as the jobs access variable noted previously. Interactions with household size and income were considered for all such variables. The full model specification with estimated coefficients is presented in Figure B-1.

Variable	HH Size effect	Income effect
hhDensity_N	-0.617885	
Is_MFH	-1.93155	-3.490095
Is_OTH		-5.367755
job_access_lag	0.480175	2.57672
jobDensity_N	-3.34148	2.57805

Figure B-1. Residential bid function.

All estimated coefficients were statistically significant. Additionally, they have intuitive interpretations, such as:

- Larger households (e.g. families as opposed to singles and couples) prefer less dense neighborhoods, i.e. places with higher average lot sizes and/or more public open space;
- The likelihood of living in multi-family housing decreases with both household size and income, as families seek larger, more private housing options; higher-income households have more financial means to achieve this goal;
- Other housing units (neither single-family nor multi-family housing, e.g. low-cost mobile homes) are less likely to be inhabited by higher-income households;
- Access to jobs via auto increases in priority with both household size and income, both of which imply a higher potential number of workers; and
- Higher-income households actually prefer greater job density, likely because it implies a more urban environment with more services, dining, and entertainment activities which require disposable income; however, larger households (e.g. families) do not share this preference, possibly due to more constrained time budgets or overall preference for less dense neighborhoods (as mentioned previously).

Note that alternative-specific constants were also estimated for household income groups one, two, three, and four, leaving the highest income group as the “reference alternative”. These are not presented here since they lack any real intrinsic interpretation, but are simply a necessary part of any multinomial logit model specification.

Appendix C. Land Use Change Screening Model

Though deliberately scaled-down in parameter count and optimized to the maximum extent possible, the land use change language model represents a computationally intensive process that would take a prohibitive amount of time to evaluate if every parcel were considered, even using a computer with a fast GPU. Thus we developed a screening model to select only parcels with a high chance of changing *at all*, since most parcels do not change their land use type in any given five-year period. A binomial logit (logistic regression) model was fitted using historical ELUSE data, based in the current land use classification as well as household and job densities in the immediate vicinity. This model is presented in Figure C-1; higher and/or positive coefficients imply greater likelihood of land use transition.

Term	Meaning	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	Bias constant	-1.56454	0.012936	-120.947	< 2e-16	***
hh_density	Neighborhood household density	0.962146	0.017661	54.478	< 2e-16	***
is_Residential	Parcel is residential (any density)	-2.52896	0.03135	-80.668	< 2e-16	***
is_Industrial	Parcel is industrial	-1.34502	0.26574	-5.061	4.16E-07	***
is_Commercial	Parcel is commercial	-1.62668	0.158686	-10.251	< 2e-16	***
is_Lodging	Parcel is lodging	2.344859	0.10996	21.325	< 2e-16	***
is_Office	Parcel is office	-1.00268	0.216564	-4.63	3.66E-06	***
is_Mixed	Parcel is mixed	-2.05671	0.564637	-3.643	0.00027	***
is_Institutional	Parcel is institutional	-1.13638	0.127022	-8.946	< 2e-16	***
hh_density:is_Residential	Residential interaction with HH density	-0.94471	0.034163	-27.653	< 2e-16	***
hh_density:is_Industrial	Industrial interaction with HH density	-1.1743	0.246286	-4.768	1.86E-06	***
hh_density:is_Commercial	Commercial interaction with HH density	-1.60729	0.172385	-9.324	< 2e-16	***
hh_density:is_Lodging	Lodging interaction with HH density	-1.32617	0.142867	-9.283	< 2e-16	***
hh_density:is_Office	Office interaction with HH density	-1.03178	0.135836	-7.596	3.06E-14	***
hh_density:is_Institutional	Institutional interaction with HH density	-0.9426	0.08525	-11.057	< 2e-16	***
is_Residential:emp_density	Residential interaction with job density	0.076173	0.006551	11.628	< 2e-16	***
is_Lodging:emp_density	Lodging interaction with job density	-0.35076	0.09297	-3.773	0.000161	***
is_Institutional:emp_density	Institutional interaction with job density	0.055377	0.019714	2.809	0.004969	**

Figure C-1. Land use change screening model.

All coefficients are statistically significant. Due to interactions, the interpretation of some effects may be a bit complex; for example:

- All else being equal, lodging is the likeliest land use type to change to something else; however density of either housing or employment discounts this effect significantly. This implies that downtown hotels, for example, are much less likely to change land use classification than bed and breakfast or cabin lodgings.
- Parcels with a mixture of uses are least likely to change, probably because their definition already implies a combination of uses—hence addition of more employment or housing wouldn't necessitate a reclassification.

- Density—of either housing or employment—tends to increase the likelihood that residential parcels will change to something else. This aligns with the consumer preferences identified via residential bid function estimation.
- Institutional land uses seem to follow a pattern of sensitivity to density that is similar to residential uses, being more likely to change into something else with higher housing or employment density.

Note that the density values are updated at each time step in response to outputs of the Allocation sub-model. This means that there is a feedback loop in place, giving the land use change model some sensitivity to market dynamics and demand.

Appendix D. Input Data Dictionary

The ARLUM software can be found on Github ([ManhanGroup/ARLUM: Asheville Region Land Use Model](#)). Input data for the four scenarios described herein are also stored in the project Github repository, within the “inputs” folder. Beneath this folder, there are sub-folders named “allocation”, “bid_model”, “luc_model”, and “travel_model”, each of which contains inputs specific to the referenced portion of the model system. This section attempts to provide an inventory of important input data, focusing on those which might need to be updated for a future model of an alternative or revised scenario.

Allocation Working Table Format

The input data file for the Allocation sub-model is a flat table stored in comma-separated-values (CSV) format. Versions of this file can be found in the allocation sub-folder of the input folder in the ARLUM repository. Each record represents a 2020 Census block, such that the table can be joined to a block shapefile for mapping. The output of this model is also stored in the same format, with updated values for some fields. A listing of fields and their source is given in Figure D-1.

<i>Field</i>	<i>Description</i>	<i>Source</i>
<i>GEOID20</i>	Census geographic identifier for block	
<i>NEWTAZ</i>	TAZ number (v.1)	FBRMPO "final TAZ" layer (March 2024)
<i>hh_init</i>	Initial estimate of housin units located in the block	U.S. Census 2020
<i>HH_add</i>	Households added to the block's inventory (initially zero)	
<i>HH_del</i>	Households removed from to the block (initially zero)	
<i>HH_final</i>	Final estimate of households located in the block	
<i>SF_init</i>	Initial estimate of single-family housing units	

Field	Description	Source
<i>SF_add</i>	Single-family housing units added (initially zero)	
<i>SF_del</i>	Single-family housing units deleted (initially zero)	
<i>SF_final</i>	Final estimate of single-family housing units	
<i>MF_init</i>	Initial estimate of multi-family housing units	
<i>MF_add</i>	MULTi-family housing units added (initially zero)	
<i>MF_del</i>	Multi-family housing units deleted (initially zero)	
<i>MF_final</i>	Final estimate of multi-family housing units	
<i>OR_init</i>	Initial estimate of other residential units	
<i>OR_add</i>	Other residential units added (initially zero)	
<i>OR_del</i>	Other residential units deleted (initially zero)	
<i>OR_final</i>	Final estimate of other residential units	
<i>ws_init</i>	Initial estimate of total workspaces/jobs in block	InfoUSA and U.S. Census LEHD LODS data
<i>EMP_final</i>	Final estimate of total block employment	
<i>IND_init</i>	Initial estimate of industrial jobs	
<i>IND_add</i>	Industrial jobs added	
<i>IND_del</i>	Industrial jobs deleted	
<i>IND_final</i>	Final estimate of industrial jobs	
<i>RET_init</i>	Initial estimate of retail jobs	
<i>RET_add</i>	Retail jobs added	
<i>RET_del</i>	Retail jobs deleted	
<i>RET_final</i>	Final estimate of retail jobs	
<i>HTRET_init</i>	Initial estimate of retail jobs	
<i>HTRET_add</i>	Retail jobs added	
<i>HTRET_del</i>	Retail jobs deleted	
<i>HTRET_final</i>	Final estimate of retail jobs	
<i>SER_init</i>	Initial estimate of service jobs	
<i>SER_add</i>	Service jobs added	
<i>SER_del</i>	Service jobs deleted	
<i>SER_final</i>	Final estimate of service jobs	

Field	Description	Source
<i>OFF_init</i>	Initial estimate of office jobs	
<i>OFF_add</i>	Office jobs added	
<i>OFF_del</i>	Office jobs deleted	
<i>OFF_final</i>	Final estimate of office jobs	
<i>hh_cap</i>	Build-out capacity for housing units	FBRMPO parcel ELUSE dataset & LUC-LM
<i>ws_cap</i>	Build-out capacity for workspaces/jobs	FBRMPO parcel ELUSE dataset & LUC-LM
<i>intdens_N</i>	Intersection density, normalized to 0-1 range	EPA Smart Location Database (SLD) v.3.0
<i>ma_mc_shr</i>	Share of block area near major arterials and minor collectors	FBRMPO GIS
<i>ramp_shr</i>	Share of block area within a buffer distance of freeway ramps	FBRMPO GIS
<i>p_water</i>	Share of block area with water service	FBRMPO GIS
<i>p_sewer</i>	Share of block area with sewer service	FBRMPO GIS
<i>p_transit</i>	Share of block area within transit walkshed	FBRMPO GIS
<i>elevation</i>	Average block elevation (normalized)	Derived from USGS DEM
<i>avg_slope</i>	Average block slope (normalized)	Derived from USGS DEM
<i>p_steep</i>	Share of block area with steep slopes	Derived from USGS DEM
<i>mkt_access_lag</i>	Access to labor and consumer markets (households) via auto, AM peak period	Derived from NCDOT FBRMPO travel model
<i>job_access_lag</i>	Access to jobs via auto, congested AM peak period	Derived from NCDOT FBRMPO travel model
<i>logsum_SFH</i>	Residential market pressure index, single-family housing	Residential bid model
<i>logsum_MFH</i>	Residential market pressure index, multi-family housing	Residential bid model
<i>logsum_OTH</i>	Residential market pressure index, other housing units	Residential bid model
<i>shr_Commercial</i>	Share of block area classified as Commercial	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Industrial</i>	Share of block area classified as Industrial	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Institutional</i>	Share of block area classified as Institutional	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Lodging</i>	Share of block area classified as Lodging	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Mixed</i>	Share of block area classified as Mixed	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Office</i>	Share of block area classified as Office	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Residential</i>	Share of block area classified as Residential	FBRMPO parcel ELUSE dataset & LUC-LM
<i>shr_Special</i>	Share of block area classified as Special	FBRMPO parcel ELUSE dataset & LUC-LM
<i>p_floodway</i>	Share of block area in designated floodway area	FEMA

Field	Description	Source
<i>flood_impact_res</i>	Probability of catastrophic Helene flooding damage, residential properties	Modeled based on FINMAN and ICEYE data
<i>flood_impact_oth</i>	Probability of catastrophic Helene flooding damage, other properties	Modeled based on FINMAN and ICEYE data
<i>pipeshr_sqft_mean</i>	Region-wide percent of non-residential square footage "pipeline"	2045 Land Use Forecast Study
<i>pipeshr_units_mean</i>	Region-wide percent of housing unit "pipeline" (planned projects)	2045 Land Use Forecast Study
<i>ln_sqft_pipeshr</i>	Logarithm of above (transform used to adjust allocation in 2020-2025)	2045 Land Use Forecast Study
<i>ln_units_pipeshr</i>	Logarithm of above (transform used to adjust allocation in 2020-2025)	2045 Land Use Forecast Study
<i>WalkUP</i>	Walkable urban place classification (short)	2045 Land Use Forecast Study
<i>WalkUPType</i>	Walkable urban place type classification (long)	2045 Land Use Forecast Study
<i>is_walkup</i>	Indicator variable equal to one if WalkUPType="Walkable"	2045 Land Use Forecast Study
<i>SF_commit</i>	Committed single-family housing units added in 2025-2030	Data submitted by local planning agencies
<i>MF_commit</i>	Committed multi-family housing units added in 2025-2030	Data submitted by local planning agencies
<i>tractID</i>	Census Tract identifier	
<i>ln_hu_factor</i>	Adjustment factor based on comparison to 2024 observed housing units	Derived from ESRI Updated Demographics dataset
<i>ln_emp_factor</i>	Adjustment factor based on 2024 daytime population (workers)	Derived from ESRI Updated Demographics dataset
<i>ModelTAZ</i>	Travel model transportation analysis zone number	NCDOT travel model (Fall 2024)

Figure D-1. Allocation working dataset (both input & output).

Note that:

- The total amounts of housing units and workspace by type to be added and/or removed within a given five-year forecast time step are input in a separate control file defining parameters for the model run.
- Some attributes are estimated at a higher level of geographic resolution, such as block group, tract, or TAZ, and associated with each block within that larger area. Those values may be updated by ARLUM models.
- For all time steps except the base year, the initial housing unit and workspace count fields are set to the previous time step's output value, and the added/deleted fields are initialized to zero before a model run. The final estimate is based upon the initial value, plus the added amount, minus the deleted amount, truncated at zero.
- The block-level estimates may vary slightly from TAZ-level amounts. In cases where they disagree, the TAZ numbers should be viewed as more authoritative. The block-level file mostly serves as a record of the allocation process.

Residential Bid Model Data

A copy of the calibrated bid model is stored in the `bid_model` sub-folder. This “`calibration_run`” directory has its own input sub-folder which has multiple input files of its own, all of which are semicolon-delimited text CSV tables:

- *agents*: characteristics of household types
- *agents_zones*: location-specific characteristics of agents (not used)
- *bids_adjustments*: calibration adjustments to make the bid model match observed distribution of households by size & income category (according to U.S. Census ACS block group data)
- *bids_functions*: bid function parameters, as described in Appendix B, in a tabular format
- *demand*: regional control totals of households by type
- *demand_exogenous_cutoff*: a table controlling which households may compete for which housing units (not used)
- *real_estates_zones*: location-specific characteristics of housing units
- *rent_adjustments*: adjustments to output rent values (not used)
- *rent_functions*: a function to convert bid logsums or market pressure indices into money “rents” (not used)
- *subsidies*: a table that can be used to input taxes or developer subsidies (not used)
- *supply*: real estate supply stock inventory by type and zone
- *zones*: zone-specific variables

In general, directly editing these tables is not recommended; the ARLUM python scripts will copy the calibration run and automatically update fields that need to be modified based upon changes in other input data sources. For example, to run scenarios involving alternative demand forecasts, such as a different households size and income group distribution, the user may work with the tables found in the `demand_series` sub-folder of the `bid_model` directory. Each table in this folder is a copy of the *demand* table found in the calibrated residential bid model input folder, with two fields:

- *H_IDX*: housing type index
- *DEMAND*: region-wide number of households (total, not incremental)

The *H_IDX* values are related to size and income categories according to the table in Figure D-2

INCOME GROUP	ONE-PERSON	TWO PEOPLE	THREE PEOPLE	FOUR OR MORE
UNDER \$10,000	1	2	3	4
\$10,000 TO \$35,000	5	6	7	8
\$35,000 TO \$75,000	9	10	11	12
\$75,000 TO \$125,000	13	14	15	16
\$125,000 OR MORE	17	18	19	20

Figure D-2. Allocation working dataset (both input & output).

The values in the agent-specific characteristics table are generally derived from base year 2020 U.S. Census American Community Survey Public Use Microdata Samples (ACS PUMS). In very specific cases, such as modeling changing trends in vehicle ownership, it may be desirable to modify these values manually. If needed, the fields in the *agents.csv* table are defined as follows:

- *IDAGENT*: ID number of household type (see above)
- *IDMARKET*: ID number of real estate market (always 1 since only residential real estate is modeled for FBRMPO)
- *IDAGGRA*: aggregate household type (in this case, income group number)
- *UPPERBB*: internally-used upper bound on bid values

- *avg_size*: average household size
- *ln_income*: natural logarithm of average household income (the log transformation is used because income distributions have a more log-normal than normal shape; hence the transformed values have a normal shape)
- *vehicles*: average number of vehicles per household (used to create travel model socioeconomic input data)
- *k12enroll*: average number of children aged 5-18 years old in household (used in creating travel model input data)

Travel Model Data

The travel_model sub-folder contains data imported from the FBRMPO TransCAD model for the base year and forecast years in which the travel model is run to obtain updated highway accessibility (job and household access) measures. For the base year, this includes the 2020 input socio-economic data file prepared by NCDOT for model calibration purposes. These data are replaced by the output from ARLUM to calculate accessibility for 2025 and future years. Origin-destination travel times, however, must be manually imported from future year TransCAD model runs. The “skim” table format used by ARLUM is a comma-separated-values (CSV) text file with at least the following three columns:

- Origin zone
- Destination zone
- Highway travel time

Additional columns for other skim values may be included but are not used by the highway accessibility calculation scripts. Note that there are no column headings in the table format as exported from TransCAD, and none are required by ARLUM.

A decay curve based upon National Household Travel Survey (NHTS) data is used to convert travel times to “weights” applied to jobs or households reachable from a given origin. This lookup table can be found at the root of the inputs folder and should generally not be modified unless new travel survey are supplied. The decay curve is related to the trip length distribution for a given trip purpose (in this case, home-based work or commute travel).

Land Use Change Working Data

Similar to the Allocation model, the inputs and outputs to LUC-LM are in essentially the same format; whereas Allocation modifies existing columns, LUC-LM appends additional columns. The core working data file is a subset of the 2020 ELUSE parcel GIS data exported to CSV format (to conserve storage space and facilitate file versioning). The initial set of fields in this file are listed in Figure D-3. The land use change model appends RLUC_20XX and TYPE_20XX fields corresponding to the new, predicted land use classification (XX being the last two digits in the year being evaluated). Input fields for the land use change screening model (i.e. housing and employment density variables as well as indicator variables for aggregate land use types) are also updated as the land use change model moves from one time step to another.

A parcel adjacency matrix was built using PostGIS to store relationships between neighboring parcels, so that these do not have to be calculated on the fly. This CSV-format table has only two fields:

- *OBJECTID*: the geographic identifier for a parcel
- *NEIGHBOR_ID*: the geographic identifier for a neighboring parcel

Note that values may be repeated in either of these fields, based on the many-to-many spatial relationships between parcels. If ARLUM is updated in the future with new base year parcel data, including changes to parcel boundaries, then the adjacency matrix must be updated as well in order to run the LUC-LM component.

In addition, because the ELUSE parcels do not nest perfectly within U.S. Census blocks, a geographic crosswalk file is included based upon the output of a spatial intersection between these two layers, containing the following fields:

- *OBJECTID*: the geographic identifier for a parcel
- *GEOID20*: the geographic identifier for a 2020 U.S. Census Block
- *area*: the amount of overlap between the two geographic entities

This table must also be updated if the base year parcel or block boundaries change.

Build-out capacity is determined using a lookup table that may include other columns, but must at minimum include the RLUC_Code field, an HH_Cap_Acre field, and a Job_Cap_Acre field giving maximum household and job values per acre for the specified land use type.

Field	Description
OBJECTID	Geographic ID of the parcel
COUNTY	County name
City	City name
Zoning	Zoning code
Managed	Managed wildlife area status
Watershed	Watershed status (Yes/No)
Slope	Average slope
Sewage	Sewer service status (Yes/No)
Water	Water service status (Yes/No)
Transit	Transit service status (Yes/No)
ACRES_2015	Parcel acreage in 2015
ACRES_2020	Parcel acreage in 2020
RLUC_2020	Regional land use classification code in 2020
RLUC_2015	Regional land use classification code in 2015
RLUC_2010	Regional land use classification code in 2010
GEOID	U.S. Census block geographic identifier
hh_density	Neighborhood household density
emp_density	Neighborhood job density
Type_2020	Aggregate land use classification
is_Residential	Indicator variable for residential land use
is_Industrial	Indicator variable for industrial land use
is_Commercial	Indicator variable for commercial land use
is_Lodging	Indicator variable for lodging
is_Office	Indicator variable for office land use
is_Mixed	Indicator variable for mixed-use parcels
is_Institutional	Indicator variable for institutional land uses
is_Special	Indicator variable for "special" land uses
is_Undeveloped	Indicator variable for undeveloped land
NEWTAZ	TAZ number (based on March 2024 FBRMPO shapefile)
ModelTAZ	TAZ number (based on Fall 2024 NCDOT travel model)

Figure D-3. Initial Land Use Change model input dataset.