The Geospatial Small-Area Population Forecasting (GSAPF) Model Methodology Used by the Southwest Florida Water Management District

Prepared For

Southwest Florida Water Management District
Agreement 19CN0002133
TWA 21TW0003642

Prepared By

GIS Associates, Inc.
2158 NW 5th Avenue
Gainesville, Florida 32603

Under Subcontract With

WRA Engineering, LLC
4260 West Linebaugh Ave
Tampa, FL 33624

December 21, 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This document describes the methodologies used by the Southwest Florida Water Management District (SWFWMD) to develop small-area population projections in support of water supply planning and water use permitting. The county-level projections available from the Bureau of Economic and Business Research (BEBR) at the University of Florida are the official projections for the State of Florida and the generally accepted standard throughout the state, but accurately projecting future water demand for water utility potable service areas requires more precision than is offered by county-level projections. In addition, the Census Population Cohort projected by BEBR does not include important non-permanent populations, such as seasonal residents, tourists, or commuters. For these reasons, SWFWMD contracted GIS Associates, Inc. (GISA) through WRA Engineering to provide small-area population projections for the 17 SWFWMD counties. (Note that there are only 16 counties located partly or entirely within SWFWMD, but forecasts were also made for two utility service areas in the northern part of Lee County that have water use permits with SWFWMD.) This was achieved by implementing GIS Associates' Geospatial Small-Area Population Forecasting Model (GSAPF Model), which makes Census Population Cohort projections at the 2020 Census Tract level, and spatially distributes those projections to individual land parcels to facilitate aggregation by utility or other boundaries. In addition, GISA applied SWFWMD methods for projecting non-permanent population to the Census Population Cohort projections derived from the GSAPF Model. This document describes these projection methodologies and their use to project future populations. Ultimately, these small-area population projections are used as a basis for making future water demand projections for SWFWMD.

GEOSPATIAL SMALL-AREA POPULATION FORECASTING MODEL OVERVIEW

The geographic information system (GIS) based GSAPF Model projected future Census Population Cohort population growth at the parcel level and controlled those projections to BEBR's latest county-level forecasts. Figure 1 on the following page is a process flow chart of the population projection and distribution methodology. First, County Build-out Submodels were developed using property parcel data for each of the 17 counties that are entirely or partly within SWFWMD. The purpose of the County Build-out Submodel is to develop maximum residential development potential at the parcel level. A detailed description of this model is included in the chapter titled "County Build-out Submodels". Current permanent population was estimated and then the maximum population to which a county can grow was modeled by the County Build-out Submodels. Areas which cannot physically or lawfully allow residential development (built-out areas, water bodies, public lands, commercial areas, etc.) were excluded from the County Build-out Submodel. Conversely, the model identified areas where growth is more likely to occur based on proximity to existing infrastructure. This is discussed in detail in chapter titled "Growth Drivers Submodel".

Next, future population change was forecasted from the 2020 launch year estimates to a horizon of 2045 in five-year increments. Projections are based on a combination of historic growth trends (using an approach similar to what BEBR uses for its county-level projections), and spatial constraints and influences, which both restrict and direct growth. Population growth calculations were controlled to BEBR's 2021 medium growth projections (BEBR's latest population forecasts for the years 2025 through 2045), which were available in five-year increments. Because the 2020 Census counts were released during this project, they were used instead of the 2020 BEBR estimates for the county growth controls. The source of the projections is the BEBR publication *Projections of Florida Population by County, 2025-2045, with Estimates for 2020* (Florida Population Studies, Volume 54, Bulletin 189, April 2021), and the source for the 2020 estimates is the 2020 Census. The process is described in the chapter titled "Geospatial Small-Area Population Forecasting Model".

The launch year for the version of the model described in this document was 2020, which was calibrated to the 2020 Census population counts. Projections were made for April 1, 2025 to April 1, 2045 in five-year increments.

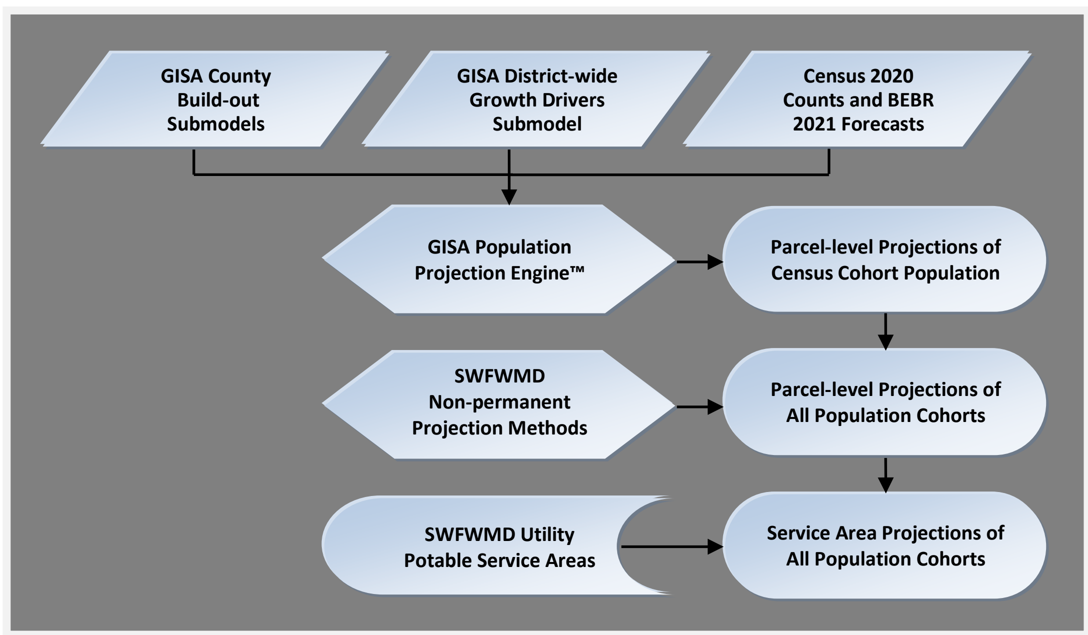

Figure 1. SWFWMD population projection process flowchart

Finally, the parcel level projections were summarized by water utility service area boundaries that SWFWMD maintains in a spatial (GIS) database. These summaries were exported to a Microsoft Excel spreadsheet with separate tabs for each county to facilitate the review and distribution of the results.

COUNTY BUILD-OUT SUBMODELS

The County Build-out Submodels were composed of multiple GIS data elements. Each model was based on each county property appraiser's GIS parcel database, including the associated tax roll information. Other GIS data elements included the 2020 Decennial Census data, SWFWMD wetland data, local government future land use maps, and planned developments.

Parcels

GIS parcel layers and county tax roll databases were obtained from each county property appraiser's office. Parcel geometry was checked for irregular topology, particularly overlaps and fragments. Parcel tables were checked for errors, particularly non-unique parcel identifiers and missing values. Required tax roll table fields include actual year built, Florida Department of Revenue (DOR) land use code, and the total number of existing residential units for each parcel. In cases where values or fields were missing, other data were used. For example, data reported by the State of Florida, military bases, and colleges and universities were used to identify the number of residential units (and population) in some large group quarters facilities.

2020 US Census Data

Some of the essential attribute information to translate parcels to population in the County Build-out Submodels were derived from data from the 2020 Decennial Census. Average housing unit occupancy and population per household by census tract were calculated and then transferred to each county's parcel data. When combined with parcel-level housing units from the property appraisers, these were used to estimate 2020 population at the parcel level. Those estimates were then controlled at county and place levels to the 2020 Census population counts.

In cases where property appraiser data were missing or incomplete, other data were used. For example, because mobile home parks without individually platted parcels may not contain the number of units within the property appraiser data, the number of residential units for some of the parks had to be estimated using information on their websites or hand counts from recent imagery.

Water Management District Boundaries

Each parcel in the County Build-out Submodels was also attributed with water management district boundaries (from SWFWMD's GIS boundary file), which enabled the county submodels for any counties split between two or more water management districts to be summarized by district.

Wetlands

Wetlands (including surface water) are an important consideration when modeling a county's build-out. SWFWMD maintains a detailed wetlands GIS data layer. This data layer contains the location and spatial extent of the wetlands, as well as the specific types of wetlands as defined by SWFWMD land cover classifications. Certain wetland types were identified that would be difficult and expensive to convert to residential development. These areas were identified in the SWFWMD wetland database and applied to the appropriate County Build-out Submodel. The wetland types are listed in Table 1.

Table 1. Wetland land cover codes and descriptions used in the County Build-out Submodels

Code	Description	Code	Description
5100	Streams and waterways	5600	Slough waters
5200	Lakes	6110	Wetland Hardwood Forests
5250	Marshy Lakes	6120	Mangrove swamp
5300	Reservoirs	6170	Mixed wetland hardwoods
5400	Bays and estuaries	6180	Cabbage palm wetland
6181	Cabbage palm hammock	6410	Freshwater marshes
6200	Wetland Coniferous Forest	6420	Saltwater marshes
6210	Cypress	6430	Wet prairies
6220	Pond pine	6440	Emergent aquatic vegetation
6250	Hydric pine flatwoods	6460	Mixed scrub-shrub wetland
6300	Wetland Forested Mixed	6500	Non-vegetated Wetland

Wetland GIS data (using the above classifications) were overlaid with a county's land parcels. The area of wetlands within parcels were calculated and recorded as the water area for that parcel. If the area covered by water within a parcel exceeded ½ acre, it was subtracted from the total area of the parcel feature to determine the relative developable area in that parcel. There were exceptions to this rule. In some cases, parcels with little or no developable area after wetlands were removed were already developed, thus the estimated unit total was not reduced by the wetland acreage. In other cases, inaccurate wetland delineations were overridden, such as when a newly platted residential parcel was shown to be covered by a wetland (Figure 2). In such a case, the parcel was considered developable by the submodel.

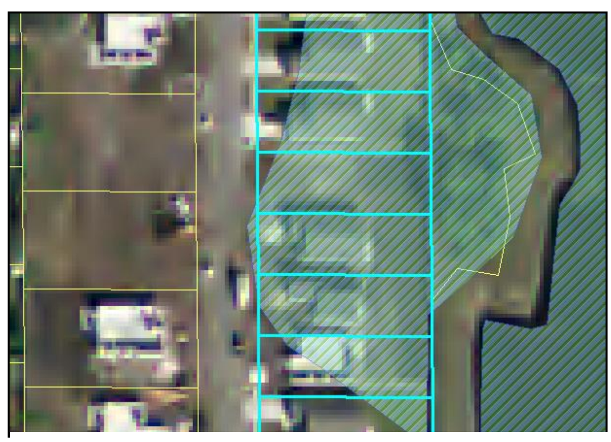

Figure 2. Example of inconsistencies between wetland delineation and residential parcels (outlined here in light blue)

Future Land Use

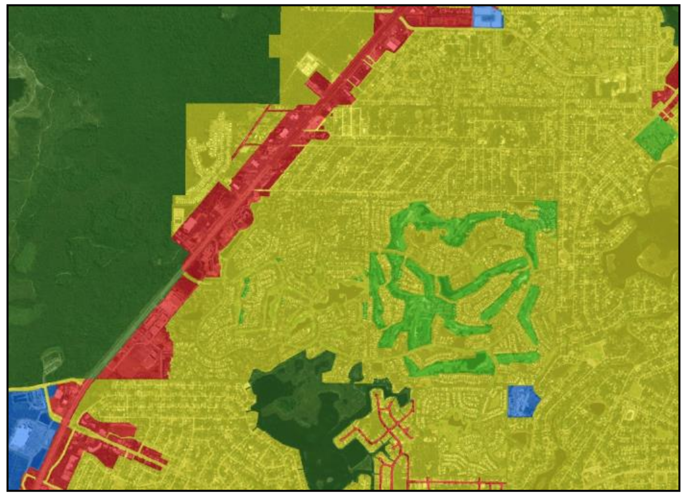

Figure 3. Future land use helps identify future residential areas (here shaded in yellow)

Future land use maps are essential elements of the County Build-out Submodels. These maps help guide where and at what density residential development could occur within a county (Figure 3). Future land use maps are a part of the local government comprehensive plans required for all local governments by Chapter 163, Part II, Florida Statutes. They are typically developed by the local government's planning department, or, in some cases, a regional planning council on behalf of the local government. The planning horizons for these are a minimum of 10 years, and they often extend for 15 to 20 years into the future. Although these future land use maps may be revised over time, they reflect the most up-to-date plan for future growth areas and densities. The latest available future land use maps were obtained and applied to the County Build-out Submodels.

Each land parcel in the County Build-out Submodels received a future land use designation. In places where parcels overlapped multiple future land use areas, the parcel was assigned the future land use class within which its centroid fell. Build-out population was modeled only for future land use classes that allow residential development (which include agriculture and mixed use). Table 2 shows which future land use map classes are typically assigned residential densities in the County Build-out Submodels. There are exceptions to this, as some commercial, institutional and other uses allow residential development in some cases. These are determined on a case-by-case basis, driven largely by each jurisdiction's comprehensive plan.

Table 2. Generalized future land use classes allowing future residential development

Generalized Future Land Use Classes	New Residential Development Allowed by Model
Agricultural	Yes
Low Density Residential	Yes
Medium Density Residential	Yes
High Density Residential	Yes
Mixed Use	Yes
Commercial	No
Recreation / Open Space	No
Conservation / Preservation	No
Industrial	No
Institutional	No
Right of Way	No
Water	No

Development typically does not occur at the maximum densities allowed for each future land use category, so recent development densities were considered a better proxy for future densities than the maximum allowable density. For this reason, the County Build-out Submodels reflect the median density of recent development for each future land use category in the specific incorporated place. For example, if a city's medium density residential future land use designation allows up to 8 housing units per acre, but the median density of units built over the last 20 years is 5.7 housing units per acre, the submodel assumed future densities at 5.7 housing units per acre for that future land use designation in that city. Typically, the median density calculation was limited to the last 20 years of development within each unique combination of land use and jurisdiction, as more recent development was deemed a better proxy for future densities than older development.

In some cases, limiting the historical data to the last 20 years resulted in too small a sample, so either county average values were used (extended beyond the jurisdiction) or all historical development was used (not limited to the last 20 years). In those cases, the determination of which sample to use depended upon the heterogeneity of the category across county jurisdictions, the heterogeneity of historical densities prior to the last 20 years, and our professional judgement. Also, vacant or open parcels less than one acre in size were typically considered single family residential, with one housing unit as the maximum allowable density.

Build-out Density Calculation

Using GIS overlay techniques, attributes of the census, political boundary, wetlands, and future land use data were attributed to each county's parcel map to develop the County Build-out Submodels. These submodels forecast the maximum residential population by parcel at build-out. Figure 4 depicts an example in Highlands County, showing lower densities at build-out in yellow and higher ones in brown.

Figure 4. Example of Build-out Density Submodel for Highlands County shaded by housing units per acre

Census tracts where the 2020 population was zero, and therefore the average persons per housing unit was zero, were assigned the county's average persons per housing unit. Also, if there were tracts with 2020 census values for persons per housing unit greater than zero that were based on a small number of homes with greater than five persons per housing unit, the county's average persons per housing unit was typically used.

Planned Developments

The final step in the development of the County Build-out Submodels was adjusting build-out densities within planned developments to correspond with approved development plans wherever their boundaries are available in a GIS format. Although planned developments often do not develop as originally planned by the developer, the total number of units planned (regardless of timing) is likely to be a better forecast of the units at build-out than one based on the median historic densities. Therefore, in each of the County Build-out Submodels, parcels with centroids within a planned development were attributed with the name of the development. The build-out densities for those parcels were adjusted so that the total build-out for the development was consistent with the development plan, and the build-out population for that area was recalculated.

GROWTH DRIVERS SUBMODEL

The Growth Drivers Submodel is a district-wide, raster (cell-based) GIS model representing development potential. The submodel is a continuous surface of 10-meter cells containing values of 0-100, with '100' having the highest development potential and '0' having the lowest development potential. It influences the GSAPF Model by factoring in the attraction of certain spatial features, or growth drivers on development. These drivers were identified from transportation and land use/land cover data. They included the following:

1. Proximity to roads and interchanges prioritized by level of use (with each road type modeled separately)
2. Proximity to planned developments
3. Proximity to existing commercial development (based on parcels with commercial land use codes deemed attractors to residential growth)
4. Proximity to existing residential development
5. Proximity to coastal and inland waters

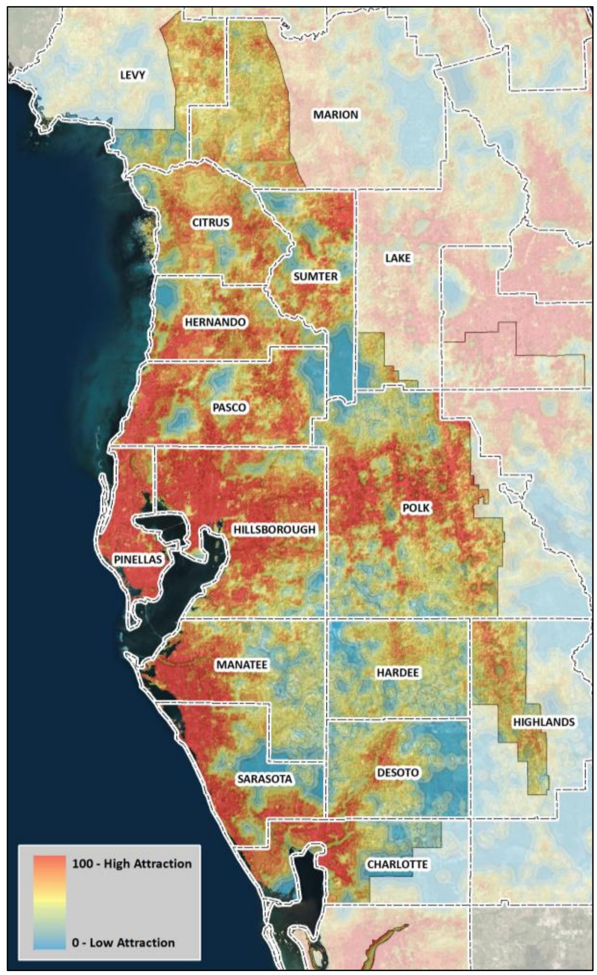

Figure 5. Growth Drivers Submodel

Figure 5 depicts the Growth Drivers Submodel for SWFWMD, with high development potential in red, moderate development potential in yellow and low development potential in blue.

Data used for generating the Growth Drivers Submodel and their sources are listed in Table 3 below.

Table 3. GIS datasets used in the Growth Drivers Submodel

Growth Driver	Data Source
Roads and Limited Access Road Interchanges	Florida Department of Transportation (FDOT) Major Roads: Functional Classification (FUNCLASS), and FDOT Limited Access Road Interchanges
Existing Residential Land Uses	County Property Appraiser Parcel Data
Selected Existing Commercial Land Uses	County Property Appraiser Parcel Data
Coastal and Inland Waters	SWFWMD Land Cover Data, and Florida Geographic Data Library (FGDL) Coastline Data
Planned developments	Multiple sources, including Regional Planning Councils, local governments, SWFWMD and GIS Associates

Each of the drivers listed in Table 3 were used as independent variables in a logistic regression equation. Dependent variables included existing residential units built during or after 1995 as the measure of "presence", and large undeveloped vacant parcels outside of large, planned developments were used to measure "absence". The resulting equation could then be applied back to each of the regional grids resulting in a single regional grid with values 0 through 100, for which a value of 0 represented the lowest relative likelihood of development, and a value of 100 represented the highest relative likelihood of development.

This seamless, "regional" submodel covers all the counties that are all or partially within the Southwest Florida Water Management District, plus a one-county buffer to account for growth drivers outside the District that could influence growth within the District. Each county was then individually reclassified to stretch its mean driver values between 0 and 100, to better differentiate the highest values relative to the rest of the county.

This submodel was then used by the GSAPF Model to rank undeveloped parcels based on their development potential, which is explained in the "Growth Calculation Methodology" section. Note that growth may still occur in areas assigned relatively low values from this model based on the historical growth trends. This model only helps guide growth when the county Model projection totals are below the BEBR targets.

GEOSPATIAL SMALL-AREA POPULATION FORECASTING MODEL

GISA's Geospatial Small-Area Population Forecasting Model (GSAPF Model) integrates the County Build-out Submodels and the Growth Drivers Submodel with the GISA Population Projection Engine™, which makes the projection calculations using a combination of those submodels, historic growth trends, and growth controls from BEBR's county-level forecasts.

Historic Growth Trends

The historic growth trends were based on historic population counts from the 1990, 2000, 2010, and 2020 censuses. For 1990, 2000, and 2010, census block population counts from the Florida House of Representatives Redistricting Data (available at the Office of Economic and Demographic Research's website http://edr.state.fl.us/Content/population-demographics/redistricting/2010-redistricting/index.cfm) were summarized at the 2020 tract level, and combined with the 2020 tract population estimates. These estimates were used to produce twelve tract level projections using five different demographic extrapolation methods. The three highest and three lowest calculations were discarded to moderate the effects of extreme projections (Smith and Rayer 2004). The remaining six projections were then averaged.

The five demographic extrapolation methods for projecting population utilized by the model were:

1. Linear
2. Exponential
3. Constant Share
4. Share-of-Growth
5. Shift-Share

The Linear and Exponential techniques employ a bottom-up approach, extrapolating the historic growth trends of each census tract with no consideration for the county's overall growth. The Constant Share, Share-of-Growth and Shift-Share techniques employ a ratio allocation, or top-down approach, allocating a portion of the total projected county population or growth to each census tract based on that census tract's percentage of county population or growth over the historical period. Each of the five methods is a good predictor of growth in different situations and growth patterns, so using a combination of all five was the best way to avoid the largest possible errors resulting from the least appropriate techniques for each census tract within the 17-county area (Sipe and Hopkins 1984). This approach is similar to the one BEBR uses for its county population forecasts, but the base periods and the number of projections are somewhat different because annual estimates are not available at the tract level.

The calculations associated with the five statistical methods are described on the following pages. The launch year was 2020, and the projections were made for 2025, 2030, 2035, 2040, and 2045.

1. **Linear Projection Method:** The Linear Projection Method assumes that the change in the number of persons for each census tract will be the same as during the base period (Rayer and Wang, 2020). Three linear growth rate calculations were made, 1990 to 2000, 2000 through 2020, and 2010 through 2020. In the three Linear methods (LIN), population growth was calculated using the following formulas:

$$LIN1 = \frac{(TractPop2020 - TractPop1990)}{30} * 5$$

$$LIN2 = \frac{(TractPop2020 - TractPop2000)}{20} * 5$$

$$LIN3 = \frac{(TractPop2020 - TractPop2010)}{10} * 5$$

2. **Exponential Projection Method:** The Exponential Projection Method assumes that population will continue to change at the same percentage rate as during the base period (Rayer and Wang, 2020). One calculation was made from 2010 through 2020. In the Exponential method (EXP), population growth was calculated using the following formula:

$$EXP = (TractPop2020 * e^{5r}) - TractPop2020$$

$$Where, r = \frac{ln\frac{TractPop2020}{TractPop2010}}{10}$$

3. **Constant Share Projection Method:** The Constant Share Projection Method assumes that each census tract's percentage of the county's total population (CntyPop) will be the same as over the base period (Rayer and Wang, 2020). One Constant Share (CS) calculation was made based on 2020 shares, but the weight was doubled for this method. Population growth was calculated using the following formula (using 2020–2025 as an example):

$$CS = \frac{TractPop2020}{CntyPop2020} * (CntyPop2025 - CntyPop2020)$$

4. **Share-of-Growth Projection Method**: The Share-of-Growth Projection Method assumes that each census tract's percentage of the county's total growth will be the same as over the base period (Rayer and Wang, 2020). However, if population change is negative at the tract level and positive at the county level (or vice versa), higher county-level projections would result in larger declines in tract projections. This is counterintuitive, so the "Plus-minus" variant of the Share-of-Growth Method was used (Rayer, 2015). Three Share-of-Growth calculations were made, 1990 through 2020 2000 through 2020, and 2010 through 2020.

In the three Share-of-Growth (SOG) calculations, population growth was calculated using the following formulas if the changes in growth over the base period for the tract and county were both positive or both negative (using 2020–2025 as an example):

$$SOG1 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right] * \left[\begin{array}{l} \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right] \end{array}\right] \div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,

$$SOG = \frac{(TractPop2020 - TractPop1990)}{(CntyPop2020 - CntyPop1990)} * (CntyPop2025 - CntyPop2020)$$

$and\ ABS = Absolute\ Value$

$$SOG2 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right] * \left[\begin{array}{l} \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right] \end{array}\right] \div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,

$$SOG = \frac{(TractPop2020 - TractPop2000)}{(CntyPop2020 - CntyPop2000)} * (CntyPop2025 - CntyPop2020)$$

$and\ ABS = Absolute\ Value$

$$SOG3 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right] * \left[\begin{array}{l} \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right] \end{array}\right] \div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,

$$SOG = \frac{(TractPop2020 - TractPop2010)}{(CntyPop2020 - CntyPop2010)} * (CntyPop2025 - CntyPop2020)$$

$and\ ABS = Absolute\ Value$

If the changes in growth over the base period were negative at the tract level and positive at the county level or vice versa, the population growth was calculated using the following formulas (using 2020–2025 as an example):

$$SOG1 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}/5\right] * \left[\begin{array}{l}\left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right]\end{array}\right]$$
$$\div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,
$$SOG = \frac{(TractPop2020 - TractPop1990)}{(CntyPop2020 - CntyPop1990)} * (CntyPop2025 - CntyPop2020)$$
$and\ ABS = Absolute\ Value$

$$SOG2 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}/5\right] * \left[\begin{array}{l}\left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right]\end{array}\right]$$
$$\div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,
$$SOG = \frac{(TractPop2020 - TractPop2000)}{(CntyPop2020 - CntyPop2000)} * (CntyPop2025 - CntyPop2020)$$
$and\ ABS = Absolute\ Value$

$$SOG3 = \left[\frac{SOG}{CntyPop2025 - CntyPop2020}/5\right] * \left[\begin{array}{l}\left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] \\ + \left[1 - CntySum\left[\frac{SOG}{CntyPop2025 - CntyPop2020}\right]\right]\end{array}\right]$$
$$\div \left[CntySum\left[\frac{ABS(SOG)}{CntyPop2025 - CntyPop2020}\right]\right] * (CntyPop2025 - CntyPop2020)$$

Where,
$$SOG = \frac{(TractPop2020 - TractPop2010)}{(CntyPop2020 - CntyPop2010)} * (CntyPop2025 - CntyPop2020)$$
$and\ ABS = Absolute\ Value$

6. **Shift-Share Projection Method:** The Shift-Share Projection Method assumes that each census tract's percentage of the county's total annual growth will change by the same annual amount as over the base period (Rayer and Wang, 2020). Three Shift-Share calculations were made, 1990 through 2020, 2000 through 2020, and 2010 through 2020. In the three Shift-Share Projection Method (SSH) calculations, population growth was calculated using the following formulas (using the five years from 2020–2025 as an example):

$$SSH_1 = \left[\frac{TractPop2020}{CntyPop2020} + \left[\frac{\left(\frac{TractPop2020}{CntyPop2020} - \frac{TractPop1990}{CntyPop1990}\right)}{30} * 5\right]\right] * \begin{pmatrix} CntyPop2025 \\ -CntyPop2020 \end{pmatrix}$$

$$SSH_1 = \left[\frac{TractPop2020}{CntyPop2020} + \left[\frac{\left(\frac{TractPop2020}{CntyPop2020} - \frac{TractPop2000}{CntyPop2000}\right)}{20} * 5\right]\right] * \begin{pmatrix} CntyPop2025 \\ -CntyPop2020 \end{pmatrix}$$

$$SSH_2 = \left[\frac{TractPop2020}{CntyPop2020} + \left[\frac{\left(\frac{TractPop2020}{CntyPop2020} - \frac{TractPop2010}{CntyPop2010}\right)}{10} * 5\right]\right] * \begin{pmatrix} CntyPop2025 \\ -CntyPop2020 \end{pmatrix}$$

7. **Average of the Projection Extrapolations:** The three minimum and three maximum of the twelve calculations for each census tract were removed to eliminate the most extreme results of the thousands of heterogeneous census tracts within the 17-county area. The six remaining calculations were averaged to account for the considerable variation in growth rates and patterns over all of the census tracts within the 17-county area. All four remaining methods were weighted equally, and the average was calculated using the following formula:

$$AVG = \frac{\begin{array}{c}(LIN1 + LIN2 + LIN3 + EXP + CS + CS + SOG1 + SOG2 + SOG3 + SSH1 + SSH2 + SSH3) \\ -(MIN1 + MIN2 + MIN3 + MAX1 + MAX2 + MAX3)\end{array}}{6}$$

$$\begin{array}{l} \textit{Where,} \\ MIN1 - MIN3 \textit{ are the four lowest growth calculations for each tract, and} \\ MAX1 - MAX3 \textit{ are the four highest growth calculations for each tract.} \end{array}$$

Growth Calculation Methodology

After the development of the County Build-out Submodels, the Growth Drivers Submodel, and the historic growth trends, the GISA Population Projection Engine™ was used to make the growth calculations. The methodology for calculating growth for each projection increment included the following steps:

1. Applying the tract-level projected growth to parcels within a particular tract.
2. Checking growth projections against build-out population, and reducing any projections exceeding build-out to the build-out numbers.
3. After projecting growth for all census tracts within a particular county, summarizing the resulting growth and comparing it against countywide BEBR target growth. This step led to two scenarios:
    a. If the Small-Area Population Forecasting Model's projections exceeded the BEBR target growth, projected growth for all tracts was reduced by the percentage that the projections exceeded the BEBR target.
    b. If the Small-Area Population projection model's projections were less than the BEBR target, the model would continue growing the county using the Growth Drivers Submodel until the BEBR target growth for each five-year increment was reached. This process involved developing parcels with growth driver values in the highest decile that had available capacity for growth.

Counties that are partially within another water management district were processed in their entirety and controlled to the BEBR-based target growth.

NON-PERMANENT POPULATION PROJECTIONS

In addition to the Census Population Cohort projections generated by the GSAPF Model, projections of non-permanent population were also made. Those projections include peak seasonal population, functionalized seasonal population, tourist population and net commuter population. The methods were developed by SWFWMD and implemented by GIS Associates for projecting those population types are described in this section.

Peak Seasonal Population Cohort

Seasonal population was estimated by SWFWMD using a combination of 2010 census data and emergency room admissions data, both at the Zip Code Tabulation Area (ZCTA) level. Average 2009 - 2011 emergency room admissions data was utilized for the 45-74 age cohort, which is typical of seasonal residents. A "Seasonal Resident Ratio" was calculated by ZCTA to estimate the proportion of peak (including seasonal) to permanent population.

The Seasonal Resident Ratio was derived using the following steps:

1. Subtract total 2009 - 2011 total third quarter (Q3, or July, August and September) hospital admissions from first quarter (Q1, or January, February and March) admissions.
2. Calculate the average annual difference between Q1 and Q3 by dividing above result by three.
3. Calculate a seasonal population estimate for each ZCTA by dividing the above difference by the probability of the population in the 45-74 age cohort being admitted to the emergency room (approximately 2.23%).
4. Calculate the Seasonal Resident Ratio by adding the seasonal population to the permanent population and dividing that total by the permanent population.

The number of seasonal households was then estimated using the following steps:

1. Multiply the permanent population in households (from the 2010 census) by the Seasonal Resident Ratio.
2. Subtract the permanent population in households from above result.
3. Divide above result by the lesser of SWFWMD's seasonal persons per household (1.95) or the census permanent persons per household for each ZCTA.

The ratio of seasonal to total households was then calculated by dividing seasonal households by the sum of seasonal and permanent households. Seasonal peak population was then calculated using the following steps:

1. Subtract vacant housing units for reasons other than seasonal, recreational, or occasional use from total housing units (from the 2010 census).
2. Multiply above result times the seasonal to total household ratio.
3. Multiply above result times the lesser of SWFWMD's seasonal persons per household (1.95) or the census permanent persons per household for each ZCTA.

Because the Census Population Cohort contains some non-permanent residents who complete the census forms in Florida but reside for part of the year outside of Florida, it was also necessary to calculate the permanent population. Permanent population was calculated using the following steps:

1. Subtract vacant housing units for reasons other than seasonal, recreational, or occasional use from total housing units (from the 2010 census).
2. Multiply above result times one minus the seasonal to total household ratio.
3. Multiply above result times the census permanent persons per household for each ZCTA.

The ratio of total unadjusted peak population to total census population was then calculated by dividing the sum of the seasonal peak population, the permanent population in households, and the group quarters population (from the 2010 census) by the total census population. This ratio was then applied by GISA to the future projections of the Census Population Cohort from the GSAPF Model to derive parcel level peak population projections (which also includes permanent residents).

Functionalized Seasonal Population Cohort

The functional population is the peak seasonal resident population reduced to account for the percentage of the year seasonal residents typically reside elsewhere, and the absence of indoor water use during that time. It was calculated using the following generalized steps:

1. Utilize the following metrics previously derived by SWFWMD data and surveys:
    a. The appropriate proportion of the year seasonal residents spend in Florida, which varies from beach destination counties (44.2%) to non-beach destination counties (56.7%).
    b. The seasonal resident adjustment based on average per capita water use.
        i. The five-year District-wide average per capita use is 132 gallons per person per day, and 69.3 gallons is estimated indoor use and 62.7 gallons for outdoor use.
        ii. The adjustment factor is calculated using the following equation for "beach destination" counties (Charlotte, Manatee, Pinellas and Sarasota):
        ((0.442 x 132 gpd) + ((1 – 0.442) x 62.7 gpd)) / 132 gpd = 0.707
        iii. The adjustment factor is calculated using the following equation for "non-beach destination counties":
        ((0.567 x 132 gpd) + ((1 – 0.567) x 62.7 gpd)) / 132 gpd = 0.773
2. Calculate "functionalized" seasonal population by multiplying the seasonal peak population by the appropriate seasonal resident adjustment factor for the particular county (0.707 or 0.773).

The ratio of total functional to total census population was then calculated by dividing the sum of the functionalized seasonal population, the permanent population, and the group quarters population (from the 2010 census) by the total census population. This ratio was then applied by GISA to the future projections of the Census Population Cohort from the GSAPF Model to derive parcel level functionalized seasonal population projections (which also includes permanent residents).

Tourist Population Cohort

The tourist population projections were based on 25 years (1996-2020) of county-level lodging room data from the Florida Department of Business and Professional Regulation (DBPR). The SWFWMD methodology for projecting future tourist rooms by county utilizes two different methods and averages the two results for each county.

The first method projects the increase in rooms by county by extrapolating the linear trend using the least squares method derived from the last 25 years of county total room estimates.

A second method projects future rooms based on projections of employment in the Accommodation and Food Services industries (from data from Woods and Poole). This is also an extrapolation of a linear trend using the least squares method, but rooms by county are projected as a function of a county's employment projections rather than time.

SWFWMD staff previously tested both methods by projecting values for the years 2007-2013 using room estimates from 1996-2006. Based on the differences between actual room estimates and projected values for 2007-2013, neither method was clearly superior to the other. For that reason, SWFWMD staff opted to use both methods. The results of both methods were averaged, but only after adjusting for the average 2007-2013 error for each projection in each county.

These projections of future rooms were then converted to "functionalized" tourist population by applying various county-level average unit occupancy and party size ratios derived from SWFWMD surveys. SWFWMD also updated the values associated with locations identified as short term rentals for this projection set based on SWFWMD research.

These projections of tourist population were joined to the existing lodging facility locations, which were geocoded by SWFWMD. No attempt was made to forecast new locations of lodging facilities, as:

1. The precise locations would be highly speculative.
2. It was assumed that lodging facilities often are built in the general vicinity of existing lodging facilities, or at least in close enough proximity to be within the same utility service area.

Net Commuter Population Cohort

The net commuter population projections were based on net commuter data provided by SWFWMD. A census tract ratio was developed of net commuters to total census population. This ratio was then applied to the future projections of the Census Population Cohort from the GSAPF Model to derive parcel level projections for net commuter population. That population was then "functionalized" with the following ratios:

1. 8 / 24 (typical working hours per day)
2. 5 / 7 (typical working days per week)

By applying both of these ratios to the net commuter population, the resulting functional net commuter population is 23.8% of the actual net commuter population. This functional number better reflects the water use that is expected for net commuters.

Note that the net commuter population projection summaries by utility service area were often negative, as many utilities serve "bedroom communities" and other areas where more residents work outside the utility service area than the population (residents and non-residents) employed within it. Only positive net commuter populations were included in a utility's total functional population.

UTILITY SERVICE AREA POPULATION SUMMARIES

The parcel level population projections for all population cohorts discussed above were then summarized by water utility service area boundaries for all utilities mapped by SWFWMD, which typically are those that average more than 0.1 million gallons per day (mgd) of total water use. These service areas, maintained by SWFWMD, were overlaid with each county's parcel level results, and each parcel within a service area was assigned a unique identifier for that service area. The projected population was then summarized by that identifier and joined to SWFWMD's public service area boundary database to produce tabular and spatial output. Note that these service areas change over time, so for any future use of these deliverables, it is important to match this projection set only with the service areas included in the GIS deliverables for this project.

PROJECTION DELIVERABLES

The final population projections were delivered in multiple formats, including:

1. GIS – Esri file geodatabases, with individual feature classes for each county containing parcel level results, and a single feature class districtwide of utility service areas and their population summaries.
2. Tabular – Excel spreadsheet summaries by utility service area

The summaries of population outside of service areas include population with private wells for potable use (considered to be domestic self-supply, or DSS) or small utilities without a service area boundary in SWFWMD's database. Small utilities are generally defined as those utilities permitted for less than 100,000 gallons per day (gpd). However, there are some small utilities in that category that are included here because their service area boundaries are in SWFWMD's database.

Note that these service area population summaries may include some self-supplied populations (or populations with private wells) that reside within the service areas. The population projections utilized for SWFWMD's planning and permitting may vary from the raw functional population projections developed with the model due to service area boundary changes, self-supplied population, and current population served reported by utility using the required population estimation methodology in Part D of the Water Use Permitting Manual.

The tabular deliverables were parcel summaries at the utility level. Table 4 on the following page shows the service area population projection summaries table for Manatee County.

The GIS outputs are useful for quality assuring the results and inputs, for maintaining the projection inputs over time, and for graphically depicting projected patterns of future population growth. Figure 7 on Page 22 is an example of this data in the Tampa Bay area, depicting current population in white and projected 2025-2045 growth in red.

Table 4. Utility service area population projection summaries table for Manatee County

BEBR / Census Population (Permanent Resident Population)						
Utility Name	POP20	POP25	POP30	POP35	POP40	POP45
OUTSIDE SERVICE AREAS	12,014	12,744	13,357	13,882	14,350	14,779
CITY OF BRADENTON PUBLIC WORKS	55,698	58,696	60,089	61,043	61,260	61,360
CITY OF PALMETTO PUBLIC WORKS	14,225	15,215	16,187	17,220	17,928	18,574
MANATEE COUNTY UTILITIES DEPARTMENT	314,921	349,187	379,211	404,099	427,303	447,925
TOWN OF LONGBOAT KEY	2,746	2,859	2,859	2,859	2,859	2,859
PALMETTO PARK	25	26	26	26	26	26
PINES TRAILER PARK	41	40	40	39	42	46
LAZY ACRES	39	39	39	39	39	39

Peak Seasonal Population (Includes BEBR / Census)						
Utility Name	POP20_P	POP25_P	POP30_P	POP35_P	POP40_P	POP45_P
OUTSIDE SERVICE AREAS	12,334	13,081	13,707	14,245	14,723	15,162
CITY OF BRADENTON PUBLIC WORKS	59,438	62,580	64,034	65,029	65,254	65,358
CITY OF PALMETTO PUBLIC WORKS	15,829	16,930	18,012	19,161	19,949	20,667
MANATEE COUNTY UTILITIES DEPARTMENT	347,083	384,055	416,396	443,212	468,091	490,172
TOWN OF LONGBOAT KEY	5,855	6,097	6,097	6,097	6,097	6,097
PALMETTO PARK	28	28	28	29	29	29
PINES TRAILER PARK	77	75	74	73	79	85
LAZY ACRES	43	43	43	43	43	44

Functionalized Seasonal Population (Includes BEBR / Census)						
Utility Name	PERMSEAS20	PERMSEAS25	PERMSEAS30	PERMSEAS35	PERMSEAS40	PERMSEAS45
OUTSIDE SERVICE AREAS	12,229	12,969	13,591	14,124	14,599	15,034
CITY OF BRADENTON PUBLIC WORKS	58,288	61,370	62,797	63,772	63,993	64,095
CITY OF PALMETTO PUBLIC WORKS	15,386	16,457	17,508	18,625	19,392	20,089
MANATEE COUNTY UTILITIES DEPARTMENT	339,792	376,049	407,774	434,057	458,510	480,217
TOWN OF LONGBOAT KEY	5,443	5,667	5,667	5,667	5,667	5,667
PALMETTO PARK	28	28	28	28	28	28
PINES TRAILER PARK	74	72	71	70	75	82
LAZY ACRES	42	42	42	42	42	42

Functionalized Tourist Population						
Utility Name	POP20_T	POP25_T	POP30_T	POP35_T	POP40_T	POP45_T
OUTSIDE SERVICE AREAS	5	6	7	8	9	10
CITY OF BRADENTON PUBLIC WORKS	1,810	2,058	2,340	2,645	2,975	3,333
CITY OF PALMETTO PUBLIC WORKS	21	24	27	30	34	38
MANATEE COUNTY UTILITIES DEPARTMENT	12,266	13,944	15,854	17,920	20,157	22,581
TOWN OF LONGBOAT KEY	1,084	1,232	1,401	1,583	1,781	1,995
PALMETTO PARK	-	-	-	-	-	-
PINES TRAILER PARK	-	-	-	-	-	-
LAZY ACRES	-	-	-	-	-	-

Functionalized Net Commuter Population						
Utility Name	POP20_NC	POP25_NC	POP30_NC	POP35_NC	POP40_NC	POP45_NC
OUTSIDE SERVICE AREAS	(717)	(762)	(799)	(831)	(860)	(886)
CITY OF BRADENTON PUBLIC WORKS	1,255	1,505	1,728	1,825	1,849	1,846
CITY OF PALMETTO PUBLIC WORKS	(362)	(405)	(450)	(502)	(546)	(587)
MANATEE COUNTY UTILITIES DEPARTMENT	(7,627)	(8,497)	(9,290)	(10,085)	(10,799)	(11,409)
TOWN OF LONGBOAT KEY	129	133	133	133	133	133
PALMETTO PARK	(2)	(2)	(2)	(2)	(2)	(2)
PINES TRAILER PARK	-	-	-	-	-	-
LAZY ACRES	-	-	-	-	-	-

Total Functional Population (Functionalized Seasonal + Tourist + Positive Net Commuter)						
Utility Name	POP20_TF	POP25_TF	POP30_TF	POP35_TF	POP40_TF	POP45_TF
OUTSIDE SERVICE AREAS	12,234	12,975	13,597	14,131	14,607	15,043
CITY OF BRADENTON PUBLIC WORKS	61,353	64,933	66,865	68,242	68,817	69,274
CITY OF PALMETTO PUBLIC WORKS	15,407	16,481	17,535	18,655	19,426	20,128
MANATEE COUNTY UTILITIES DEPARTMENT	352,058	389,993	423,628	451,977	478,667	502,798
TOWN OF LONGBOAT KEY	6,655	7,032	7,201	7,383	7,581	7,795
PALMETTO PARK	28	28	28	28	28	28
PINES TRAILER PARK	74	73	72	71	76	82
LAZY ACRES	42	42	42	42	42	42

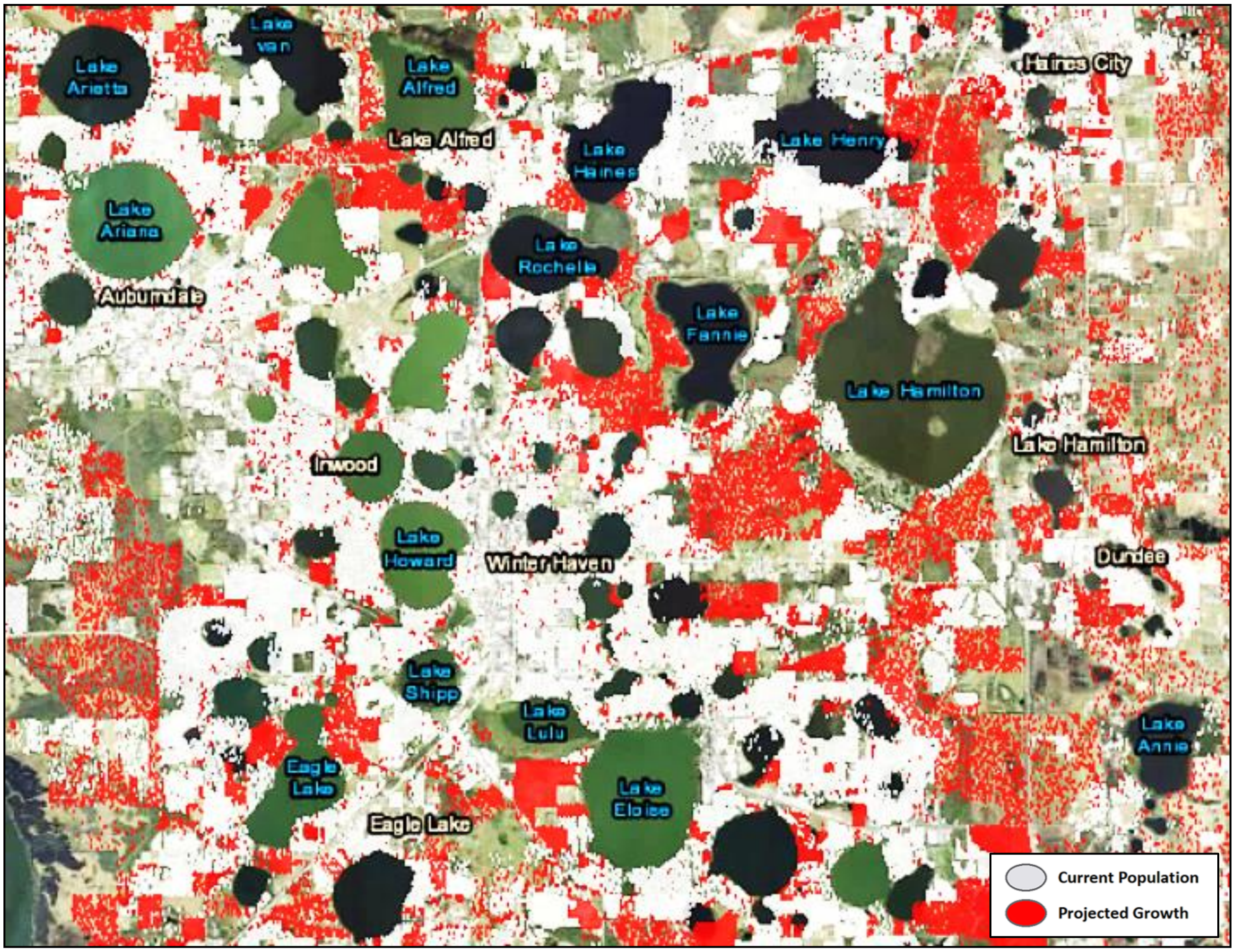

Figure 6. Parcels with 2020 population (white) and projected 2045 growth (red) in Polk County.

CONCLUSIONS

Small area population projections have become increasingly important for planning and permitting, particularly in an area experiencing rapid growth. With ever changing population dynamics and requirements for water supply planning and permitting, it is critical for SWFWMD to be able to accurately forecast population and water demand for small areas, and to be able to update these projections regularly and in a consistent and cost-effective manner. For these reasons, GIS Associates updated and implemented its Geospatial Small-Area Population Forecasting Model (GSAPF Model) and implemented SWFWMD methods for projecting seasonal, tourist and net commuter populations. The GSAPF Model was updated with current data to project population in an efficient and consistent manner throughout the entire 17-county region. Controlling the projections to BEBR’s county-level forecasts provided consistency with other projections made by state and local governments, while at the same time providing the spatial precision needed for water supply planning, water use permitting and groundwater modeling.

REFERENCES

BEBR (2020). "Florida Estimates of Population 2020". Bureau of Economic and Business Research, University of Florida, Gainesville, Florida. URL: https://www.bebr.ufl.edu/sites/default/files/Research%20Reports/estimates_2020.pdf.

Doty, Richard L., Suzanne L. Roulston-Doty, and Jason K. Teisinger (2020). "The Geospatial Small-Area Population Forecasting (GSAPF) Model Methodology Used by the Southwest Florida Water Management District". GIS Associates, Inc. Gainesville, Florida.

Florida House of Representatives Redistricting Committee (2002). "Florida Redistricting Data", http://www.floridaredistricting.org/FredsData.html (Accessed 10/07/2002), Florida House of Representatives, Tallahassee, Florida.

Rayer, Stefan, and Ying Wang (2021). "Projections of Florida Population by County, 2025-2045, with Estimates for 2020." Florida Population Studies, Volume 54, Bulletin 189, April 2021. Bureau of Economic and Business Research, University of Florida. Gainesville, Florida. URL: https://www.bebr.ufl.edu/sites/default/files/Research%20Reports/projections_2021.pdf.

Rayer, Stefan (2015). Personal communication September 9, 2015. Bureau of Economic and Business Research, University of Florida. Gainesville, Florida.

Sipe, Neil G., and Robert W. Hopkins (1984), "Microcomputers and Economic Analysis: Spreadsheet Templates for Local Government", BEBR Monographs, Bureau of Economic and Business Research, University of Florida, Gainesville, Florida.

Smith, Stanley K., and Stefan Rayer (2004). "Methodology for Constructing Projections of Total Population for Florida and Its Counties: 2003-2030." Bureau of Economic and Business Research, University of Florida. Gainesville, Florida.

U.S. Census Bureau (2021). 2020 US Census Tracts in Florida with Selected Fields from 2020 Redistricting Summary File. Vector Digital Data. U.S. Census Bureau, Suitland, MD.

U.S. Census Bureau (2011). 2010 US Census Tracts in Florida (with Selected Fields from 2010 Redistricting Summary File and Summary File 1). ETAT.CENTRACT2010_AUG11. Vector Digital Data. U.S. Census Bureau, Suitland, MD.