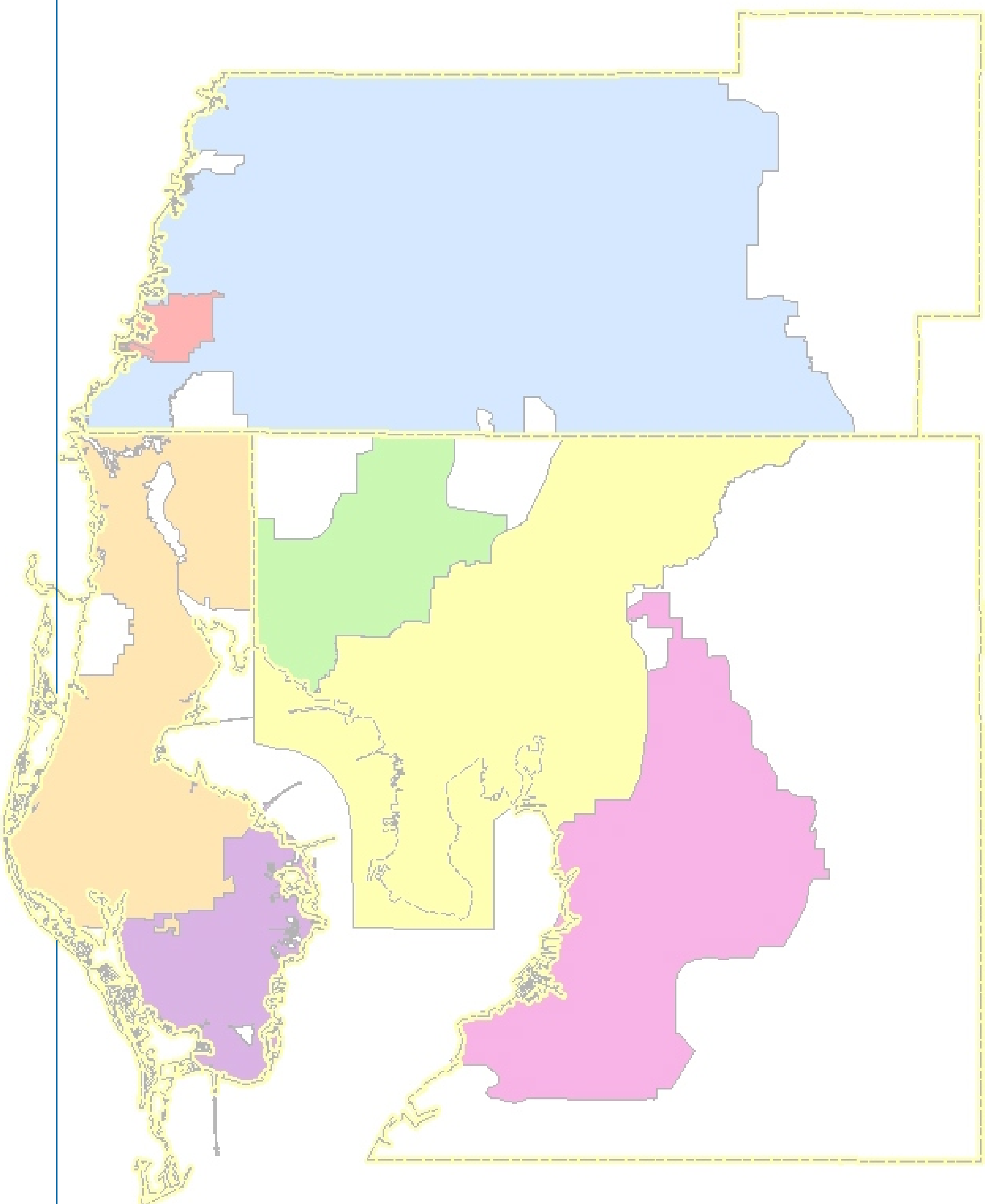
TAMPA BAY WATER
Supplying Water To The Region

The Tampa Bay Water Long-Term Demand Forecasting Model

June 2004



HAZEN AND SAWYER
Environmental Engineers & Scientists

PMCL

Acknowledgements

Hazen and Sawyer would like to acknowledge Tampa Bay Water, its member governments (cities of Tampa, St. Petersburg and New Port Richey, and the counties of Hillsborough, Pinellas, and Pasco), the Southwest Florida Water Management District and others for their support and contributions to this project. The following individuals provided comments and/or data that was helpful in developing the Long-term Demand Forecasting Model:

Tampa Bay Water
David Bracciano
Alison Adams

Hillsborough County
Jim Jeffers
Rachel Rivera

Pinellas County
Tim Wiley
James Butler

City of Tampa
Marjorie Guillory
Karl Craig

Arnie Neeman

City of St. Petersburg
Joan Bradshaw
Gregory Cram

City of New Port Richey
Mary Healey
Clark Jones

Pasco County
Annamarie O'dell
Monte Welton

Southwest Florida Water Management District
Jay Yingling

Hazen and Sawyer
Damann Anderson
Sanjay Puranik
Avinash Puranik
John Clayton

PMCL
Jack Kiefer

Table of Contents

Appendices

References

List of Figures

List of Tables

Executive Summary

Tampa Bay Water, like many other major water utilities, faces a difficult task in forecasting potable water needs over a relatively long time horizon for a large and growing service population. In many cases, these water needs dictate and are influenced by size and timing of future supply and demand expenditures, such as additions to supply and treatment capacity, implementation of water conservation programs, development of reclaimed water projects, and changes to the structure and level of water prices.

In the past, Tampa Bay Water depended mostly upon groundwater sources to meet regional demand. In the early nineties, it was determined that the continued pumping from these wellfields was contributing to depletion of wetlands and lowering of water levels in various lakes, especially in the vicinity of well locations. The Northern Tampa Bay New Water Supply and Groundwater Agreement (WCRWSA, 1998) was incorporated between Southwest Florida Water Management District (SWFWMD), Tampa Bay Water, and its member governments, wherein all parties agreed to cooperate with each other to develop new water supply and reduce pumpage from existing regional wellfields. The Consolidated Water Use Permit for 11 long-producing regional wellfields required allowable withdrawals of 158 MGD be reduced to 121 MGD or less by the end of 2003 and 90 MGD or less by the end of 2008. The Partnership Agreement required Tampa Bay Water and its member governments to continue to plan, coordinate, develop, construct and implement new water supplies, conservation and reclaimed water projects.

The Master Water Plan (MWP) developed by Tampa Bay Water provided a framework for developing alternative sources to groundwater and the related transmission, treatment, and storage components. As of this writing, Phase I of the MWP is almost complete with various alternative water sources developed including a 66 MGD surface water treatment plant, a 25 MGD desalination plant and a 15 billion gallon storage reservoir. These sources are expected to allow the Agency to meet the increasing demand over the next few years.

In order to develop a better understanding of increases in demand and its implications on supply development options (size and timing), the Agency commissioned the development of a long-term demand forecasting system (LTDFS). The LTDFS is a major initiative to quantify how socioeconomic, meteorological, and policy conditions in its service area influence potable water demand.

The initial step in achieving this Board directive was creation of a Long-term Demand Forecasting System Technical Advisory Committee (LTDFS TAC) comprised of representatives of Tampa Bay Water's member governments and the SWFWMD. The following personnel represented their respective member governments, agencies, or SWFWMD.

- Mr. Tim Wiley, Pinellas County
- Mr. Jim Jeffers, Hillsborough County
- Ms. Patti Anderson, City of St. Petersburg
- Mr. Doug Bramlett, Pasco County
- Mr. Karl Craig, City of Tampa
- Ms. Mary Healey, City of New Port Richey
- Mr. Jay Yingling, SWFWMD
- Mr. Dave Bracciano, Tampa Bay Water
- Dr. Alison Adams, Tampa Bay Water

The purpose of this TAC was to periodically review the project progress and provide pertinent comments. The TAC members played a major role in facilitating billing and rate data collection by directing project team members to appropriate personnel within each utility.

A principal product of this initiative was a regional demand model that calculated demand as a function of meteorological, socioeconomic, and policy conditions. After the model was developed, projections of socioeconomic growth, meteorological conditions, and policy were determined and applied to the demand model to generate forecasts of regional potable water demand. These forecasts will assist decision-makers in understanding and planning for future water needs. In addition, potential variability in projected socioeconomic, meteorological, and policy conditions was determined, then applied to the regional demand model to forecast a range, or interval, of possible values for demand growth. These interval-based forecasts will help decision-makers assess risk of supply shortfalls relative to demand and plan for demand and supply expenditures, while avoiding unacceptable risk.

This document chronicles the development of Tampa Bay Water's Long-Term Demand Forecasting Model and its implementation for developing long-term point and probabilistic demand forecasts. The project was completed for Tampa Bay Water by a team led by Hazen and Sawyer, P.C. of Tampa, Florida with key support in model development from Planning and Management Consultants, Ltd. (PMCL) of Carbondale, Illinois, over a period of time from 2001 to 2004.

Development of Water Demand Models for Tampa Bay Water. Tampa Bay Water's demand is currently composed of demands from member governments, including City of Tampa, Hillsborough County, Pinellas County, St. Petersburg, New Port Richey, and Pasco County. Member governments are responsible for serving geographically-distinct Water Demand Planning Areas, or WDPAs (see Figure ES.1). Member government demands are satisfied through targeted bulk deliveries of water from Tampa Bay Water. Members then use these wholesale deliveries to satisfy *retail* demand for individually-billed and tracked customers. Customers are classified into three consumer categories: single-family (SF), multi-family (MF), and non-residential (NR) sectors. In addition, some members re-sell wholesale water to other local utilities. Furthermore, each member experiences losses in water, or *unbilled* demand, as reflected by differences between deliveries from Tampa Bay Water and combined WDPA retail and wholesale distribution. In this study, these demand components were modeled at individual member government levels to reflect the true nature of demand as a set of spatially-distinct member requirements.

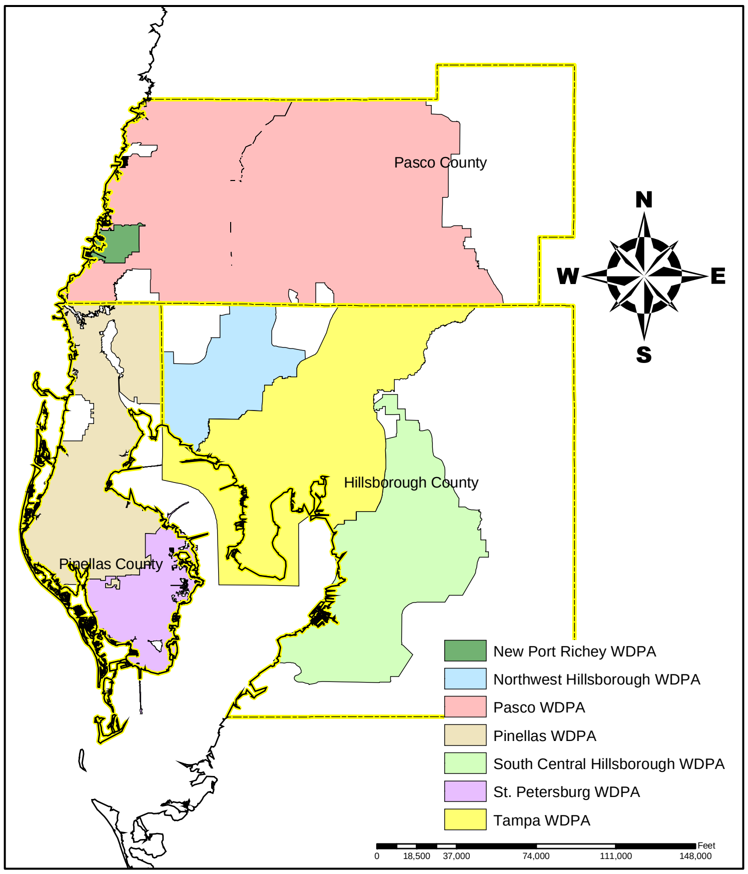


Figure ES.1 – Water Demand Planning Areas (WDPAs) for Tampa Bay Water

Tampa Bay Water's member retail demand was modeled using three sector-specific econometric models. These *sectoral models* related demand for a given time and geographical area to meteorological and socioeconomic variables, or *explanatory variables*, in that time and area. Each sectoral model calculated demand on a "per-driver-unit" basis. For example, the single-family (SF) sectoral model calculated retail demand per household for a specific geography and time. To obtain total single-family use for that geography and time, per-household use was multiplied by number of single family households in that area at that time. Number of single-family households served as a *driver variable* for total single-family demand. Likewise, the multi-family (MF) sectoral model calculated retail demand per multi-family dwelling unit, and the non-residential (NR) sectoral model calculated retail demand per employee, with number of multi-family dwelling units and number of employees serving as driver variables, respectively.

A comprehensive water use database was developed to support construction of sector-specific models. This database included several thousand observations of explanatory variable data, including socioeconomic, land use, weather, and price variables, and sector-classified water use billing information for Water Years 1999-2002. Data were specified for historical months and nearly 1,500 small geographies, termed Traffic Analysis Zones (TAZs)[1], which were mapped to WDPAs (see Figure ES.2). An Ordinary Least Squares regression analysis was used to develop regression equations between TAZ-specific billed water use (i.e. demand) and explanatory variables within the database. The wealth of available modeling data allowed the estimation of models with precise coefficient estimates relating socioeconomic and other factors to variation in water use.

Regional demand forecasts were developed in this project by determining projected values of explanatory and driver variables for WDPAs at future times, calculating WDPA demand corresponding to these assumptions using the aforementioned sectoral models, and summing results across WDPAs. Two types of demand forecasts, deterministic and probabilistic forecasts, were produced in this manner. In deterministic forecasts, or *point forecasts*, each explanatory and driver variable is assigned a single projected value for each WDPA at each identified future point in time. This value is then applied to the demand model to produce a single-valued forecast of corresponding demand for each WDPA and time. In *probabilistic forecasts*, probability distributions are assumed for explanatory and driver variables for each WDPA at each future point in time. These distributions characterize the expected range of, and uncertainty in, future values for explanatory and driver variables. Probabilistic demand forecasts are calculated by propagating uncertainty in explanatory and driver variables through the demand model, producing a probability distribution of demand for each WDPA at each future time. Demand distributions produced in probabilistic forecasts reflect ranges of potential future water use given uncertainty in factors generating that demand.

1 *TAZs are geographical units defined by Florida Department of Transportation for automobile traffic analysis studies, but are also used by Metropolitan Planning Organizations as a basis for characterizing and projecting other socioeconomic information, such as population or income.*

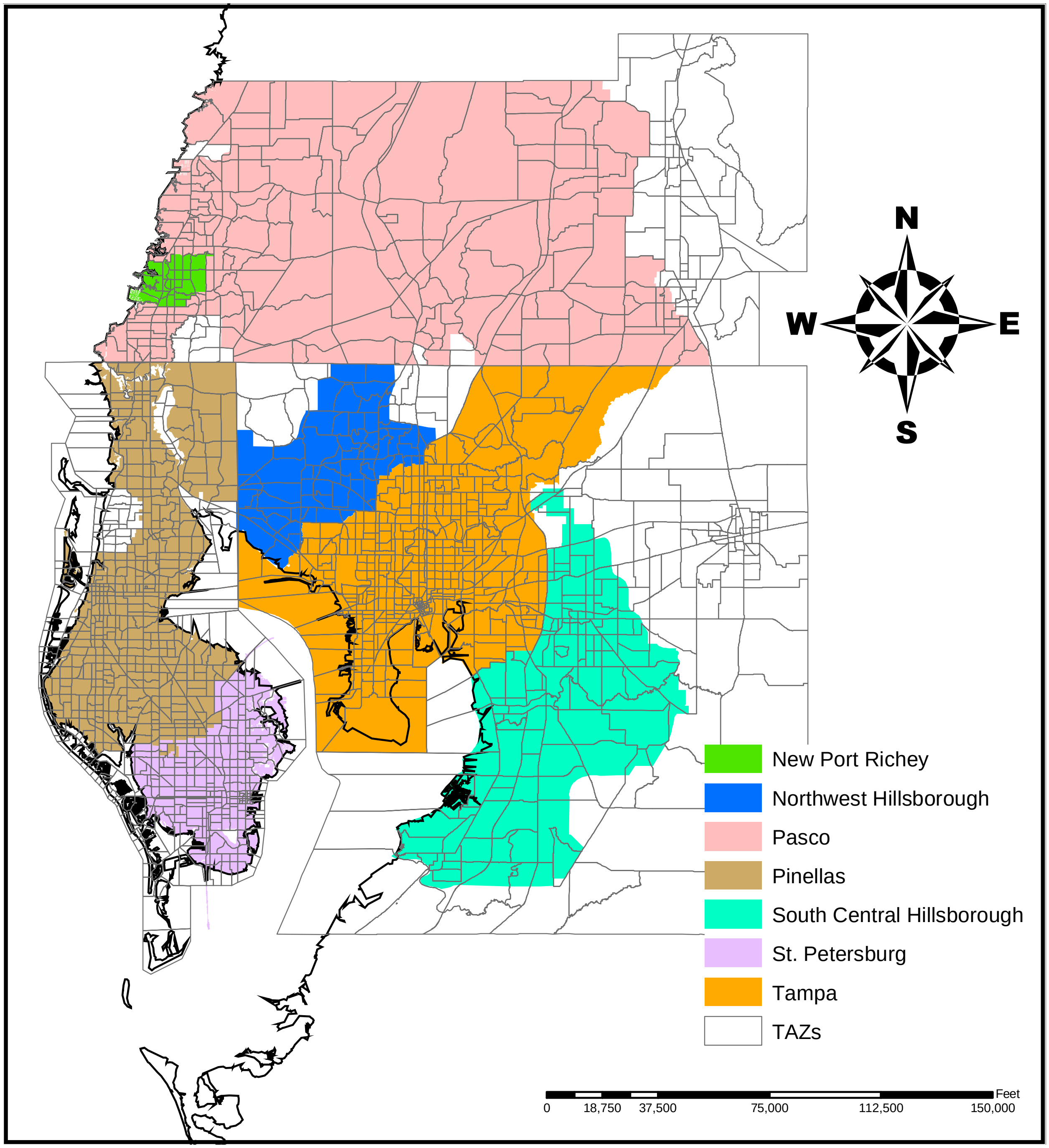


Figure ES.2 – Traffic Analysis Zones (TAZs) used in developing demand models.

Generation of point and probabilistic forecasts from projected explanatory and driver data involved a stepwise accounting process. First, projections for explanatory/driver variables and wholesale/unbilled quantities were specified over the forecast horizon for all WDPAs. These projections were used in sectoral models to determine per-unit demand forecasts for all sectors and WDPAs over the forecast period. Per-driver-unit forecasts were then multiplied by projected driver variables, producing sector, time, and WDPA-specific retail demand forecasts. These retail demand forecasts were aggregated across sectors and WDPAs and adjusted to include wholesale and unbilled quantities, producing demand forecasts for the entire Tampa Bay Water service area.

Development of Long-Term Point Forecast for Tampa Bay Water Demand. To generate a point forecast of water demand, single-valued monthly explanatory and driver variable projections were obtained for each WDPA across the 2003-2025 forecast period. Demand forecasting models were applied to these projections.

First, demand models were applied to observed WDPA-level econometric and driver values for Water Year 2002 (October 1, 2001 to September 30, 2002), the most recent year with observed driver, explanatory, and demand data. The resulting demand calculations were compared to observed Water Year 2002 demand by WDPA and sector. This comparison produced a high degree of predictive accuracy, in terms of absolute water demand, for forecasting at the WDPA and Tampa Bay Water regional service area levels (Table ES.1). Accuracy was lower when comparing observed and calculated demand in specific sectors within specific WDPAs, primarily due to increased variability in observed data at these smaller scales. This increased variability was canceled out when predicting demand at aggregated regional and sectoral levels, a critical strength of the disaggregated approach to forecasting.

Table ES.1
Comparison of WY 2002 Observed to Point Forecasted Demands*

Predicted Demand Portion	WDPA Pinellas	St. Pete	NPR[1]	Pasco	Tampa	NW Hills[2]	SC Hills[3]	Total TBW Service Area
Single-Family Difference	24.22	13.37	1.34	12.42	27.84	10.21	15.62	105.02
	1.18	0.50	-0.19	-0.16	2.07	-1.16	-0.12	2.12
Multi-Family Difference	10.98	6.37	0.64	1.06	10.85	1.72	5.44	37.06
	0.00	-1.04	0.07	-0.34	-1.87	-0.80	2.19	-1.80
Non-Residential Difference	8.96	6.66	0.57	2.67	22.11	2.46	3.86	47.29
	0.02	-0.17	-0.12	0.39	1.65	0.48	-0.03	2.21
Wholesale Difference	20.60	2.14	0.31	0.52	0.23	0.00	0.00	23.80
	-1.77	-0.08	-0.05	-0.18	-0.07	0.00	0.00	-2.14
Unbilled Difference	4.19	2.77	0.37	2.00	12.37	0.97	2.50	25.17
	-0.05	-0.17	-0.01	-0.15	-1.91	0.07	0.19	-2.02
Total Predicted Demand	68.95	31.32	3.22	18.67	73.40	15.36	27.41	238.33
	-0.63	-0.97	-0.30	-0.44	-0.12	-1.42	2.23	-1.64

* - Forecasted demands have white backgrounds, differences between forecasted and observed demands (forecasted minus observed) have shaded backgrounds. All values in million gallons per day (MGD).

1 - New Port Richey

2 - Northwest Hillsborough

3 - South Central Hillsborough

After assessing accuracy of forecasted water usage, point forecasts were calibrated. In each WDPA, WY2002 model-predicted demand was adjusted by the difference between observed and predicted demand, making observed and predicted values equal. These same adjustment values were then applied to each forecasted demand value after WY2002.

Total demand in the Tampa Bay Water service area was projected to grow at an annual average rate of 0.88 percent over a 23-year forecast horizon, from an initial value of 238 MGD in 2002 to a final value of about 300 MGD in 2025. The greatest annual average changes in demand were projected to occur in multi-family and non-residential sectors, though demand in the single-family sector was projected to remain the largest component of retail use (Figure ES.3). Projected rates of growth in water demand components differed considerably among WDPAs, due to variations in projected housing, employment, and other socioeconomic factors between WDPAs (Table ES.2).

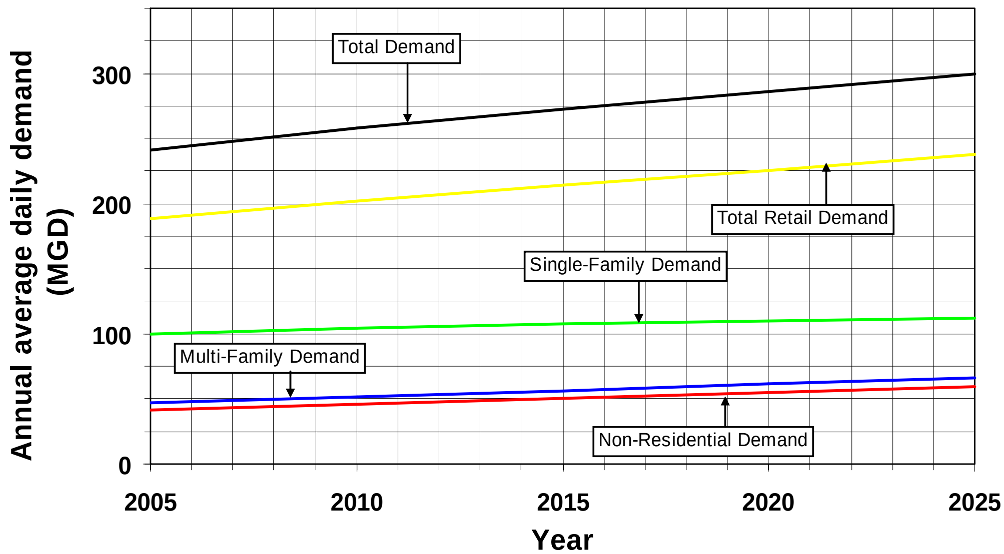


Figure ES.3 – Point Forecast of Tampa Bay Water Regional Demand, Including Total Sectoral Demand

Table ES.2
Summary of Point Forecast Demand in the Tampa Bay Water Service Area

WDPA	Forecasted Demand, MGD 2005	2010	2015	2020	2025	Percent Change 2005-2025	Average Annual Change
Pinellas	68.88	71.29	72.66	73.48	74.20	7.73%	0.30%
St. Petersburg	31.94	33.11	33.86	34.39	34.86	9.13%	0.35%
New Port Richey	3.38	3.43	3.49	3.54	3.60	6.40%	0.25%
Pasco	19.40	21.55	23.40	25.19	26.86	38.45%	1.31%
Tampa	74.22	79.90	85.24	89.68	94.26	27.00%	0.96%
NW Hillsborough	17.01	18.96	20.67	22.60	24.62	44.71%	1.49%
SC Hillsborough	26.40	30.25	33.66	37.62	41.63	57.70%	1.84%
Total TBW	241.23	258.50	272.97	286.51	300.02	24.37%	0.88%

Development of Long-Term Probabilistic Forecast for Tampa Bay Water Demand. Following development of the point forecast, the Demand Forecasting Model was embedded in a Monte Carlo simulation to generate a probabilistic forecast. The probabilistic forecast represented effects of uncertainty in future values of explanatory and driver variables on demand forecasts, depicting a range of likely future demand values over the forecast period.

First, uncertainty was quantified for explanatory and driver variables by specifying probability distributions for future variable values. The project team conducted a workshop with Tampa Bay Water staff (October 4, 2002) on development of probabilistic water demand models, during which a consensus was reached regarding the appropriate probability distribution for each model variable. Specified distributions were based on data analysis, published recommendations, and experience.

Following specification of distributions, a Monte Carlo simulation procedure[2] was used to develop a conditional probabilistic water demand forecast based on the point model. Each iteration of the Monte Carlo simulation randomly selected a value for each explanatory and driver variable based on the distribution specified for that variable, then used the complete set of values to produce a water demand forecast. The simulation procedure performed numerous independent iterations (between 5,000 and 10,000), each generating an independent forecast. All forecasts were then pooled and forecasted demand values were ranked at each forecast month, yielding distributions of estimated water demand for each month over the forecast horizon.

Figure ES.4 illustrates average annual probabilistic forecasted demand results. Interpreting the forecast median as an expected value, expected average annual water demand in the Tampa Bay Water service area is forecasted to reach approximately 298 MGD in 2025. Compared to the expected forecast value in 2003, this represents about a 26 percent increase over 23 years, or roughly a one-percent increase in annual average daily demand per year. This result is in good agreement with the 2025 point forecasted demand of 300 MGD. This agreement is not and should not be exact, as the point forecast is based on deterministic assumptions of explanatory and driver variables while the expected value of the probabilistic forecast arises from random samplings of these variables.

2 *@Risk, produced by The Palisade Corporation (www.palisade.com), was used to perform Monte Carlo Analysis. This software is a Microsoft Excel plug-in that operates on a spreadsheet version of the model of interest and on in-spreadsheet specifications of input variable probability density functions.*

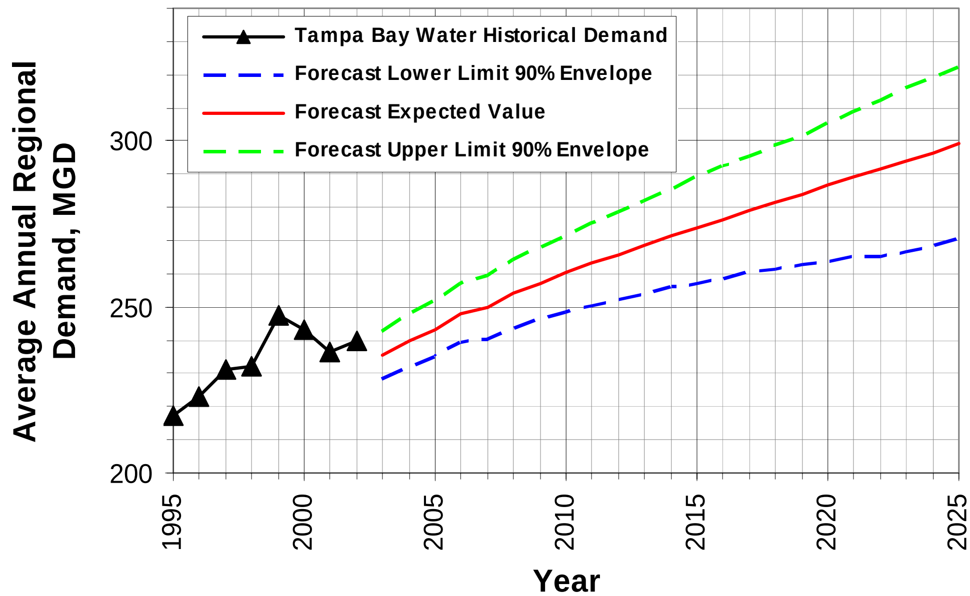


Figure ES.4 – Forecasted annual average daily water demand in the Tampa Bay Water service area, with 90% confidence interval and historical demand for comparison.

The 5th and 95th percentiles of simulated demand were used to define the 90 percent confidence interval of forecast demand and are shown as lines above and below the median forecast in Figure ES.4. For example, 90 percent of total demand values are predicted to fall between 232 MGD and 240 MGD in 2003, whereas in 2025 the 90 percent forecast envelope is between 278 MGD and 315 MGD. As illustrated in Figure ES.4, increases in uncertainty over the forecast period are evident in growing standard deviations of forecast distributions (e.g., a standard deviation of 2.56 MGD in 2003 compared with a standard deviation of 11.12 MGD in 2025). Forecasted demand intervals are larger in future years because uncertainties in projections grow over time. Intervals shift upward over time because expected values for driver variables generally increase over time, similarly to the point forecast.

A similar pattern of upward-sloping expected demand with widening uncertainty is visible in monthly forecasted demand (Figure ES.5). The monthly forecast interval also shows seasonality in water use, which arises predominantly from seasonal variations in explanatory variables influencing single-family residential demands.

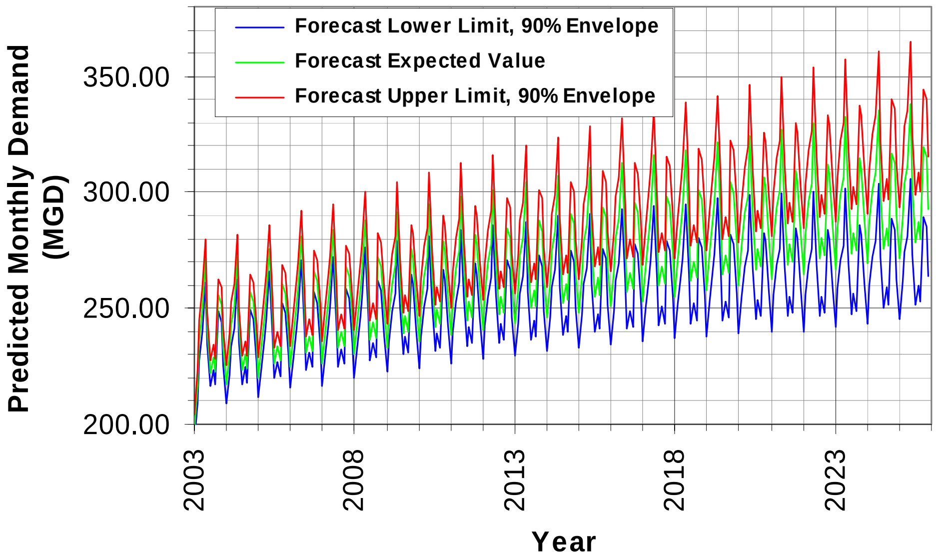


Figure ES.5 – Forecasted monthly average of daily water demand in the Tampa Bay Water service area, with 90% confidence interval

Applicability of Demand Models. This project represents an important step in accounting for inherent uncertainties in the water demand forecasting efforts of Tampa Bay Water. With these results, Tampa Bay Water has obtained a more robust understanding of potential uncertainty associated with future demand predictions. The demand models and forecasting approaches developed in this project are currently being applied in ongoing projects.

- The point forecasting methodology is being implemented in a custom computer application, the Tampa Bay Water Long-Term Demand Forecasting System (LTDFS). This application will allow users to browse historical demand geographically and to generate new point demand forecasts based on modified projections of model variables.

- The probabilistic demand model is being leveraged in the Tampa Bay Water Future Need Analysis (FNA) project. FNA will involve coupling the probabilistic demand model with a probabilistic water supply model to forecast future need for additional water supply facilities. This tool will serve to identify timing and risk of surface water supply shortages. FNA will assist decision-makers in efficient planning of future water supply expenditures and risk mitigation efforts.

Recommendations. It is recommended that Tampa Bay Water work to further refine and maintain the demand model, databases, and uncertainty assumptions, thereby continually improving its future demand projections.

- **Maintain explanatory/driver variable projections and update forecasts.** As new projection data and sources of projections become available, they should be evaluated for use as inputs for demand forecasts. Forecasts should be reevaluated any time new projections for per capita income, real marginal price of water for member governments, fraction of accounts using reclaimed water, number of single-family households, number of multi-family dwelling units, and number of employees become available. The LTDFS is currently being developed to streamline forecast re-evaluation. In addition, planning organizations may begin projecting new types of data and cease projecting existing types and the values of projections and methods by which projections are generated can change. Sources and methods for developing projections of driver and explanatory variables must be periodically revisited and updated as necessary.

- **Recommend new projection methods to planning organizations.** Projected data and distributions were not directly available from planning organizations for several variables or were not available at desired scales of geography and time. Projections of single-family households, multi-family dwelling units, single- and multi-family persons per household and housing density were estimated using most recent observations and/or available projections of associated variables. Needs for direct projections of these variables should be communicated to planning organizations, and methods for directly projecting desired values should be suggested if possible.

- **Maintain modeling database.** It will be necessary to periodically recalibrate forecasts and refit the demand model. The modeling database must be maintained to support these tasks. New time series observations covering a longer period of record for existing explanatory and driver variables should be entered into the modeling database as they become available, and the model should be periodically refit based on this data. It is particularly important that water consumption, real marginal price of water for member governments, fraction of accounts using reclaimed water, number of single-family households, number of multi-family dwelling units, and number of employees be kept updated. Tampa Bay Water is already developing the mechanisms to update this data as part of the Enterprise database and LTDFS application database. In addition to adding new data, the association between billing accounts and TAZs must be updated whenever TAZ boundaries are redefined by the Florida Department of Transformation. Data aggregated by new TAZ definitions must be used for subsequent model refits.

- **Periodic recalibration of demand forecasts and refitting of the demand model.** Based on an actively maintained modeling database, forecasts should be recalibrated for each new year of water use data, and the demand forecasting model should be refitted (i.e., regression should be repeated using the updated modeling database) at least once every five years. Refitting should be performed any time it is suspected that a change has occurred in mechanisms by which explanatory variables influence demand, such as a change in irrigation restrictions, and any time a significant number of TAZs (50 or more) have been redefined by Florida DOT.

- **Develop and implement collection methods for data types not currently available.** In some cases, key data for model development were not available and required estimation. New data collection methods must be developed and implemented to allow actual data to be used in place of the estimates. Data should be collected in billing databases for number of multi-family dwelling units served by each multi-family account. Within billing databases, non-residential accounts should be classified into more precisely-defined categories that group these accounts by well-defined water use characteristics. As these new observations become available, the model must be refitted to accommodate that data. A billing methodology study is currently underway to address these issues and to synchronize data collection techniques among member governments.

- **Disaggregation and detailed modeling of wholesale demand.** In the current demand model, wholesale demand is treated as a "black box" component of demand, and is not calculated as a function of explanatory and driver variables within wholesale service areas. This treatment of wholesale demand was necessary, as customer-level billing data were not available for wholesale utilities. Billing data should be obtained from wholesale utilities (through member governments) and included in the modeling database. Corresponding exploratory and driver variable data should be obtained for TAZs in wholesale service areas. The model should be refitted to reflect the new data from wholesale regions. The resulting model would greatly facilitate integration of wholesale demand into member government retail demand as these non-member utilities are acquired.

- **Evaluate potable demand offset produced by reclaimed water use.** One or more studies should be performed to evaluate the decrease in potable demand per amount of reclaimed water used, or potable demand offset. Studies should compare demand in similar neighborhoods with and without reclaimed water connections. These studies would include metering and collection of reclaimed water use data in the neighborhoods of interest and determination of potable demand offset for those neighborhoods. Offset rates could potentially be extrapolated to other similar neighborhoods. Such studies would be valuable as a reclaimed water system planning aid, identifying areas where reclaimed water development would produce greatest demand reduction. If possible, results of these studies should be used to model reclaimed water use effects in the demand model as a function of TAZ-level socioeconomic characteristics.

- **Maintain uncertainty assumptions for projected explanatory and driver variables and update probabilistic forecasts.** Explanatory and driver variable projections and uncertainties may change in the future, due to release of new projection data, changes in projection methods, and changes of sources for projection data (i.e., the organizations providing the projections). When such changes arise, probability density functions should be redefined for exploratory and driver variables and new probabilistic forecast simulations should be performed. Furthermore, new probabilistic forecasts should be produced whenever the point model is refitted or recalibrated.

- **Integration of cost in evaluating forecasted results of demand and supply projects.** Reducing probability of future potable water need and increasing system reliability could involve additions to supply at potentially significant financial cost. Probabilistic supply and demand forecasting should be performed for each supply alternative, assuming the alternative is implemented in the supply system and assessing resulting change in forecast need. These results can then be coupled with cost assessments of alternatives and allow relationships between capital costs and risk mitigation to be evaluated. With this relationship defined, decision-makers could identify projects offering the greatest improvement in reliability per unit cost.

Together with the methodologies and results developed thus far, pursuit of these recommendations would improve further upon Tampa Bay Water's comprehensive long-term water supply reliability planning.

Introduction

Tampa Bay Water, like many other major water resources agencies, faces a difficult task in forecasting potable water needs over a relatively long time horizon for a large and growing service population. In many cases, these water needs dictate and are influenced by size and timing of future supply and demand expenditures, such as additions to supply and treatment capacity, implementation of water conservation programs, development of reclaimed water programs, and changes to the structure and level of water prices.

In the past, Tampa Bay Water depended mostly upon groundwater sources to meet regional demand. In the early nineties, it was determined that the continued pumping from these wellfields was contributing to depletion of wetlands and lowering of water levels in various lakes, especially in the vicinity of well locations. The Northern Tampa Bay New Water Supply and Groundwater Agreement (WCRWSA, 1998) was incorporated between Southwest Florida Water Management District (SWFWMD), Tampa Bay Water, and its member governments, wherein all parties agreed to cooperate with each other to develop new water supply and reduce pumpage from existing regional wellfields. The Consolidated Water Use Permit for 11 long-producing regional wellfields required allowable withdrawals of 158 MGD be reduced to 121 MGD or less by the end of 2003 and 90 MGD or less by the end of 2008. The Partnership Agreement required Tampa Bay Water and its member governments to continue to plan, coordinate, develop, construct and implement new water supplies, conservation and reclaimed water projects.

The Master Water Plan (MWP) developed by Tampa Bay Water provided a framework for developing alternative sources to groundwater and the related transmission, treatment, and storage components. As of this writing, Phase I of the MWP is almost complete with various alternative water sources developed including a 66 MGD surface water treatment plant, a 25 MGD desalination plant, and a 15 billion gallon storage reservoir. These sources are expected to allow the Agency to meet the increasing demand over the next few years.

In order to develop a better understanding of increases in demand and its implications on supply development options (size and timing), the Agency commissioned the development of a long-term demand forecasting system (LTDFS). The LTDFS is a major initiative to quantify how socioeconomic, meteorological, and policy conditions in its service area influence potable water demand.

The initial step in achieving this Board directive was creation of a Long-term Demand Forecasting System Technical Advisory Committee (LTDFS TAC) comprised of representatives of Tampa Bay Water's member governments and the SWFWMD. The following personnel represented their respective member governments, agencies, or SWFWMD.

- Mr. Tim Wiley, Pinellas County
- Mr. Jim Jeffers, Hillsborough County
- Ms. Patti Anderson, City of St. Petersburg
- Mr. Doug Bramlett, Pasco County
- Mr. Karl Craig, City of Tampa
- Ms. Mary Healey, City of New Port Richey
- Mr. Jay Yingling, SWFWMD
- Mr. Dave Bracciano, Tampa Bay Water
- Dr. Alison Adams, Tampa Bay Water

The purpose of this TAC was to periodically review the project progress and provide pertinent comments. The TAC members played a major role in facilitating billing and rate data collection by directing project team members to appropriate personnel within each utility.

A principal product of this initiative was a regional demand model that calculated demand as a function of meteorological, socioeconomic, and policy conditions. This model was combined with projections of socioeconomic growth and meteorological/policy conditions to generate forecasts of water demand, which will assist decision-makers in understanding how growth affects future water needs. In addition, the model was developed on a geographic basis, such that water demand was forecasted for distinct locations in the Tampa Bay Water service area. Estimations were determined for potential variations in future socioeconomic, meteorological, and policy conditions. When applied to the demand model, these variations portrayed future demand growth as a resulting range of possible demand values. These model demand forecasts will help decision-makers assess the risk of supply shortfalls relative to demand and plan for demand and supply expenditures, while avoiding unacceptable risk.

This document chronicles the development of Tampa Bay Water's Long-Term Demand Forecasting Model and use of the model for long-term point and probabilistic demand forecasts. The project was completed for Tampa Bay Water by a team led by Hazen and Sawyer, P.C. of Tampa, Florida with key support in model development from Planning and Management Consultants, Ltd. (PMCL) of Carbondale, Illinois, over a period of time from 2001 to 2004.

The document contains three main chapters, each describing a step in development of the Long-Term Demand Forecasting Model and its probabilistic capabilities.

- Chapter 1: Development of Water Demand Models and Long-Term Forecasts for Tampa Bay Water – describes development of a demand forecasting architecture, point models for demand prediction, and implementation of these models in generating a point forecast of Demand for 2003-2025

- Chapter 2: Development of Uncertainty Assumptions for the Tampa Bay Water Long-Term Water Demand Forecast – describes assumptions of quantitative uncertainty characteristics for demand-influencing input variables in demand models and outlines a Monte Carlo strategy for incorporating these assumptions into a probabilistic demand forecast

- Chapter 3: Demand Forecast Report – describes development of a Monte Carlo simulation and presents probabilistic demand forecasting results.

Supplementary information in Appendices includes:

- Details of model development procedures

- Mathematical descriptions of equations composing final demand models,

- Details concerning the implementation of the monte carlo simulation in a spreadsheet (using @Risk, a Microsoft Excel plug-in offered by the Palisade Corporation), and

- Detailed results of probabilistic forecasts by Water Demand Planning Area (WDPA).

The results of this project are being automated in a custom computer application, the Tampa Bay Water Demand Forecasting System. This application will allow users to browse historical demand geographically and to generate new point demand forecasts based on modified projections of model variables.

The probabilistic demand model will be leveraged in several subsequent projects. Tampa Bay Water's Future Need Analysis (FNA) project will involve coupling the probabilistic demand model with a probabilistic Water Supply model to forecast future need for additional water supply facilities. Any shortfalls in supply from variable surface water sources must be met by consolidated wellfields, which are under regulatory withdrawal constraint. FNA will serve to identify timing and risk of surface water supply exhaustion and groundwater permit exceedence, assisting decision-makers in efficient planning of future water supply expenditures. As half of the analysis required for FNA, the Long-Term Demand Forecasting Model is immediate applicable in critical Tampa Bay Water supply planning initiatives.

1.0 Development of Water Demand Models and Long-Term Point Forecasts

This chapter describes the development and use of water demand prediction models to prepare long-term point estimates of water demand in the Tampa Bay Water service area. Separate water use models were developed for single-family residential, multi-family residential, and the non-residential (combined commercial, industrial, and service) utility customer categories, or *sectors*. These three sectoral models were built using geocoded water billing records for 1999-2002, along with observed meteorological and socioeconomic data for geographies corresponding to billing records. Together with projected future values for model variables, the sectoral models were used to develop point forecasts of average monthly and yearly future water use. Forecasts were prepared for seven Water Demand Planning Areas (WDPAs) that comprise the Tampa Bay Water service area and for projected expanded service areas out to the year 2025.

The sections in this chapter describe all aspects of developing the Demand Forecasting Model and point demand forecast, including;

1. Examining the structure of Tampa Bay Water's potable demand,
2. Developing a demand model architecture to correspond to demand structure,
3. Collecting historical data for model development,
4. Developing demand models and accounting procedures to predict demand at various spatial, temporal, and sectoral aggregation levels,
5. Assessing model accuracy with respect to historical data, and
6. Developing point forecasts using the fitted model.
7. In subsequent chapters, it will be illustrated how this demand model was expanded to provide probabilistic forecasts.

1.1 Overview of Demand Structure, Accounting, and Terminology

To forecast Tampa Bay Water's potable demand, it was necessary to delineate water demand structure by water use sector and *member government*. Forecast models were assembled to simulate member-specific, sector-classified components of demand. Accounting procedures were developed to aggregate demand components into forecast scales and metrics of interest to planners. This section describes Tampa Bay Water's demand structure as applied to development of the Demand Forecasting Model and accounting procedures. Common symbology is also introduced for describing demand quantities and aggregation procedures.

1.1.1 Demand Structure and Terminology

Tampa Bay Water's demand is currently composed of demands from geographically distinct member governments, or members. These member demands are satisfied through targeted bulk deliveries of water from Tampa Bay Water at sixteen (16) points of delivery. Members then use these bulk deliveries to satisfy *retail* demand for individually-billed and tracked customers. In addition, some members resell water on a *wholesale* basis to other local utilities. Furthermore, each member experiences losses in water, or *unbilled* demand, reflected by differences between deliveries from Tampa Bay Water and combined retail and wholesale distribution. Demand components were modeled at the level of specific members to reflect the true nature of demand as a set of piecewise spatially-distinct member requirements.

Tampa Bay Water currently divides its service area into seven geographic sub-areas associated with distinct members:

- New Port Richey
- Pasco County
- Pinellas County
- St. Petersburg
- Northwest Hillsborough[1]
- South Central Hillsborough[1]
- City of Tampa

These sub-areas are termed Water Demand Planning Areas, or *WDPAs* (see Figure 1.1). In this study, each individual retail and wholesale customer was classified according to the WDPA which served it, allowing aggregate consumption histories to be used for geographically-specific modeling and forecasting.

1 *Northwest Hillsborough and South Central Hillsborough are parts of the Hillsborough County service area but are geographically split by the City of Tampa service area.*

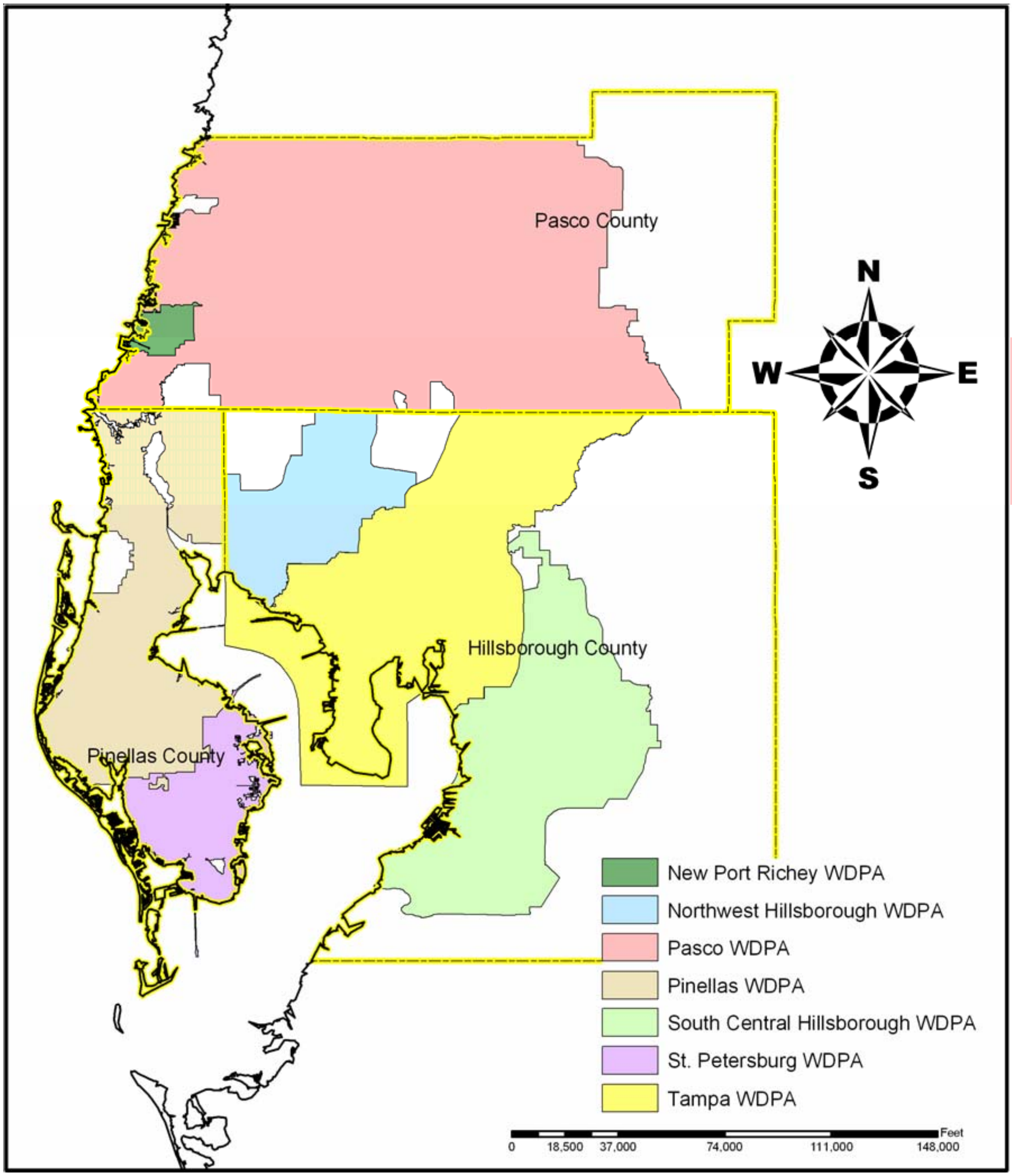


Figure 1.1 – Water Demand Planning Areas (WDPAs) for Tampa Bay Water

Each member currently has various classes, or *sectors*, of retail customers. A total of three sectors were adopted for this study, including single-family, multi-family, and non-residential customer classifications.

- The single-family (abbreviated "SF") sector was composed of dwellings typically associated with single families, such as houses on individual accounts.

- The multi-family ("MF") sector was composed of dwellings typically associated with multiple families, such as units within apartment buildings all on one account.

- The non-residential ("NR") sector was composed of commercial, industrial, and institutional water customers, such as office buildings, shopping centers, industrial plants, and public utility facilities.

This sectoral classification was chosen because retail customers in these three classifications are traditionally considered to have distinct water consumption patterns. For example, both SF and MF facilities tend to have higher water use for sanitary purposes per capita than NR facilities, and SF facilities tend to have higher irrigation uses per capita than MF facilities. There tends to be a much broader range of water use among NR facilities than SF and MF facilities, as NR water consuming activities are diverse. It would be desirable to classify non-residential water consumption in greater detail according to more precisely-defined water-consuming activities. However, there were not sufficient and consistent data to define a more detailed classification system.

Each member's (and retail/wholesale customer's) demand was classified according to *time*, specifically distinct year and calendar month. In this study, 276 forecast months were adopted, spanning 23 forecast years (Jan 2003-Dec 2025). Description of demand as monthly averages allowed both annual trends due to regional growth and within-year variation due to seasonal changes to be described historically and in forecasts.

1.1.2 Demand Accounting and Aggregation

Tampa Bay Water's retail demand was modeled using three sector-specific econometric models. Each model generated demand forecasts based on WDPA-specific meteorological and socioeconomic projections. Sector-specific models therefore satisfied the need for modeling retail demand on a member-by-member basis. From these results, sector-specific results can be aggregated as needed by the analyst.

Each sector-specific model calculated demand per water consuming entity, or *driver unit*. A different driver unit was defined for each sector. The SF sectoral model calculated retail demand per household, with households serving as a driver unit. Likewise, the MF sectoral model calculated retail demand per dwelling unit, with number of dwelling units serving as a driver unit. The NR sectoral model calculated retail demand per employee, with number of employees serving as a driver unit.

Given the "per-driver-unit" approach to modeling, forecasted demand was aggregated as follows:

1. Per-driver-unit demand forecasts were determined for specific geography, sector, and time using sectoral models.

2. Per-driver unit forecasts were then multiplied by driver units within the corresponding geographic, sectoral, and temporal specification, producing sector, time, and geography-specific total retail demand.

3. Retail demand forecasts were then aggregated upwards across sectors and/or WDPAs.

Figure 1.2 illustrates this aggregation conceptually. In addition to sectoral and geographic aggregations, demand at any aggregation level may be rolled to annual totals by summing across months within a calendar year (Jan-Dec) or water year (Oct- Sept).

In Figure 1.2 and throughout this report, the following general notation is used:

- Q = Total retail water use at some level of sectoral and spatial aggregation (e.g., single-family water use in St Petersburg WDPA in August 2020)

- N = Number of driver units (e.g., number of occupied single-family housing units)

- q = Per-unit rate of use as determined by sectoral demand models (e.g., average water use per single-family unit per day)

The above symbols may be subscripted and superscripted to specify various levels of aggregation. Subscripts will be clarified as they arise.

As will be discussed later, there were no retail-customer-specific data available for utilities accepting wholesale deliveries from members. Also, no attempt was made to disaggregate system losses beyond the member level. Therefore, wholesale and unbilled demands were specifiable only by member and time, not by sector.

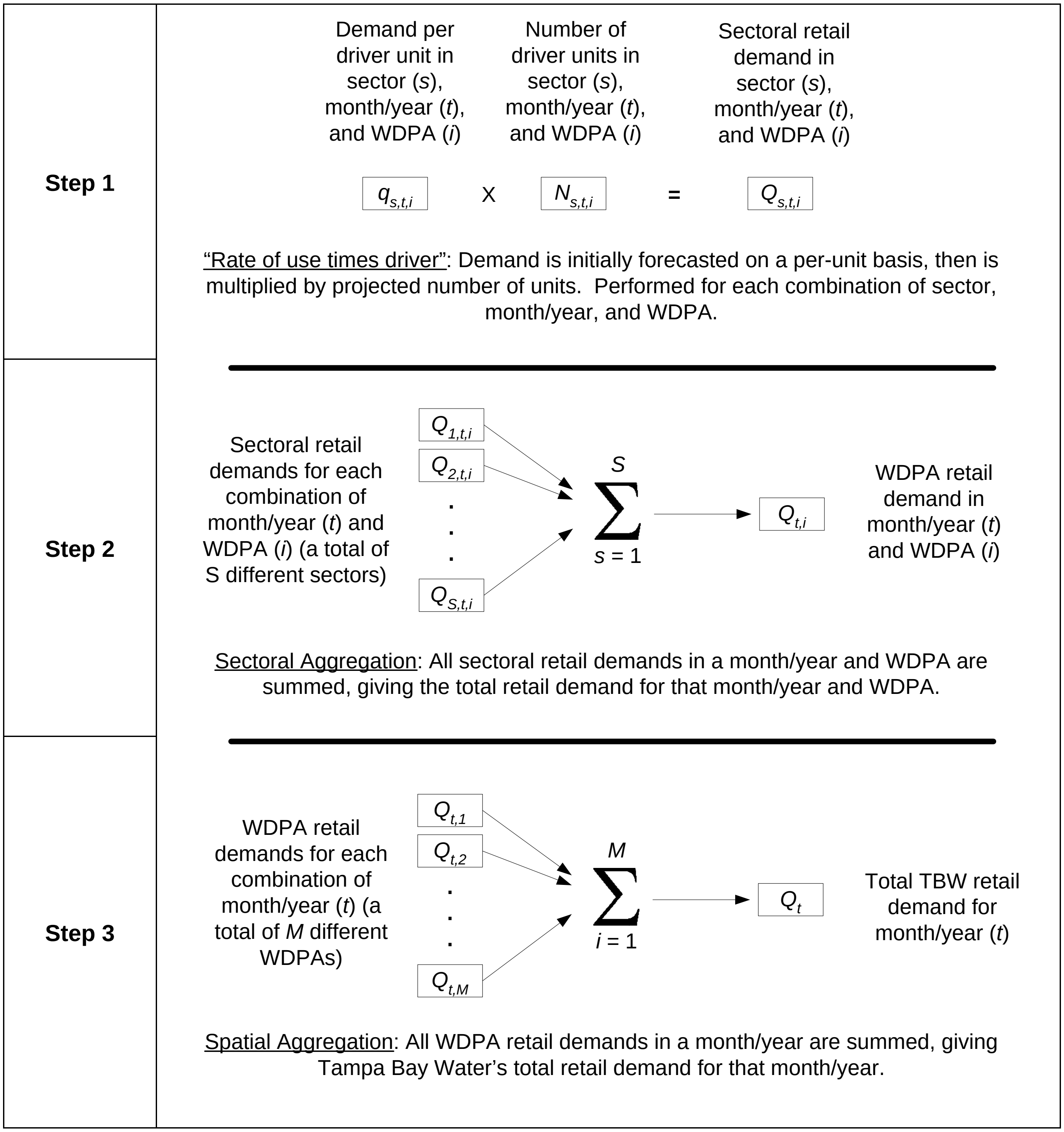


Figure 1.2 – Demand predictions originated at the per-driver-unit level (specific to sector, month/year, and WDPA), and were then aggregated by sector and WDPA.

1.2 Water Use Modeling Database

To provide a data basis for developing demand models within the previously described architecture, a historical water use modeling database was constructed. This database resulted from a comprehensive data collection effort undertaken to develop a geographical information system (GIS) for Tampa Bay Water. The final modeling database consisted of observed and estimated values of several variables by geographic region and time.

1.2.1 Geographical, Sectoral, and Temporal Basis for Data

Monthly account-level water use (monthly gallons used divided by number of days per month) and billing data were used in conjunction with tax assessor databases to identify and geocode demand by land parcels. Geocoding identified water customers with parcels and assigned each customer to one of the three retail sectors (single-family residential, multi-family residential, or non-residential).

Parcels were identified as being located within broader spatial units, known as traffic analysis zones or TAZs (see Figure 1.3). TAZs are geographical units defined by Florida Department of Transportation for automobile traffic analysis studies, but are also used by MPOs as a basis for characterizing and projecting other socioeconomic information, such as population or income. Approximately 1,400 TAZs were demarcated within the Tampa Bay Water service area. As described below, TAZs were ultimately assigned to WDPAs, completing the geocoding of consumption by WDPA.

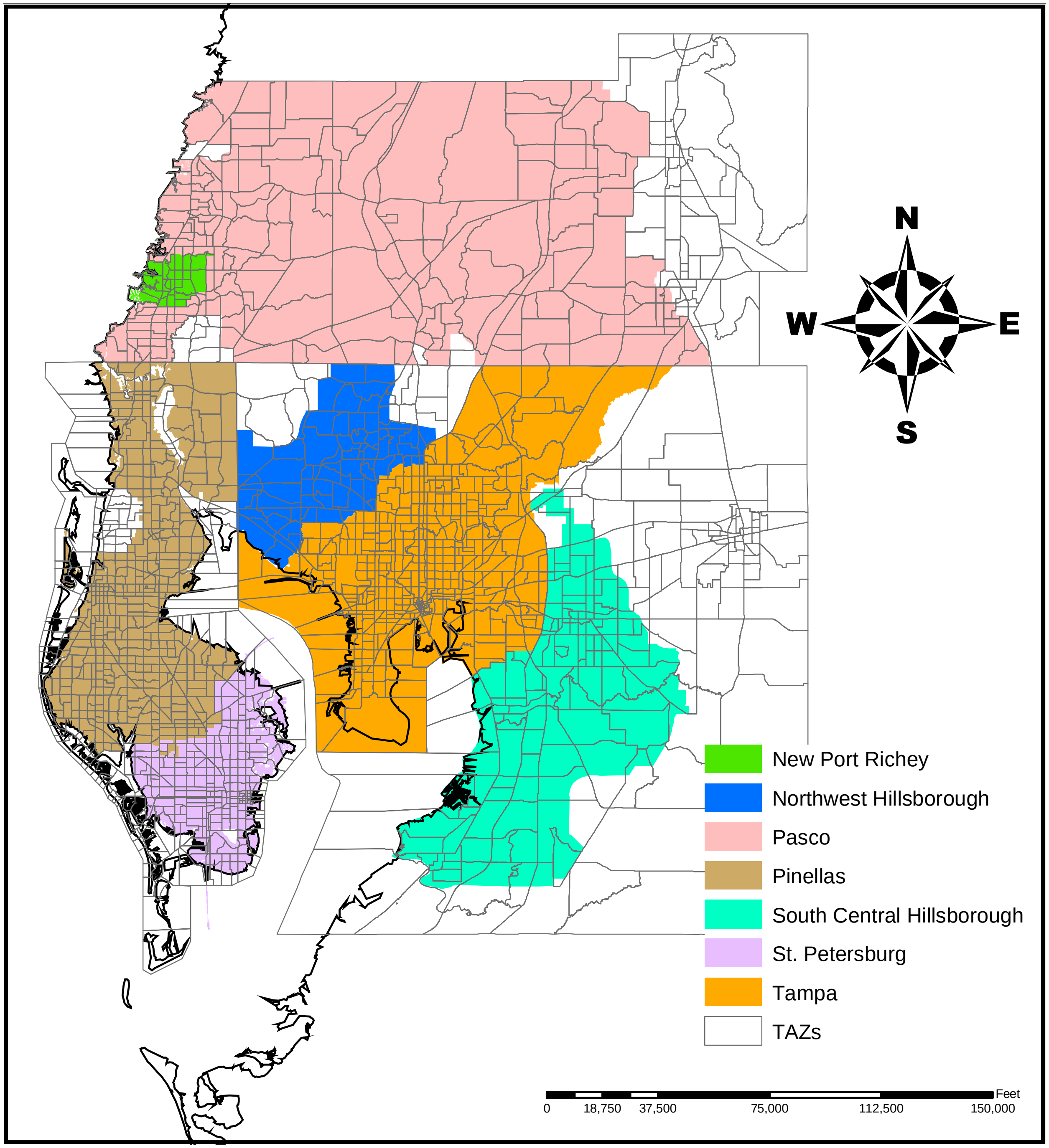


Figure 1.3 – Locations and boundaries of Traffic Analysis Zones (TAZs) used in developing demand models.

Water use data for parcels within each TAZ spanned a maximum of 48 months (January 1999 to December 2002) for each of three primary water use sectors. The modeling database contained average water use, socioeconomic, and meteorological observations within each TAZ, month, and sector. Sectoral models were subsequently developed to relate average per-unit water use to socioeconomic conditions and weather for each month, TAZ, and sector. The following sections describe how modeling database observations were derived for the period 1999-2002.

1.2.2 Per-Unit Water Use and Driver Units

TAZ-average sectoral per-unit demand observations were used as the dependent variable in sectoral model development. These data were obtained by dividing total billed monthly water use within each TAZ, sector, and month by the number of days in the month and number of units within that TAZ, sector, and month. It was necessary to obtain data for number of SF households, MF dwelling units, and employees by TAZ and month to derive per-unit demand data.

Except for a few cases, billing data did not contain number of units per account for single-family or multi-family customers, so unit data was estimated. It was assumed that the number of single-family households in each parcel was equal to the corresponding number of single-family accounts, since single-family households generally have only one account each. By this assumption, number of single-family households in each TAZ was determined by counting the number of SF accounts in each TAZ.

A similar 1-to-1 account-to-housing unit assumption was not made for multi-family dwelling units, since each multi-family account can generally contain more than one dwelling unit[2]. At the time of the study, only Pinellas and St. Petersburg member utilities collected monthly multi-family housing unit data by account in billing databases. Number of multi-family units in Pinellas and St. Petersburg TAZs were determined by summing MF unit data for accounts within those TAZs.

Multi-family unit data for TAZs in other WDPAs were estimated. First, number of dwelling units *per multi-family account* was estimated for each TAZ not in St. Petersburg or Pinellas WDPAs. A simple regression model was developed relating TAZ-level multi-family unit observations in Pinellas and St. Petersburg to number of multi-family accounts in those TAZs and months and total number of multi-family housing units within those WDPAs[3]. Data for the latter two variables were available for all WDPAs and TAZs. This regression equation was used to estimate yearly average dwelling units per multi-family account in each TAZ outside Pinellas and St. Petersburg. Finally, number of multi-family dwelling units was estimated for each TAZ and month by multiplying number

2 *Multi-family dwelling units was selected as a driver variable over multi-family accounts because water use per account could vary considerably with the number of dwelling units served. Multi-family dwelling units thus allowed more rational modeling of demand on the basis of discrete water consumers.*

3 *Data obtained from Experian via GIS Solutions, Inc., July 2002.*

of accounts in that TAZ – month by average number of units per account. Appendix A describes this estimation procedure in greater detail.

Yearly average data for total number of employees by non-residential street address was obtained from Experian. These data were geocoded to parcels by GIS Solutions, then summed within TAZs. Yearly TAZ-total number of employees was then stored in the modeling database, assigning yearly TAZ totals to each month within a TAZ and calendar year.

1.2.3 Socioeconomic Data

Socioeconomic data were derived using commercially available sources. Experian provided yearly socioeconomic data for a sampling of addresses in each WDPA, including household income, persons per household, and developed residential acreage for residential addresses and number of employees in commercial, industrial and service business categories for non-residential addresses[4]. GIS Solutions, Inc. mapped these addresses to TAZ and averaged the Experian data by TAZ, producing yearly TAZ averages of household income, persons per household, housing density, and fraction of employment in commercial, industrial, and service businesses. Yearly TAZ-average data was then stored in the modeling database, assigning yearly TAZ averages to each month within a TAZ and year.

The billing database identified accounts with reclaimed water connections. For each TAZ, month, and sector, fraction of reclaimed accounts was determined by dividing the number of accounts with reclaimed connections by the total number of accounts.

1.2.4 Meteorological Data

Historical weather data, including daily rainfall totals and maximum daily temperatures, were collected for six NOAA weather recording stations dispersed across the Tampa Bay Water service area (see Figure 1.4). The six weather stations are listed in Table 1.1.

---

4 *Experian reported employment data within nine business categories, or SIC's. Projection data for forecasting was only available at broader categorizations of Service, Industrial, and Commercial entities. The latter categorization was thus used in the modeling database. SIC categories were mapped to Service, Industrial, and Commercial groups; SICs 1, 4, 5, 6, and 9 were assigned to Commercial employment, SICs 2 and 3 were assigned to Industrial employment, and SICs 7 and 8 were assigned to Service employment.*

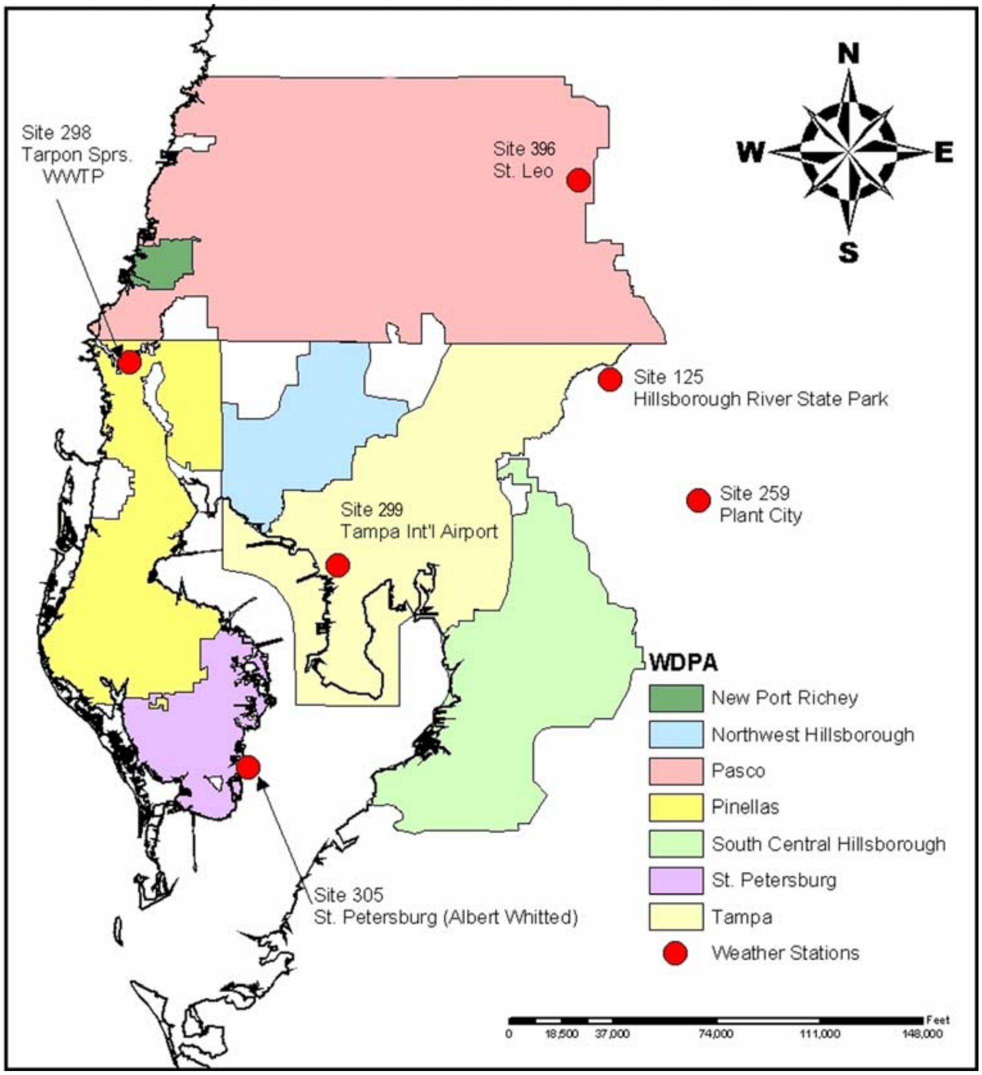


Figure 1.4 – NOAA weather stations used for meteorological data in demand model development (with corresponding SWFWMD site numbers).

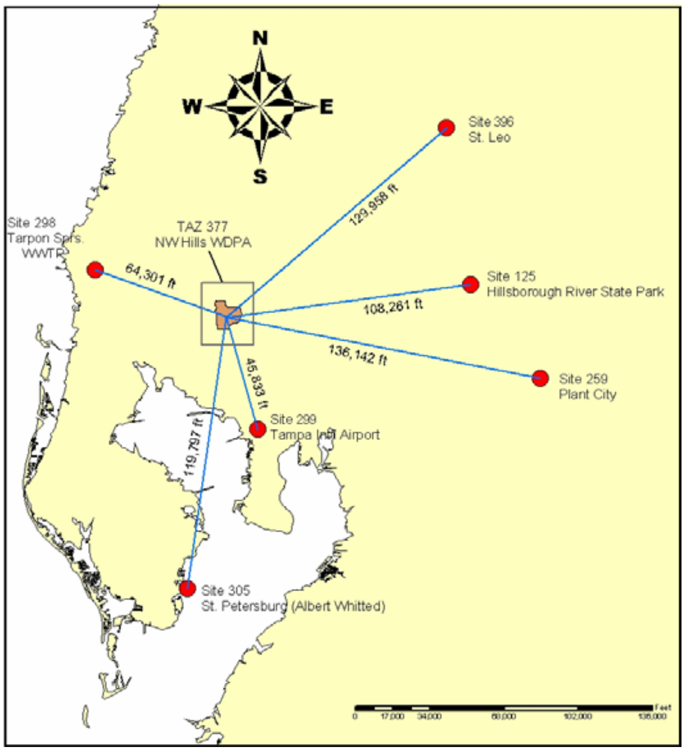
Table 1.1
NOAA Stations Used For Weather Measurements In Modeling

Station*	COOP #*	SWFWMD #	Latitude*	Longitude*	Period of Record
Hillsborough River State Park	083986	125	28º09’ N	82º14’ W	1943-2002
Plant City	087205	259	28º01’ N	82º09’ W	1901-2002
Saint Leo	087851	396	28º20’ N	82º16’ W	1895-2002
St Petersburg Albert Whitted	087886	305	27º46’ N	82º38’ W	1914-2002
Tampa International Airport	088788	299	27º58’ N	82º32’ W	1901-2002
Tarpon Springs Swg Plant	088824	298	28º09’ N	82º45’ W	1901-2002

* as listed by NOAA

Data from these stations were employed to construct tables of meteorological variables for model development, including monthly and long-term monthly normal observations on average daily maximum temperature, precipitation, and number of days per month with greater than 0.01 and 1.00 inch of rain. These terms are eventually included in the per unit demand model. Values for weather variables were calculated for each TAZ and month by inverse-squared distance-weighted average of weather station data, with weights corresponding to inverse squared distance from the geographical centroid of each TAZ to each of the stations.

Figure 1.5 illustrates the inverse-squared distance-weighted averaging procedure. In the figure, total rainfall in TAZ 377 is estimated for February 1996 using total rainfall observations at each station for that month. Distances are calculated from the centroid of TAZ 377 to each weather station (upper left map and topmost table, upper right), using standard functionality in GIS software. A raw weight is determined for each station by squaring, then inverting, the distance to TAZ 377. Normalized weights are determined by dividing each raw weight by the sum of raw weights. Finally, TAZ 377 rainfall for February 1996 is estimated by multiplying each weather station's observation of total rainfall for February 1996 by the normalized weight for the corresponding station. Normalized weights thus reflect the relative contribution of weather at each station to weather in TAZ 377. Higher weights correspond to weather stations that are closer to TAZ 377 and presumably experience more similar weather conditions.



TAZ 377- To-Station Distances and Weights

Weather station	Dist from TAZ 377, ft	Raw weight ($1/dist^2$)
Hills Riv St Pk	108,261	$8.53x10^{-11}$
Plant City	136,142	$5.40x10^{-11}$
St. Leo	129,958	$5.92x10^{-11}$
St. Petersburg	119,797	$6.97x10^{-11}$
TPA Int'l Airport	45,833	$4.76x10^{-10}$
Tapron Sprs WWTP	64,301	$2.42x10^{-10}$

February 1996 Total Monthly Rainfall at Weather Stations

Weather station	Total Precip. Feb 1996 (inches)
Hills Riv St Pk	3.08
Plant City	3.22
St. Leo	3.48
St. Petersburg	1.00
TPA Int'l Airport	3.04
Tapron Sprs WWTP	3.86

Determine February 1996 Total Rainfall For TAZ 377 by Inverse-Squared Weighting of Station Data:

1. Calculate sum of raw weights

   sum of raw weights = $8.53x10^{-11} + 5.40x10^{-11} + ... + 2.42x10^{-10} = 9.86x10^{-10}$

2. Calculate normalized weight for each station

   normalized weight = weight / sum of weights:

Weather station	Normalized weight
Hills Riv St Pk	$8.53x10^{-11} / 9.86x10^{-10}$ = 0.087
Plant City	0.055
St. Leo	0.060
St. Petersburg	0.071
TPA Int'l Airport	0.483
Tapron Sprs WWTP	0.245

3. Calculate TAZ 377 rainfall for Feb 1996

   Rainfall (TAZ 377, Feb1996) = sum of station rain-falls in Feb 1996 times normalized weights for those stations:

   Rainfall (TAZ 377, Feb1996) =
   0.087 x 3.08 + 0.055 x 3.22 +
   0.060 x 3.48 + 0.071 x 1.00 +
   0.483 x 3.04 + 0.245 x 3.86

   = 3.14 inches

Figure 1.5 – Inverse-squared distance-weighted averaging for estimation of TAZ weather measurements.

1.2.5 Wholesale and Unbilled Water Use Data

In addition to retail water deliveries, several members currently sell water on a wholesale basis to other utilities, and some water delivered to members by Tampa Bay Water ends up unaccounted for (unbilled) within member systems. Historical data were derived for these components by WDPA. For wholesale demand, historical monthly wholesale delivery data were directly obtained from individual members. To determine historical unbilled demand, data were first obtained from Tampa Bay Water for historical deliveries to members. Historical retail and wholesale deliveries by each member was then summed and compared to Tampa Bay Water delivery quantities to each member, with differences representing historical unbilled water demand by members.

1.2.6 Marginal Price of Water Data

Member-specific data for marginal price of water and sewer were obtained from Tampa Bay Water. Each TAZ was assigned the marginal price of the WDPA within which it was located.

1.3 Specification of Econometric Models

The TAZ-level historical water use and socioeconomic database was used to estimate econometric models of water use in single-family, multi-family, and non-residential sectors. The sections below describe the general linear estimating framework and the specific model estimation procedures that were employed to build the final models used for forecasting.

1.3.1 General Linear Model

Multiple regression analysis was used to estimate models of per unit rates of use. Regression is commonly used to estimate a direct and quantifiable relationship between a variable of interest, the *dependent variable*, and a set of independent variables, or *explanatory variables*, that are hypothesized to explain changes in the dependent variable. The general linear regression model, as applied to per-unit demand forecasting, may be expressed as a sum of constants times explanatory variables, as illustrated in Figure 1.6.

General regression model form:

dependent variable = a fitted constant + (an observed explanatory variable x another fitted constant) + (another observed explanatory variable x another fitted constant) + etc.

Example:

predicted per-unit demand in a sector, WDPA, and month = intercept for sector, WDPA, and month + (per capita income in sector, WDPA, and month x coefficient for income) + (rainfall in WDPA and month x coefficient for rainfall) + etc.

Figure 1.6 – General form of linear regression models and an example in the context of sectoral per-unit demand models.

In linear regression models, values for fitted constants are determined to bring predicted per-unit demand (on the left side of Figure 1.6) as close as possible to observed per-unit demand. Mathematically, the linear regression model form is expressed as a regression equation:

$$q_{s,t,i} = \alpha_{s,t,i} + \sum_{m} \beta_{m,s} X_{m,s,t,i} \qquad (1.1)$$

where

- $q_{s,t,i}$ = sectoral water demand for sector s (SF, MF, or NR), month-year t, and WDPA i

- $X_{m,s,t,i}$ = the value at month-year *t* in WDPA *i* of the *m*th explanatory or independent variable (e.g., household size, temperature) relevant to demand prediction for sector s

- $\alpha_{s,t,i}$ = an estimated model intercept term for sector *s*, WDPA *i*, and the calendar month of month-year *t*

- $\beta_{m,s}$ = estimated model parameter (specific to the model for sector s) that measures the relationship between $q_{s,t,i}$ and an explanatory variable $X_{m,s,t,i}$ (that is, if the same variable appears in more than one sector model, that variable can have a different model parameter in each model)

Parameter values were determined for per-unit demand models using TAZ-level $q_{s,t,i}$ and X values from the modeling database as datapoints for regression. Ordinary Least Squares (OLS) regression determined values for $\alpha_{s,t,i}$ and $\beta_{m,s}$ that minimized the difference between observed and predicted TAZ-level demand.

Prior to determining $\alpha_{s,t,i}$ and $\beta_{m,s}$ coefficients, regression equations were subjected to natural logarithmic transformation, or *log transformation* (for mathematical details of log transformation, refer to Appendix A). Log transformation enabled interpretation of each fitted model parameter as an *elasticity*, or the percent change in per-unit demand caused by a one percent change in each parameter's associated explanatory variable. For example, if real marginal price in a log-transformed model had a coefficient of -0.3, it would directly follow that a one percent increase in price would cause a 0.3% decrease in per-unit demand. This straightforward interpretive capability greatly facilitated experiential validation of model parameters and assessment of model sensitivity.

In addition, log transformation facilitated several mathematical aspects of the regression process. These benefits are discussed in Appendix A.

1.3.2 Model Specification and Estimation Procedures

Separate regression models were created to account for metered single-family, multi-family, and non-residential uses. Only those explanatory variables contained within the TAZ-level modeling database were examined for potential inclusion in these models. However, the modeling database was designed to maximize, to the extent practicable, the availability of data on variables that have been shown to affect water use among the three primary sectors under evaluation.

An important objective of the modeling process was to increase explanatory power of the models while achieving rational estimates for model parameters consistent with expected directions and magnitudes of influences found in the literature. This dual objective was approached through an iterative process of specifying alternative variables, screening of outlying data, and analyzing model residuals. Various robust estimation methods were employed in the process of iteratively specifying models, in attempts to *"dampen the influence of outlying cases...in an effort to provide a better fit for the majority of cases"* (Neter, 1996).

In general, and because of the very large number of available observations, the model estimation process encountered a great deal of variance that could not be explained. This result was attributed to the relative sizes of spatial and temporal dimensions spanned by the database. For each sector, between-TAZ variance, or cross-sectional variance, dominated time-series variance (there were typically 1,000 or more spatial cross-sections, *i.e.*, TAZs, and only a maximum of 48 time-series observations on water sales for each TAZ). In many cases it was also observed that independent variable variance exceeded relative variance in water use (a common situation when there are more cross-sections than time-periods).

Data screening measures were instituted to omit suspicious and incomplete observations. Incomplete observations were TAZs for which some explanatory variable values were missing. Suspicious outlier TAZs with abnormally high or low per-unit sectoral water use were screened. Additional screening was accomplished by analyzing values for independent variables and correcting or omitting observations when values of explanatory values seemed extreme or nonsensical (*e.g.*, nonzero residential water use but no housing units). Outlier screening procedures are described in Appendix A. After removal of incomplete data points, outlier TAZs generally composed less than 5% and in some cases less than 1% of remaining observations.

In addition to socioeconomic and weather influences on demand, the models were also specified to account for unique location characteristics (*i.e.*, WDPA in which a TAZ was located) and other monthly-varying systematic behavior that could not be attributed to explanatory variables. These specifications took the form of adjustments to model intercepts (*e.g.*, the α term in Equation 1.1) by month and WDPA. Implementing these adjustments made it possible for rational coefficient estimates (*i.e.*, estimates that displayed proper signs and expected numeric magnitudes) to be obtained.

Table 1.2 lists the variables and adjustments that were found to be statistically significant and insignificant in each of the three per-unit models. Per-unit models differed by sector of applicability (whether single-family, multi-family, or non-residential), driver units (occupied living units or number of employees), and model coefficients (as shown in Table 1.2 and Tables A.1 – A.5 of Appendix A). However, they all shared the same rate of use times driver forecasting methodology and some demographic, price, and weather model inputs. Specific data screening conventions and parameter estimation processes varied among the three models. Appendix A discusses the particular actions that were necessary to estimate each sector model. Concise model equations are presented in Appendix B.

Table 1.2
Statistically Significant Explanatory Variables and Associated Parameter Values in Sectoral Per-Unit Demand Models

Variable	SF Model	MF Model	NR Model
Intercept	2.91062	1.47578	1.6167
Monthly Intercept Adjustments:			
January	N/A	N/A	N/A
February	-0.0007992	o	o
March	0.03439	o	o
April	0.0974	o	o
May	0.09819	o	o
June	0.05658	o	o
July	-0.00854	o	o
August	0.00004935	o	o
September	0.03774	o	o
October	0.09224	o	o
November	0.08565	o	o
December	0.04198	o	o
WDPA Intercept Adjustments:			
Pinellas	-0.14491	o	o
St. Petersburg	-0.36252	0.48567*	o
New Port Richey	-0.30522	0.48567*	o
Pasco	-0.17354	0.48567*	o
Tampa	-0.1223	o	o
NW Hillsborough	-0.23185	0.48567	o
SC Hillsborough	-0.26407	o	o
Ln of income	0.26199	0.37054	0.12075
Ln of SF/MF housing density	-0.11679	-0.35254	N/A
Ln of SF/MF persons per household	0.55785	o	N/A
Ln of real marginal price	-0.24779	o	o
Ln of Commercial (SIC 4-6,9) fraction	N/A	N/A	1.01109
Ln of Industrial (SIC 2,3) fraction	N/A	N/A	0.34798
Ln of Services (SIC 7,8) fraction	N/A	N/A	1.19036
Departure of maximum Ln temperature	0.99185	o	o
Departure of max Ln temp, 1 month lag	0.90542	o	o
Departure of max Ln temp, 2 month lag	0.81999	o	o
Departure of max Ln temp, 3 month lag	0.73256	o	o
Ln rainfall	o	-0.01717	o
Departure of Ln rainfall	-0.02799	o	-0.04958

Table 1.2
Statistically Significant Explanatory Variables and Associated Parameter Values in Sectoral Per-Unit Demand Models

Variable	SF Model	MF Model	NR Model
Departure of Ln rainfall, 1 month lag	-0.0208	o	-0.03609
Departure of Ln rainfall, 2 month lag	-0.0136	o	-0.01708
Departure of Ln rainfall, 3 month lag	-0.00641	o	o
Departure of Ln days with 0.01 inch of rain	-0.0151	o	o
Ln of days with 1.0 inch of rain, 1 month lag	-0.0246	o	o
Ln Fraction of SF/MF/NR accounts with reclaimed water	-0.36585	-0.38540	o

o – not statistically significant

* – statistically insignificant for TAZ-level modeling, later modified for WDPA-level modeling

1.4 Interpretation of Model Estimates

This section describes results of fitting the three sectoral per-unit models. Model coefficients listed in Table 1.2 were used in log-transformed per-unit demand equations for the remainder of the model development effort. Application of these coefficient values in per-unit equations is described in Appendix B. The coefficients can also be interpreted as elasticities, which will assist experiential validation of the models in the following subsections.

1.4.1 Single-Family Residential Model Estimates

Almost 43,000 cross-sectional and time-series observations were used in estimation of the final single-family residential model shown in the first column of Table 1.2 (see previous section). The R-square value for the single family per-unit model was 0.39 (see Table A.3 of Appendix A), implying that about 39 percent of variation in TAZ-level average daily water use per housing unit was explained by the model.

The high volume of cross-sectional and time-series observations used in fitting the single-family model introduced a substantial amount of variance into the modeling process, thereby producing a low R-square value. The fine-scale TAZ geographies contained significant unmodeled heterogeneity that was not completely described by dependent and independent variables. Though this high variance served to dampen explanatory power at the TAZ level, the large volume of data helped lower standard errors of parameter values (see Table A.3 of Appendix A) and contributed to confidence in those values. Given the size and pooled nature of the modeling database, an R-square of 0.39 was therefore fairly high. Most importantly, predictive accuracy obtained by the model when applied at the WDPA level (summarized in Section 1.5) confirmed the reliability of the regression model and its parameters at larger spatial scales.

Given that the single-family per-unit model was based on a large number of pooled time-series cross-sectional observations (with dominant cross-sectional variation), the explanatory power of this model was considered relatively high and indicative of a fairly homogeneous customer class. This was true in comparison to multi-family and non-residential sectors which, as discussed in the following subsections, contained even greater heterogeneity.

Parameter values for monthly indicator variables suggest a significant degree of seasonal variation in single-family water use. Everything else remaining the same, single-family use was relatively bi-modal over a calendar year, with peak average demands occurring in the April/May time frame and a small peak in the October/November time frame.

Aside from normal seasonal fluctuations over the calendar year, single-family use was sensitive to deviations from normal weather conditions. Higher-than-normal maximum daily temperatures led to higher demand and greater-than-normal precipitation led to lower demand. Parameter estimates for prior-month weather conditions, or *lagged* weather, suggest that deviations from normal weather influence demand up to three months in the future. In addition, frequency of rainfall events was shown to affect single-family water use. More-frequent-than-normal rainfall, measured as number of days in a month with more than 0.01 inch of precipitation, led to lower water demand. A similar effect was observed for number of days in a month with greater than 1 inch of precipitation. These estimated effects followed the logic that watering in the current month is influenced by how often it rains in the current month, and with the number of days with larger rainfall events in the preceding month.

Coefficients of socioeconomic variables in the single-family model retained expected signs and relative magnitudes. Higher incomes and larger households led to higher single-family demands. Alternatively, increased single-family housing development density led to lower water use, which followed from less irrigated acreage and other outdoor use with more single-family units per acre. Marginal (or volumetric) price for water and sewer services was also found to be a significant variable. The price elasticity of single-family water use was estimated at about –0.25, which suggested that a one percent increase in marginal price of water and sewer reduced single-family use by about one-quarter of one percent.

Finally, the analysis suggested that coverage of reclaimed accounts had an impact on single-family water use. The estimated coefficient for reclaimed water coverage suggested that a one percent increase in fraction of total single-family accounts with reclaimed water service led to a decrease in water use of approximately 0.37 percent.

1.4.2 Multi-Family Residential Model Estimates

Column 2 of Table 1.2 displays model estimates for the multi-family residential model. The multi-family sectoral model had an R-squared of 0.30 (see Table A.4 of Appendix A), suggesting about 30 percent of variation in multi-family water use was explained by the model. Unlike the single-family sector, multi-family use did not demonstrate a significant seasonal variation over the calendar year. Furthermore, total monthly precipitation was the sole weather variable with any significance. As expected, more rainfall in any particular month reduced multi-family water use.

Statistically significant and rational coefficient estimates were obtained for two socioeconomic variables, household income and multi-family housing density. TAZs with higher household incomes displayed higher multi-family use. Similar to the single-family model, TAZs with denser multi-family housing development had lower water use. A 1 percent increase in average household income was estimated to produce a 0.37 percent increase in multi-family demand. Furthermore, a 1 percent increase in number of multi-family units per acre was estimated to result in about a 0.35 percent decrease in multi-family water use. Parameter values for all other socioeconomic variables were statistically insignificant.

Fraction of accounts with reclaimed water connections had a statistically significant effect on multi-family water use. The estimated coefficient for the reclaimed water variable was similar to the coefficient estimated for the single-family sector. A one percent increase in fraction of total multi-family accounts with reclaimed water service was estimated to lead to a decrease in water use of approximately 0.39 percent.

1.4.3 Non-Residential Model Estimates

Column 3 of Table 1.2 presents the estimated non-residential model for daily per-employee water use. The model explained only 2 percent of variation in per-employee water use found in modeling data (see Table A.5 of Appendix A), which contained 39,727 observations. This explanatory power was low relative to single- and multi-family models, as expected, given the typically heterogeneous nature of non-residential use (discussed in Section 1.1.1).

Like the multi-family sector, non-residential use per employee did not display a significant and systematic seasonal trend across calendar months. However, non-residential water use was influenced by precipitation. Greater than normal rainfall reduced per-employee water use for up to two lagged monthly periods.

The non-residential model indicated higher household incomes led to higher per-employee water use. This result provided a link between economic affluence, business activity, and associated water uses. The model also quantified sensitivity of per-employee water use to types and distribution of employment. The model contained variables reflecting fraction of total employment in a geographical area belonging to each of three broad employment categories: commercial, industrial, and services. In general, the model cannot be used directly to estimate the impact of changing proportions of employment in a single employment category without knowing corresponding changes in proportions among the remaining two categories, because this group of variables works together to create a distribution of total employment. However, from coefficient estimates it was possible to judge relative sensitivity of per-employee water use to employment among the broad business activities.

Average water use per employee was most sensitive to proportion of total jobs found in the services category. Characteristic customers in the services category normally include businesses such as banks and office complexes, which often have extensive landscaping and irrigation demands, aside from common domestic uses of water. Proportion of employment in services is vitally important, since employment in this category in the Tampa Bay region typically accounts for a higher proportion of jobs than any other aggregate category. In addition, services employment in the U.S. is anticipated to grow at a faster rate than other business activities.

Per-employee water use was also very sensitive to employment in the commercial category. This category is typically a broad composite of retail and wholesale trade, construction, and other activities involving a wide range of businesses or stores. The prevalence of eating and drinking establishments, shopping malls, and large warehouses in urban areas made this category an important determinant of non-residential use as well.

Average per-employee use was less sensitive to proportion of total employment in the industrial category. In general, this was likely due to the relatively small proportions of industrial employment among TAZs, since it is generally true that industrial processes can use significant quantities of water. In other words, as one allocates a greater proportion of jobs to services and commercial activities, the impact of such industrial uses on a per-employee average basis should diminish.

No statistically significant correlation was found between fraction of non-residential accounts with reclaimed water service and potable non-residential water demand. This is not to say that use of reclaimed water in the non-residential sector is generally inefficient in offsetting potable water demand. Rather, it is more likely that data available for modeling was insufficient to establish a reliable estimate of potable water demand related to reclaimed use, due primarily to wide variation in per-employee use data.

1.4.4 Model Summary

The abundance of spatial data by TAZ allowed up to 43,000 cross-sectional time series observations for water use modeling for each primary water use sector (single-family residential, multi-family residential, and combined non-residential). The high volume of spatial cross-sections allowed for development of precise estimates of effects of demographic factors on sectoral water use.

The models explained between 2 and 39 percent of time-series cross-sectional variation in water use, depending mainly on degree of homogeneity of customers within each sector under consideration. The single-family (high homogeneity) model displayed the greatest explanatory power, followed by the multi-family (moderate homogeneity) model and the non-residential model (low homogeneity). Despite differences in explanatory power, each model produced relatively low standard errors of estimated parameter values relative to observed TAZ-level water use in each sector. As shown in Section 1.5, these models provided an excellent basis for projecting WDPA-level demands in the Tampa Bay region.

There were several modeling and database elements that could be refined in the future through additional data collection and analysis. Though the richness of spatial data provided high confidence in the parameter estimates for demographic factors, more time-series data would likely improve the specification and estimation of weather effects and seasonal trends. Furthermore, presence of watering restrictions throughout the modeling period may have introduced difficulties in modeling seasonal and weather effects, effectively counteracting normally-expected relationships between demand and weather. Heterogeneity in the non-residential sector was difficult to overcome in a modeling context. Due to differences in utility accounting practices it was not possible to differentiate non-residential customers into more disaggregate and homogeneous classes, nor was it possible to obtain statistically significant measurements of the impact of reclaimed water service on non-residential demand.

1.5 Aggregation of Model Predictions and Verification of Predictive Accuracy

Predictive accuracy of the sectoral models and aggregation procedures were assessed in concert with the iterative process of developing the models. Model verification involved a backcast of WDPA and total Tampa Bay Water demands by user sector for Water Year 2002. This backcast was produced by multiplying per-unit predicted demands for each TAZ, sector, and month by corresponding driver unit values in the modeling database, then summing the resulting TAZ-level sectoral total demands by WDPA. WDPA total demand (the sum of WDPA sectoral demands) was adjusted to account for wholesale and unbilled water demand. Predictions of demand were then compared to observed WY2002 use by WDPA and sector to evaluate predictive performance.

1.5.1 Wholesale and Unbilled Demand Adjustments

Projection of total metered water demand for each WDPA included monthly sectoral demand predictions plus water deliveries to wholesale facilities. Total demand predictions consisted of total metered demand plus other/unbilled use. Wholesale and unbilled demands are described in Section 1.1, and data collected for their assessment are described in Section 1.2.

Wholesale deliveries were assumed to be a constant percent of total retail demands. Wholesale fraction (WS) for each WDPA was derived from observed wholesale deliveries by each member government for Water Year 2002. Other/unbilled use was assumed to be a constant percent of total water deliveries (total metered demand plus unbilled use) by Tampa Bay Water to its members. The other/unbilled water factor (OUW) was derived from the observed difference between total metered end use demand and total water delivered from Tampa Bay Water to its members for Water Year 2002 by WDPA and month. Total WDPA metered water demand was then adjusted by these fractions. Appendix B contains actual equations for performing WS and OUW adjustments to retail demand.

1.5.2 Comparison of Observed and Model-Estimated Demand for WY 2002

Table 1.3, Figure 1.7, and Figure 1.8 show assessments of model predictive accuracy for Water Year 2002. Predictions were compared to Tampa Bay Water's 2002 deliveries. The large-scale accuracy benefits of forecast disaggregation (i.e., bottom-up summation of predictions) are clearly shown in the table and figures.

Figure 1.7 compares observed and predicted total demand for each WDPA and the total Tampa Bay Water service area. Sectoral predictions for individual WDPAs was highly accurate in absolute terms[5.] According to this figure, the point model yielded prediction of total regional demand within 1.65 MGD, or 0.69 percent, of the observed value for Water Year 2002.

---

5 *Multi-family and non-residential models were originally estimated to provide the best fit for TAZ-level demand observations. However, these models, when used to simulate WDPA-level demand based on WDPA-level inputs, were slightly inaccurate. Predictive accuracy for the multi-family and non-residential sector models were subsequently improved by modifying some WDPA intercept adjustments. For the multi-family model, intercept adjustments for Northwest Hillsborough, St. Petersburg, and Pasco were changed from zero (statistical insignificance) to 0.48567, the intercept adjustment for TAZs within New Port Richey (as indicated in Table 1.5). For the non-residential sector, all WDPA intercept adjustments were disregarded.*

Table 1.3
Evaluation of Predictive Accuracy of Water Use Models for Water Year 2002 (Values In MGD Unless Otherwise Noted)

Sector	Pinellas	St. Pete	NPR	Pasco	Tampa	NW Hills	SC Hills	Total
Single-family observed	23.04	12.87	1.53	12.58	25.77	11.38	15.73	102.90
Single-family predicted	24.22	13.37	1.34	12.42	27.84	10.21	15.62	105.02
Difference (Pred - Obs)	1.18	0.50	-0.19	-0.16	2.07	-1.16	-0.12	2.12
Percent Difference	5.1%	3.9%	-12.4%	-1.3%	8.0%	-10.2%	-0.7%	2.1%
Multi-family	10.99	7.42	0.57	1.40	12.73	2.52	3.25	38.87
Multi-family predicted	10.98	6.37	0.64	1.06	10.85	1.72	5.44	37.06
Difference (Pred - Obs)	0.00	-1.04	0.07	-0.34	-1.87	-0.80	2.19	-1.80
Percent Difference	0.0%	-14.1%	12.4%	-24.6%	-14.7%	-31.7%	67.2%	-4.6%
Non-residential	8.95	6.84	0.69	2.28	20.46	1.98	3.89	45.08
Non-residential predicted	8.96	6.66	0.57	2.67	22.11	2.46	3.86	47.29
Difference (Pred - Obs)	0.02	-0.17	-0.12	0.39	1.65	0.48	-0.03	2.21
Percent Difference	0.2%	-2.5%	-17.9%	17.1%	8.1%	24.1%	-0.9%	4.9%
Total Retail observed	42.97	27.12	2.78	16.27	58.95	15.87	22.88	186.85
Total Retail Predicted	44.17	26.41	2.54	16.15	60.80	14.39	24.91	189.37
Difference (Pred - Obs)	1.19	-0.72	-0.24	-0.12	1.85	-1.48	2.03	2.52
Percent Difference	2.8%	-2.6%	-8.7%	-0.7%	3.1%	-9.3%	8.9%	1.3%
Wholesale Observed	22.37	2.22	0.36	0.70	0.29	0.00	0.00	25.94
Wholesale Predicted	20.60	2.14	0.31	0.52	0.23	0.00	0.00	23.80
Difference (Pred - Obs)	-1.77	-0.08	-0.05	-0.18	-0.07	0.00	0.00	-2.14
Percent Difference	-7.9%	-3.5%	-13.9%	-25.1%	-22.6%	--	--	-8.3%
Other/Unbilled Observed	4.24	2.94	0.38	2.14	14.28	0.90	2.31	27.19
Other/Unbilled Predicted	4.19	2.77	0.37	2.00	12.37	0.97	2.50	25.17
Difference (Pred - Obs)	-0.05	-0.17	-0.01	-0.15	-1.91	0.07	0.19	-2.02
Percent Difference	-1.2%	-5.9%	-1.4%	-6.8%	-13.4%	7.5%	8.4%	-7.4%
Total Demand Observed	69.58	32.29	3.52	19.11	73.52	16.77	25.18	239.98
Total Demand Predicted	68.95	31.32	3.22	18.67	73.40	15.36	27.41	238.33
Difference (Pred - Obs)	-0.63	-0.97	-0.30	-0.44	-0.12	-1.42	2.23	-1.64
Percent Difference	-0.9%	-3.0%	-8.5%	-2.3%	-0.2%	-8.4%	8.8%	-0.7%

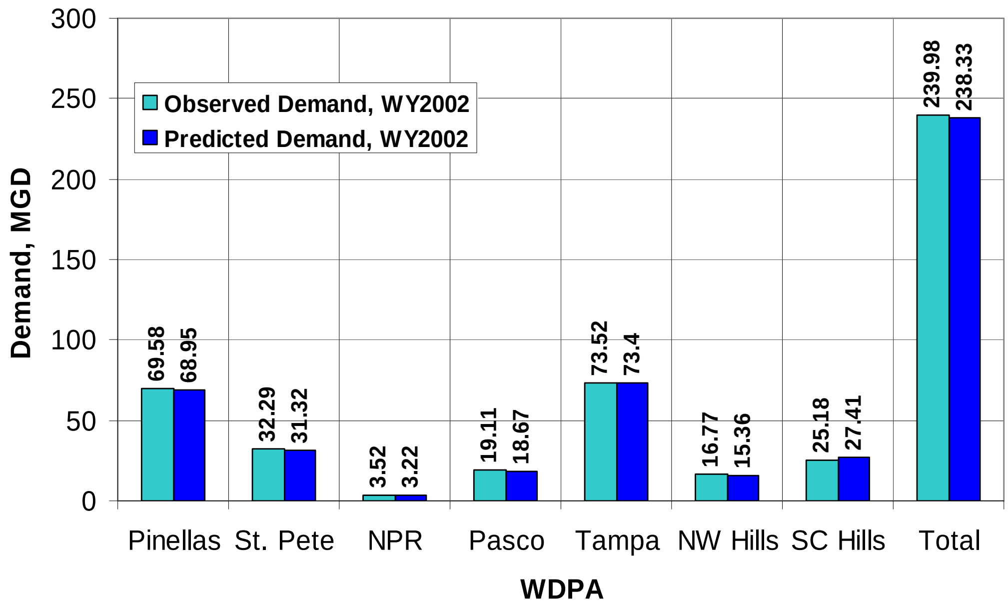


Figure 1.7 – Comparison of Observed and Point-Model-Predicted Water Year 2002 Demand for Each WDPA and the Tampa Bay Water Service Area.

By WDPA, the largest prediction error was an overprediction of only 2.23 MGD in South Central Hillsborough, and other than an underprediction of 1.42 MGD in Northwest Hillsborough, all errors were less than 1 MGD. On a percentage basis, some of these discrepancies appeared large (over 8% of observed demand for New Port Richey, Northwest Hillsborough, and South Central Hillsborough), but these large percentages corresponded to small total observed demands and absolute demand discrepancies. Given the small absolute predictive errors by WDPA and for the Tampa Bay Water service area, predictive accuracy appeared excellent for purposes of understanding total demand on a member and regional basis.

Figure 1.8 compares observed and predicted total demand Tampa Bay Water for each sector. Absolute predictive errors were slightly higher for total sectoral demands than for WDPA demands, with all error magnitudes between 1.8 and 2.5 MGD. These errors were small, however, in comparison to observed demand. On a percentage basis, errors were less than 5% for SF, MF, and NR sectors. Understandably, percent errors in wholesale and unbilled demand segments were higher (7.4 and 8.3 percent, respectively), as these demand segments were not modeled in terms of any explanatory or driver variables. In addition, these two segments represented the smallest portions of total regional demand, which can inflate absolute errors when expressed as percentages. Given the small absolute predictive error by sector, it was concluded that

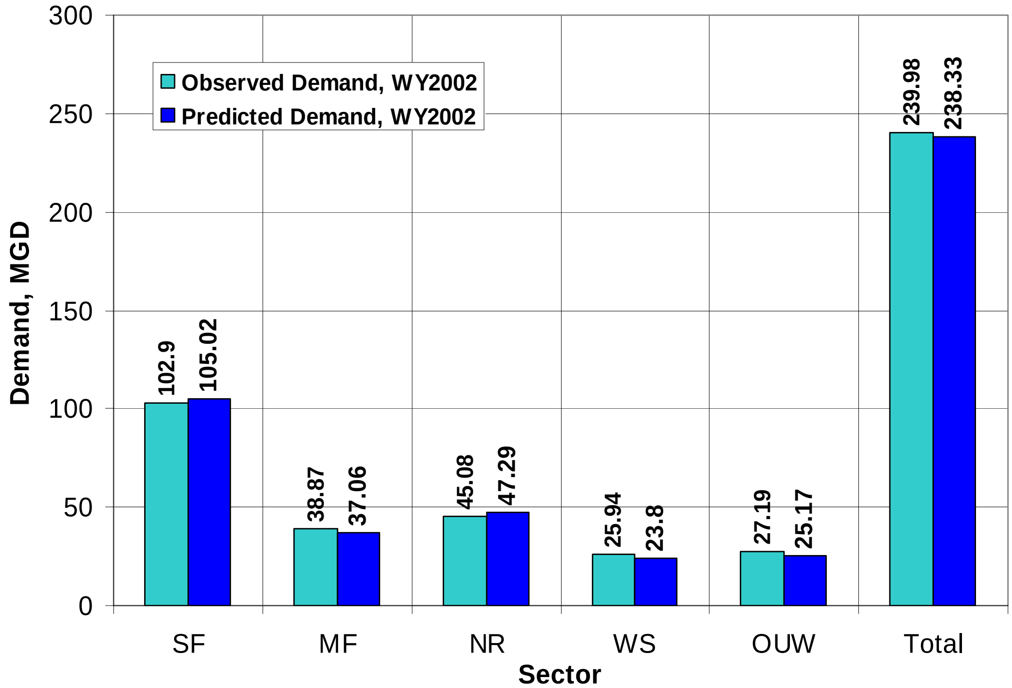


Figure 1.8 – Comparison of Observed and Point-Model-Predicted Water Year 2002 Demand for Each Sector, Wholesale Use, Unbilled Use, and the Tampa Bay Water Service Area Color Polymer System

the point model produced very good predictive accuracy for purposes of understanding total demand on a sectoral basis, though predictive accuracy by WDPA was slightly higher.

Examination of Table 1.3 reveals that, when comparing observed and predicted demand for specific sectors within specific WDPAs, accuracy deteriorated on a percentage basis though absolute errors remained the same magnitude or dropped. This drop-off in accuracy and increase in noise with higher resolution was also observed for TAZ-level data and model development in Section 1.4. It was therefore concluded that the point model, while superb for describing WDPA-level demand and sectoral demand across the entire Tampa Bay Water service area, must be applied cautiously at finer scales with recognition of an increasing potential for error.

1.5.3 Model Calibration

Based on the relatively high degree of accuracy demonstrated during model verification, the sectoral models and aggregation procedures formed a reasonable basis for forecasting future demands in the Tampa Bay region. To eliminate the possibility of measurement errors for model inputs for the base year, water use models were calibrated to exactly match sectoral water use reported by month and sector for each WDPA for Water Year 2002.

The calibration procedure consisted of determining a calibration factor for each sector, WDPA, and calendar month. Each calibration factor was defined as the ratio of predicted to Water Year 2002 observed total demand in a sector, WDPA, and calendar month. Subsequent demand forecasts were adjusted by these calibration factors; forecasted demand in each sector, WDPA, and calendar month was multiplied by the corresponding Water Year 2002-based calibration factor. In this manner, calibration uniformly corrected forecasts by removing the effects of measurement errors detected in Water Year 2002 observed-versus-predicted demand comparisons.

Actual calibration equations are described in Appendix B.

1.6 Point Forecasts of Water Demand

Monthly projections of water demand were derived for each WDPA and the Tampa Bay Water service area as a whole in annual increments to 2025. These projections required information on projected future values of explanatory variables for the three sectoral models and projected values for driver variables. These projection data were derived for the seven WDPAs and fed into the general accounting framework/sequence shown previously in Figure 1.2, where

1. Per-unit daily water use in each month and water use sector was predicted as a function of the values of explanatory variables,

2. Predicted per unit daily water use in each month and sector was multiplied by projected number of housing units or employment, as well as the fixed calibration factor (see section 1.5),

3. Adjustments for wholesale deliveries and other/unbilled use were applied to generate calculations of total water demand, and

4. Monthly sectoral demands were summed over a water/fiscal year to arrive at projections of annual use for each sector and WDPA.

The following sections describe the derivation of required projection data and other assumptions. Resultant forecasts were summarized over the 2005-2025 forecast horizon[6].

6 Corresponding spreadsheet files that provide projections and associated calculations for all years and months over the 2005-2025 time period are provided on the CD attached to this report.

Data used for forecasting differed in spatial aggregation from data used for building the models. Explanatory variable data for forecasts were either defined at or rolled up to WDPA level, whereas TAZ-level data were used for building the models. This change in geographic focus was based on the additivity assumption inherent in demand accounting (Sections 1.1 and 1.3.3) and was justified by apparent accuracy at the WDPA level (Section 1.5). TAZ-level forecasts would have been computationally expensive, and increasing error at these smaller geographic scales (arising from increases in heterogeneity between geographic units) would have rendered such forecasts highly imprecise. However, TAZ-level forecasting errors very likely would have offset one another upon aggregation to WDPAs, producing similar forecasts to those obtained by using WDPA-level explanatory and driver variable projections. These considerations were ample justification for applying the demand model to forecasts at WDPA levels rather than TAZ levels.

1.6.1 Model Input Projection Data

Demand forecasts for the Tampa Bay Water service area required WDPA-level projected values of driver and explanatory variables contained in the overall demand model. These projected driver and explanatory variable values were obtained for the 2005-2025 forecast horizon. Tables 1.4-1.13 and Figures 1.9-1.18 present assumed input values for each WDPA and totals/averages implied for the Tampa Bay Water service area.

TAZ-level observations and projections of permanent population, total dwelling units, and percent vacancy rates, as well as industrial, commercial, services, and total employment were obtained for 1999/2000 and forecast years 2005, 2015, and 2025. TAZ-level values for these variables were consolidated to form WDPA projections, and along with values of other projection inputs, were interpolated as necessary to derive consistent annual socioeconomic series spanning the 2005-2025 time period. Details for deriving these projections are presented in Appendix E.

1.6.1.1 Single-Family and Multi-family Housing Units

Housing unit projections are presented in Table 1.4. Published housing projection data (BEBR, 2001a) was presented only as total housing units (with no differentiation between single- and multi-family sectors). However, housing start projection data was available by single-family and multi-family sectors (BEBR, 2001a), as was TAZ-level projections of total number of housing units (Hillsborough County, 2002; Pasco County, 2001; Pinellas County, 2001).

Table 1.4
Historical and Projected Occupied Housing by Type[7]

	2002 (observed)	2005	2010	2015	2020	2025
Single-Family Units						
Tampa	104,806	108,674	115,069	121,429	127,016	132,667
Pinellas	102,903	105,089	107,156	108,415	109,212	109,722
St. Petersburg	79,613	80,212	80,953	81,491	81,916	82,241
New Port Richey	8,314	8,371	8,426	8,530	8,646	8,742
Pasco	58,929	61,818	66,168	70,486	74,839	78,941
NW Hillsborough	44,076	45,532	47,936	50,317	52,959	55,608
SC Hillsborough	62,758	66,927	73,829	80,638	87,755	94,876
TBW Total	461,399	476,623	499,538	521,306	542,342	562,798
Multi-Family Units						
Tampa	93,545	96,184	100,622	105,229	108,976	112,805
Pinellas	118,512	119,921	121,471	122,547	123,230	123,667
St. Petersburg	51,106	51,492	52,047	52,506	52,871	53,150
New Port Richey	5,403	5,410	5,422	5,452	5,485	5,513
Pasco	6,523	6,976	7,931	9,145	10,415	11,612
NW Hillsborough	6,160	7,153	8,822	10,546	12,318	14,113
SC Hillsborough	31,680	34,524	39,314	44,247	49,018	53,843
TBW Total	312,929	321,660	335,630	349,672	362,313	374,704
Total Housing Units						
Tampa	198,351	204,857	215,691	226,657	235,992	245,472
Pinellas	221,415	225,010	228,628	230,962	232,442	233,389
St. Petersburg	130,719	131,704	133,000	133,997	134,787	135,392
New Port Richey	13,717	13,781	13,848	13,981	14,131	14,255
Pasco	65,452	68,794	74,099	79,631	85,254	90,553
NW Hillsborough	50,236	52,685	56,758	60,863	65,277	69,722
SC Hillsborough	94,438	101,451	113,143	124,885	136,773	148,720
TBW Total	774,328	798,283	835,168	870,978	904,655	937,501

The ratio of annual average single-family households to multi-family units was derived from number of SF and MF accounts in each WDPA for Water Year 2001, the first year in the modeling database for which data were available for all WDPAs. In each WDPA, number of SF households was assumed to equal number of SF accounts. Number of

7 *Derived by applying BEBR SF and MF county-level new housing start rates (BEBR, 2001a) to estimated SF and MF housing units by WDPA in WY 2002.*

MF households in each WDPA was estimated by multiplying number of MF accounts by WDPA-level estimates of the number of multi-family units per account. Table 1.5 contains multi-family units per account estimates for each WDPA. These estimates were provided by member governments for a previous project (Ayres Associates, 1997). Growth in single-family and multi-family housing units over the forecast horizon was then derived from BEBR county-level projections of SF and MF new housing starts and MPO-projected total number of new housing units by TAZ, where each WDPA was assigned the growth rate from its corresponding county. Appendix E.1 explains derivation of SF and MF driver units in detail.

Table 1.5
Number of MF Units per Account by WDPA[8]

WDPA	# MF Units/account
Tampa	53
Pinellas	16
St Petersburg	11
New Port Richey	11
Pasco	2.5
NW Hillsborough	55
SC Hillsborough	55

Figures 1.9 and 1.10 present projected values of single-family and multi-family housing units, respectively. Total single-family housing units in the Tampa Bay Water service area were projected to grow at an average annual rate of 0.90%, while total multi-family units were projected to grow at a rate of 0.82%. Single-family housing growth was projected to be fastest in the WDPAs of Tampa (average 1.10% per year), Pasco (1.38%), Northwest Hillsborough (1.11%), and South Central Hillsborough (2.09%). Multi-family housing growth was projected to be less uniform than single-family growth; multi-family growth was concentrated in Pasco (3.32%), Northwest Hillsborough (4.87%), and South Central Hillsborough (2.80%), but was counteracted by almost no growth in Pinellas, St. Petersburg, and New Port Richey.

8 *Ayres Associates, 1997.*

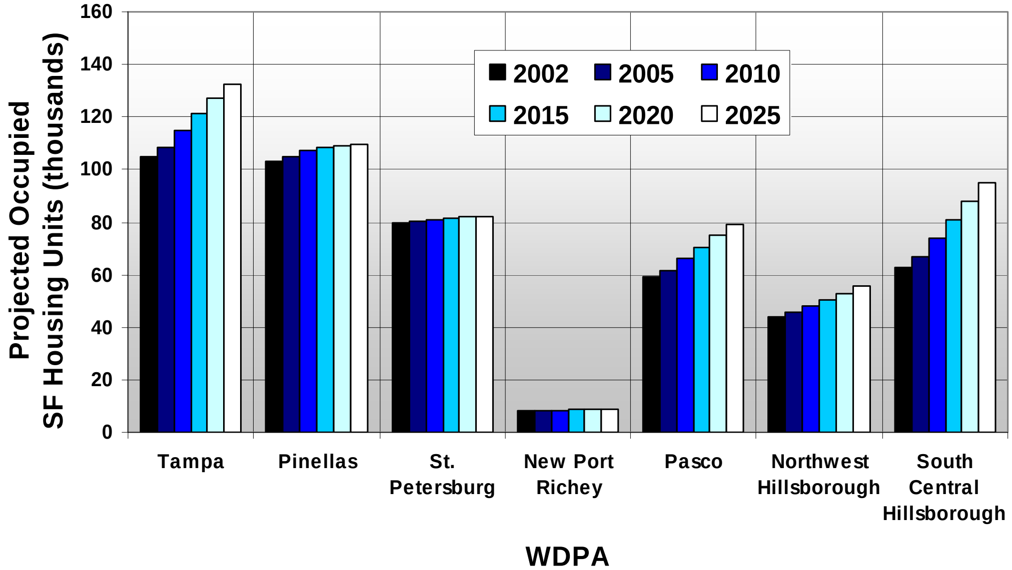


Figure 1.9 – Historical and Projected Single-Family Households by WDPA for Selected Forecast Years[9]

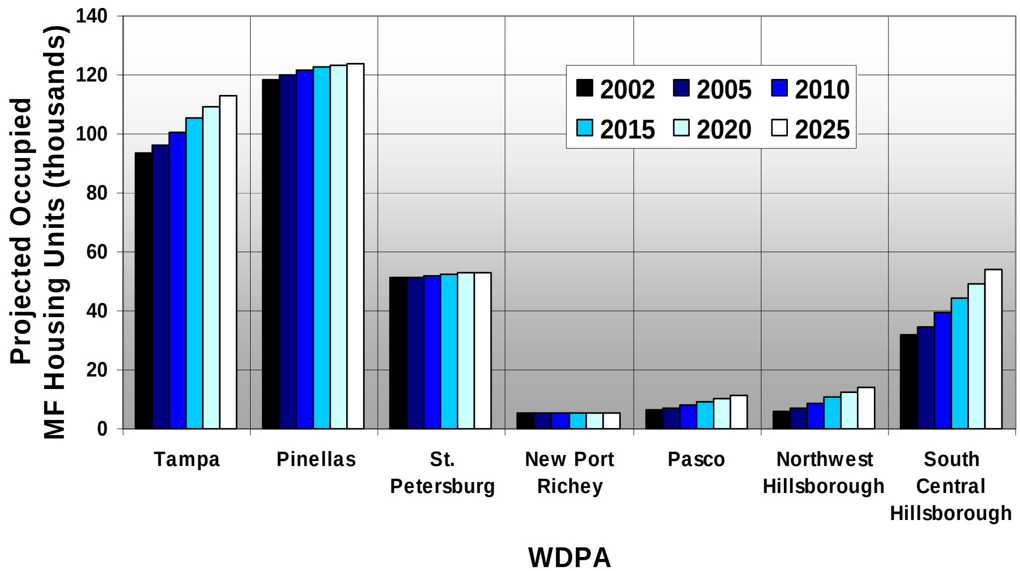


Figure 1.10 – Historical and Projected Multi-family Households by WDPA for Selected Forecast Years[10]

1.6.1.2 Median Per-Household Income

Base-year median household income was calculated by WDPA by averaging across TAZs from the modeling data[11]. Projections of household income were then derived by

9 Data corresponds to Table 1.4.

10 Data corresponds to Table 1.4.

11 Data obtained from Experian.

applying BEBR county-level year-to-year growth rates in real per capita income to base year household income of each WDPA (BEBR, 2001a). Appendix E.3 explains details of deriving income projections.

Projected household income is shown in Table 1.6 and displayed in Figure 1.11. Growth in income was projected to be fairly uniform across WDPAs, (2.73% - 3.15% per year over the forecast period). Household incomes were projected to be highest in Northwest and South Central Hillsborough and lowest in New Port Richey.

Table 1.6
Historical and Projected Median Household Income (1999 Dollars)[12]

	2002 (observed)	2005	2010	2015	2020	2025
Tampa	$61,341	$66,963	$75,319	$84,059	$93,264	$103,477
Pinellas	$47,551	$52,162	$59,826	$67,224	$74,586	$82,753
St. Petersburg	$48,398	$53,092	$60,893	$68,423	$75,915	$84,228
New Port Richey	$22,593	$24,832	$28,453	$32,255	$36,138	$40,490
Pasco	$47,122	$51,791	$59,344	$67,272	$75,372	$84,448
NW Hillsborough	$80,500	$87,878	$98,845	$110,315	$122,395	$135,797
SC Hillsborough	$69,841	$76,241	$85,756	$95,707	$106,188	$117,815

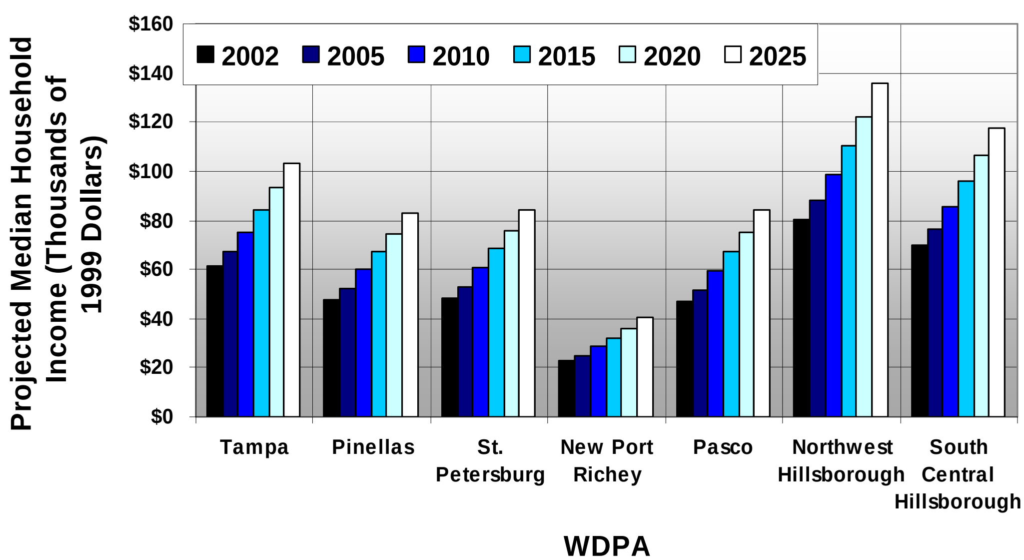


Figure 1.11 – Historical and Projected Median Household Income by WDPA for Selected Forecast Years[13]

12 *Derived by applying BEBR county-level growth rates in per capita income (BEBR, 2001a) to observed average per capita income by WDPA in WY 2002 (from the modeling database).*

13 *Data corresponds to Table 1.6.*

1.6.1.3 Single-Family Persons per Household

Single- and multi-family persons per household projections were derived from overall projections of persons per household. Total persons per household for each WDPA was first determined from BEBR projections of total population and total dwelling units for the county containing the WDPA (BEBR, 2001b). Single-family persons per household was then projected by assuming the ratio of single-family persons per household to multi-family persons per household was equal to the corresponding annual and WDPA-average ratio from Water Year 2002 (in the modeling database) and remained constant over time. Appendix E.4 details these calculations.

Projected single-family persons per household is given in Table 1.7 and displayed in Figure 1.12. While persons per household varied considerably between WDPAs (roughly 2-2.5 pph), minimal growth in this quantity was projected. Persons per household was highest in Tampa, Northwest and South Central Hillsborough, and Pinellas, and was lowest in New Port Richey and Pasco.

Table 1.7
Historical and Projected Single-Family Persons Per Household[14]

	2002 (observed)	2005	2010	2015	2020	2025
Tampa	2.47	2.47	2.46	2.45	2.45	2.44
Pinellas	2.51	2.51	2.51	2.51	2.51	2.51
St. Petersburg	2.29	2.29	2.28	2.28	2.28	2.28
New Port Richey	1.97	1.97	1.97	1.97	1.98	1.98
Pasco	1.96	1.97	1.99	2.01	2.03	2.04
NW Hillsborough	2.44	2.45	2.47	2.48	2.50	2.51
SC Hillsborough	2.56	2.56	2.55	2.55	2.55	2.55

14 *Derived from projections of total population (BEBR 2001b) and total dwelling units (See Section 1.6.1.1) aggregated by WDPA and year. Sectoral persons per household was then estimated from total persons per household using the ratio of single-family persons per household to multi-family persons per household from WY 2002 (in the modeling database).*

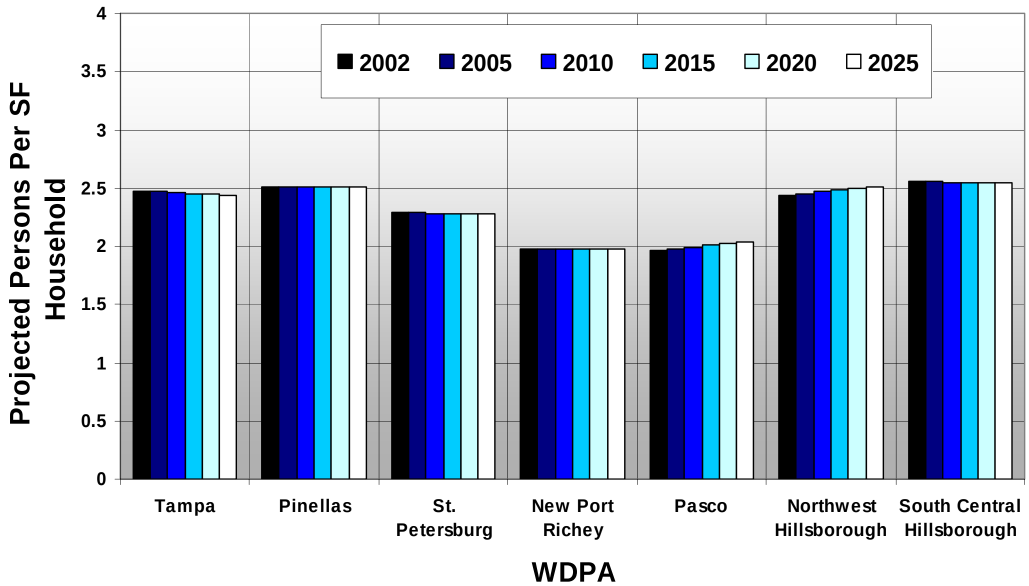


Figure 1.12 – Historical and Projected Single-Family Persons Per Household by WDPA for Selected Forecast Years[15]

1.6.1.4 Single- and Multi-Family Housing Density

Developed single-family and multi-family acreage per account for Water Year 2002 was derived from the modeling database. Average annual housing units by sector for each TAZ (from the modeling database) were divided by total developed single-family and multi-family acreage data in Water Year 2002 for each TAZ[16], producing average Water Year 2002 housing density (units per acre) for each TAZ. These TAZ estimates were averaged by WDPA. Projected housing density was assumed to remain at Water Year 2002 values for each WDPA in all forecast years. This calculation resulted in a rough measure of housing density, particularly for the multi-family sector, as number of multi-family dwelling units per account were difficult to estimate and highly variable. These projections indicate relative trends in housing density and should not be interpreted as direct quantitative estimates of density. Appendix E.5 explains these calculations.

Projected single- and multi-family housing density is shown in Table 1.8 and displayed in Figure 1.13. Density varied considerably across the region. Single-family and multi-family density was highest for St Petersburg. Northwest and South Central Hillsborough have the lowest density.

---

15 Data corresponds to Table 1.7.

16 Data provided by Experian.

Table 1.8
Projected Residential Housing Densities (Units/Acre)[17]

	SF	MF
Tampa	5.24	9.18
Pinellas	4.93	13.00
St. Petersburg	6.23	22.58
New Port Richey	5.07	12.19
Pasco	3.85	10.16
Northwest Hillsborough	3.66	4.02
South Central Hillsborough	2.31	3.46

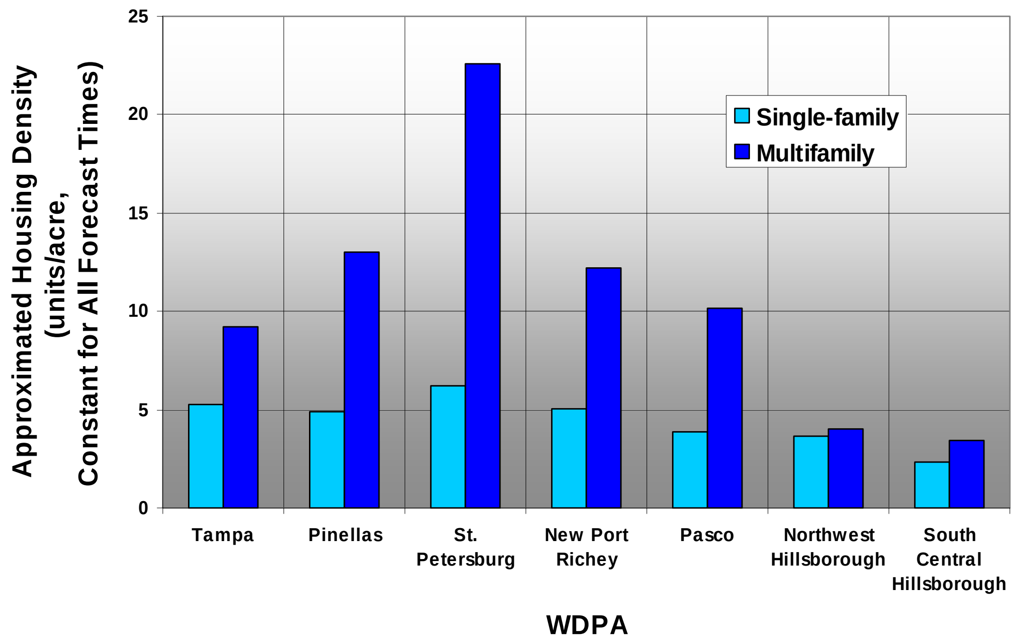


Figure 1.13 – Projected Single- and Multi-Family Housing Density by WDPA (Assumed Constant In All Forecast Years)[18]

1.6.1.5 Marginal Price of Water and Sewer

Projections of real marginal price of water and sewer were based on water and sewer rates obtained from Tampa Bay Water member governments for Water Year 2002. Future values of real marginal price were developed by applying the projected change in

17 *Average annual housing units by sector for each TAZ (from the modeling database) were divided by total developed single-family and multi-family acreage data in WY 2002 for each TAZ (from Experian). These TAZ estimates were then averaged by WDPA. Projected housing density was then assumed constant for each WDPA in all forecast years.*

18 *Data correspond to Table 1.8.*

nominal average unit cost of water to 2008. Expected changes in nominal unit costs were obtained from a consultant report (Black & Veatch, 2002). Projected marginal price of water and sewer for each WDPA was converted to real inflation-adjusted terms assuming a 3 percent annual rate of inflation through 2008. After 2008, real marginal price of water and sewer was assumed to increase by 4 percent annually, based on information from the Tampa Bay Water Finance Department.

Table 1.9 and Figure 1.14 show projected real marginal price. Real marginal price was projected to vary across the region, with Northwest and South Central Hillsborough having the highest values and New Port Richey having the lowest. Growth in real marginal price for WDPAs was projected at average rates of roughly 5-6%, with Pasco (4.96%) and Northwest and South Central Hillsborough (4.94%) having the lowest growth and New Port Richey (6.06%) having the highest.

Table 1.9
Historical and Projected Real Marginal Price of Water and Sewer (1999 Dollars)[19]

	2002 (observed)	2005	2010	2015	2020	2025
Tampa	$5.11	$6.32	$7.34	$8.94	$10.87	$13.23
Pinellas	$5.35	$7.05	$8.31	$10.11	$12.30	$14.97
St. Petersburg	$5.08	$6.67	$7.83	$9.52	$11.58	$14.09
New Port Richey	$3.69	$5.62	$6.90	$8.40	$10.22	$12.43
Pasco	$4.69	$6.41	$7.09	$8.62	$10.49	$12.77
NW Hillsborough	$6.55	$8.51	$9.39	$11.42	$13.90	$16.91
SC Hillsborough	$6.55	$8.51	$9.39	$11.42	$13.90	$16.91

19 Expected changes in nominal unit costs were obtained from a consultant report (Black & Veatch, 2002) and converted to real inflation-adjusted terms assuming a 3 percent annual rate of inflation through 2008. After 2008, real marginal price of water and sewer was assumed to increase by 4 percent annually (as suggested by the Tampa Bay Water Finance Department).

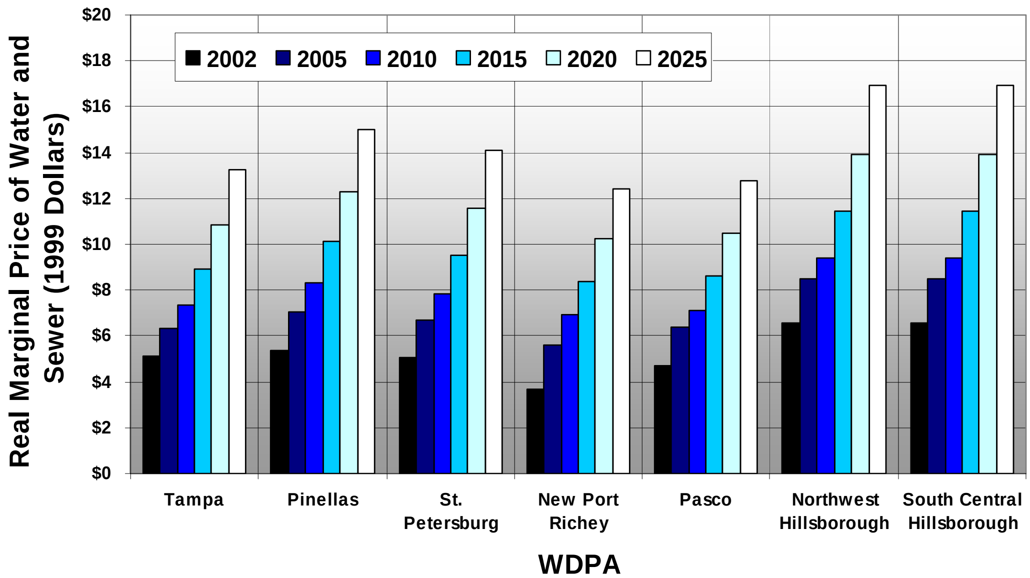


Figure 1.14 – Historical and Projected Real Marginal Price of Water And Sewer (1999 Dollars)[20]

1.6.1.6 Total Employment and Percent Employment in Commercial, Industrial, and Service Classifications

Projections of number of employees in commercial, industrial, and service entities were obtained by TAZ from Metropolitan Planning Organizations (Pasco County FL, 2001; Pinellas County FL, 2001; Hillsborough County FL, 2002)[21]. Total employment projections for a WDPA were thus derived by aggregating all projected employment values for each TAZ in the WDPA. Projected fractions of employment in each category and WDPA were then derived by summing employment values for that category for each TAZ in the WDPA, then dividing by total WDPA employment. Appendix E.2 explains employment-related calculations.

Projected total employment is shown in Table 1.10 and Figure 1.15. Number of employees was projected as highest in Tampa and lowest in New Port Richey. Growth in employment for the Tampa Bay Water service area was projected at an average annual rate of 1.59%, with an increase of more than 360,000 employees over the 2005-2025 time period. By far, the fastest growth in employment was projected in South Central Hillsborough (average 5.3% per year). This result agreed well with growth in number of

20 Data correspond to Table 1.9.

21 MPO employment projections for commercial and service categories were further segmented into “regional” and “local” subclasses. “Regional” employment refers to employees that work in a given TAZ but reside in another TAZ, while “local” employment refers to employees that work and live in the same TAZ. For developing demand forecasts, projected employment in service and commercial categories was taken as the total of regional and local employment for each category.

households which was highest in South Central Hillsborough for single-family and high for multi-family.

Table 1.10
Historical and Projected Total Employment[22]

	2002 (observed)	2005	2010	2015	2020	2025
Tampa	513,601	539,554	582,709	626,352	659,805	693,739
Pinellas	211,385	216,229	222,197	226,584	229,088	231,293
St. Petersburg	151,726	157,780	165,005	170,020	173,032	175,038
New Port Richey	14,134	14,225	14,386	14,531	14,661	14,854
Pasco	62,302	67,418	77,594	88,293	98,175	106,822
NW Hillsborough	54,659	58,493	64,876	71,327	80,360	89,477
SC Hillsborough	87,667	100,908	122,945	145,127	176,396	207,805
TBW Total	1,095,474	1,154,607	1,249,712	1,342,234	1,431,517	1,519,028

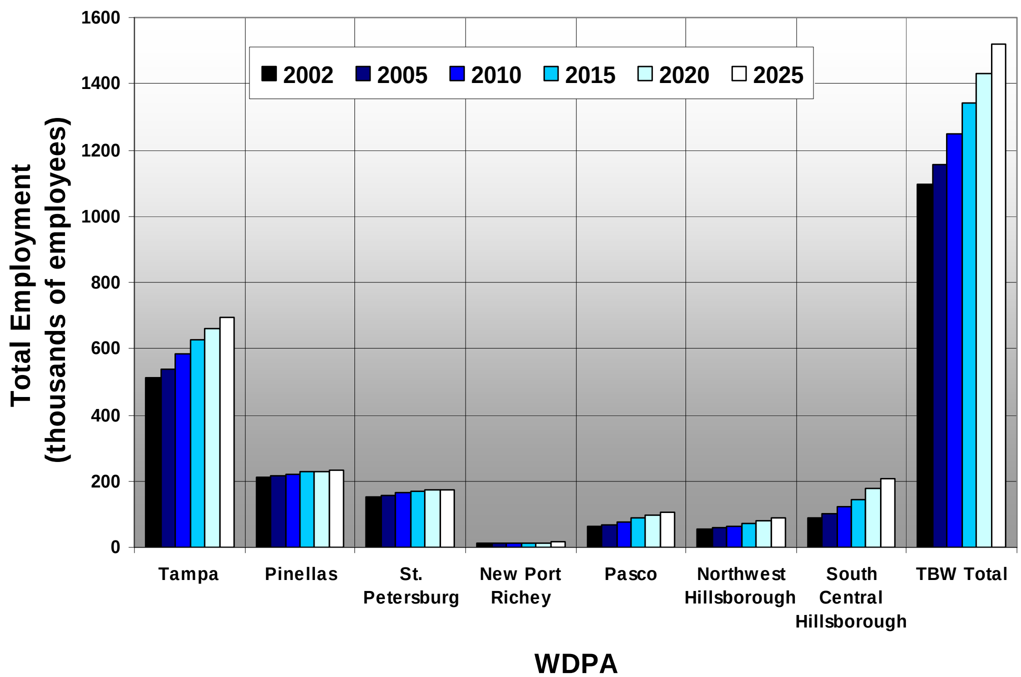


Figure 1.15 – Historical and Projected Total Employment by WDPA for Selected Forecast Years[23]

---

22 *Source: MPO Socioeconomic Forecast Reports (Pasco County FL, 2001; Pinellas County FL, 2001; Hillsborough County FL, 2002)*

23 *Data correspond to Table 1.10.*

Projected percent employment in commercial, industrial, and service categories is shown in Table 1.11 and Figures 1.16-1.18. Services composed the largest portion of total employment, followed by commercial, then industrial. Additionally, percent of total employment in service entities was projected to increase in all cases, while percent employment in industrial and commercial settings was generally projected to decrease. This trend was strongest in Northwest and South Central Hillsborough WDPAs, which also showed the strongest projected growth in total employment.

Table 1.11
Historical and Projected Distribution of Total Employment[24]

Employment Percentages	2002 (observed)	2005	2010	2015	2020	2025
Tampa						
Commercial	16.46%	16.47%	16.49%	16.52%	16.32%	16.16%
Industrial	19.17%	19.03%	18.81%	18.62%	18.39%	18.17%
Services	64.37%	64.50%	64.70%	64.86%	65.30%	65.68%
Pinellas						
Commercial	24.77%	24.69%	24.65%	24.74%	24.95%	25.13%
Industrial	17.99%	17.82%	17.53%	17.37%	17.25%	17.19%
Services	57.24%	57.50%	57.82%	57.89%	57.80%	57.69%
St. Petersburg						
Commercial	23.18%	22.89%	22.59%	22.48%	22.57%	22.76%
Industrial	11.32%	11.18%	11.22%	11.26%	11.22%	11.26%
Services	65.51%	65.93%	66.20%	66.27%	66.21%	65.99%
New Port Richey						
Commercial	24.44%	24.32%	24.13%	23.93%	23.72%	23.61%
Industrial	10.42%	10.36%	10.29%	10.22%	10.14%	10.01%
Services	65.13%	65.32%	65.57%	65.85%	66.14%	66.38%
Pasco						
Commercial	32.77%	32.21%	31.81%	31.80%	31.54%	30.96%
Industrial	14.44%	13.76%	13.61%	13.19%	13.06%	12.85%
Services	52.79%	54.03%	54.57%	55.01%	55.40%	56.19%
Northwest Hillsborough						
Commercial	29.68%	29.17%	28.46%	27.89%	26.21%	24.90%
Industrial	19.08%	18.64%	18.03%	17.52%	15.93%	14.66%
Services	51.24%	52.19%	53.51%	54.59%	57.86%	60.45%

24 Percentages derived from employment data by category in MPO Socioeconomic Forecast Reports (Pasco County FL, 2001; Pinellas County FL, 2001; Hillsborough County FL, 2002)

Table 1.11 (continued)
Historical and Projected Distribution of Total Employment[24]

Employment Percentages	2002 (observed)	2005	2010	2015	2020	2025
South Central Hillsborough						
Commercial	30.12%	30.04%	29.95%	29.89%	27.75%	26.26%
Industrial	19.01%	18.48%	17.84%	17.40%	15.72%	14.54%
Services	50.86%	51.48%	52.21%	52.71%	56.54%	59.20%
Tampa Bay Water Total						
Commercial	21.78%	21.71%	21.71%	21.78%	21.52%	21.28%
Industrial	17.46%	17.30%	17.07%	16.89%	16.46%	16.10%
Services	60.77%	60.99%	61.22%	61.33%	62.01%	62.62%

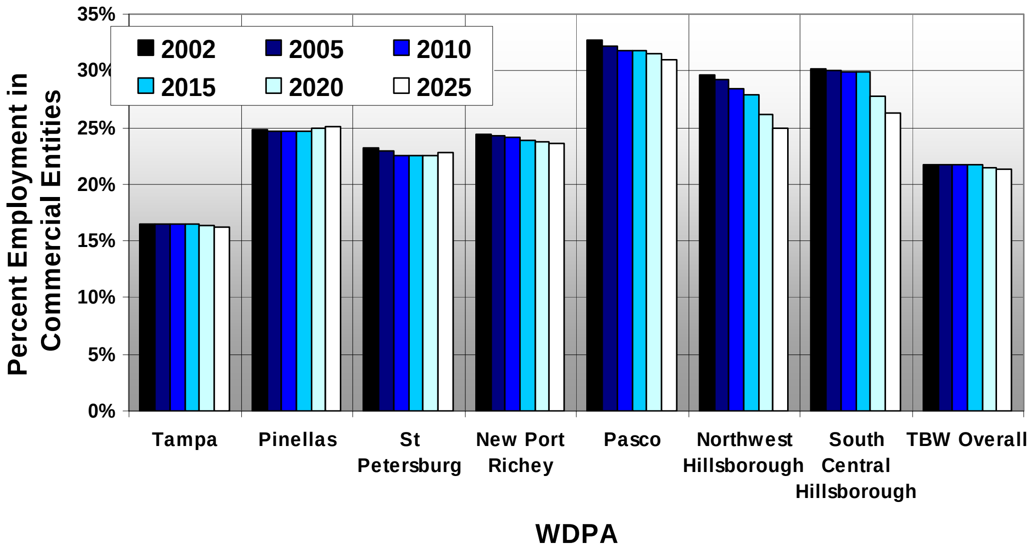


Figure 1.16 – Historical and Projected Percent Employment in Commercial Entities by WDPA for Selected Forecast Years[25]

25 Data correspond to Table 1.11.

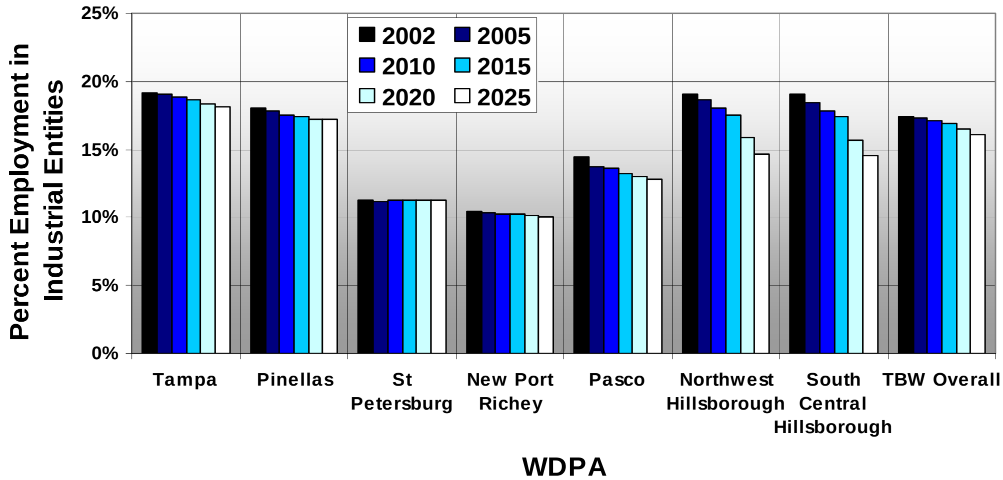


Figure 1.17 – Historical and Projected Percent Employment in Industrial Entities by WDPA for Selected Forecast Years[26]

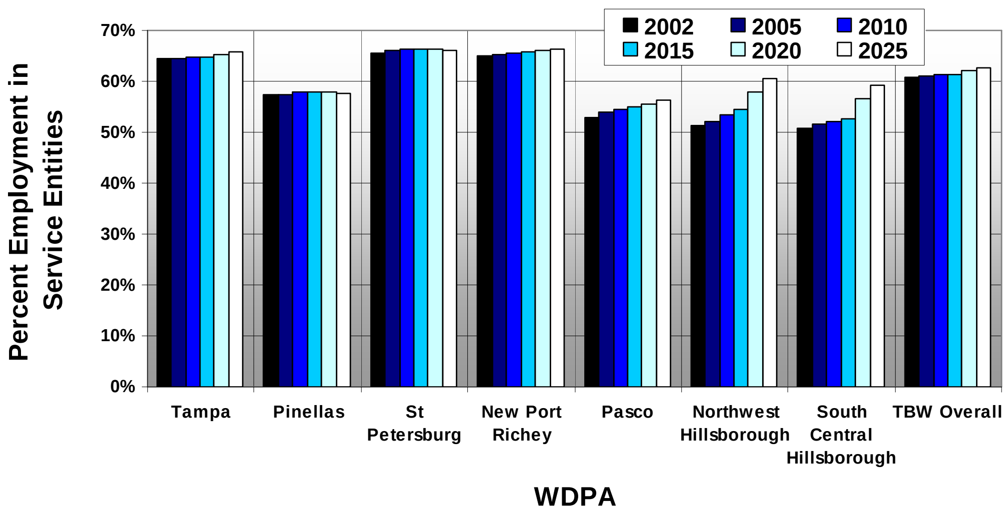


Figure 1.18 – Historical and Projected Percent Employment in Service Entities by WDPA for Selected Forecast Years[27]

1.6.1.7 Fraction of Accounts Accepting Reclaimed Water

Future reclaimed water service projections were not available for most member utilities. It was assumed that percentage of residential units connected to reclaimed water service would remain at Water Year 2002 levels (listed in Table 1.12, calculated from billing data) over the entire forecast period. For example, 5.41 percent of Pinellas County single-family customers were connected to reclaimed water service by the end of Water Year 2002. It was therefore assumed that 5.41 percent of single-family customers would

26 Data correspond to Table 1.11.

27 Data correspond to Table 1.11.

be connected to reclaimed water service throughout the forecast period. This assumption implied that, as additional customers join the single-family sector of Pinellas County, 5.41 percent of those new customers will accept reclaimed water service. As members develop more detailed reclaimed water implementation plans, this assumption can be updated.

Table 1.12
Assumptions for Percent of Accounts Reclaimed[28]

	Sector		
WDPA	**Single-Family**	**Multi-Family**	**Non-Residential**[29]
Pinellas	5.41%	2.37%	4.71%
St. Petersburg	12.03%	2.76%	4.95%
Pasco	9.39%	4.27%	0%
New Port Richey	0.29%	0.20%	4.34%
NW Hillsborough	20.24%	0%	0%
SC Hillsborough	1.42%	0%	0.7%
Tampa[30]	5%	1%	0.5%

1.6.1.8 Temperature, Rainfall, and Number of 0.01” and 1” Rainy Days

Weather conditions were assumed to follow historic long-term normal seasonal patterns for all forecast years. Each weather variable was defined by weighting daily/monthly weather observations (from weather stations listed in Section 1.2) by distance to each station from the geographical centroid of each WDPA. The distance weighting procedure was identical to that illustrated in Figure 1.5, except that distances between WDPA centroids and weather stations were used for weights. Long-term normal weather values were then produced by aggregating weather values for each WDPA and calendar month. These values are listed in Table 1.13. It should be noted that it is not strictly necessary to assume long-term normal weather for forecasts; weather values can be adjusted in forecasts to reflect dry or wet conditions in any given year. Appendix E.6 explains the mathematics of deriving long-term normal WDPA weather.

28 *Projections of percent of accounts with reclaimed connections were obtained by calculating percent values for WY 2002 from billing data, then assuming these percentages remained constant over all forecast years.*

29 *Note that the fraction of reclaimed accounts in the non-residential sector was not included as a model variable due to the inability to measure rational and statistically significant impacts.*

30 *The City of Tampa had no reclaimed water connections at the time of this analysis, but several large reclaimed projects were in construction and scheduled to be operational by 2004. Therefore, for demand forecasting purposes, FR for the City of Tampa was based on the number of reclaimed connections in the 5-year conservation plan submitted to Tampa Bay Water in 2002. These values were 5%, 1%, and 0.5% for the SF, MF, and NR sectors respectively.*

Table 1.13
Long-Term Weather Averages[31]

Average Max. Temperature	Tampa	Pinellas	St. Petersburg	N. Port Richey	Pasco	N.W. Hillsborough	S.C. Hillsborough
January	70.84	70.73	69.89	71.22	71.50	70.84	71.40
February	72.66	72.46	71.51	72.94	73.55	72.63	73.30
March	77.50	77.23	76.20	77.66	78.40	77.46	78.08
April	81.72	81.45	80.61	81.77	82.59	81.67	82.21
May	87.45	87.06	86.27	87.16	88.13	87.35	87.72
June	90.15	90.07	89.55	90.35	90.84	90.18	90.36
July	90.92	91.04	90.64	91.47	91.73	91.01	91.22
August	90.83	90.95	90.22	91.55	91.59	90.97	91.01
September	89.46	89.49	88.61	90.09	90.12	89.57	89.63
October	84.54	84.52	83.62	85.05	85.08	84.62	84.72
November	78.40	78.35	77.23	78.99	78.94	78.48	78.64
December	72.87	72.86	71.94	73.47	73.53	72.93	73.32
Annual Average	**82.28**	**82.18**	**81.36**	**82.64**	**83.00**	**82.31**	**82.63**

Total Rainfall	Tampa	Pinellas	St. Petersburg	N. Port Richey	Pasco	N.W. Hillsborough	S.C. Hillsborough
January	2.49	2.75	2.75	3.09	3.06	2.61	2.67
February	2.80	2.92	2.87	3.10	3.14	2.86	2.93
March	3.08	3.36	3.28	3.75	3.70	3.22	3.28
April	1.90	1.92	1.92	1.97	2.12	1.91	2.03
May	3.01	2.96	2.81	3.06	3.41	3.00	3.21
June	5.88	5.88	6.08	5.87	6.48	5.83	6.45
July	6.75	6.83	6.73	7.06	7.28	6.79	7.09
August	7.72	8.05	8.24	8.32	7.82	7.84	7.81
September	6.68	7.04	7.56	7.16	6.80	6.79	6.80
October	2.43	2.75	2.64	3.19	2.79	2.59	2.48
November	1.83	2.07	2.04	2.40	2.30	1.94	2.01
December	2.42	2.62	2.59	2.88	2.67	2.51	2.52
Annual Total	**46.99**	**49.15**	**49.51**	**51.85**	**51.57**	**47.89**	**49.28**

31 Data derived by weighting daily/monthly weather observations (from weather stations listed in Section 1.2) by distance to each station from the geographical centroid of each WDPA, then aggregating weighted average weather values for each WDPA and calendar month. The weighting procedure was identical to that illustrated in Figure 1.5, except that WDPA centroids were used instead of TAZ centroids.

Table 1.13 (continued)
Long-Term Weather Averages[31]

Number of Rainy Days with >0.01" of Precipitation	Tampa	Pinellas	St. Petersburg	N. Port Richey	Pasco	N.W. Hillsborough	S.C. Hillsborough
January	7.45	8.00	6.50	9.44	8.51	7.90	7.35
February	6.66	6.95	6.04	7.82	7.44	6.91	6.69
March	6.71	7.10	6.14	8.11	7.63	7.02	6.67
April	4.89	4.95	4.30	5.45	5.52	4.99	5.07
May	6.25	5.90	5.05	6.02	6.62	6.18	6.31
June	11.81	11.22	10.50	11.08	12.11	11.58	12.21
July	15.14	14.87	13.78	15.28	15.70	15.11	15.49
August	15.92	15.43	14.28	15.56	16.02	15.82	15.93
September	12.46	12.33	12.02	12.39	12.58	12.43	12.52
October	6.60	6.70	5.93	7.24	6.92	6.74	6.56
November	5.77	6.07	5.04	7.04	6.67	6.03	5.90
December	6.25	6.46	5.42	7.31	7.09	6.48	6.24
Annual Total	**105.91**	**105.98**	**95.00**	**112.74**	**112.81**	**107.19**	**106.94**

Number of Rainy Days with >1.0" of Precipitation	Tampa	Pinellas	St. Petersburg	N. Port Richey	Pasco	N.W. Hillsborough	S.C. Hillsborough
January	0.61	0.74	0.70	0.92	0.91	0.67	0.69
February	0.80	0.81	0.86	0.81	0.89	0.80	0.85
March	0.96	1.12	1.03	1.36	1.27	1.04	1.06
April	0.48	0.49	0.47	0.52	0.66	0.49	0.53
May	0.74	0.80	0.80	0.89	0.95	0.76	0.85
June	1.76	1.74	1.69	1.78	1.97	1.75	1.92
July	1.94	2.00	1.97	2.13	2.25	1.96	2.19
August	2.46	2.60	2.85	2.62	2.44	2.50	2.43
September	2.21	2.20	2.33	2.08	2.02	2.20	2.11
October	0.83	0.86	0.71	0.97	0.90	0.86	0.80
November	0.41	0.48	0.59	0.50	0.53	0.43	0.51
December	0.67	0.69	0.64	0.75	0.73	0.68	0.71
Annual Total	**13.87**	**14.53**	**14.64**	**15.33**	**15.52**	**14.14**	**14.65**

There was little difference in projected maximum daily temperatures, total rainfall, or number of 0.01"/1" rainy days for WDPAs. Pasco had the highest long-term average maximum daily temperatures, and St. Petersburg had the lowest, perhaps due to relative proximity to large bodies of water. Pasco also had the highest long-term average rainfall except for some wet months, but even then it was among the wettest WDPAs. Tampa was the driest WDPA in most cases. Ordering of WDPAs by number of rainy days was less clear. New Port Richey and Pasco had the highest average number of 0.01" rainy days in many cases and were among the highest in all cases, while St. Petersburg had the fewest number of 0.01" rainy days in all cases, but there was no discernable ordering of WDPAs by number of 1" rainy days.

1.6.2 Forecast Results

Table 1.14 and Figure 1.19 provide a breakdown of historical and point forecast demand by WDPA and sector as a whole in five-year increments over the 2005-2025 time period. Forecast results were derived by applying the sectoral per-unit demand models and accounting framework to projections of explanatory and driver variables listed in Tables 1.4-1.13. In the subsections that follow, the results presented in Table 1.14 are examined in greater detail.

Table 1.14
Actual and Point-Forecasted Water Use by WDPA and Sector

		2002 (obs)	2005	2010	2015	2020	2025	% Change 2005-2025	Avg Ann % Change
Pinellas	Single-family	22.76	22.00	22.32	22.18	21.87	21.51	-2.26%	-0.09%
	Multi-family	10.95	11.47	12.23	12.88	13.46	14.04	22.36%	0.81%
	Non-residential	8.82	9.06	9.48	9.81	10.05	10.28	13.45%	0.51%
	Total Retail	42.53	42.54	44.03	44.87	45.38	45.83	7.73%	0.30%
	Total Gross Demand	68.86	68.88	71.29	72.66	73.48	74.20	7.73%	0.30%
St. Petersburg	Single-family	12.83	12.12	12.18	12.04	11.84	11.64	-4.03%	-0.16%
	Multi-family	7.40	7.71	8.20	8.63	9.04	9.44	22.47%	0.81%
	Non-residential	6.76	7.00	7.44	7.77	8.01	8.21	17.24%	0.64%
	Total Retail	26.98	26.83	27.82	28.45	28.89	29.28	9.13%	0.35%
	Total Gross Demand	32.11	31.94	33.11	33.86	34.39	34.86	9.13%	0.35%
New Port Richey	Single-family	1.50	1.40	1.39	1.39	1.38	1.37	-2.33%	-0.09%
	Multi-family	0.56	0.59	0.62	0.65	0.68	0.72	22.15%	0.80%
	Non-residential	0.68	0.68	0.70	0.72	0.74	0.76	10.85%	0.41%
	Total Retail	2.74	2.67	2.71	2.76	2.80	2.84	6.40%	0.25%
	Total Gross Demand	3.47	3.38	3.43	3.49	3.54	3.60	6.40%	0.25%
Pasco	Single-family	12.56	12.50	13.61	14.35	15.02	15.62	24.89%	0.89%
	Multi-family	1.40	1.56	1.86	2.25	2.67	3.11	99.50%	2.80%
	Non-residential	2.25	2.45	2.87	3.32	3.75	4.14	68.79%	2.12%
	Total Retail	16.21	16.51	18.35	19.92	21.44	22.86	38.45%	1.31%
	Total Gross Demand	19.04	19.40	21.55	23.40	25.19	26.86	38.45%	1.31%
Tampa	Single-family	25.57	24.65	25.90	26.75	27.36	27.94	13.34%	0.50%
	Multi-family	12.70	13.46	14.71	16.02	17.24	18.54	37.80%	1.29%
	Non-residential	20.18	21.40	23.46	25.58	27.30	29.09	35.95%	1.24%
	Total Retail	58.46	59.51	64.06	68.34	71.90	75.57	27.00%	0.96%
	Total Gross Demand	72.91	74.22	79.90	85.24	89.68	94.26	27.00%	0.96%
NW Hills	Single-family	11.26	10.98	11.68	12.06	12.47	12.85	17.08%	0.63%
	Multi-family	2.51	3.01	3.88	4.83	5.87	6.99	131.82%	3.42%
	Non-residential	1.95	2.11	2.38	2.66	3.05	3.46	63.99%	2.00%
	Total Retail	15.72	16.10	17.94	19.56	21.39	23.30	44.71%	1.49%
	Total Gross Demand	16.61	17.01	18.96	20.67	22.60	24.62	44.71%	1.49%

Table 1.14
Actual and Point-Forecasted Water Use by WDPA and Sector

		2002 (obs)	2005	2010	2015	2020	2025	% Change 2005-2025	Avg Ann % Change
SC Hills	Single-family	15.70	15.85	17.58	18.82	20.06	21.23	33.96%	1.18%
	Multi-family	3.25	3.66	4.35	5.10	5.87	6.70	83.25%	2.45%
	Non-residential	3.84	4.47	5.55	6.65	8.25	9.88	120.89%	3.22%
	Total Retail	22.79	23.98	27.48	30.57	34.18	37.82	57.70%	1.84%
	Total Gross Demand	25.09	26.40	30.25	33.66	37.62	41.63	57.70%	1.84%
TBW Overall	Single-family	102.17	99.51	104.67	107.59	110.00	112.15	12.70%	0.48%
	Multi-family	38.77	41.45	45.84	50.37	54.83	59.54	43.62%	1.46%
	Non-residential	44.48	47.18	51.88	56.51	61.16	65.82	39.51%	1.34%
	Total Retail	185.42	188.15	202.40	214.47	225.98	237.51	26.24%	0.94%
	Total Gross Demand	238.09	241.23	258.50	272.97	286.51	300.02	24.37%	0.88%

Figure 1.19 provides a line plot of projected water use over the 2005-2025 time period. Total Tampa Bay Water regional (gross) demand was projected to grow at an annual average rate of 0.88 percent over the forecast horizon to a value of about 300 MGD in 2025. The greatest annual average changes in demand were projected to occur in multi-family and non-residential sectors (1.46% and 1.34% per year, respectively), though demand in the single-family sector was projected to remain the largest component of retail use.

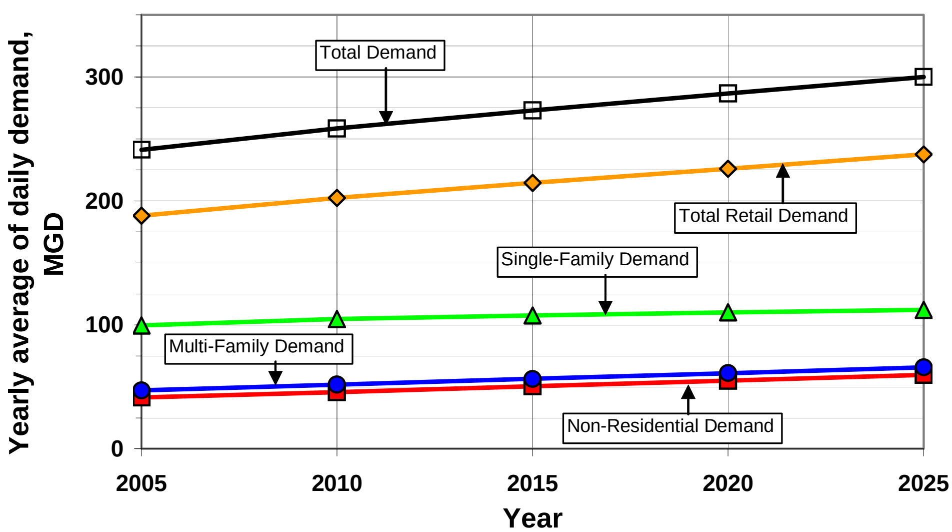


Figure 1.19 – Point Forecast of Tampa Bay Water Demand

Table 1.14 shows that total multi-family demand was forecasted to grow at a higher rate than total single-family demand and SF demand was forecasted to decrease slightly over time in some WDPAs. As will be explained in the following subsections, these trends occurred because marginal price of water was a significant factor in the single-family sectoral model but was statistically insignificant in the multi-family equation. Projected increases in real marginal price produced reducing effects on forecasted single-family per-household demand, but had no effect on multi-family per-unit demand. The single-family marginal price effect more than counteracted increasing effects on demand from projected income increases, leading to reductions over time in single-family per-household forecasted demand in all WDPAs. While number of SF households was projected to increase, producing an increasing trend in total SF demand, forecasted per-unit SF demand decreases counteracted growth in number of households and led to sluggish growth or slight decreases in WDPA-level forecasted single-family total demand.

1.6.2.1 Total Forecasted Demand by WDPA

Table 1.15 and Figure 1.20 show historical and forecasted total demand by WDPA. The highest rate of growth was forecasted to occur in Tampa and South Central Hillsborough, while the lowest rate of growth was in New Port Richey. Projected growth in water demands for the City of Tampa WDPA roughly mimicked projected growth for Tampa Bay Water, with an average annual increase of about 1 percent over the forecast horizon. Total demand projections for Pinellas, St. Petersburg, and New Port Richey WDPAs were relatively flat. However, these WDPAs showed decreases in single-family demands and growth in multi-family and non-residential demands (Table 1.14). This result is rationalized by the aforementioned effect of real marginal price on single-family per household demand. In some WDPAs, per-unit decreases in single-family demand were strong enough to cancel or even reverse total single-family demand growth from increasing number of households. This will be discussed in more detail in the next sections.

Table 1.15
Historical and Forecasted Annual Average Total Demand by WDPA for Selected Forecast Years[32]

	2002 (observed)	2005	2010	2015	2020	2025
Tampa	72.91	74.22	79.9	85.24	89.68	94.26
Pinellas	68.86	68.88	71.29	72.66	73.48	74.2
St. Petersburg	32.11	31.94	33.11	33.86	34.39	34.86
New Port Richey	3.47	3.38	3.43	3.49	3.54	3.6
Pasco	19.04	19.4	21.55	23.4	25.19	26.86
NW Hillsborough	16.61	17.01	18.96	20.67	22.6	24.62
SC Hillsborough	25.09	26.4	30.25	33.66	37.62	41.63
TBW Total	238.09	241.23	258.5	272.97	286.51	300.02

32 Data correspond to Table 1.14.

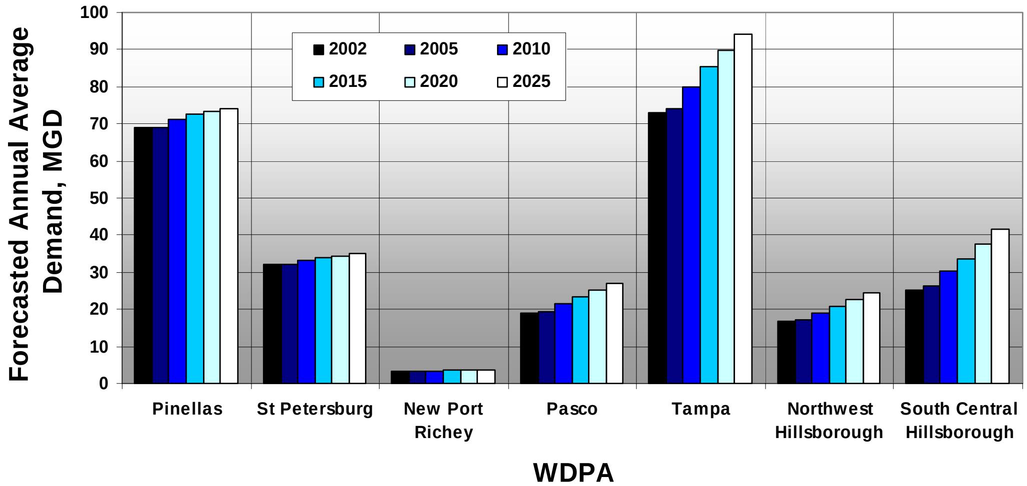


Figure 1.20 – Historical and Forecasted Annual Average Total Demand by WDPA for Selected Forecast Years[33]

Total water demands for Pasco, Northwest Hillsborough, and South Central Hillsborough WDPAs were projected to grow at average annual rates exceeding 1 percent. South Central Hillsborough demand was forecasted to grow sufficiently fast as to overtake St. Petersburg's demand by 2015.

1.6.2.2 Regional Per-Unit Forecasted Demand

Forecasted rates of growth in water demand components differed considerably among WDPAs, due to variations in projected housing, employment, and other socioeconomic factors between WDPAs. Figure 1.21 and Table 1.16 show forecasted average annual per-unit demand for the Tampa Bay Water Service Area. Single-family use per household was forecasted to decrease, while multi-family use per dwelling unit was projected to increase. Non-residential use per employee remained essentially constant.

Table 1.16
Observed and Forecasted Annual Average Per-Unit Demand in the Tampa Bay Water Service Area (Selected Forecast Years)

Year	SF Demand Per Household	MF Demand Per Dwelling Unit	NR Demand Per Employee
2002 (observed)	223.02	124.21	41.15
2005	208.71	128.88	40.86
2010	209.45	136.59	41.52
2015	206.31	144.04	42.11
2020	202.75	151.34	42.72
2025	199.20	158.89	43.33

33 Data correspond to Table 1.15

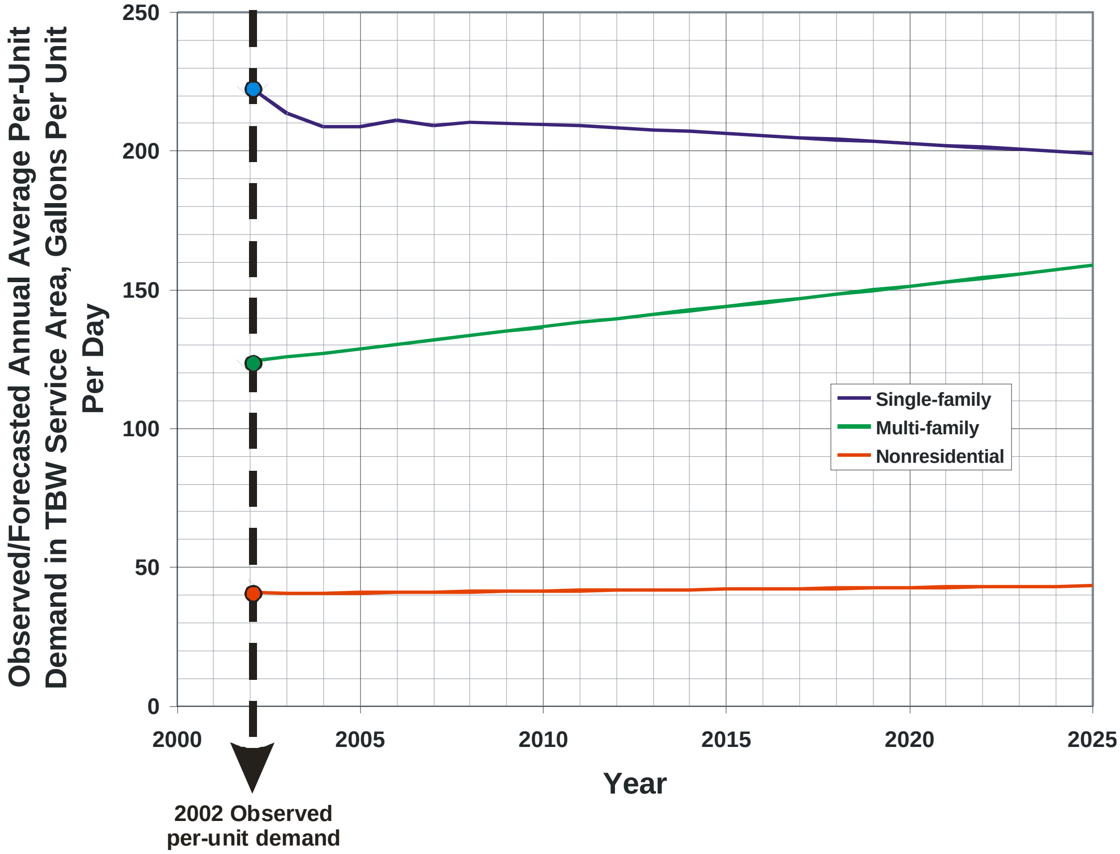


Figure 1.21 – Observed and Forecasted Annual Average Per-Unit Demand in the Tampa Bay Water Service Area[34]

Single- and multi-family regional per-unit demand trends. Single-family regional per-household demand decreased primarily because growth in real marginal price of water is projected to outpace growth in income. While all other variables in the single-family model changed little from year to year, income and real marginal price across the region were projected to grow at average rates of roughly 2.8% and 5.3% (non-compounded) per year, respectively. Single-family elasticities for income and price were 0.26 and -0.24, implying that equal percent increases in income and price had opposite and nearly equal influences on single-family demand per household. Thus, regional single family demand per household was forecasted to decrease because growth in real marginal price was projected to increase at nearly twice the rate of growth in income. This trend was considered reasonable, given projections of income and marginal price and implications of these variables on single-family per-unit demand inherent in the observed modeling data.

34 Data correspond to Table 1.16 and contain additional values for forecast years between Table 1.16 listings.

In contrast to the single-family case, multi-family regional per-unit demand was forecasted to grow because there was no influence of real marginal price to offset projected regional increases in income. Multi-family elasticity for income was 0.37, implying that increases in income led to increases in demand per dwelling unit. All other variables in the multi-family per-unit model changed little from year to year. As a result, regional per-unit multi-family demand grew with projected growth in income.

It was considered reasonable for significance of price in multi-family demand to be lower than in single-family demand. While most single-family units are generally billed on an individual basis, the same is not true for most multi-family dwellings; many have water costs included in rent and distributed evenly among all units in a complex. Thus, people living in multi-family units may not directly realize the personal costs of increased water use, dampening the effect of real marginal price on use.

Any indirect effects of increasing water price on water use through rental rates might only be evident over many years, for two reasons.

1. Apartment tenants might respond to rental rate increases by moving to apartments with lower rental rates rather than by changing water consumption behavior. Were this the case, water consumption behavioral responses would be aliased with consumers' housing selections. For marginal price to influence consumption through rental rates, it would be necessary for the entire rental market to respond to increasing water price and for public awareness of the effect of water use on rental rates to be heightened.

2. In many cases, rental rates are adjusted only annually or semi-annually at the end of leasing periods, spreading any signal of water price increases through rental rates over several years.

In light of these two considerations, a small marginal price influence on multi-family demand might not be detected given the short time span covered by the modeling database. More than the four years of water billing data available for this modeling study might be required to identify an effect of price on multi-family use, if such an effect exists. This requirement is an impetus for sustaining data collection and periodically updating the demand forecasting model based on new data. Furthermore, as price increases, it may become more economically viable for multi-family complexes to sub-meter water use for individual units. Wide-scale sub-metering would represent a new multi-family price signal and would change the relationship between price and multi-family per unit demand. Only through sustained data collection and model maintenance could such an emerging trend be identified.

Non-residential per-unit demand trend. Non-residential demand per employee was forecasted to increase minimally, only averaging 0.32% per year over the 2005-2025 forecast period. This consistency was produced by small counteracting effects in four non-residential model variables: income and fraction of employment in industrial, commercial, and service entities. Elasticities for these four variables were 0.12, 0.23, 1.01, and 1.19, respectively. Average projected annual percent changes in these variables for the Tampa Bay Water service area were 2.8%, -0.35%, -0.1%, and 0.13%, respectively. All other variables in the non-residential model changed little from year to year. Annual percent change in total Tampa Bay Water demand arising from these variables was estimated (by multiplying elasticities by corresponding percent changes and summing) as 0.27%. This estimated change was appropriately low and was consistent with forecasted non-residential per-employee demand growth.

Effects of per-unit demand on total demand. Despite these varied characteristics in per-unit demand, all sectors exhibited regional growth in forecasted total demand, primarily because of across-the-board projected growth in driver variable values. Slow growth in forecasted single-family total demand was attributable to decreases in forecasted per-household demand, which offset growth in total demand from increases in the number of households. By comparison, increases in multi-family per-dwelling use reinforced growth in total demand from increases in number of multi-family units. Thus, relative rates of growth for single-family and multi-family total demand can be rationalized by interaction of forecast trends in per-unit demand and number of units.

1.6.2.3 Per-Unit Forecasted Demand by WDPA

Figure 1.22 and Table 1.17 illustrate per-household historical and forecasted demand by WDPA for the single-family sector. Tampa, Northwest Hillsborough, and South Central Hillsborough had the highest forecasted single-family demand per household, followed closely by Pinellas and Pasco. St. Petersburg and New Port Richey had the lowest single-family per-household demand forecasts. This ordering agreed with Water Year 2002 observed single-family per-household demand.

Table 1.17
Actual and Forecasted Annual Average Single-Family Per-Household Demand By WDPA for Selected Forecast Years

	2002 (observed)	2005	2010	2015	2020	2025
Pinellas	223.92	209.32	208.25	204.51	200.18	195.94
St. Petersburg	161.65	151.09	150.42	147.68	144.51	141.42
New Port Richey	183.70	167.63	165.13	162.65	159.75	156.79
Pasco	213.53	202.21	205.63	203.47	200.63	197.76
Tampa	245.86	226.75	224.96	220.20	215.32	210.52
NW Hillsborough	258.15	240.99	243.55	239.63	235.35	231.03
SC Hillsborough	250.67	236.70	238.10	233.34	228.47	223.67

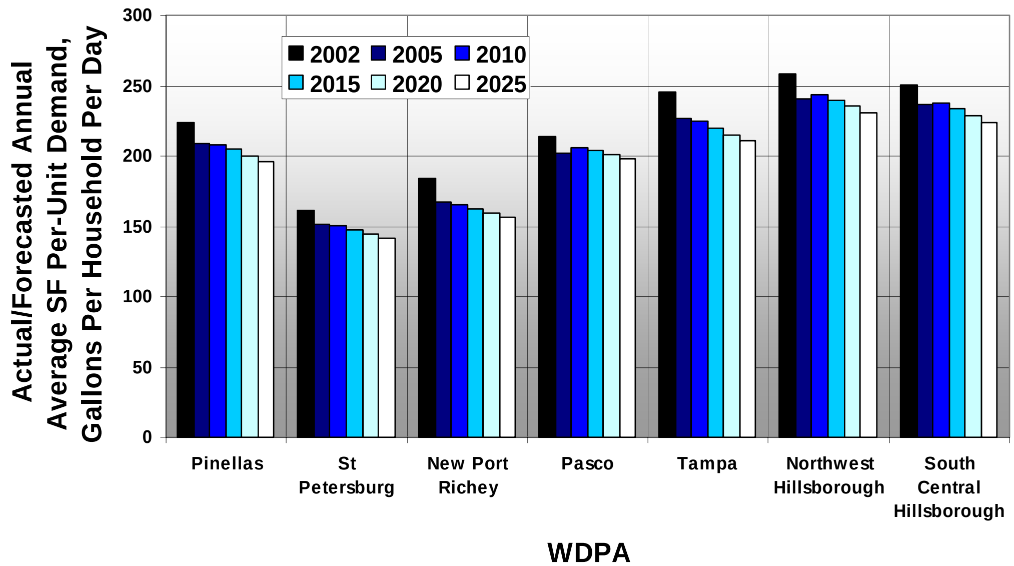


Figure 1.22 – Actual and Forecasted Annual Average Single-Family Per-Household Demand by WDPA for Selected Forecast Years[35]

As seen in Figure 1.22, single-family per-household demand dropped sharply from the 2002 baseline year to 2005. This sudden drop occurred partly because 2002 observed demand reflected 2002 observed weather, which was drier than normal. All subsequent years were simulated using long-term normal weather which is comparatively wetter than 2002, resulting in lower demand forecasts. Also, real marginal price climbed sharply from 2002 to 2003 for most WDPAs, such that the change in per-household de-

35 Data correspond to Table 1.17.

mand between 2002 and 2005 reflects a larger real marginal price change than between subsequent five-year periods.

Per-household demand differences between WDPAs arose in part from variations in single-family persons per household. Persons per household had the largest positive elasticity (0.56) in the single-family model. South Central Hillsborough, Tampa, Pinellas, and Northwest Hillsborough had the highest projected persons per household (roughly 2.5), while St. Petersburg had an intermediate value (2.3) and Pasco and New Port Richey had the lowest (roughly 2.0). These relative values also agree with the ranking of single-family per-household demand except for the relationship between St. Petersburg and Pasco, where Pasco had higher per-unit demand than St. Petersburg but lower persons per household.

Table 1.18 and Figure 1.23 show per-unit observed and forecasted demand by WDPA for the multi-family sector. In most WDPAs, multi-family per-unit demands were generally lower than single-family per-household demands, as expected. Both observed and forecasted multi-family per-unit demand was much higher than single-family demand for Northwest Hillsborough and about the same as single-family demand for Pasco. Variables that most heavily influenced multi-family per-unit demand included multi-family housing density (elasticity of -0.35), income (0.37), and fraction of multi-family reclaimed accounts (-0.39). Northwest Hillsborough had the lowest multi-family reclaimed account projections, the highest projected income, and extremely low projected multi-family household density, resulting in increases in per-unit demand.

Table 1.18
Historical and Forecasted Annual Average Multi-Family Per-Dwelling-Unit Demand By WDPA for Selected Forecast Years

	2002 (observed)	2005	2010	2015	2020	2025
Pinellas	92.70	95.68	100.66	105.11	109.23	113.52
St. Petersburg	145.15	149.68	157.48	164.43	170.89	177.60
New Port Richey	104.79	108.25	113.85	119.26	124.40	129.75
Pasco	215.08	223.35	234.91	246.08	256.67	267.71
Tampa	136.04	139.91	146.14	152.21	158.18	164.39
NW Hillsborough	408.34	421.39	440.16	458.44	476.43	495.14
SC Hillsborough	102.68	105.97	110.69	115.29	119.81	124.52

Trends in observed and forecasted multi-family per-unit demand among WDPAs were difficult to verify because directly observed base-year data for number of multi-family units was not available.

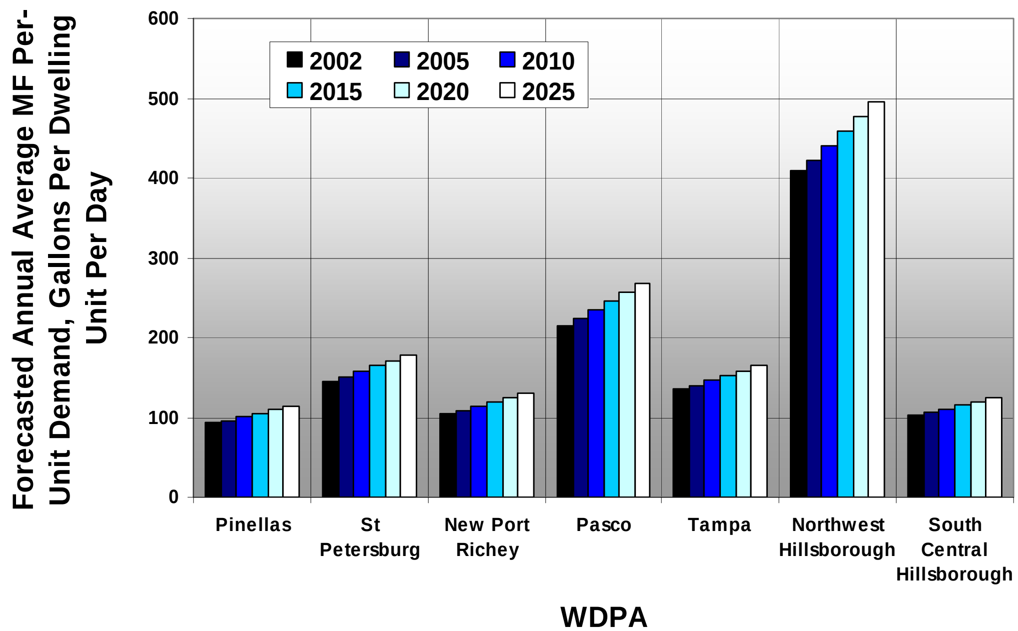


Figure 1.23 – Historical and Forecasted Annual Average Multi-Family Per-Dwelling-Unit Demand By WDPA for Selected Forecast Years

- As discussed in Section 1.6.1.4, projected values of multi-family housing density were assumed to remain constant at Water Year 2002 levels, with the 2002 level defined as number of multi-family units in a WDPA divided by multi-family developed acreage.

- Number of multi-family dwelling units was, in turn, estimated from number of MF accounts (Section 1.6.1.1).

- Multi-family housing density (units per acre) was a significant variable in the multi-family per-unit model, with an elasticity of -0.35. Because the multi-family housing density elasticity was high, large percentage errors in estimated units per acre would have generated considerable per-unit demand variations.

- Within the Northwest Hillsborough WDPA, there are a small number of multi-family units. Small absolute variations in number of units would have led to high percentage variations in base year and projected units per acre and subsequently observed and forecasted per-unit demand.

Thus, multi-family per-unit demand in Northwest Hillsborough was highly sensitive to estimation errors in number of multi-family housing units per account, and resultantly high per-unit observations and forecasts for this WDPA and sector may have resulted from this sensitivity. To remove the need for multi-family unit estimation, billing methodologies of members should be modified to record direct observations of number of multi-family units in each multi-family account.

High forecasted Northwest Hillsborough multi-family per-unit demand did not adversely impact total regional forecasts. Underestimation of MF dwelling units in Northwest Hillsborough would have led to underestimated MF housing density and overestimated MF per-unit demand. However, this per-unit demand would have been multiplied by an underestimated number of units when determining total MF demand, counteracting the per-unit error. This is a general strength of the per-unit-times-driver approach to demand modeling. Furthermore, Northwest Hillsborough's total multi-family demand (Table 1.14) was forecast as 3 to 7 MGD over the forecast period, such that any error would be only a fraction of those totals and small in comparison to regional and Northwest Hillsborough WDPA total demand forecasts.

Figure 1.24 and Table 1.19 show historical and forecasted non-residential per-employee demand for each WDPA. Per-employee demands varied little by WDPA and over time. The minimal variation between WDPAs corresponded to variation in income, a significant variable in per-employee equations. The only deviation in this trend was Tampa, which had higher projected income than Pinellas, St Petersburg, and Pasco but similar per-employee demand. However, Tampa also had the highest fraction of employees in industrial entities, influencing per-employee demand downward relative to the other WDPAs and counteracting the deviation in the above income trend.

Table 1.19
Historical and Forecasted Annual Average Multi-Family
Per-Dwelling-Unit Demand by WDPA for Selected Forecast Years

	2002 (observed)	2005	2010	2015	2020	2025
Pinellas	42.32	41.90	42.66	43.30	43.87	44.45
St. Petersburg	45.06	44.37	45.09	45.72	46.30	46.89
New Port Richey	48.77	48.02	48.82	49.57	50.26	50.98
Pasco	36.62	36.39	37.02	37.65	38.20	38.76
Tampa	39.83	39.66	40.26	40.84	41.38	41.93
NW Hillsborough	36.21	36.08	36.70	37.28	38.01	38.68
SC Hillsborough	44.40	44.35	45.12	45.82	46.75	47.57

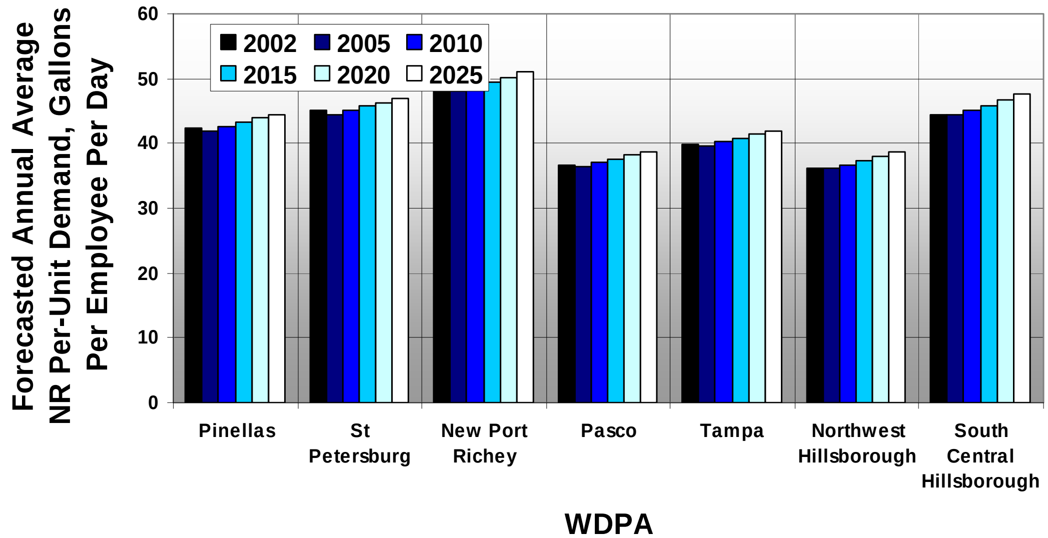


Figure 1.24 – Historical and Forecasted Annual Average Non-residential Per-Employee Demand by WDPA for Selected Forecast Years

Forecasted per-unit demands were thus rationalized in terms of explanatory variables. In many cases, this rationalization showed how regional variations in explanatory variables affect use characteristics for water-demanding entities. It was also shown that there is a significant need to collect data on actual number of multi-family units in Pasco, New Port Richey, Northwest and South Central Hillsborough, and Tampa.

1.7 Summary and Recommendations

This chapter has described the development of sectoral water demand models and demand accounting structures to produce deterministic water demand forecasts for Tampa Bay Water's water service area. Sectoral models were used to generate forecasts of future per-unit water use in the Tampa Bay region within a disaggregated water demand accounting framework.

A GIS application enabled development of a comprehensive water use modeling database including socioeconomic, land use, weather, and price information for nearly 1,500 geographical areas. Together with monthly demand data spanning the 1999-2002 time period, this database provided several thousand observations of water use and factors that affect water use across time and space. Using this database, Ordinary Least Squares Regression water use models were prepared for single-family, multi-family, and non-residential water use sectors. Table 1.20 summarizes variables that were significant for each sectoral model and driver units assumed for each model. The wealth of modeling data allowed estimation of models with precise coefficient estimates that related socioeconomic and other factors to variation in water use. Validation of these models against observed WDPA-level demand in Water Year 2002 indicated a high degree of predictive accuracy for forecasting at the WDPA and Tampa Bay Water regional service area levels.

Table 1.20
Summary of Explanatory Variables Examined in Developing Predictive Water Demand Models

Single-Family Per-Household Model	Multi-Family Per Dwelling-Unit Model	Non-Residential Per Employee Model
Driver Variable	**Driver Variable**	**Driver Variable**
Total SF Housing Units	Total MF Housing Units	Total Employment
Explanatory Variables	**Explanatory Variables**	**Explanatory Variables**
■ Avg. household income ■ Housing density ■ Persons per household ■ Marginal price of water and sewer ■ Departure of log maximum daily temperature from normal ■ Departure of log total monthly rainfall from normal ■ Departure of log number of rainy days (rain > 0.01”) from normal ■ Number of rainy days (rain > 1”) ■ Fraction of reclaimed accounts	■ Avg. household income ■ Housing density ■ Total monthly rainfall ■ Fraction of reclaimed accounts	■ Avg. household income ■ Fraction of employment among major industry groups ■ Departure of log total monthly rainfall from normal

Projection data for model determinants and other assumptions were input into the demand model to derive deterministic forecasts of sectoral demands for each WDPA and for Tampa Bay Water as a whole. These forecasts suggested that water demand in the Tampa Bay region would grow by a little less than 1 percent per year over the next 25 years.

It was forecasted that regional single-family demand would remain larger than multi-family and non-residential demand through 2025. Overall growth in regional demand was forecasted to be concentrated in multi-family and non-residential sectors, with relatively sluggish growth in the single-family sector. This forecasted growth pattern arose because projected increases in marginal price of water outpaced projected growth in income, resulting in forecasted decreases in single-family per-household demand. Decreasing per-household demand offset projected increases in number of single-family households, resulting in sluggish regional growth in total single-family demand. Regional multi-family and non-residential per-unit demand was forecasted to increase, reinforcing total demand growth trends in those sectors.

It is strongly recommended that procedures be developed within billing databases to directly observe number of housing units served by multi-family accounts. This would make projections of multi-family demand, particularly on a per-unit basis, more robust and could improve the fit of future models.

It is recommended that refinements be made to the predictive models and associated water use modeling database. More time series data for monthly water demand covering a larger period of record should be added to the database. Demand models be updated periodically using this new data to monitor for emergence of price influences on multi-family demand. Following these recommendations would better define seasonal variation and weather effects within sectoral models and identify subtle influences of marginal price on multi-family demand, should these influences exist.

As more water users in the Tampa Bay region connect to reclaimed water systems, there should be a measurable impact on potable retail demands. It is recommended that one or more studies be performed to evaluate the decrease in potable demand per amount of reclaimed water used, or potable demand offset. Studies should compare demand in similar neighborhoods with and without reclaimed water connections. These studies would include metering and collection of reclaimed water use data in the neighborhoods of interest and determination of *potable demand offset* for those neighborhoods. Offset rates could potentially be extrapolated to other similar neighborhoods. Such studies would be valuable as a reclaimed water system planning aid, identifying areas where reclaimed water development would produce the greatest demand reductions. If possible, results of these studies should be used to model reclaimed water use effects in the demand model as a function of TAZ-level socioeconomic characteristics.

The heterogeneous nature of the non-residential sector was, and may continue to be, problematic because a significant amount of variation in non-residential water use existed at a TAZ level but could not be explained by the non-residential model. Future data collection and modeling efforts should focus on obtaining water use and economic data for more detailed non-residential categorizations, preferably such that the resulting categories group non-residential accounts by well-defined water use characteristics.

The point demand model will be implemented in a custom computer application. This application will allow users to browse historical demand geographically and to generate new point demand forecasts based on modified projections of model variables.

2.0 Development of Risk and Uncertainty Assumptions

A water demand forecast typically forms the basis of many decisions concerning expected amount and timing of supply and demand expenditures, such as additions to supply and treatment capacity, implementation of water conservation programs and changes to the structure and level of water prices. Because water demand forecasts are often portrayed by a series of point estimates, external parties can erroneously perceive them as wholly accurate and certain to come true. By their very nature, such *deterministic* forecasts do not inform decision-makers of real uncertainties inherent to the forecasting process. The presence and magnitude of these uncertainties can and should be explicitly represented within forecasts, allowing decision-makers to factor uncertainty into judgments concerning future water supply capital improvements.

Understanding the key role that forecast uncertainty might play in project planning, Tampa Bay Water undertook an initiative to expand its water demand forecasting procedures to include uncertainty in long-term demand forecasts. The purpose of this chapter is to document the development of the Tampa Bay Water Probabilistic Demand Forecasting Model to support this initiative.

The probabilistic model was developed as an extension of the point forecasting model described in Chapter 1. Development of the probabilistic model consisted of identifying or estimating the quantitative nature of uncertainty for each explanatory and driver variable in the point demand model. These specifications of uncertainty were developed at a workshop with Tampa Bay Water staff (October 4, 2002). The point demand model and these specified variable uncertainties were subsequently nested in a Monte Carlo simulation, forming a probabilistic demand model. This probabilistic model is capable of generating probabilistic forecasts conditioned on input variable uncertainty.

2.1 Risk and Uncertainty Concepts and Workshop Review

The project team conducted a workshop with Tampa Bay Water on the development of probabilistic water demand models. It was explained at this workshop that a probabilistic demand model could be derived from the point model by specifying uncertainty in explanatory and driver variables, then propagating this uncertainty through the point model to describe uncertainty in demand predictions. The goal of the workshop was to address assumptions about uncertainty in future values of explanatory and driver variables composing the point demand models. An overview of the point demand model was presented at the workshop, as well as a discussion of uncertainty concepts and techniques to incorporate uncertainty into water demand models.

The workshop served as a means to define differences between probabilistic and deterministic forecasting. Unlike a deterministic forecast, which is defined by a series of connected single points of predicted water use, a probabilistic forecast provides a range (or interval) of demands within which most future outcomes should fall. The underlying idea behind interval forecasting is that one should be more confident in a range of possible outcomes than a single predicted outcome.

A useful analogy relates to hurricane storm-tracking as shown in Figure 2.1. The middle of a forecast storm track may form the overall basis for predicting the future path of the storm, but upper-level winds, approaching fronts, water temperature and other influential factors are variable, all of which lead to uncertainty around the series of points represented by the most likely path. Thus, it is critical to warn people that there is a chance of a direct hit some distance away from the most likely storm track. Similarly, uncertainty in future weather conditions, incomes, prices, employment and many other factors can also lead to variability in a forecast of water demand. Therefore, it is important for decision-makers to understand how uncertainties in future weather and socioeconomic conditions influence water use forecasts when evaluating subsequent capital improvement and policy decisions.

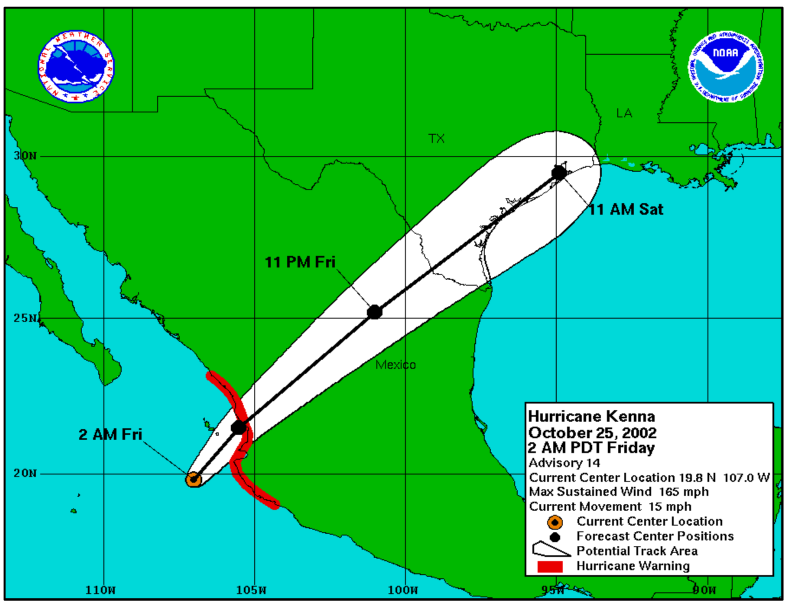


Figure 2.1 – Hurricane analogy for probabilistic forecasting.

The process of defining uncertainty assumptions included selection of a probability density function (pdf) for each explanatory and driver variable in the point demand model. It is convenient to consider a pdf as a formula that draws an assigned curve; once specific numbers, or *parameters*, are inserted into the formula, the area under the curve integrates to a value of one. These parameters usually represent a pdf's location (e.g., mean) and shape (e.g., standard deviation or variance). The chosen parameters for a particular pdf are generally intended to match a historical distribution of values or fulfill theoretical or practical concerns about how values of a particular variable are distributed. The workshop included a discussion of variables composing the point demand forecasting model, pdf recommendations for each variable based on data analysis, and concurrence or selection by the workshop attendees of appropriate distributions and parameters for each model variable.

Methods for combining uncertainty and the point forecast model were also discussed at the workshop. Given project resources and scheduling constraints, it was decided that a *conditional* probabilistic forecast would be developed from the water demand models and ranges of model input values[39]. Therefore, model error was not incorporated into the probabilistic water demand forecast.

Subsequent to the workshop, several applications of the probabilistic model were run to judge influence and sensitivity of agreed-upon uncertainty assumptions on the long-term water demand forecast. The uncertainty assumptions that were defined in preparation of draft and final probabilistic forecasts are defined in the remainder of this chapter. Resulting forecasts are presented in the next chapter.

2.2 Review of Point Models, Uncertain Variables, and Monte Carlo Simulation

Water demand models were previously developed for single-family (SF), multi-family (MF) and non-residential (NR) water use sectors (Chapter 1). These models were estimated from historical water use, socioeconomic, and weather data and generated a single deterministic point forecast. The point demand models were originally used to forecast water demand for each of the water demand planning areas (WDPAs) based upon projected values of driver and explanatory variables. Table 2.1 summarizes explanatory and driver variables that were significant determinants (by their presence in point demand models) of Tampa Bay Water's demand.

39 *A <u>conditional probabilistic</u> forecast portrays variation in predicted water demand by assuming either (a) no error in the forecast model or (b) no error in model inputs. The forecast is based on the condition that either (a) the model is correctly specified or (b) the model inputs are accurate. An unconditional probabilistic forecast assumes no conditions, such that the model and its inputs can both contain error. An unconditional probabilistic forecast is difficult and more costly to estimate since any single set of values of model inputs will produce a range of estimated water use values and there are many possible sets of inputs.*

Table 2.1
Summary of Explanatory Variables Examined in Developing Predictive Water Demand Models

SF SF Housing Units	MF MF Housing Units	NR Total Employment
Driver Variable	**Driver Variable**	**Driver Variable**
Total SF Housing Units	Total MF Housing Units	Total Employment
Explanatory Variables	**Explanatory Variables**	**Explanatory Variables**
■ Avg. household income ■ Housing density ■ Persons per household ■ Marginal price of water and sewer ■ Departure of log maximum daily temperature from normal ■ Departure of log total monthly rainfall from normal ■ Departure of log number of rainy days (rain > 0.01”) from normal ■ Number of rainy days (rain > 1”) ■ Fraction of Reclaimed Accounts	■ Avg. household income ■ Housing density ■ Total monthly rainfall ■ Fraction of Reclaimed Accounts	■ Avg. household income ■ Fraction of employment among major industry groups ■ Departure of log total monthly rainfall from normal

In addition, wholesale deliveries and other/unbilled demands were estimated for each WDPA. Wholesale was estimated as a percentage of retail demand. Other/unbilled demand was estimated as a percentage of total metered demand (*i.e.*, retail plus wholesale). Percentage wholesale and percentage other/unbilled were fixed at values observed for Water Year 2002.

In developing the probabilistic model, probability distributions were stipulated (or invariability was assumed) for each of these quantities. These distributions are described in Section 2.3. In some cases, these distributions varied by WDPA and/or calendar month.

A Monte Carlo simulation procedure[40] was subsequently used to develop a conditional probabilistic water demand model based on a combination of the point model and distributions for model variable inputs. Each iteration of the Monte Carlo simulation randomly selected a value for each input variable based on the probability density function speci-

40 *@Risk, produced by The Palisade Corporation (www.palisade.com), was used to perform Monte Carlo Analysis. This software is a Microsoft Excel plug-in that operates on a spreadsheet version of the model of interest and on in-spreadsheet specifications of input variable pdfs.*

fied for that variable, then used the complete set of values to produce a complete water demand forecast. The simulation procedure performed numerous independent iterations (between 5,000 and 10,000), each generating independent forecast curves. Forecast curves were then pooled and forecasted demand values were ranked at each forecast time point, yielding a distribution of estimated water demand for each month and year over the forecast horizon. The process is illustrated in the diagram of Figure 2.2.

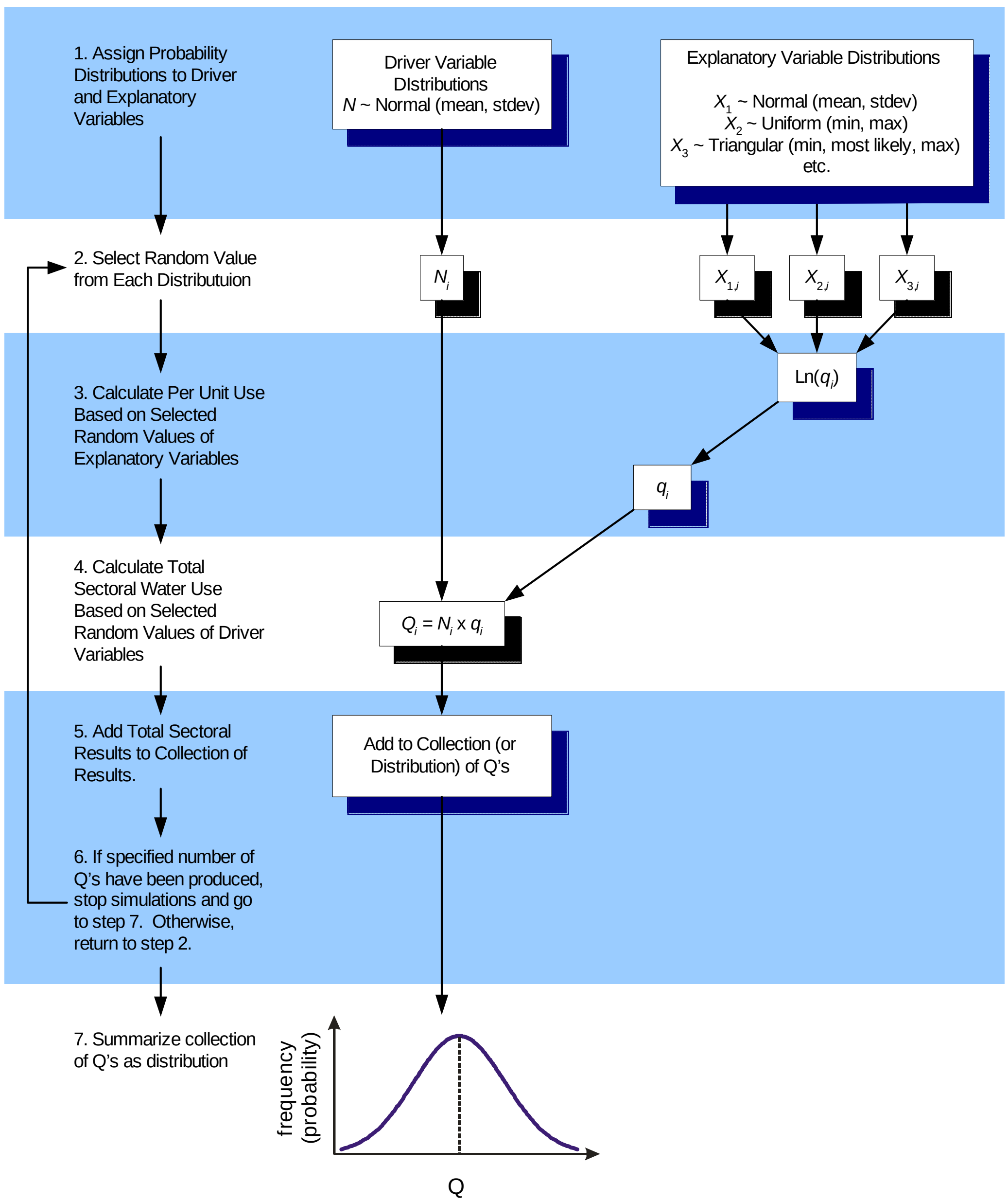


Figure 2.2 – Conditional Probabilistic Simulation Process

2.3 Assignment of Probability Density Functions

This section describes the development and stipulation of pdfs for explanatory and driver variables in the probabilistic demand forecast. Assumptions about distributions for most model variables were derived from actual data where possible, and by reference to external sources and experience where necessary. In general, probability distributions for weather variables were defined through analysis of historical data. Socioeconomic variables were assigned based on readily available data and/or experience.

Three families of probability distributions are commonly used in cases where actual distributions are unknown or cannot readily be estimated: normal, uniform and triangular distributions.

- *Normal* distributions are bell-shaped and symmetrical. The distribution is defined by two parameters: the mean (μ, mu) and the standard deviation (σ, sigma). The normal distribution is easy to interpret, in that 68 percent of all possible cases lay within $\pm 1\sigma$, 95 percent of all cases lie within $\pm 2\sigma$, and 99 percent of all cases fall within $\pm 3\sigma$.

- *Uniform* distributions are rectangular in shape and defined by two parameters, the minimum and maximum possible values that a particular variable can take. In this distribution all possibilities between the minimum and maximum share an equal likelihood of occurrence.

- *Triangular* distributions are similar to normal distributions but are not required to be symmetrical. The triangular distribution is defined by three parameters, a minimum value, a most likely value, and a maximum value.

Figure 2.3 illustrates the basic shapes and characteristics of these families of density functions.

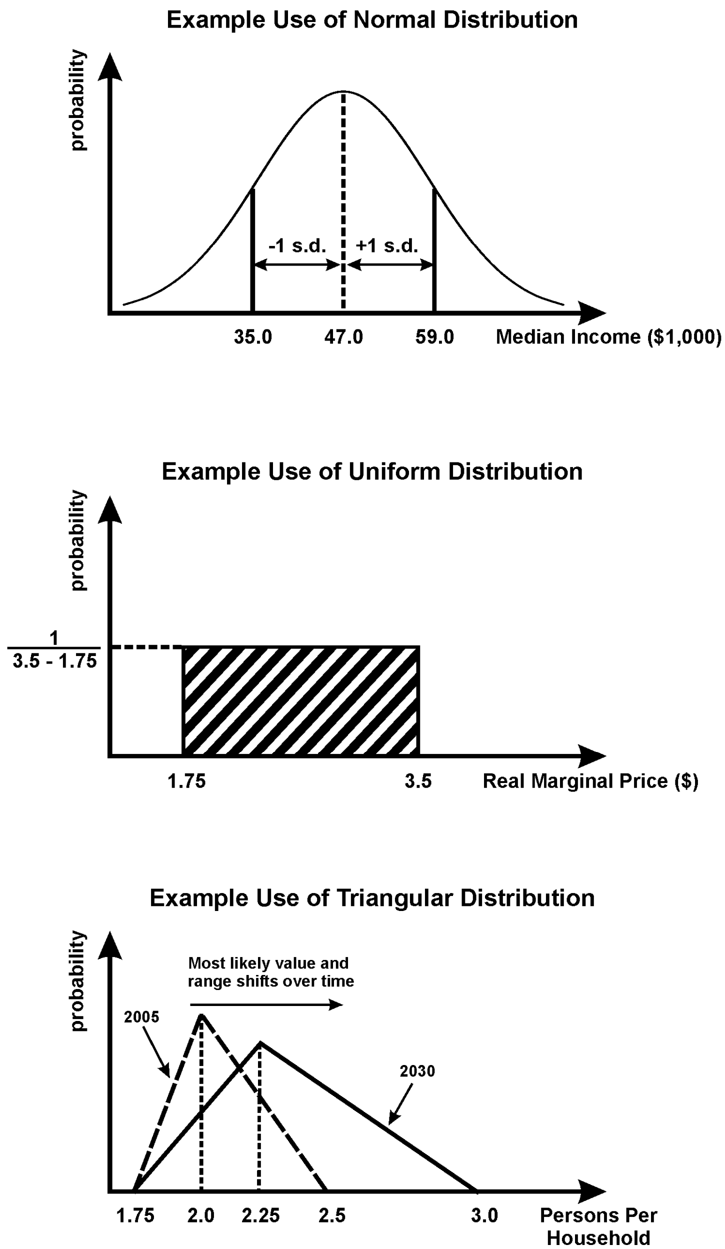


Figure 2.3 – Example Uses of Three Common Probability Density Functions

The following sections describe probability distributions assigned to uncertain explanatory and driver variables. Table 2.2 summarizes these assignments. In many cases, point projections of explanatory and driver variables were incorporated in pdf definitions, such as in defining the mean for a distribution. Point projections for variables were described in Chapter 1.

Table 2.2
Tampa Bay Water Demand Model Inputs

Variable	Sector	PDF	Remarks
Driver Variables:			
Number of Occupied Housing Units	SF, MF	Normal	Strong correlation is enforced between housing units and total employment and income and marginal price across all agencies.
Total Employment	NR	Normal	Total employment reflects the sum of employment in Commercial, Industrial, and Services groups.
Explanatory Variables:			
Real Average Household Income	SF, MF, NR	Normal	Expressed in 1999 dollars.
Real Marginal Price of Water and Sewer	SF	Normal	Expressed in 1999 dollars per thousand gallons.
Persons per Household	SF, MF	Uniform	Min and max defined as 0.9 and 1.1 times average household size in year for each WDPA
Housing Density	SF, MF	Uniform	Calculated from TAZ data, measured in units per acre.
Employment Group Proportion of Employment	NR	Constant	Percentages of Commercial, Industrial, and Services employment.
Average Monthly Maximum Temperature	SF, MF	Normal	Expressed in degrees Fahrenheit.
Total Monthly Rainfall	SF, MF, NR	Gamma	Expressed in inches per month. Strong correlation is enforced between rainfall and rainy days within each WDPA.
Number of Monthly Rainy Days (More than 0.01 or 1 inches)	SF	Gamma, Extreme Value	Expressed as days per month.
Reclaimed Water Percentage	SF, MF	Constant	Expressed as percentage of total accounts in sector.
Month and WDPA Intercept Adjusters	SF, MF, NR	Constant	No information available to justify uncertainty assumptions for these parameters.

Table 2.2
Tampa Bay Water Demand Model Inputs

Variable	Sector	PDF	Remarks
		Other Variables:	
Sector Calibration Coefficients	SF, MF, NR	Constant	These annual and monthly agency multipliers are derived from Water Year 2002 (October 2001-September 2002) data.
Wholesale Percentage	M&I	Triangular	Most likely value defined as value for WY 2002, min and max defined as 0.9 and 1.1 times WY 2002 value
Other Percentage	Total	Triangular	

Note: The model inputs are derived from MPO and BEBR data unless otherwise indicated.

2.3.1 Single-Family Households, Multi-Family Dwelling Units, and Total Employment

All driver variables were assumed to follow Normal distributions that differed by WDPA and year. The mean SF households and MF dwelling units parameters (μ) were taken as point projections of corresponding units over the 2005-2025 time period. These point projections were derived from the base year housing count and the assumed Florida Bureau of Economic and Business Research (BEBR) growth rate in SF and MF new housing starts by county (Chapter 1). Corresponding values for the standard deviation (σ) were based upon BEBR county-level population projections (BEBR, 2001c); these population projections were assumed to correlate with number of SF and MF housing units. BEBR population projections contained expected (“medium”) estimates along with “low” and “high” estimates indicating the range within which the middle two thirds of actual future county populations would fall. While population projection ranges were not completely symmetric, they corresponded closely to one standard deviation of a normal distribution because they contained two-thirds of expected population outcomes. To construct a standard deviation for housing units, the larger of the differences between low-and-medium and medium-and-high population projections was assumed to represent one standard deviation on population. The larger half-interval was expressed as a fraction of total population. This fraction was multiplied by mean housing unit projections to derive housing unit standard deviations. Means and standard deviations were independently derived in this manner for SF and MF housing units in each WDPA for each forecast year.

There was no information available regarding employment projection methodology adopted by BEBR, so the same approach was adopted for total employment as for housing units. Re-application of this approach was based on an assumption that total employment and total population were strongly correlated. Mean employment counts by year and WDPA were defined by previous point projections of total employment. Corresponding standard deviations for employment by year and WDPA were defined as one-half the maximum percent difference between either the BEBR medium-to-high population projection or the medium-to-low population projections for the corresponding county and year.

2.3.2 Median Household Income

Future values of median household income were assumed to follow a normal distribution for each WDPA and forecast year. The mean projected WDPA income was based upon base year income and BEBR county-level projections of annual percent growth in per capita personal income. The BEBR source did not include a range of future income growth, so based upon assumptions derived by a prominent planning agency for a similar project in Southern California (Kiefer and Porter, 2000), the corresponding standard deviations started at zero for 1998 (the first reporting year in the BEBR source) and were assumed to increase at a rate of 1.5 percent of the mean per year over the forecast horizon.

2.3.3 Real Marginal Price of Water and Sewer

Future values of marginal price were assigned normal distributions which varied by year and WDPA. For each WDPA, mean price represented the point projection of inflation-adjusted marginal price (in 1999 dollars), which was assumed to increase at rates equal to changes in the projected unit cost of water to member governments from Tampa Bay Water until 2008 (Black and Veatch, 2002) and then to increase throughout the rest of the forecast period at a rate of 4 percent per year. The corresponding standard deviations were set at 3 percent of the mean marginal price. Allowing uncertainty in the real price of water and sewer allowed for the very real possibility that rate adjustments may over- or under-compensate for inflationary effects.

2.3.4 Housing Density

Housing density was assigned a separate uniform distribution for each housing type and WDPA. In absence of available data regarding future acreage development, assigned distributions for housing density were fixed across all forecast years. Minimum and maximum values for the distributions were defined as 0.9 and 1.1, respectively, times the WDPA-average SF and MF units per acre.

2.3.5 Persons Per Household

Persons per household was assigned a uniform distribution for each year and WDPA. Persons per household projections were calculated from projected population and total housing units while holding the SF to MF unit ratio constant. Minimum and maximum values for distributions were defined as 0.9 and 1.1, respectively, times the average SF household size in each year and WDPA.

2.3.6 Proportion of Employment Among Non-Residential Classes

Distribution of total employment was allowed to vary over time according to projections provided at the TAZ level by Metropolitan Planning Organizations (MPOs). Distribution of future employment among service, commercial and industrial employment categories was treated as certain. However, distribution of employment was applied to total employment, which were assigned probability distributions as defined above.

2.3.7 Average Maximum Temperature

In the original point forecast, long-term average weather variable projections (*i.e.,* maximum daily temperature, total rainfall, and number of 0.01" and 1" rainy days for a month) were calculated for each WDPA and calendar month as a weighted average of monthly readings for six weather stations (listed in Chapter 1) over the 1971-2000 historical period. For each WDPA and calendar month, each station's weather values for that month were weighted by inverse squared distance between the station and the WDPA centroid.

Analysis of monthly values for average daily maximum temperature in WDPAs indicates this variable was normally distributed. A separate normal distribution was fitted to temperature data for each calendar month and WDPA. Figure 2.4 provides an illustration of the distribution of monthly maximum daily temperature for the month of January.

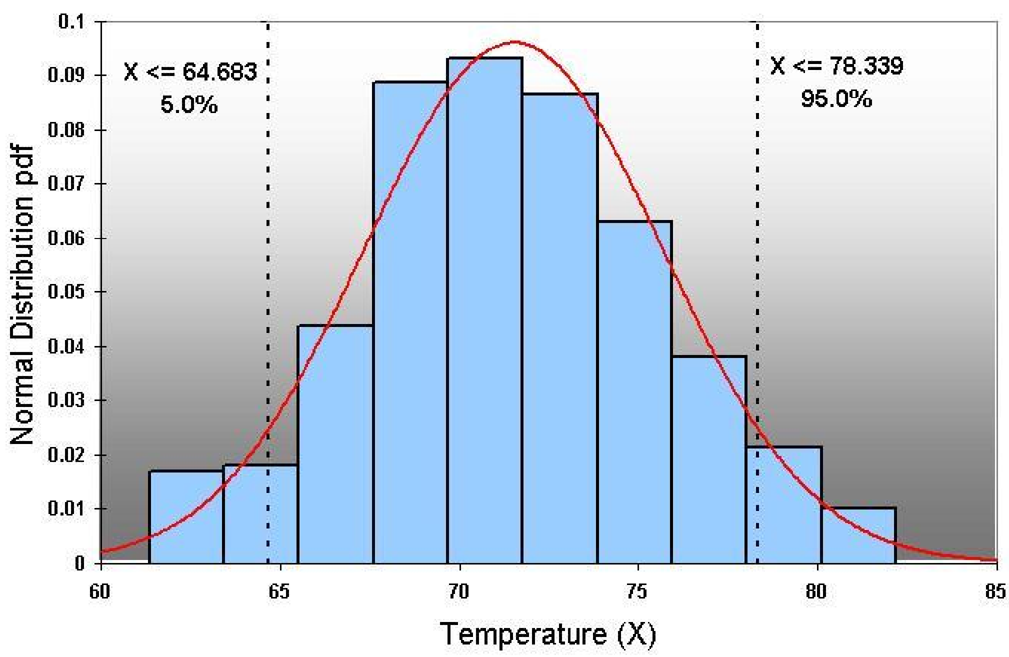


Figure 2.4 – Example of Normal Curve Fit to Average Maximum Daily Temperature (January)

2.3.8 Precipitation

Historical normal monthly rainfall exhibited a skewed distribution. Thus, determination of calendar monthly total rainfall distributions required selecting a skewed distribution (such as Gamma, Weibull or log-normal), then parameterizing that distribution. To select a distribution, screening studies were performed wherein each skewed distribution type was ranked according to how well it fit monthly precipitation data. The Gamma distribution consistently had the best rank and was chosen. Next, Gamma distributions were estimated for each calendar month in each WDPA. Unlike normal distribution parameters, Gamma parameters generally do not compare to a mean or standard deviation, and their determination is not as straightforward as normal parameters. Parameters were estimated using BestFit[41], a Microsoft Excel plug-in designed to fit a wide variety of pdfs to spreadsheet data. Figure 2.5 shows an example distribution of rainfall for the month of March.

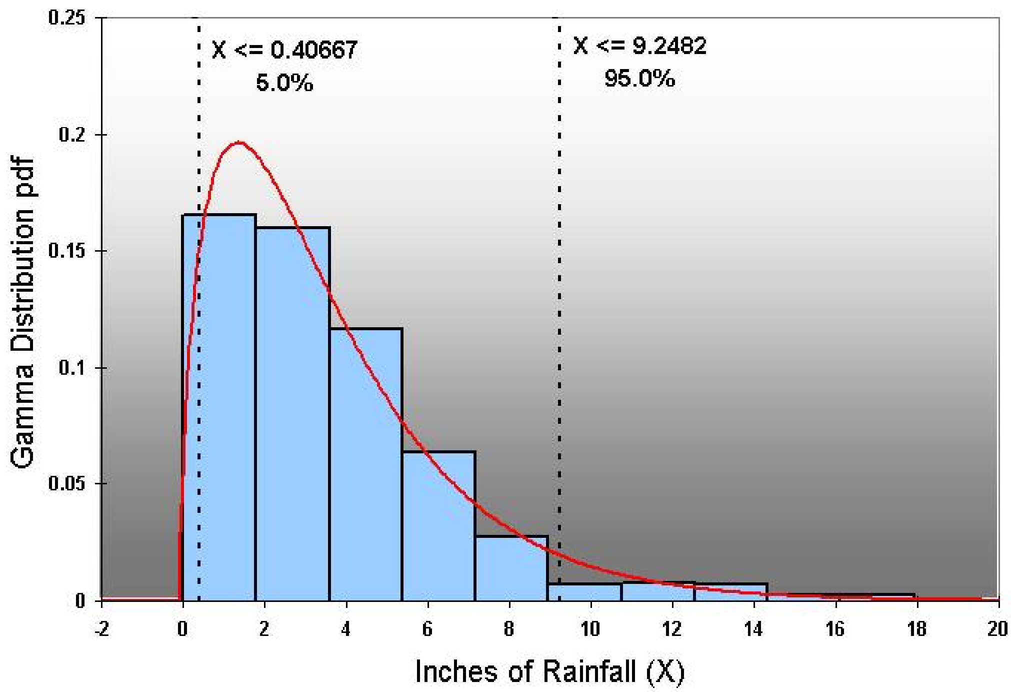


Figure 2.5 – Example Gamma Distribution Fit for Precipitation (March)

2.3.9 Number of Rainy Days

Two rainy-day variables were involved in demand forecasting: number of days with more than 0.01 inch of precipitation and number of days with more than 1.00 inch of precipitation. Gamma and extreme value distributions were considered for these two variables.

41 BestFit is produced by The Palisade Corporation: www.palisade.com.

Under general conditions, the total number of rainy days in a month can be considered as the maximum value of a sequence of observations (*i.e.*, days), indicating an extreme value distribution, or the monthly rate of exceeding a particular threshold (e.g., number of days exceeding some measurement), indicating a Gamma distribution[42]. As indicated by BestFit, the Gamma distribution provided a better fit for 0.01" rainy days. This distribution was consistent with the Gamma pdf used with total precipitation. BestFit indicated that the extreme value distribution was a better fit for 1" rainy days. Gamma and extreme value distributions were thus determined for rainy day variables for each calendar month in each WDPA. Figure 2.6 depicts an extreme value distribution for the number of days exceeding 1 inch of rain in the City of Tampa WDPA for August.

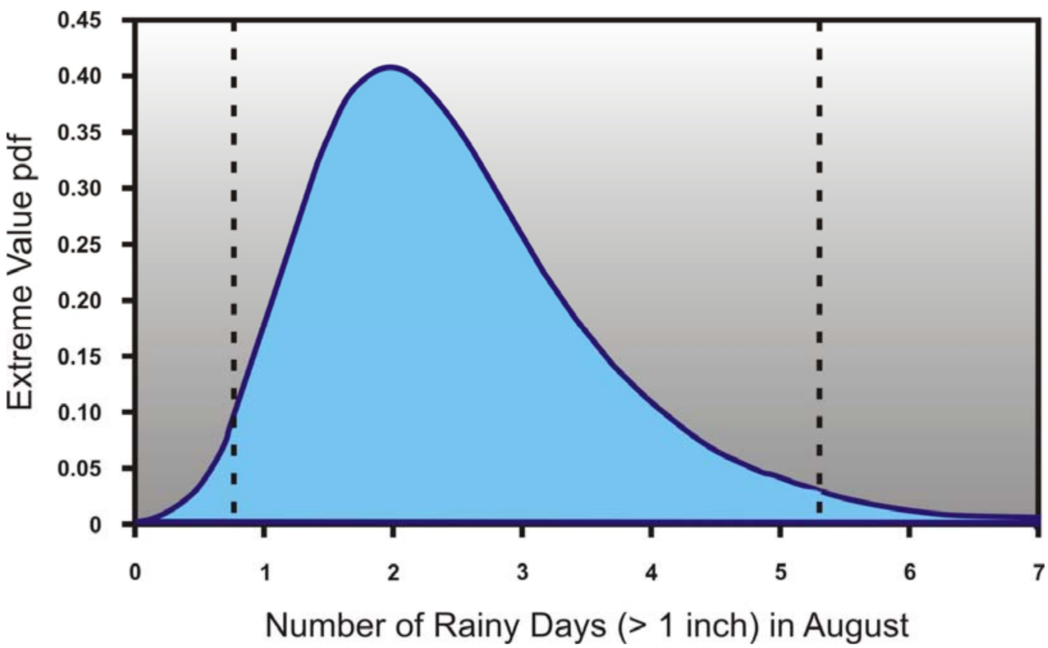


Figure 2.6 – Distribution of Number of Days with >1.0" Precipitation in August for the City of Tampa WDPA

2.3.10 Fraction of Reclaimed Accounts

Fraction of total accounts in each WDPA and sector with reclaimed water connections, or "Fraction of Reclaimed Accounts", in any WDPA was held constant at 2002 values and treated as a certain variable over the forecast horizon.

42 *A summary of extreme value theory in climate analysis is located at http://www.esig.ucar.edu/extremevalues/extreme.html [Statistics of Weather and Climate Extremes].*

2.3.11 Correlation Among Variables and Distributions

Correlations among model input variables were analyzed to determine whether and in what cases it was appropriate to have pdfs conditionally relate to each other within the Monte Carlo simulation procedure.

Precipitation and number of rainy days, by the nature in which they are measured, were expected to have dependent values. This definitional dependence was reinforced by a Pearson coefficient (r) of about 0.9, not only for each WDPA, but also between WDPAs for any designated forecast year. This estimated correlation coefficient of 0.9 was assumed for Monte Carlo simulations[43]. Pair-wise monthly correlations among other weather variables were found to be negligible or statistically insignificant.

Spatial correlation was assumed for housing units by type, total employment, median household income and price variables. Though these variables were assumed independent of one another (*e.g.*, independence of income and price), values for each of these variables were assumed to be correlated across WDPAs (*e.g*., the income in Pinellas was assumed to be correlated with income in Tampa and all other WDPAs). For each forecast year, distributions for these variables were assumed to be perfectly correlated between WDPAs. It seemed much more feasible to have higher or lower than expected numbers for these variables across the entire region as a whole than for the future values to be spatially independent of each other.

2.3.12 Percent Wholesale and Other/Unaccounted

Wholesale water distributions were included in Pinellas, St. Petersburg and New Port Richey demand forecasts and were treated as percentages of retail consumption on a monthly basis. *Wholesale percentage* for each WDPA and month was assigned a separate triangular distribution, where “most likely” values were assumed to be monthly percentages calculated for Water Year 2002 (October 2001-September 2002). Minimum and maximum values were assumed as 0.90 and 1.10 of Water Year 2002 values, respectively.

All WDPAs had unbilled water use, defined as the difference between total Tampa Bay Water deliveries and total WDPA-metered use (as described in Chapter 1). Each WDPA’s unbilled use was expressed as a percentage of total deliveries to that WDPA from Tampa Bay Water, or *other/unbilled percentage*. Unbilled percentage in each WDPA and month was assigned a separate triangular distribution, where “most likely” values were assumed to be monthly percentages calculated for Water Year 2002. Minimum and maximum values were assumed as 0.90 and 1.10 of Water Year 2002 values, respectively.

43 *@Risk enforces correlation during Monte Carlo selection using Rank Order Correlation, whereby sampling percentiles are generated to reproduce a user-specified rank correlation coefficient.*

2.4 Summary

Development of probabilistic water demand forecasts for Tampa Bay Water required explicit quantitative specification of uncertainties in future values of variables influencing water use. Through research, analysis and a workshop with Tampa Bay Water staff, probability density functions were assigned to nearly all of the variables within Tampa Bay Water's forecasting model.

Probability distributions described in this chapter were used in a Monte Carlo simulation of future water demands in the Tampa Bay region. Resultant interval forecasts of water use are documented in Chapter 3.

3.0 Summary of Water Demand Simulations

This chapter documents the results of the probabilistic water demand forecast. Following a brief review of the probabilistic water demand simulation methodology, an interval forecast of water use for Tampa Bay Water out to the year 2025 is presented and discussed. Interval forecasts for each WDPA are presented in Appendix C. This chapter concludes with a summary and look forward to a probabilistic water needs analysis that explicitly accounts for both demands and supply uncertainty over the forecast period.

3.1 Demand Simulation Methodology

In Chapter 2, explanatory and driver variables were identified that would be considered uncertain for probabilistic forecasts. Probability distribution functions were determined for those variables, and degrees of rank correlation between certain distribution pairs were specified. Treating variables as uncertain was identical to recognizing that, for any given WDPA and future point in time, many variable values would be possible. It was further recognized that variable uncertainty would propagate to demand forecasts via relationships contained in the deterministic water demand models. Figure 3.1 conceptually illustrates propagation of meteorological and socioeconomic uncertainty through demand relationships to water demand projections.

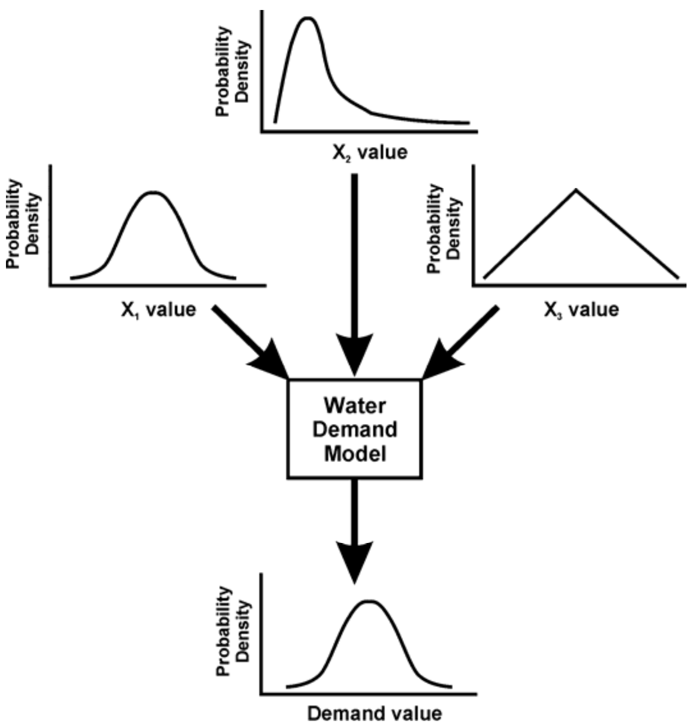


Figure 3.1 – Propagation of socioeconomic and meteorological uncertainty through the point demand model to probabilistic water demand forecasts.

Because there were literally an infinite number of combinations of variable values, all with different chances of occurring, it was necessary to enact a process of *simulating* these multiple scenarios. Iterative sampling of values of model variables ultimately produced a distribution of demand forecasts for any given time period.

The adopted probabilistic forecasting methodology was a Monte Carlo simulation (Figure 3.2). The simulation involved multiple independent, iterative calculations of point demand conditioned on randomly-selected driver and explanatory variables. Each iteration consisted of:

1. Random selection of a complete set of values for explanatory variables (X) and driver variables (N) from assigned probability density functions (pdfs), enforcing correlation where specified,

2. Calculation of retail water use by sector and WDPA corresponding to variable value selections using the water demand model of Chapter 1,

3. Random selection and application (for relevant WDPAs) of percent wholesale values from assigned pdfs, and calculation of total metered demand for each WDPA, and

4. Random selection and application of percent other/unbilled values from assigned pdfs, and calculation of total water demand for each WDPA.

Similarly to the point forecast, this process generated forecasts at WDPA, sectoral, and monthly levels. Each selection (or *iteration*) of data produced a complete, independent Tampa Bay Water forecast by WDPA and month over the 2003-2025 forecast period. Forecast demand results for each WDPA and month were then pooled and forecasted demand values were ranked at each forecast month, yielding a distribution of estimated water demand for each WDPA and month over the forecast horizon.

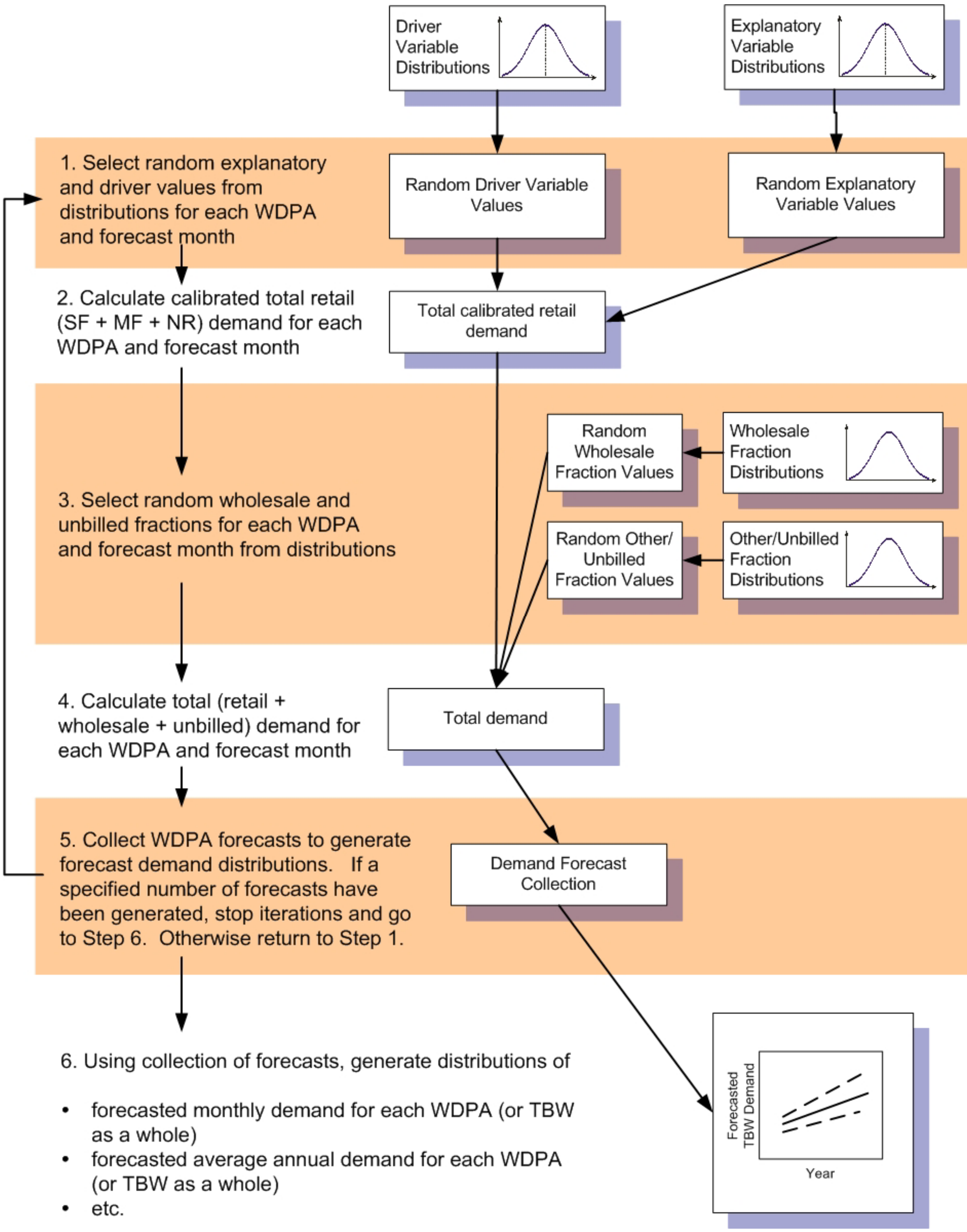


Figure 3.2 – Schematic of Conditional Probabilistic Simulation Process

The Monte Carlo procedure generally requires a large number of sampling iterations to converge to stable output distributions (*i.e.*, unchanging with additional iterations). Such requirements are often intensive on computing resources. A total of 10,000 iterations of the Monte Carlo demand simulation process were performed, so each month, sector, and geographical area had a distribution of 10,000 predictions of water use. The selection of 10,000 iterations was arbitrary but sufficiently high to assure convergence of the simulation process without undue computational intensity.[44] Appendix B contains specifics regarding implementation of the Monte Carlo simulation using @Risk and a spreadsheet version of the demand models.

3.2 Interval Demand Forecast for Tampa Bay Water

This section focuses upon the annual average demand forecast for the Tampa Bay Water service area, which was aggregated from all sectors, WDPAs, and calendar months within each forecast year. These results are listed in Table 3.1. Annual average probabilistic forecasts for each WDPA are provided in Appendix C.[45] Furthermore, monthly forecast results are available in spreadsheet format in the CD accompanying this document.

---

44 In this context, convergence refers to achieving stability in the general characteristic of a distribution of outcomes. Typical indicators of convergence are the degree of change in the mean and variance of a sample. The selected number of 10,000 iterations was sufficient in achieving mean and variances that were generally unaffected by adding additional iterations.

45 Due to the independence of sectors and WDPAs inherent in the forecasting methodology the probabilistic results for the WDPAs will not sum to the probabilistic Tampa Bay Water service area totals.

Table 3.1
Tampa Bay Water Service Area Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	238.1	241.2	258.5	273.0	286.5	300.0
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	236.1	242.5	259.5	273.3	285.7	297.9
Standard Deviation	2.6	3.2	4.5	6.4	8.7	11.1
Median	236.1	242.5	259.5	273.5	286.2	298.8
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	232.0	237.2	252.0	262.4	270.6	278.5
10%	232.9	238.3	253.6	265.1	274.5	283.6
15%	233.5	239.1	254.8	266.8	276.9	286.8
20%	233.9	239.8	255.7	268.1	278.9	289.3
25%	234.4	240.3	256.5	269.1	280.4	291.1
30%	234.7	240.8	257.2	270.1	281.6	292.8
35%	235.1	241.2	257.8	271.1	282.9	294.3
40%	235.4	241.6	258.4	271.9	284.1	295.9
45%	235.8	242.0	259.0	272.8	285.2	297.3
50%	236.1	242.5	259.5	273.5	286.2	298.8
55%	236.5	242.9	260.1	274.3	287.3	300.2
60%	236.8	243.3	260.7	275.1	288.3	301.5
65%	237.1	243.7	261.3	276.0	289.4	302.8
70%	237.5	244.2	261.9	276.8	290.6	304.2
75%	237.9	244.6	262.6	277.7	291.8	305.6
80%	238.3	245.2	263.3	278.7	293.0	307.2
85%	238.8	245.8	264.2	279.8	294.4	309.1
90%	239.4	246.6	265.3	281.3	296.4	311.4
95%	240.3	247.8	266.9	283.2	298.9	314.7

3.2.1 Interpretation of Forecast Results

Simulation results were summarized in the form of percentiles, which characterize the cumulative distribution of demand for each WDPA and time period. The top portion of Table 3.1 (and the tables in Appendix C) display measures of central tendency, *i.e.* the mean and median of simulated forecast distributions, either of which could act as point estimates of future demand similarly to the point forecast developed in Chapter 1. Unlike the point forecast, these central tendency estimates refer to expectations and/or characteristics of a distribution of potential outcomes and not to deterministic estimates calculated by means of a single vector of values for model inputs.

Tampa Bay Water service area forecast distributions were relatively symmetric, meaning that distributions approximated normal distributions. This approximation allows interpretation of the median forecast (50^{th} percentile in Table 3.1) as an estimate of the forecasted mean, since mean and median are identical in a normal distribution.

The bottom portion of Table 3.1 (and the tables in Appendix C) provides percentile values for average annual demand at selected forecast years for the corresponding geography. These percentiles were derived by ranking forecast demand values generated by all iterations at each specified point in time. Percentiles reflected a cumulative distribution of demand values at each forecast time, thus quantifying uncertainty in forecast water demand over time. One-half of all observations were expected to lie above the median, or the 50^{th} percentile. Similarly, 30 percent of possible values were expected to fall below the 30^{th} percentile, 20 percent above the 80^{th} percentile, and so on.

One may generally use pairs of percentile values to delineate particular forecast intervals. For example, 50 percent of future demand values would be expected to fall between the 25^{th} and 75^{th} percentile values, or between the 10^{th} and 60^{th} percentiles. In the discussion of results below, a median-rank-centered 90 percent confidence interval for forecast demand trajectories was delineated using 5^{th} and 95^{th} percentile demand values at each forecast time point. This interval was interpreted as containing 90 percent of all forecast possibilities for the corresponding geography over the forecast horizon, conditional on assumed explanatory and driver variable uncertainties. Alternatively, the interval was interpreted to exclude the highest 5% and lowest 5% of possible demands, conditional on assumed explanatory and driver variable uncertainties.

3.2.2 Probabilistic Forecast Results

Using the forecast mean as expectation, Table 3.1 shows expected average annual water demand in the Tampa Bay Water service area was forecast to reach approximately 298 MGD in 2025. Compared to the expected forecast value in 2003, this represents about a 26 percent increase over 23 years, or roughly a one-percent increase in annual average daily demand per year.

The 5^{th} and 95^{th} percentiles of simulated demand were used to define the 90 percent confidence interval of forecast demand. For example, 90 percent of total demand values were predicted to fall between 232 MGD and 240 MGD in 2003, whereas in 2025 the bounds of the 90 percent forecast envelope were 278 MGD and 315 MGD. Increase in uncertainty over the forecast period was evident in the growing standard deviation of forecast distributions (e.g., a standard deviation of 2.56 MGD in 2003 compared with a standard deviation of 11.12 MGD in 2025).[46] The annual forecast interval is shown in Figure 3.3, illustrating the characteristic widening of the forecast envelope with increasing uncertainty over time in variables that influence water use.

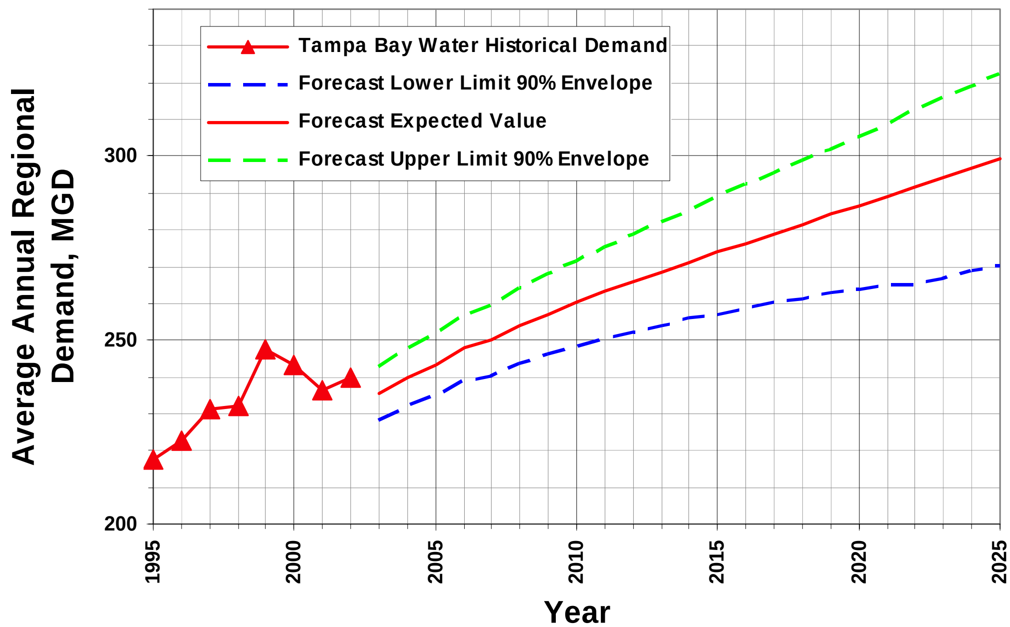


Figure 3.3 – Forecasted annual average of daily water demand in the Tampa Bay Water service area, with 90% confidence interval and historical demand for comparison.

Figure 3.4 depicts the 90 percent forecast interval on a monthly time step. Monthly forecasts represented raw results of Monte Carlo simulations, which were originally performed at monthly time steps (annual results in Table 3.1 and Figure 3.3 were subsequently derived by averaging annual demands for each year within each Monte Carlo iteration). The monthly forecast interval showed widening over time similarly to the annual forecast, as well as seasonality in water use arising predominantly from single-family residential demands.

Interpreting the percentiles of Table 3.1 in terms of estimated exceedence probabilities can be useful for supply planning. Percentile exceedence probabilities can be used to establish thresholds for water demand that may occur with a particular estimated likelihood. For example, 5 percent of simulated outcomes in 2025 were projected to exceed about 315 MGD and 50 percent of outcomes to exceed about 299 MGD. Meanwhile, 95

46 Earlier forecast years had relatively tight bands, mostly because assumed error in model inputs was lower at an earlier forecast time point than later times. For future years, these intervals expanded because assumed input uncertainties were expressed over time as increasing percentages of usually increasing point estimates of explanatory variables.

percent of simulated outcomes for 2025 were projected to exceed 278 MGD. Depending on risk tolerance, one could use results such as these to decide upon a threshold that would represent a supply reliability target.

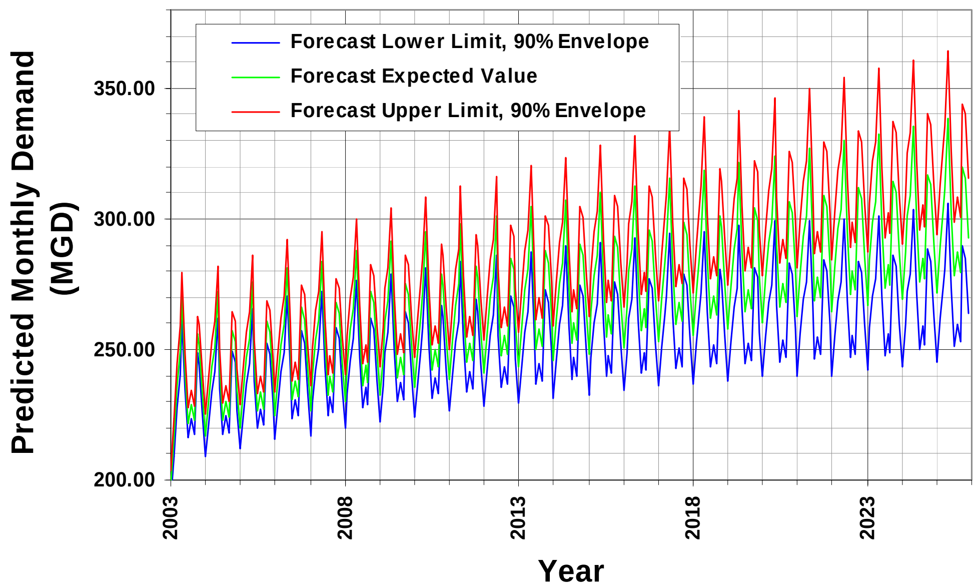


Figure 3.4 – Forecasted monthly average of daily water demand in the Tampa Bay Water service area, with 90% confidence interval.

3.3 Summary and Recommendations

Chapters 2 and 3 have established and implemented a methodology for portraying and quantifying uncertainty in Tampa Bay Water's demand forecast. With the help of simulation and statistical methods, Tampa Bay Water can now understand and view its future demand within a probabilistic framework. This framework reflects genuine uncertainties about future water demands and allows assessment of risks in efforts to provide reliable water supplies in the region.

The probabilistic annual regional demand forecast showed two critical forecasted trends. First, expected value of demand was forecasted to steadily increase, to approximately 298 MGD in 2025. This represents about a 26 percent increase over 23 years, or roughly a one-percent increase in annual average daily demand per year. Growth in the expected value of demand was in excellent agreement with the point forecasting results. Second, demand uncertainty was forecasted to steadily widen for times further in the future. Increase in uncertainty over the forecast period was evident in the growing standard deviation of forecast distributions (e.g., a standard deviation of 2.56 MGD in 2003 compared with a standard deviation of 11.12 MGD in 2025.) This increase in uncertainty for forecasted demand arose from widening uncertainty in explanatory and driver variables at future times. Such increasing uncertainty over time in driver and explanatory variables was reasonable for two reasons, a longer time span generally offers more opportunities for unexpected changes in values to accumulate, and as the expected value of many variables increases, it is often the case, uncertainty in those variables increases as well.

By the very nature of the forecasting process, uncertainty in inputs propagates to uncertainty in outputs. Forecast accuracy and reliability are generally direct functions of accuracy and reliability of model inputs. Demand forecasts presented in this chapter should be recognized as *living analysis*, indicative of the current understanding of demand influences in the Tampa Bay Water service area and subject to revision as understanding of these influences improves.

Explanatory and driver variable projections and uncertainties may change, due to release of new projection data, changes in projection methods, and changes of sources for projection data (*i.e.*, the organizations providing the projections). When such changes arise, probability density functions should be redefined for exploratory and driver variables and new probabilistic forecast simulations should be performed.

Furthermore, the point forecasting model may be recalibrated or refitted periodically due to recommendations presented at the end of Chapter 1 and in Chapter 4. Any time the point forecasting model is revised in this manner, new probabilistic forecasts should be produced using the revised model.

4.0 Conclusions and Recommendations and Future Model Applications

Following implementation of its Master Water Supply Plan in 2001, Tampa Bay Water began a major, advanced initiative to quantitatively understand how socioeconomic, meteorological, and policy conditions in its service area influence potable water demand. The Tampa Bay Water Long-Term Demand Forecasting Model was the product of this initiative. The model is composed of a set of mathematical relationships that quantify geographical, socioeconomic, and meteorological influences on potable water demand. By combining the demand model with single-valued, or *point*, projections of socioeconomic growth and meteorological conditions, point demand forecasts were generated. These forecasts assist decision-makers in visualizing impacts of regional growth on future water needs.

Furthermore, Tampa Bay Water set out to understand the potential likelihood and degree of variation in projected socioeconomic and meteorological conditions. Uncertainty in socioeconomic and meteorological conditions was quantified and applied to the probabilistic demand model, resulting in forecasted distributions of demand. With these forecasted distributions, decision-makers can assess the risk of supply shortfalls relative to demand and plan for capital expenditures to avoid unacceptable risk.

4.1 Future Model Development and Forecasting Efforts

The demand model may be incrementally improved in several ways. These issues involve collecting new data as it becomes available, modifying the types of data collected, and updating model estimates based on these new data. Specific recommendations for future model development efforts include the following.

- **Maintain modeling database and include time-series data from a longer period of record for explanatory and driver variables**. Variables for which additional time-series observations are critical include water consumption, weather variables, real marginal price of water for member governments, fraction of accounts using reclaimed water, number of single-family households, number of multi-family dwelling units, and number of employees. Associations between billing accounts and TAZs must be maintained as new accounts are added and TAZ boundaries are redefined.

- **Develop and implement collection methods for data types not currently available.** New data collection methods for each member government must be developed to record actual number of multi-family dwelling units served by each multi-family account. Also, the Service/Commercial/Industrial categories into which non-residential customers are classified should be divided into more precisely-defined categories corresponding closely to common water use rates and consumption activities. This new categorization will require collection of data necessary to classify non-residential customers. A billing methodology study is currently underway to address these issues, and to synchronize data collection techniques among member governments.

- **Refit the demand model paramaters.** Based on an actively maintained modeling database, the demand forecasting model parameters should be refitted (i.e., regression should be repeated using the updated modeling database) at least once every five years. Refitting should also be performed any time it is suspected that a change has occurred in mechanisms by which explanatory variables influence demand, such as a change in irrigation restrictions, and any time a significant number of TAZs (50 or more) have been redefined. The model should also be refit when retail service areas are expanded and new corresponding TAZs are added to WDPAs.

- **Disaggregate wholesale demand model in similar detail as member retail demand.** Billing data should be obtained from wholesale customers and included in the modeling database. Corresponding exploratory and driver variable data should be obtained for TAZs in wholesale service areas. The model should be refitted to reflect the new data from wholesale regions. The resulting model would greatly facilitate integration of wholesale demand into member government retail demand as these utilities are acquired.

- **Evaluate potable demand offset produced by reclaimed water use.** One or more studies should be performed to evaluate the decrease in potable demand per amount of reclaimed water used, or *potable demand offset*. Studies should compare demand in similar neighborhoods with and without reclaimed water connections. These studies require metering and collection of reclaimed water use data in the neighborhoods of interest. Offset studies would establish potable demand offset for those neighborhoods, which could potentially be extrapolated to other similar neighborhoods. Such studies would be valuable as a reclaimed water system planning aid, identifying areas where reclaimed water development would produce the greatest demand reductions. If possible, results of these studies should be used to model reclaimed water use effects in the demand model with greater precision regarding TAZ- and WDPA-level socioeconomic characteristics.

The benefits of obtaining time series data for a larger period of record are twofold.

1. Time series data for a longer period of record will reveal whether real marginal price has a slow, indirect effect on multi-family water use or whether there truly is no effect. The presence and strength of this effect may change in the future, particularly if rising marginal price induces multi-family complexes to begin billing individual dwelling units for water consumption. To identify emergent relationships between price and multi-family demand, time series data from those periods would be necessary.

2. Additional time series data would allow representation of a broader set of weather effects within sectoral models. The period of record for the modeling database contained an abnormally intense drought mitigated by stringent irrigation restrictions. As the billing period of record is extended, contemporaneous weather data should be maintained in the modeling database, such that refitted models can be based on a broader set of weather conditions.

As the demand model is refined, point and probabilistic forecasts should be redeveloped to reflect the current version of the model. Forecasting recommendations include the following.

- **Maintain explanatory/driver variable projections and update forecasts.** Methods for developing projections of driver and explanatory variables must be periodically revisited and updated as necessary. In the future, planning organizations may begin projecting new types of data and cease projecting existing types. Also, values of projections and methods by which projections are generated can change. As new projection data and sources of projections become available, they should be evaluated for use as inputs for demand forecasts. Forecasts should be re-evaluated any time new projections for per capita income, real marginal price of water for member governments, fraction of accounts using reclaimed water, number of single-family households, number of multi-family dwelling units, and number of employees become available. The LTDFS is currently being developed to streamline forecast re-evaluation.

- **Yearly recalibration of demand forecast models.** Based on an actively maintained modeling database and projection data, models should be recalibrated yearly as new water use observations become available. Recalibration involves developing new calibration factors for the most recent data and application in new demand forecasts. Recalibration should be performed at least once a year. Recalibration of models does not require refitting of model parameters through regression analysis.

- **Recommend new projection methods to planning organizations.** Projected data and distributions were not directly available from planning organizations for several variables or were not available at desired scales of geography and time, including projections of single-family households, multi-family dwelling units, single- and multi-family persons per household and housing density. In these cases, projection values were estimated using most recent observations and/or available projections of associated variables. Needs for projections of these variables should be communicated to planning organizations, and methods for directly projecting desired values should be suggested if possible.

- **Maintain uncertainty assumptions for projected explanatory and driver variables and update probabilistic forecasts.** Explanatory and driver variable projections and uncertainties may change in the future, due to release of new projection data, changes in projection methods, and changes of sources for projection data (*i.e.*, the organizations providing the projections). When such changes arise, probability density functions should be redefined for exploratory and driver variables and new probabilistic forecast simulations should be performed. Furthermore, new probabilistic forecasts should be produced whenever the point model is refitted or recalibrated.

4.2 Future Model Application Efforts

This project represents an important step in accounting for the factors influencing demand as well as inherent uncertainties in future demand arising from uncertainties in those factors. With these results, Tampa Bay Water has derived a robust understanding of how demand should grow in the future, both in terms of expected growth and potential variation in that expectation. Tampa Bay Water has initiated several projects to leverage this new understanding in supporting resource planning decisions. Several additional applications of this model are recommended below.

4.2.1 Tampa Bay Water Demand Forecasting System (DFS)

The demand model developed in this project is currently being implemented in a custom computer application, the Tampa Bay Water Demand Forecasting System (DFS). This application will enable historical demand to be browsed geographically. In addition, the DFS will enable new point demand forecasts to be built based on modified projections of model variables. Identified users will be able to evaluate the sensitivity of point forecasts under different assumptions regarding projected explanatory and driver values.

4.2.2 Future Need Analysis (FNA)

Shortfalls in supply from surface water sources relative to total Tampa Bay Water demand must be met by the Consolidated Wellfields (under regulatory constraint). The potential need to use the Consolidated Wellfields drives the necessity for new water supply facilities. Tampa Bay Water's FNA project is currently underway to probabilistically assess future need as represented by wellfield use. In the FNA, the probabilistic demand model is coupled with a probabilistic water supply model to forecast future water supply needs. FNA results will serve to identify timing and risk of surface water supply shortfall and wellfield use, assisting decision-makers in efficient planning of future water supply capital expenditures. As half of the analysis required for FNA, the Long-Term Demand Forecasting Model thus provides immediate applicability in critical Tampa Bay Water planning initiatives.

4.2.3 Evaluation of Future Need Mitigation Costs

FNA should be applied during evaluation of new supply project alternatives, and the relationship between capital costs and risk mitigation should be evaluated. The probabilistic approach to FNA provides demand, supply, and need forecasts in terms of water quantity to support supply planning and capital investment decisions. Reducing probability of future need and increasing system reliability could involve additions to supply at potentially significant financial cost. FNA should be performed for each supply alternative, assuming the alternative is implemented in the supply system and assessing the resulting change in forecast need. These results can then be coupled with cost assessments of alternatives to allow the relationship between capital costs and risk mitigation to be evaluated. With this relationship defined, decision-makers would be able to identify projects offering the greatest improvement in reliability per unit cost.

Appendix A
Specification of Sectoral Per-Unit Models

The TAZ-level water use and socioeconomic database was used to estimate models of per-unit water use in single-family, multi-family, and non-residential sectors. The following sections describe the general linear estimating framework and specific model estimation procedures employed to build the final per-unit models.

A.1 General Linear Model

Multiple regression analysis was used to estimate models of per unit rates of use. Multiple regression is commonly used to estimate a direct and quantifiable relationship between a variable of interest (the dependent variable) and a set of independent variables that are hypothesized to explain changes in the variable of interest. The general linear regression model, as applied to per-unit demand forecasting, may be expressed as:

$$q_{s,t,i} = \alpha_{s,t,i} + \sum_m \beta_{m,s} X_{m,s,t,i} + \varepsilon_{s,t,i} \tag{A.1}$$

where

- $q_{s,t,i}$ = sectoral water demand for sector s (SF, MF, or NR), month-year t, and WDPA i

- $X_{m,s,t,i}$ = the value at month-year t in WDPA i of the mth explanatory or independent variable (e.g., household size, temperature) relevant to demand prediction for sector s

- $\alpha_{s,t,i}$ = an estimated model intercept term for sector s, WDPA i, and the calendar month of month-year t

- $\beta_{m,s}$ = estimated model parameter (specific to the model for sector s) measuring the relationship between $q_{s,t,i}$ and an explanatory variable $X_{m,s,t,i}$ (if the same variable appears in more than one sector model, that variable can have a different model parameter in each model)

- $\varepsilon_{s,t,I}$ = a random error term that denotes the difference between actual $q_{s,t,\mathrm{i}}$ and $q_{s,t,\mathrm{i}}$ as estimated from the model

In building per-unit demand models, $q_{s,t,i}$ and X values specific to TAZs were taken from the TAZ level database and treated as datapoints for regression. Generally, regression procedures select values for $\alpha_{s,t,i}$ and $\beta_{m,s}$ that best explain changes in $q_{s,t,i}$, or in statistical terms, those estimates of $\alpha_{s,t,i}$ and $\beta_{m,s}$ that minimize the sum of squared errors (also known as ordinary least squares (OLS) regression).

Often, the regression relationship uses natural logarithmic transformations of data such that Equation A.1 can be written as:

$$\ln q_{s,t,i} = \alpha_{s,t,i} + \sum_m \beta_{m,s} \ln X_{m,s,t,i} \qquad \text{(A.2)}$$

where the term "ln" denotes the natural logarithmic transformation. After estimating this type of transformed equation, the relationship would retain the following mathematical form after it is re-transformed from the logarithmic to raw scale:

$$q_{s,t,i} = e^{\alpha_{s,t,i}} \prod_m \left(X_{m,s,t,i}\right)^{\beta_{m,s}} \qquad \text{(1.3)}$$

Log-transformed equations and data were used in this study for several reasons. Log transformation enables interpretation of each model parameters as an elasticity, or the effect on the dependent variable of a one percent change in the value of an independent variable. For example, assuming marginal price in a month were log-transformed and, for a given log-transformed sectoral model containing only the marginal price of water, the β_s coefficient for price was -0.3:

$$q_{s,t,i} = e^{\alpha} X_{s,t,i}^{-0.3} \qquad \text{(1.4)}$$

where $X_{s,t,i}$ is price and α is the intercept term. If a one percent increase in price were enacted, say from \$2.00 to \$2.02 per 1000 gallons, the resulting percent difference in predicted demand would be

$$\Delta q_{s,t,i} = \frac{e^{\alpha}}{e^{\alpha}} \left(\frac{2.02^{-0.3} - 2^{-0.3}}{2^{-0.3}} \right) \times 100\% = -0.3\% \qquad \text{(1.5}$$

Thus, the percent change in demand resulting from a 1% increase in price is equal to the model coefficient for log-transformed price, β_s. This interpretation capability greatly facilitates experiential validation of model parameters and assessment of model sensitivity.

In addition to imparting interpretability was performed for the following reasons.

- Log transformation produced scaling benefits, such that the variance of the transformed variable values were within a shorter range

- For any particular time period, water use exhibited a skew to the right. Such a skew is a property of log-normally distributed variables.

- Past experience in selecting the most appropriate transformation of values (i.e., in using the Box-Cox technique for evaluating functional form) lent support to the natural log transformation.

A.2 Model Specification and Estimation Procedures

Separate log-transformed models were created to account for metered single-family, multi-family, and non-residential uses. The specification of variables used to define the statistical relationships was confined to variables that were available from the TAZ-level database. However, the database was designed to maximize, to the extent practicable, the availability of data on variables shown to affect water use among the three primary sectors under evaluation.

An important objective of the modeling process was to increase explanatory power of the models while achieving rational estimates for model parameters that were consistent with expected directions and magnitudes of influences found in the literature. This dual objective was approached through an iterative process of specifying alternative variables, screening of outlying data, and analyzing model residuals. Various robust estimation methods were employed in the process of iteratively specifying models, in attempts to *"dampen the influence of outlying cases...in an effort to provide a better fit for the majority of cases"* (Neter, 1996).

In general, and because of the very large number of available observations, the model estimation process encountered a great deal of variance that could not be explained. This result was attributed to the relative spatial and temporal extents spanned by the database, which was comprised of time-series and cross-sectional (or pooled) data. For each sector, cross-sectional variance dominated time-series variance. There were typically 1,000 or more cross-sections and only a maximum of 48 time-series observations on water sales for each cross-section. In many cases it was observed that independent variable variance exceeded relative variance in water use, a common situation when there are more cross-sections than there are time-periods.

Data screening measures were instituted to omit suspicious observations. In most cases, data points could be screened according to per-unit sectoral demand. Additional screening was accomplished by analyzing values for independent variables and correcting or omitting observations when values of explanatory values seemed extreme or nonsensical (*e.g.*, nonzero residential water use but no housing units).

In addition to socioeconomic and weather influences on demand, the models were also specified to account for unique location characteristics (*i.e.*, WDPA in which a TAZ was located) and other monthly-varying systematic behavior that could not be attributed to explanatory variables. These specifications took the form of adjustments to model intercepts (the α term in Equation A.3) by month and WDPA. Implementing these adjustments made it possible for rational coefficient estimates (i.e., estimates that displayed proper signs and expected numeric magnitudes) to be obtained. These adjustments also served to introduce member-specificity for predictions, a requirement of demand forecasting architecture described in Section 1.1.

Per-unit models differed by sector of applicability (whether single-family, multi-family, or non-residential), accounting units (occupied living units or number of employees), and model coefficients (see Tables A.3-A.5 in the next three sections). However, they all shared the same rate of use times driver forecasting methodology and some demographic, price, and weather model inputs. Specific data screening conventions and parameter estimation processes varied among the three models. The following sections discuss the particular actions that were necessary to estimate each sector model. Concise model equations are presented in Appendix B.

A.2.1 Estimation of Single-Family Model

The driver unit assumed for the single-family model was number of single-family households, such that the dependent variable for the single-family sectoral model was average single-family use per household in each TAZ and month. For estimation of the single-family model, TAZs whose average monthly single-family water use per household fell below 50 gallons per day (gpd) or exceeded 2,500 gpd were omitted. Generally, this eliminated TAZs that displayed average single-family use falling below reasonable levels of indoor sanitary use given prevailing household sizes. This action screened out very large values for TAZ-average water use considered clear outliers, while maintaining the characteristic right-hand skew normally found when analyzing single-family demand. Beyond removal of TAZs with incomplete data, outlier omissions amounted to less than 1% of the total number of remaining TAZ data points.

In the estimation process a value of 1 was added to precipitation, rainy day, and reclaimed fraction variables. These variables had several observations of zero and were likely to have zero-valued projections (note that ln(0) is undefined). Adding 1 to all observations of these variables preserved the relative magnitude of observations while allowing use of observations with 0 values.

Initial regression estimates for the single-family sector displayed inconsistent and uneven weather variable impacts and seemingly-dampened socioeconomic effects. As evidence, several statistically significant parameters for weather and socioeconomic variables were found, but they did not conform to expectations regarding sign or magnitude.

For the single-family residential model, it was necessary to isolate time-series variation from cross-sectional variation to capture more precise and rational weather effects. A 3-stage estimation process was adopted to accomplish this task, as shown in Tables A.1 to A.3. First, all observations (40,000+) were modeled as a function of time of year and weather variables alone (See Table A.1). Weather impacts were estimated successfully using 1^{st} degree polynomial distributed lags for temperature and precipitation such that lagged values had a gradually diminishing effect on contemporaneous use. Though other polynomial forms were tested, only a 1^{st} degree polynomial form generated coefficient estimates that were both statistically significant and rational in terms of sign, magnitude, and expected taper over lagged periods.

The second stage regressed historic average water use per month for each TAZ on socioeconomic variables only (Table A.2). This stage eliminated time-related variance and took advantage of cross-sectional data richness to isolate impacts of socioeconomic characteristics on demand.

In the third stage (Table A.3), all observations were re-pooled and parameter coefficient values were restricted to values obtained in the first and second stages of the process. The third stage also reintroduced both time and location as independent variables.

Table A.1
Stage 1 of Development of Single-Family Residential Per Unit Use Model:
Polynomial Distributed Lag Regression Parameter Estimates

Independent Variables	Parameter Estimate	Std Error	t-value	Pr > |t|
Ln (Max. Temp) – avg Ln (Max. Temp)	0.991853	0.0837	11.84	<.0001
1-month lag of Ln (Max. Temp) departure	0.905423	0.0499	18.14	<.0001
2-month lag of Ln (Max. Temp) departure	0.818993	0.0452	18.12	<.0001
3-month lag of Ln (Max. Temp) departure	0.732563	0.0753	9.73	<.0001
Ln (Precip. + 1) – avg Ln (Precip. + 1)	-0.02799	0.0069	-4.04	<.0001
1-month lag of Ln (Precip + 1) departure	-0.020797	0.0045	-4.56	<.0001
2-month lag of Ln (Precip + 1) departure	-0.013604	0.0034	-3.9	<.0001
3-month lag of Ln (Precip + 1) departure	-0.006411	0.0047	-1.36	0.1726
Ln (1” rainy days + 1) – avg Ln (1” rainy days)	-0.0151	0.0074	-2.03	0.0422
1-month lag of Ln (0.01” rainy days + 1)	-0.0246	0.0072	-3.39	0.0007
Number of Observations	42,899			
Regress R-Square	0.0449			
Root MSE	0.36977			

Table A.2
Stage 2 of Development of Single Family Per Unit Use Model:
Non-Temporal Socioeconomic Parameter Estimates

Independent Variables	Parameter Estimate	Std Error	t-value	Pr > |t|
Intercept	2.78671	0.20904	13.33	<.0001
Ln (Avg. household Income)	0.26199	0.02025	12.94	<.0001
Ln (Avg. Single-family housing units per acre)	-0.11679	0.01396	-8.36	<.0001
Ln (Avg. # of persons per single-family housing unit)	0.55784	0.05752	9.7	<.0001
Ln (Avg. marginal price of water and sewer)	-0.24779	0.07506	-3.3	0.001
Ln (fraction of reclaimed accounts in TAZ + 1)	-0.36585	0.08935	-4.1	<.0001
Regression Statistics				
Adjusted R-Square	0.3333			
Root MSE	0.303			
Dependent Mean	5.44409			
F-Value	104.38			
Prob > F	<0.0001			
Number of Observations	1,035			

Table A.3
Stage 3 of Development of Single-Family Residential
Per Unit Use Model: Final Parameter Estimates

Variable	Parameter Estimate	Std Error	t-value	Pr > |t|
Intercept	2.91062	0.00959	303.62	<.0001
Seasonal Intercept Adjustments*				
Intercept adjustment for February	-0.0007992	0.00721	-0.11	0.9118
Intercept adjustment for March	0.03439	0.00721	4.77	<.0001
Intercept adjustment for April	0.0974	0.00719	13.55	<.0001
Intercept adjustment for May	0.09819	0.00717	13.69	<.0001
Intercept adjustment for June	0.05658	0.00716	7.9	<.0001
Intercept adjustment for July	-0.00854	0.00717	-1.19	0.2333
Intercept adjustment for August	0.00004935	0.00716	0.01	0.9945
Intercept adjustment for September	0.03774	0.00716	5.27	<.0001
Intercept adjustment for October	0.09224	0.00776	11.89	<.0001
Intercept adjustment for November	0.08565	0.00776	11.04	<.0001
Intercept adjustment for December	0.04198	0.00774	5.42	<.0001
Weather Components				
Ln (Max. Temp) – avg Ln (Max. Temp)	0.99185	0	Infty	<.0001
1-month lag of Ln (Max. Temp) departure	0.90542	0	Infty	<.0001
2-month lag of Ln (Max. Temp) departure	0.81999	0	Infty	<.0001
3-month lag of Ln (Max. Temp) departure	0.73256	0	Infty	<.0001
Ln (Precip. + 1) – avg Ln (Precip. + 1)	-0.02799	0	-Infty	<.0001
1-month lag of Ln (Precip + 1) departure	-0.0208	0	-Infty	<.0001
2-month lag of Ln (Precip + 1) departure	-0.0136	0	-Infty	<.0001
3-month lag of Ln (Precip + 1) departure	-0.00641	0	-Infty	<.0001
Ln (0.01” rainy days + 1) – avg Ln (0.01” rainy days)	-0.0151	0	-Infty	<.0001
1-month lag of Ln (1” rainy days + 1)	-0.0246	0	-Infty	<.0001
Socioeconomic Components				
Ln (avg. household income)	0.26199	0	Infty	<.0001
Ln (avg. single-family housing units per acre)	-0.11679	0	-Infty	<.0001
Ln (avg. persons per single-family dwelling unit)	0.55785	0	Infty	<.0001
Ln (real marginal price of water and sewer)	-0.24779	0	-Infty	<.0001
Ln (fraction of reclaimed accounts + 1)	-0.36585	0	-Infty	<.0001

Table A.3
Stage 3 of Development of Single-Family Residential
Per Unit Use Model: Final Parameter Estimates

Variable	Parameter Estimate	Std Error	t-value	Pr > |t|
WDPA Intercept Adjustments				
Intercept adjustment for Pasco	-0.17354	0.0146	-11.88	<.0001
Intercept adjustment for City of Tampa	-0.1223	0.00857	-14.27	<.0001
Intercept adjustment for Northwest Hillsborough	-0.23185	0.00984	-23.55	<.0001
Intercept adjustment for City of St. Petersburg	-0.36252	0.00881	-41.13	<.0001
Intercept adjustment for South-Central Hillsborough	-0.26407	0.0093	-28.4	<.0001
Intercept adjustment for New Port Richey	-0.30522	0.01221	-24.99	<.0001
Intercept adjustment for Pinellas	-0.14491	0.00869	-16.68	<.0001
Regression Statistics				
Adjusted R-Square	0.389			
Root MSE	0.30927			
Dependent Mean	5.40784			
F-Value	1516.58			
Prob > F	<0.0001			
Number of Observations	42,857			

* January was omitted for intercept adjustment. It was mathematically necessary to omit one month.

The third stage served at least two purposes:

1. It allowed estimation of systematic seasonal and location effects, and

2. It provided a set of variables accounting for systematic errors in the first two stages, assuming that these errors in the first 2 stages related to unique location or seasonal effects.

A.2.2 Estimation of Multi-Family Model

Generally, the most difficult multi-family sector modeling issue is that a multi-family account can house more than one individual dwelling unit. Unlike the single-family sector where the number of accounts closely match housing units, multi-family accounts can differ considerably in the number of housing units served. Therefore, number of multi-family dwelling units was selected as a driver of multi-family demand because it was an easier and more meaningful variable to project into the future than number of multi-family accounts. The corresponding dependent variable for the multi-family sectoral model was multi-family use per housing unit averaged across each month and TAZ.

To derive the dependent variable for the multi-family model, it was necessary to divide total multi-family use in each TAZ and month by number of multi-family dwelling units in the corresponding TAZ and month. At the time of the study, only Pinellas and St. Petersburg members included number of housing units served by multi-family accounts in their billing data. For TAZs in other WDPAs, estimates of the number of multi-family units were obtained by estimating number of units per multi-family account, then multiplying this ratio by number of multi-family accounts in each TAZ and month. Units per multi-family account was estimated by deriving and applying a simple regression relationship. In the regression, number of multi-family dwelling units in a TAZ and month was treated as a dependent variable, while number of multi-family accounts for each TAZ and month and average number of multi-family housing units for each year and WDPA were treated as independent variables. Dependent variable data were available for only St. Petersburg and Pinellas WDPAs, while independent variable data were available for all WDPAs[1]. The resulting regression model[2] was used to estimate monthly numbers of multi-family units in TAZs that did not belong to Pinellas and St. Petersburg WDPAs.

Because number of multi-family accounts varied each month due to billing cycles, predicted number of multi-family units from the regression was not used directly as TAZ-level estimates. Instead, these predictions were used to estimate average number of multi-family units per account in each TAZ. Monthly units per account for TAZs outside St. Petersburg and Pinellas were estimated by dividing regression predictions by number of accounts in each TAZ. Observed (St. Petersburg and Pinellas TAZs) and predicted (TAZs in all other WDPAs) monthly values of units per account were then averaged across all months for 1999, the first year in the water use database. These year-1999 units per account estimates were used for multi-family unit estimations in all months and years from 1999 to 2002. Number of multi-family units in each TAZ and month was estimated by multiplying year-1999 average multi-family units per account for that TAZ by the number of multi-family accounts in that TAZ and month.

The requirement to estimate number of units by TAZ and month, together with the wide observed variance in water use per account in the multi-family sector, made screening of observations more complicated than in the single-family sector. Omission of TAZs displaying less than 25 gallons per day per unit and more than 1,000 gallons per day per unit provided the best and most rational modeling results. These omissions led to less than 5% of the total number of TAZ-level observations. The multi-family model was estimated in one step; coefficients for the multi-family model are provided in Table A.4.

---

1 *Number of accounts per TAZ and month was available from the geocoded billing database. Number of multi-family units per WDPA was obtained from Experian and geocoded by GIS Solutions, Inc.*

2 *Number of multi-family units in any particular TAZ outside Pinellas and St. Pete was estimated as: Ln(# MF units in TAZ and month) = -0.18001 + 0.41613 Ln(# MF accounts in TAZ and month) + 0.65014 Ln(avg # of MF units in WDPA in year).*

Table A.4
Multi-Family Residential Per Unit Use Model: Final Parameter Estimates

Independent Variables	Parameter Estimate	Standard Error	t-value	Pr > |t|
Intercept	1.47578	0.11831	12.47	<.0001
WDPA Intercept Adjustments				
Intercept adjustment for New Port Richey	0.48567	0.03035	16.00	<.0001
Intercept adjustment for City of St. Petersburg[3]	0.48567			
Intercept adjustment for Pasco[4]	0.48567			
Intercept adjustment for Northwest Hillsborough[5]	0.48567			
Socioeconomic and Weather Components				
Ln (avg. number of multi-family housing units per acre)	-0.35254	0.00440	-80.13	<.0001
Ln (avg. household income)	0.37054	0.01093	33.90	<.0001
Ln (fraction of reclaimed accounts + 1)	-0.38540	0.05561	-6.93	<.0001
Ln (precipitation + 1)	-0.01717	0.00643	-2.67	0.0076
Regression Statistics				
Adjusted R-Square	0.2976			
Root MSE	0.75188			
Dependent Mean	4.72633			
F-Value	1731.01			
Prob > F	<0.0001			
Number of Observations	20,418			

In the estimation process a value of 1 was added to precipitation and reclaimed fraction variables. These variables had several observations of zero and were likely to have zero-valued projections (note that ln(0) is undefined). Adding 1 to all observations of these variables preserved the relative magnitude of observations while allowing use of observations with 0 values.

A.2.3 Estimation of Non-Residential Model

Non-residential utility customers can use water for cooling, outside watering, and as a direct input for the production of goods and services in addition to traditional domestic

---

3 *These intercept adjustments were originally found to be insignificant (i.e. equal to zero) for TAZ-level modeling. However, during validation (Section 1.5), it was found that assigning New Port Richey's adjustment parameter to Northwest Hillsborough, St Petersburg, and Pasco WDPAs improved prediction of WDPA-level data for 2002.*

4 *See footnote above.*

5 *See footnote above.*

purposes. The non-residential sector therefore generally displays more heterogeneous water use than single-family or multi-family sectors. Similarly to the multi-family sector, number of non-residential accounts does not provide adequate information for projecting future non-residential use. Number of employees is a pertinent driver variable. Therefore, average water use per employee in each TAZ was adopted as the measure of per unit use for this study.

Total employment data were available only annually at the TAZ level for the 1999-2002 period[6]. The SAS Expand procedure was used to interpolate annual employment values into a series of monthly employment values for each TAZ over the 1999-2002 period, removing abrupt changes in number of employees between years and producing smooth monthly transitions between observed yearly employment values. Monthly water sales per TAZ were divided by these employment values to derive the dependent variable, average monthly water use per employee per day in each TAZ.

Because of the heterogeneous nature of non-residential demand, there is a general lack of benchmarking data for typical non-residential water use per employee. Therefore, screening of potential outliers was qualitative. Very small or large values were possible depending on the types of business and actual water-using processes involved with a non-residential customer. Aside from omission of observations due to questionable socioeconomic data, only those observations that fell below a TAZ average of 5 gallons per employee per day were screened out prior to model estimation. As in the single-family and multi-family cases, these omissions amounted to less than 5% of the total number of TAZ-level use observations.

The non-residential model contained variables reflecting fraction of total employment in a geographical area belonging to each of three broad employment categories: commercial, industrial, and services. In the estimation process a value of 1 was added to the fraction of employment in each employment grouping. These variables had several observations of zero and were likely to have zero-valued projections (note that ln(0) is undefined). Adding 1 to all observations of these variables preserved the relative magnitude of observations while allowing use of observations with 0 values.

The non-residential model was successfully estimated in one step. Coefficients for the non-residential model are provided in Table A.5.

---

6 *Data obtained from Experian through GIS Solutions, Inc., July 2002.*

Table A.5
Non-Residential Per Unit Use Model: Final Parameter Estimates

Independent Variables	Parameter Estimate	Standard Error	t-value	Pr > |t|
Intercept	1.6167	0.16823	9.61	<.0001
Socioeconomic Components				
Ln (fraction of total employment in industrial sector + 1)	0.34798	0.15571	2.23	0.0254
Ln (fraction of total employment in commercial sector + 1)	1.01109	0.15204	6.65	<.0001
Ln (fraction of total employment in services sector + 1)	1.19036	0.18008	6.61	<.0001
Ln (average household income)	0.12075	0.01033	11.69	<.0001
Weather Components				
Ln (Precip. + 1) - avg Ln (Precip. + 1)	-0.04958	0.01009	-4.91	<.0001
1-month lag of Ln (Precip + 1) departure	-0.03609	0.00976	-3.7	0.0002
2-month lag of Ln (Precip + 1) departure	-0.01708	0.00972	-1.76	0.0789
Regression Statistics				
Adjusted R-Square	0.0221			
Root MSE	0.98107			
Dependent Mean	3.44834			
F-Value	75.66			
Prob > F	<0.0001			
Number of Observations	39,727			

A.3 Additivity of Sectoral Per-Unit Demand Models

The forecast models and accounting framework assumed temporal and geographic independence and thus the forecasts were treated as additive across geographies, sectors, and months. Aside from implicit correlations among data inputs, model predictions for any given sector, month, or location were assumed to not depend on corresponding predictions in any other sector, month, or location.

Appendix B
Specific Equations Forming the Model

This appendix details the mathematics composing the Tampa Bay Water Point Demand Forecasting Model. The model was developed jointly by Hazen and Sawyer, P.C. of Tampa, FL and PMCL of Carbondale, IL.

The Tampa Bay Water Point Demand Forecasting Model consists of econometric equations correlating historical demographics and weather characteristics for a specified month and Water Demand Planning Area, or *WDPA*, to observed average daily potable water usage within that WDPA during that month. Equations describe average daily potable water usage in a given WDPA and month for three classes of Tampa Bay Water customers, or sectors: single-family (*SF*), multi-family (*MF*), and non-residential (*NR*), as well as for wholesale deliveries (*WS*) of potable water to non-member utilities and for unbilled (*OUW*) water consumption. Summed together, these five components of water usage represent the total water usage in the corresponding WDPA and month. Demographic and weather projections for future months are then applied to the model equations to determine forecasts of average daily potable water usage as a function of geography and month.

The model equations are evaluated in a stepwise fashion, as shown in Figure B.1. First, user-supplied demographic and weather projections for a given month and WDPA are used to determine forecasted average monthly water use per household (for SF), housing unit (for MF), and employee (for NR). These per-unit demands are then multiplied by projected values for the number of units in that month and WDPA to produce forecast total water demand for each sector. Each forecasted sectoral demand is adjusted, or calibrated, using a comparison of Water Year 2002 observed and forecast demand for that month, sector, and WDPA. Calibrated total sectoral demands are then summed to produce total retail demand. Wholesale deliveries are determined as a fraction of retail demand, and unbilled consumption is determined as a fraction of total demand. Retail, wholesale, and unbilled demand are finally summed to produce the total potable water demand for the given WDPA and month.

In this appendix and the original forecast, time- and geography-specific variables are indexed at various levels of temporal and geographic specificity. This does *not* imply that all variable values should be specified in these terms in future work. Any variable could potentially be varied by month-and-year and by WDPA. Rather, the indexing is meant to reflect the specificity of the original point forecast.

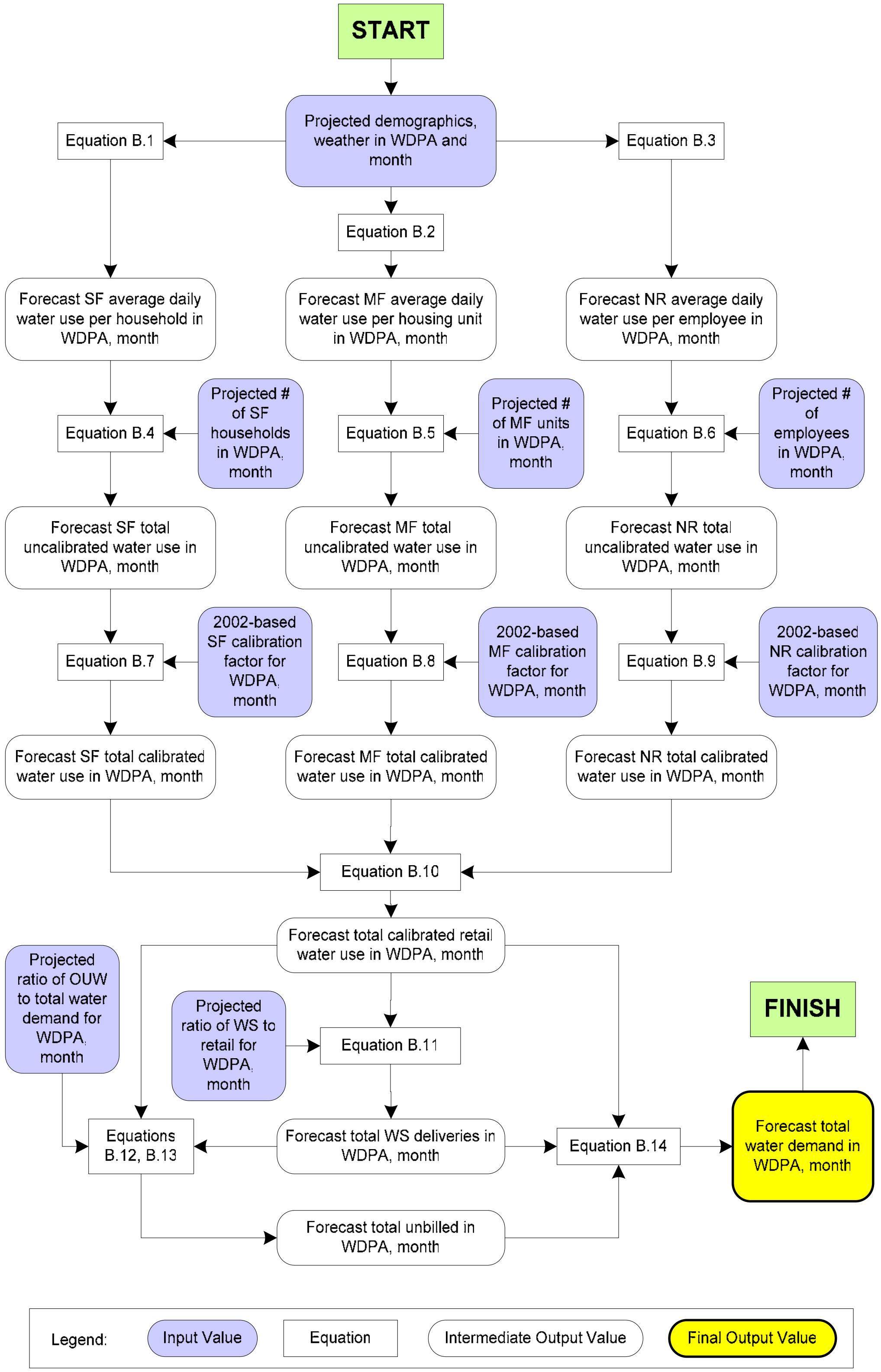


Figure B.1 – Procedural flowchart for executing the Tampa Bay Water Point Demand Forecasting Model

B.1 Sectoral Per-Unit Water Demand Calculations

To arrive at monthly MGD forecasts by WDPA, calculations are first performed to determine water use in gallons per day per driver unit for single-family (SF), multi-family (MF), and non-residential (NR) sectors within each WDPA. Driver units are number of housing units in a WDPA for SF, number of dwelling units for MF, and number of employees in a WDPA for NR.

B.1.1 Single-Family Per-Unit Water Use Equation

The single-family per-unit water use in a WDPA is calculated as follows:

$$
\begin{aligned}
\ln\left(q_{WDPA,m,y,SF}\right) = {} & \alpha_{SF} + \alpha_m + \alpha_{WDPA,SF} + \beta_{INC,SF}\ln\left(INC_{y,WDPA}\right) + \beta_{HOU,SF}\ln\left(HOU_{SF,y,WDPA}\right) \\
& + \beta_{PPH,SF}\ln\left(PPH_{SF,y,WDPA}\right) + \beta_{RMP}\ln\left(RMP_{m,y,WDPA}\right) \\
& + \beta_{\Delta T,lag0}\left[\ln\left(T_{MAX,m,y,WDPA}\right) - \overline{\ln\left(T_{MAX,m}\right)}\right] + \beta_{\Delta T,lag1}\left[\ln\left(T_{MAX,m-1,y,WDPA}\right) - \overline{\ln\left(T_{MAX,m-1}\right)}\right] \\
& + \beta_{\Delta T,lag2}\left[\ln\left(T_{MAX,m-2,y,WDPA}\right) - \overline{\ln\left(T_{MAX,m-2}\right)}\right] + \beta_{\Delta T,lag3}\left[\ln\left(T_{MAX,m-3,y,WDPA}\right) - \overline{\ln\left(T_{MAX,m-3}\right)}\right] \\
& + \beta_{\Delta R,lag0,SF}\left[\ln\left(R_{m,y,WDPA}+1\right) - \overline{\ln\left(R_m+1\right)}\right] + \beta_{\Delta R,lag1,SF}\left[\ln\left(R_{m-1,y,WDPA}+1\right) - \overline{\ln\left(R_{m-1}+1\right)}\right] \\
& + \beta_{\Delta R,lag2,SF}\left[\ln\left(R_{m-2,y,WDPA}+1\right) - \overline{\ln\left(R_{m-2}+1\right)}\right] \\
& + \beta_{\Delta R,lag3}\left[\ln\left(R_{m-3,y,WDPA}+1\right) - \overline{\ln\left(R_{m-3}+1\right)}\right] \\
& + \beta_{RD0.01,lag0}\left[\ln\left(RD0.01_{m,y,WDPA}+1\right) - \overline{\ln\left(RD0.01_m+1\right)}\right] + \beta_{RD1,lag1}\ln\left(RD1_{m-1,y,WDPA}+1\right) \\
& + \beta_{RECL,SF}\ln\left(RECL_{SF,y,WDPA}+1\right)
\end{aligned}
\qquad \text{(B.1)}
$$

where

- y = year,
- m = calendar month (Jan = 0, Feb = 1, ... Dec = 11),
- $WDPA \in$ {Pasco, NPR, SCH, NWH, Tampa, St Pete, Pinellas},

and explanatory variables include:

- $INC_{y,WDPA}$ = median per capita household income in year y and $WDPA$ ($ per year per capita)
- $HOU_{SF,y,WDPA}$ = density of single-family housing units in year y and $WDPA$ (housing units per acre)

- $PPH_{SF,y,WDPA}$ = persons per single-family housing unit in year y and WDPA (persons per housing unit)

- $RMP_{m,y,WDPA}$ = real marginal price in month m, year y and WDPA ($ per 1000 gallons)

- $T_{MAX,m,y,WDPA}$ = maximum daily temperature in month m, year y and WDPA (°F)[1]

- $\overline{\ln T_{MAX,m}}$ = average of natural log historical maximum daily temperature in month m (a constant for m, values given in Table B.1)

- $R_{m,y,WDPA}$ = total rainfall in month m, year y and WDPA (inches)[2]

- $\overline{\ln(R_m + 1)}$ = average of natural log of 1 + historical total rainfall in month *m* (a constant for m, values given in Table B.1)

- $RD0.01_{m,y,WDPA}$ = total number of rainy days (≥ 0.01 inches) in month *m*, year *y* and *WDPA* (number of days)[3]

- $\overline{\ln(RD0.01_m + 1)}$ = average of natural log of 1 + historical total number of rainy days (≥ 0.01 inches) in month m (a constant for m, values given in Table B.1)

- $RD1_{m,y,WDPA}$ = total rainy days (≥ 1 inches) in month *m*, year *y* and *WDPA* (number of days)[4]

- $RECL_{SF,y,WDPA}$ = fraction of single-family accounts using reclaimed water in year *y* and *WDPA* (fraction)

The various α and β terms are empirical constants determined by fitting model predictions to historical water use data (values listed in Table B.2).

B.1.2 Comments on Single-Family Per-Unit Water Use Terms

In the point forecast, the income term, $INC_{y,WDPA}$, housing density term, $HOU_{SF,y,WDPA}$, and persons per household term, $PPH_{SF,y,WDPA}$, were each held constant across all months of each year. $HOU_{SF,y,WDPA}$ was only varied by WDPA (and held constant with time), while $PPH_{SF,y,WDPA}$ varied across years and WDPAs. The same value of $INC_{y,WDPA}$

1 *Varied only by calendar month and WDPA in the original point forecast. In the probabilistic forecast, one distribution was specified for each calendar month, but independent random values were drawn from these distributions in different years.*

2 *See footnote above.*

3 *See footnote above.*

4 *See footnote above.*

was used in all sectoral per-unit water use equations, while $HOU_{SF,y,WDPA}$ and $PPH_{SF,y,WDPA}$ terms were specific to the single-family equation (Equation B.1). As will be shown, an analogous but differently-valued term for multi-family housing density was applied in the multi-family water use equation (Equation B.2, below). Persons per household was only significant in the single-family equation.

Table B.1
Normal Weather Values Used in Weather Deviation Terms of Equation 1

	$\overline{\ln T_{MAX,m}}$	$\overline{\ln(R_m + 1)}$	$\overline{\ln(RD0.01_m + 1)}$
Jan (m = 0)	4.281	1.210	2.09
Feb (m = 1)	4.307	1.211	1.99
Mar (m = 2)	4.370	1.309	1.98
Apr (m = 3)	4.421	0.983	1.73
May (m = 4)	4.485	1.257	1.90
Jun (m = 5)	4.515	1.924	2.55
Jul (m = 6)	4.524	2.058	2.79
Aug (m = 7)	4.522	2.157	2.82
Sep (m = 8)	4.507	1.961	2.57
Oct (m = 9)	4.451	1.104	1.92
Nov (m = 10)	4.377	0.986	1.88
Dec (m = 11)	4.308	1.070	1.89

Table B.2
Model Coefficients for Single-Family Per-Unit Water Use Equation (Equation B.1)

α_{SF}	2.91062	$\beta_{INC,SF}$	0.261989
		$\beta_{HOU,SF}$	-0.11679
α_m:		$\beta_{PPH,SF}$	0.557845
Jan (m = 0)	0	β_{RMP}	-0.24779
Feb (m = 1)	-0.0007992		
Mar (m = 2)	0.03439		
Apr (m = 3)	0.0974	$\beta\Delta_{T,lag0}$	0.99185
May (m = 4)	0.09819	$\beta\Delta_{T,lag1}$	0.90542
Jun (m = 5)	0.05658	$\beta\Delta_{T,lag2}$	0.81999
Jul (m = 6)	-0.00854	$\beta\Delta_{T,lag3}$	0.73256
Aug (m = 7)	0.00004935		
Sep (m = 8)	0.03774		
Oct (m = 9)	0.09224	$\beta\Delta_{R,lag0,SF}$	-0.02799
Nov (m = 10)	0.08565	$\beta\Delta_{R,lag1,SF}$	-0.0208
Dec (m = 11)	0.04198	$\beta\Delta_{R,lag2,SF}$	-0.0136
		$\beta\Delta_{R,lag3,SF}$	-0.00641
$\alpha_{WDPA,SF}$:			
WDPA = Pinellas	-0.14491		
WDPA = St. Pete	-0.36252	$\beta_{RD0.01,lag0}$	-0.0151
WDPA = NPR	-0.30522	$\beta_{RD1,lag1}$	-0.0246
WDPA = Pasco	-0.17354		
WDPA = Tampa	-0.1223		
WDPA = NWH	-0.23185	$\beta_{RECL,SF}$	-0.36585
WDPA = SCH	-0.26407		

Real marginal price, $RMP_{m,y,WDPA}$, was only significant in the single-family per-unit water use equation. Historically, $RMP_{m,y,WDPA}$ varies across year-and-month and WDPA. In the point forecast, however, this term was only varied by calendar month and WDPA.

Natural logs of maximum monthly temperature, $\ln T_{MAX,m,y,WDPA}$, and normal maximum monthly temperature, $\overline{\ln T_{MAX,m}}$, were used to determine deviation from normal log maximum temperature in Equation B.1. Values of $\overline{\ln T_{MAX,m}}$ differ among calendar months only. Projected values of $T_{MAX,m,y,WDPA}$ were varied across calendar months and WDPAs in the point forecast. In the probabilistic forecast, a single distribution was applied to each calendar month and WDPA, but different random values were drawn from those distributions for each month in each year. In the point forecast, projected values of

$T_{MAX,m,y,WDPA}$ were assumed as fixed normal values of monthly average maximum daily temperature.

Total monthly rainfall, $R_{m,y,WDPA}$, and log one-plus-normal total monthly rainfall, $\overline{\ln(R_m+1)}$, were used to determine deviation from normal log rainfall in Equation B.1. Values of $\overline{\ln(R_m+1)}$ differ among calendar months only. Projected values of $R_{m,y,WDPA}$ were varied across calendar months and WDPAs in the point forecast. In the probabilistic forecast, a single distribution was applied to each calendar month and WDPA, but different random values were drawn from those distributions for each month in each year. In the point forecast, projected values of $R_{m,y,WDPA}$ were assumed as fixed normal average monthly rainfall totals.

Total monthly 0.01-rainy days, $RD0.01_{m,y,WDPA}$, and log one-plus-normal total monthly 0.01-rainy days, $\overline{\ln(RD0.01_m+1)}$, were used to determine deviation from normal log 0.01-rainy days in Equation B.1. Values of $\overline{\ln(RD0.01_m+1)}$ differ among months and WDPAs. Projected values of $RD0.01_{m,y,WDPA}$ varied across calendar month and WDPA in the point forecast. In the probabilistic forecast, a single distribution was applied to each calendar month and WDPA, but different random values were drawn from those distributions for each month in each year. Projected values of $RD0.01_{m,y,WDPA}$ in the point forecast were assumed as normal average rainy day counts. It should be noted that all values of $RD0.01_{m,y,WDPA}$ and $\overline{RD0.01_m}$ contained fractional components, including historical values of $RD0.01_{m,y,WDPA}$. Noninteger values for these terms arose because they were determined by weighted average of weather stations, with weights corresponding to distance of stations from the corresponding WDPA centroid.

Equation B.1 illustrates that the one-month lag of total monthly 1-inch rainy days, $RD1_{m,y,WDPA}$, has a significant effect on single-family water use. Unlike temperature, rainfall, and 0.01-rainy days, however, 1-inch rainy days is not represented as a log deviation from normal in the equation. Projected values of $RD1_{m,y,WDPA}$ generally varied with calendar month and WDPA in the point forecast and with year-and-month and WDPA in the probabilistic forecast. Projected values of $RD1_{m,y,WDPA}$ in the point forecast were assumed as normal average rainy day counts, while in the probabilistic forecast, simulated values were selected from distributions specified for each calendar month. Similarly to 0.01-rainy days, all values for $RD1_{m,y,WDPA}$ were noninteger, reflecting weighted averaging of rainy day measurements from multiple stations.

Finally, fraction of single-family accounts using reclaimed water, $RECL_{SF,y,WDPA}$, was assumed to be constant with time for each WDPA but to vary among WDPAs. Also, the $RECL_{SF,y,WDPA}$ term is specific to the single-family equation (Equation B.1). An analogous but differently-valued term for multi-family reclaimed fraction is applied in the multi-family water use equation (Equation B.2, below).

B.1.3 Multi-Family Per-Unit Water Use Equation

The multi-family per-unit water use in a WDPA is calculated as follows:

$$\ln\left(q_{WDPA,m,y,MF}\right) = \alpha_{MF} + \alpha_{WDPA,MF} + \beta_{INC,MF}\ln\left(INC_{y,WDPA}\right) + \beta_{HOU,MF}\ln\left(HOU_{MF,y,WDPA}\right) + \beta_{R}\ln\left(R_{m,y,WDPA}+1\right) + \beta_{RECL,MF}\ln\left(RECL_{MF,y,WDPA}+1\right) \quad \text{(B.2)}$$

where

- y = year,
- m = calendar month (Jan = 0, Feb = 1, ... Dec = 11),
- $WDPA \in$ {Pasco, NPR, SCH, NWH, Tampa, St Pete, Pinellas},

and explanatory variables include:

- $INC_{y,WDPA}$ = median per capita household income in year y and *WDPA* ($ per year per capita)
- $HOU_{MF,y,WDPA}$ = density of multi-family housing units in year y and *WDPA* (housing units per acre)
- $R_{m,y,WDPA}$ = total rainfall in month m, year y and *WDPA* (inches)
- $RECL_{MF,y,WDPA}$ = fraction of multi-family accounts using reclaimed water in year y and *WDPA* (fraction)

The various α and β terms are empirical constants determined by fitting model predictions to historical water use data (values listed in Table B.3).

Table B.3
Model Coefficients For Multi-Family Per-Unit Water Use Equation (Equation A.2)

α_{MF}	1.47578	$\alpha_{WDPA,MF}$:	
		WDPA = Pinellas	0
$\beta_{INC,MF}$	0.37054	WDPA = St. Pete	0.48567
$\beta_{HOU,MF}$	-0.35254	WDPA = NPR	0.48567
β_{R}	-0.01717	WDPA = Pasco	0.48567
$\beta_{RECL,MF}$	-0.3854	WDPA = Tampa	0
		WDPA = NWH	0.48567
		WDPA = SCH	0

B.1.4 Comments on Multi-Family Per-Unit Water Use Terms

The income term, $INC_{y,WDPA}$, and rainfall term, $R_{m,y,WDPA}$, are identical to terms used in Equation B.1. Comments regarding these terms from prior discussion are applicable here as well.

In the point forecast, the multi-family housing density term, $HOU_{MF,y,WDPA}$, and the multi-family fractional reclaimed term, $REC_{LMF,y,WDPA}$, were held constant with time but varied among WDPAs. These terms are analogous to $HOU_{SF,y,WDPA}$ and $REC_{LSF,y,WDPA}$ from Equation B.1 but are specific to the multi-family case and had different values from their single-family analogs in the original forecasts.

B.1.5 Non-Residential per-Unit Water Use Equation

The non-residential per-unit water use in a WDPA is calculated as follows:

$$\begin{aligned}\ln\left(q_{WDPA,m,y,NR}\right) = \; & \alpha_{NR} + \beta_{INC,NR}\ln\left(INC_{y,WDPA}\right) + \beta_{COM}\ln\left(COM_{y,WDPA}+1\right) + \beta_{IND}\ln\left(IND_{y,WDPA}+1\right) \\ & + \beta_{SER}\ln\left(SER_{y,WDPA}+1\right) + \beta_{\Delta R,lag0,NR}\left[\ln\left(R_{m,y,WDPA}+1\right) - \overline{\ln\left(R_m+1\right)}\right] \\ & + \beta_{\Delta R,lag1,NR}\left[\ln\left(R_{m-1,y,WDPA}+1\right) - \overline{\ln\left(R_{m-1}+1\right)}\right] \\ & + \beta_{\Delta R,lag2,NR}\left[\ln\left(R_{m-2,y,WDPA}+1\right) - \overline{\ln\left(R_{m-2}+1\right)}\right]\end{aligned} \quad \text{(B.3)}$$

where

- y = year,
- m = calendar month (Jan = 0, Feb = 1, ... Dec = 11),
- $WDPA \in$ {Pasco, NPR, SCH, NWH, Tampa, St Pete, Pinellas},

and explanatory variables include:

- $INC_{y,WDPA}$ = median per capita household income in year y and WDPA ($ per year per capita)
- $COM_{y,WDPA}$ = fraction of total employment in Commercial entities in year *y* and WDPA (fraction)
- $IND_{y,WDPA}$ = fraction of total employment in Industrial entities in year *y* and *WDPA* (fraction)

- $SER_{y,WDPA}$ = fraction of total employment in Service entities in year y and WDPA (fraction)

- $R_{m,y,WDPA}$ = total rainfall in month m, year *y* and *WDPA* (inches)[5]

- $\overline{\ln(R_m + 1)}$ = average of natural log of 1 + historical total rainfall in month m (a constant for m, values given in Table B.1)

The various α and β terms are empirical constants determined by fitting model predictions to historical water use data (values listed in Table B.4).

Table B.4
Model Coefficients For Non-Residential Per-Unit Water Use Equation

α_{SF}	1.6167
$\beta_{INC,NR}$	0.12075
β_{COM}	1.01109
β_{IND}	0.34798
β_{SER}	1.19036
$\beta\Delta_{R,lag0,NR}$	-0.04958
$\beta\Delta_{R,lag1,NR}$	-0.03609
$\beta\Delta_{R,lag2,NR}$	-0.01708

B.1.6 Comments on Non-Residential Per-Unit Water Use Terms

The income term, $INC_{y,WDPA,}$ and rainfall terms, $R_{m,y,WDPA}$ and $\overline{R_m}$, are identical to terms used in Equation B.1. Comments regarding these terms from prior discussion are applicable here as well.

The fractional employment terms, $COM_{y,WDPA}$, $IND_{y,WDPA}$, and $SER_{y,WDPA}$, were varied by year and WDPA in the point forecast. Values of $COM_{y,WDPA}$, $IND_{y,WDPA}$, and $SER_{y,WDPA}$ should always sum to 1, and did so for all year/WDPA combinations in the original forecasts.

5 *Varied only by calendar month and WDPA in original point forecast, but varied by month-and-year and WDPA in the probabilistic forecast.*

B.2 Sectoral Uncalibrated Total Water Demand Calculations

To determine total water usage in each month-and-year, sector, and WDPA, projected monthly values for driver variables, $N_{Sector,m,y,WDPA}$, are specified in each sector and WDPA. For the SF sector, the driver variable is defined as the number of SF housing units in a WDPA, month, and year. The MF driver variable represents the number of dwelling units (*e.g.*, individual apartments, not apartment buildings or complexes) in a WDPA, month, and year. For the NR sector, the driver variable is the total number of employees in a WDPA, month, and year. Uncalibrated total water use for each sector, month-and-year, and WDPA is then determined by multiplying the corresponding per-unit water usage by a projected value for the driver variable in that sector and WDPA for that month and year.

B.2.1 Single-Family Uncalibrated Water Use Equation

Total uncalibrated SF water use (in million gallons per day, or MGD) in a month, year, and WDPA, $Q^{uncal}_{SF,m,y,WDPA}$, is calculated using Equation B.4:

$$Q^{uncal}_{SF,m,y,WDPA} = 10^{-6}\left(N_{SF,m,y,WDPA}\right)\left(q_{SF,m,y,WDPA}\right) \tag{B.4}$$

where:

- 10^{-6} converts from gallons per day to MGD
- $N_{SF,m,y,WDPA}$ = number of single family households in a WDPA, month, and year
- $q_{SF,m,y,WDPA}$ = single-family daily water use per household in a WDPA, month, and year (calculated using Equation B.1)

In the point forecast, the driver variable $N_{SF,m,y,WDPA}$ was assumed to vary across WDPAs and years but to be constant from month to month within each year in a WDPA.

B.2.2 Multi-Family Uncalibrated Water Use Equation

Total uncalibrated MF water use (in million gallons per day, or MGD) in a month, year, and WDPA, $Q^{uncal}_{MF,m,y,WDPA}$, is calculated using Equation B.5:

$$Q^{uncal}_{MF,m,y,WDPA} = 10^{-6}\left(N_{MF,m,y,WDPA}\right)\left(q_{MF,m,y,WDPA}\right) \tag{B.5}$$

where:

- 10^{-6} converts from gallons per day to MGD
- $N_{MF,m,y,WDPA}$ = number of multi-family units in a WDPA, month, and year
- $q_{MF,m,y,WDPA}$ = multi-family daily water use per multi-family unit in a WDPA, month, and year (calculated using Equation B.2)

In the point forecast, the driver variable $N_{MF,m,y,WDPA}$ was assumed to vary across WDPAs and years but to be constant from month to month within each year in a WDPA.

B.2.3 Non-Residential Uncalibrated Water Use Equation

Total uncalibrated NR water use (in million gallons per day, or MGD) in a month, year, and WDPA, $Q^{uncal}_{NR,m,y,WDPA}$, is calculated using Equation B.6:

$$Q^{uncal}_{NR,m,y,WDPA} = 10^{-6}\left(N_{NR,m,y,WDPA}\right)\left(q_{NR,m,y,WDPA}\right) \tag{B.6}$$

where:

- 10^{-6} converts from gallons per day to MGD
- $N_{NR,m,y,WDPA}$ = total number of employees in a WDPA, month, and year
- $q_{NR,m,y,WDPA}$ = non-residential daily water use per employee in a WDPA, month, and year (calculated using Equation B.3)

In the point forecast, the driver variable $N_{NR,m,y,WDPA}$ was assumed to vary across WDPAs and years but to be constant from month to month within each year in a WDPA.

B.3 Calibrated Total Sectoral and Total Retail Demand Calculations

All point model forecasts are calibrated to a baseline year of measured water demand such that the first forecast year's demand matches the baseline demand. Calibration is based on the assumption that predictive discrepancy in each month of the baseline year is the result of small, systematic, month-specific bias from unmodeled demand influences. If this bias is small and consistent for each month in successive years, removal of this discrepancy can improve accuracy of the overall forecast, especially for short forecast horizons.

Calibration factors, $k_{Sector,m,WDPA}$ are determined for each calendar month, sector, and WDPA. Each factor represents the ratio between observed sectoral demand and forecast uncalibrated demand for each sector, WDPA, and calendar month in a baseline year where both forecast and measured demand are available. In the point and probabilistic forecasts, Water Year 2002 was used as the baseline year.

Each monthly forecast uncalibrated sectoral water demand is then corrected by its corresponding calibration factor:

$$Q^{cal}_{SF,m,y,WDPA} = \left(k_{SF,m,WDPA}\right)\left(Q^{uncal}_{SF,m,y,WDPA}\right) \tag{B.7}$$

$$Q^{cal}_{MF,m,y,WDPA} = \left(k_{MF,m,WDPA}\right)\left(Q^{uncal}_{MF,m,y,WDPA}\right) \tag{B.8}$$

$$Q^{cal}_{NR,m,y,WDPA} = \left(k_{NR,m,WDPA}\right)\left(Q^{uncal}_{NR,m,y,WDPA}\right) \tag{B.9}$$

where $k_{SF,m,WDPA}$, $k_{MF,m,WDPA}$, and $k_{NR,m,WDPA}$ *are* calibration coefficients for SF, MF, and NR sectors in each calendar month m and WDPA, and $Q^{uncal}_{SF,m,y,WDPA}$, $Q^{uncal}_{MF,m,y,WDPA}$, and $Q^{uncal}_{NR,m,y,WDPA}$ are uncalibrated sectoral water demands from Equations B.4, B.5, and B.6, respectively.

It should be noted that $k_{Sector,m,WDPA}$ in the point forecast varied from 0.56 to 1.69, indicating, at the extremes, high fractional discrepancy between observed and forecast uncalibrated demand. However, it was always the case that extremely low or high calibration factors corresponded to relatively small water demand values and that larger water demand values had calibration values close to 1. Therefore, absolute demand discrepancies corresponding to extreme calibration coefficients never caused large absolute demand calibrations. Calibration therefore appeared to be a valid procedure in spite of the large calibration coefficients. Further analysis may be worthwhile to define when an absolute demand discrepancy would be large enough to invalidate calibration and suggest a refitting of the model.

Total calibrated forecast retail water use (in million gallons per day, or MGD) in a month, year, and WDPA, $Q^{cal}_{RET,m,y,WDPA}$, is simply calculated by summing the individual sectoral terms:

$$Q^{cal}_{RET,m,y,WDPA} = Q^{cal}_{SF,m,y,WDPA} + Q^{cal}_{MF,m,y,WDPA} + Q^{cal}_{NR,m,y,WDPA} \tag{B.10}$$

From this point forward, $Q^{cal}_{RET,m,y,WDPA}$ will be referred to as total retail water use.

B.4 Wholesale Deliveries, Unbilled Use, and Total Monthly Water Demand Calculations

In addition to total retail water use, wholesale deliveries and unbilled water use is forecast for each WDPA, month, and year. The forecast wholesale water demand in a given WDPA, month, and sector, $Q_{WS,m,y,WDPA}$, is defined as a fraction of the total retail water use for the same WDPA and time period:

$$Q_{WS,m,y,WDPA} = \left(WS_{m,y,WDPA}\right)\left(Q^{cal}_{RET,m,y,WDPA}\right) \tag{B.11}$$

where:

- $WS_{m,y,WDPA}$ = ratio of forecast wholesale water use in a month, year, and WDPA to forecast retail water use in that WDPA and time period (fraction)

- $Q^{cal}_{RET,m,y,WDPA}$ = total retail water use in a month, year, and WDPA (from Equation B.10)

In the point and probabilistic forecasts, $WS_{m,y,WDPA}$ was set equal to the ratio of observed wholesale demand to observed total retail demand in the corresponding month and WDPA for Water Year 2002. These WY2002-based monthly wholesale ratio values were then applied in each forecast year.

Unbilled water use, $Q_{OUW,m,y,WDPA}$, is defined slightly differently: as a fraction of total water demand for a WDPA, month and year (i.e., as a fraction of the sum of retail, wholesale, and unbilled demand):

$$Q_{OUW,m,y,WDPA} = \left(OUW_{m,y,WDPA}\right)\left(Q^{cal}_{RET,m,y,WDPA} + Q_{WS,m,y,WDPA} + Q_{OUW,m,y,WDPA}\right) \tag{B.12}$$

Solving for $Q_{OUW,m,y,WDPA}$:

$$Q_{OUW,m,y,WDPA} = \frac{OUW_{m,y,WDPA}}{1 - OUW_{m,y,WDPA}}\left(Q^{cal}_{RET,m,y,WDPA} + Q_{WS,m,y,WDPA}\right) \tag{B.13}$$

where:

- $OUW_{m,y,WDPA}$ = ratio of unbilled water use in a month, year, and WDPA to total calibrated water use (calibrated retail, wholesale, and unbilled) in that WDPA and time period (fraction)

- $Q_{RET,m,y,WDPA}^{cal}$= total retail water use in a month, year, and WDPA (from Equation B.10)

- $Q_{WS,m,y,WDPA}$ = wholesale deliveries in a given month, year, and WDPA (from Equation B.11)

In the point and probabilistic forecasts, $OUW_{m,y,WDPA}$ was set equal to the ratio of observed unbilled demand to observed total demand in the corresponding month and WDPA for Water Year 2002. These WY2002-based monthly unbilled ratio values were then applied in each forecast year.

Finally, total water demand in a given WDPA, month, and year, $Q_{TOT,m,y,WDPA}$, is the sum of calibrated retail, wholesale, and unbilled demand:

$$Q_{TOT,m,y,WDPA} = Q_{RET,m,y,WDPA}^{cal} + Q_{WS,m,y,WDPA} + Q_{OUW,m,y,WDPA} \tag{B.14}$$

These total water demand values may be rolled up to the entire TBW service area by summation and to an entire year by averaging across months.

Appendix C
@Risk Simulation Tool and Setup

@Risk, a Microsoft Excel add-in from Palisade Corporation, was used to enact probabilistic demand simulations within a spreadsheet containing the point model. @Risk was selected not only because it was compatible with the existing spreadsheet-based model, but also because of positive past experiences with using the software. @Risk provided all the tools needed for a complete simulation, including assignment and recognition of various probability distribution functions, random number generation, assignment of sampling rules, and scoring (or tallying) of results, as well as graphics and scenario options.

When using @Risk, the spreadsheet itself still visually displayed one value for each model input, which generally represented the set of point expectations of future values. Clicking on an input cell, however, showed that the cell formula was no longer a constant but an @Risk function that represented a probability distribution with specific parameters. This functionality was the most important element for simulating water demand probabilistically. Consistent with the discussion in Chapter 2, a variable was not represented by a single value, but rather a range of possible values implied by the shape of a probability density function, or pdf.

@Risk employs random number generators to replicate random samples from pdfs. These generators are effectively composed of extended sequences of uniform random numbers between 0 and 1, each with a reference point (also known as a seed). The user is allowed to pick a particular seed to initiate or repeat a simulation run. @Risk generates a series of random numbers for each distribution, one number for each sampling iteration. Assigned numbers are used as cumulative probabilities for selecting random variable values from pdfs. Where correlations are specified, @Risk adjusts cumulative probabilities prior to application to pdfs such that the specified rank correlation between probabilities is reproduced.

@Risk provides two types of sampling rules for picking random samples: Monte Carlo and Latin Hypercube sampling. If one could place all the possible model inputs onto a table, then Monte Carlo sampling would be equivalent to arranging the inputs into stacks (with heights of stacks determined by their probabilities) and choosing freely from the stacks. Given a sufficiently large number of decisions, the most common choices should match the set of most likely model inputs. The Latin Hypercube sampling technique is generally considered more efficient. To continue the stack analogy, it first separates the distribution values into segments of equal cumulative probability (or different tables) before stacking them and then samples an equal number of times from each segment (or table). This approach has advantages over Monte Carlo sampling in that it will ensure that rarer input values which may be considered important will be sampled. Because of these factors, Latin Hypercube sampling was chosen for selection of random variable values.

After including necessary @Risk functions to represent assigned pdfs, the spreadsheet possessed almost 15,000 input cells to derive probabilistic forecasts, of which more than 10,300 cells contained @Risk functions. Roughly 3,400 output cells were assigned to contain demand projections for a single set of input variable values. @Risk performed the Monte Carlo simulation by iteratively substituting randomly selected values into input cells, determining contents of output cells, tabulating these output values, and calculating summary statistics for each output cell to generate forecast results.

Appendix D
Summary Interval Forecasts of Annual Water Demand by Water Demand Planning Area

The following tables contain probabilistic forecast results for annually-averaged daily demand in each of the seven WDPAs. The top portion of each table contains measures of central tendency in probabilistic forecast versus time. The bottom portions contain percentiles of cumulatively-distributed demand forecasts at selected forecast years. These tables may be interpreted similarly to Table 3.1 in Chapter 3.

Table D.1
Pinellas WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	68.86	68.88	71.29	72.66	73.48	74.20
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	67.96	69.20	71.59	72.84	73.46	73.97
Standard Deviation	1.51	1.79	2.29	2.87	3.65	4.36
Median	67.91	69.17	71.60	72.94	73.71	74.39
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	65.52	66.27	67.81	67.91	67.26	66.25
10%	66.05	66.91	68.67	69.16	68.92	68.50
15%	66.39	67.33	69.25	69.99	69.89	69.80
20%	66.66	67.69	69.69	70.61	70.66	70.89
25%	66.92	67.99	70.07	71.08	71.26	71.60
30%	67.12	68.26	70.43	71.49	71.87	72.25
35%	67.32	68.50	70.73	71.89	72.37	72.83
40%	67.52	68.73	71.02	72.26	72.82	73.42
45%	67.72	68.94	71.32	72.61	73.26	73.93
50%	67.91	69.17	71.60	72.94	73.71	74.39
55%	68.11	69.39	71.86	73.29	74.15	74.88
60%	68.32	69.60	72.13	73.64	74.58	75.35
65%	68.52	69.85	72.43	74.00	75.03	75.83
70%	68.75	70.12	72.77	74.37	75.46	76.35
75%	68.98	70.38	73.09	74.78	75.91	76.91
80%	69.23	70.70	73.50	75.21	76.44	77.51
85%	69.53	71.08	73.94	75.77	77.05	78.18
90%	69.93	71.54	74.50	76.38	77.80	78.98
95%	70.51	72.22	75.33	77.33	78.95	80.25

Table D.2
St. Petersburg WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	32.11	31.94	33.11	33.86	34.39	34.86
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	31.77	32.22	33.36	34.02	34.39	34.70
Standard Deviation	0.74	0.95	1.42	1.97	2.62	3.29
Median	31.76	32.23	33.39	34.15	34.63	35.06
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	30.57	30.67	30.95	30.62	29.80	28.87
10%	30.81	31.00	31.52	31.49	31.06	30.55
15%	30.99	31.24	31.90	32.06	31.87	31.53
20%	31.14	31.43	32.18	32.47	32.42	32.32
25%	31.27	31.59	32.45	32.79	32.91	32.88
30%	31.38	31.74	32.66	33.11	33.30	33.40
35%	31.48	31.86	32.87	33.38	33.65	33.84
40%	31.57	31.99	33.05	33.65	34.00	34.28
45%	31.66	32.11	33.22	33.90	34.34	34.70
50%	31.76	32.23	33.39	34.15	34.63	35.06
55%	31.85	32.35	33.57	34.37	34.92	35.44
60%	31.94	32.47	33.74	34.62	35.22	35.82
65%	32.04	32.59	33.92	34.85	35.54	36.21
70%	32.15	32.73	34.12	35.09	35.83	36.58
75%	32.27	32.86	34.33	35.36	36.17	36.97
80%	32.40	33.01	34.55	35.67	36.52	37.36
85%	32.55	33.20	34.82	36.01	36.92	37.84
90%	32.72	33.43	35.14	36.41	37.47	38.44
95%	32.98	33.77	35.65	37.01	38.24	39.28

Table D.3
New Port Richey WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	3.47	3.38	3.43	3.49	3.54	3.60
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	3.44	3.41	3.45	3.50	3.54	3.58
Standard Deviation	0.09	0.11	0.15	0.21	0.27	0.34
Median	3.44	3.41	3.46	3.51	3.56	3.61
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	3.29	3.23	3.20	3.14	3.06	2.97
10%	3.32	3.27	3.26	3.24	3.20	3.15
15%	3.35	3.30	3.30	3.30	3.28	3.26
20%	3.36	3.32	3.33	3.34	3.33	3.33
25%	3.38	3.34	3.35	3.37	3.39	3.40
30%	3.39	3.35	3.38	3.40	3.43	3.45
35%	3.40	3.37	3.40	3.44	3.46	3.49
40%	3.42	3.38	3.42	3.46	3.49	3.54
45%	3.43	3.40	3.44	3.49	3.53	3.58
50%	3.44	3.41	3.46	3.51	3.56	3.61
55%	3.45	3.42	3.48	3.54	3.59	3.65
60%	3.46	3.44	3.49	3.56	3.63	3.69
65%	3.48	3.45	3.51	3.59	3.66	3.72
70%	3.49	3.47	3.53	3.61	3.69	3.76
75%	3.50	3.49	3.56	3.64	3.72	3.80
80%	3.52	3.50	3.58	3.67	3.76	3.85
85%	3.54	3.53	3.61	3.71	3.80	3.90
90%	3.56	3.55	3.64	3.75	3.86	3.96
95%	3.60	3.59	3.69	3.82	3.93	4.05

Table D.4
Pasco WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	19.04	19.40	21.55	23.40	25.19	26.86
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	18.81	19.51	21.63	23.42	25.10	26.65
Standard Deviation	0.61	0.73	1.04	1.43	1.96	2.55
Median	18.80	19.49	21.65	23.47	25.25	26.93
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	17.84	18.33	19.91	20.98	21.69	22.19
10%	18.05	18.58	20.30	21.58	22.56	23.43
15%	18.18	18.75	20.55	21.98	23.15	24.19
20%	18.29	18.90	20.78	22.30	23.60	24.74
25%	18.39	19.01	20.95	22.53	23.97	25.22
30%	18.48	19.11	21.10	22.74	24.27	25.61
35%	18.56	19.21	21.24	22.93	24.54	25.97
40%	18.65	19.31	21.38	23.12	24.79	26.31
45%	18.73	19.40	21.52	23.30	25.02	26.62
50%	18.80	19.49	21.65	23.47	25.25	26.93
55%	18.88	19.58	21.77	23.65	25.48	27.20
60%	18.95	19.68	21.91	23.82	25.71	27.48
65%	19.04	19.78	22.04	24.01	25.93	27.78
70%	19.13	19.88	22.18	24.20	26.17	28.07
75%	19.22	20.00	22.33	24.37	26.43	28.40
80%	19.33	20.13	22.49	24.60	26.71	28.75
85%	19.45	20.27	22.69	24.86	27.04	29.12
90%	19.61	20.45	22.95	25.19	27.45	29.58
95%	19.84	20.75	23.30	25.68	27.97	30.24

Table D.5
City of Tampa WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	72.91	74.22	79.90	85.24	89.68	94.26
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	72.76	74.65	80.26	85.37	89.47	93.63
Standard Deviation	1.61	2.07	3.13	4.58	6.30	8.19
Median	72.72	74.66	80.34	85.61	89.99	94.58
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	70.13	71.19	75.00	77.46	78.51	79.05
10%	70.71	72.00	76.28	79.56	81.52	83.50
15%	71.11	72.52	77.02	80.84	83.39	85.88
20%	71.40	72.92	77.66	81.79	84.76	87.60
25%	71.66	73.29	78.21	82.58	85.91	89.16
30%	71.89	73.59	78.70	83.29	86.92	90.52
35%	72.11	73.87	79.15	83.96	87.80	91.62
40%	72.31	74.14	79.56	84.54	88.56	92.67
45%	72.52	74.40	79.97	85.09	89.28	93.65
50%	72.72	74.66	80.34	85.61	89.99	94.58
55%	72.92	74.92	80.74	86.17	90.66	95.56
60%	73.14	75.19	81.12	86.76	91.44	96.36
65%	73.37	75.44	81.53	87.29	92.22	97.27
70%	73.59	75.75	81.97	87.83	92.97	98.16
75%	73.84	76.05	82.43	88.46	93.78	99.17
80%	74.13	76.41	82.92	89.15	94.66	100.28
85%	74.45	76.78	83.44	89.93	95.66	101.49
90%	74.82	77.32	84.19	90.94	96.87	102.96
95%	75.48	78.02	85.17	92.43	98.61	105.18

Table D.6
NW Hillsborough WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	16.61	17.01	18.96	20.67	22.60	24.62
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	16.38	17.05	18.97	20.60	22.42	24.30
Standard Deviation	0.50	0.61	0.90	1.29	1.84	2.46
Median	16.37	17.04	18.99	20.67	22.56	24.58
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	15.58	16.07	17.47	18.38	19.22	19.92
10%	15.76	16.28	17.82	18.98	20.08	21.16
15%	15.87	16.42	18.05	19.33	20.62	21.91
20%	15.96	16.54	18.22	19.58	21.02	22.48
25%	16.04	16.64	18.38	19.82	21.35	22.94
30%	16.11	16.73	18.52	20.01	21.64	23.33
35%	16.18	16.80	18.64	20.19	21.90	23.68
40%	16.25	16.89	18.76	20.35	22.14	24.01
45%	16.31	16.96	18.87	20.51	22.34	24.31
50%	16.37	17.04	18.99	20.67	22.56	24.58
55%	16.43	17.12	19.10	20.82	22.77	24.86
60%	16.50	17.19	19.20	20.97	22.99	25.11
65%	16.57	17.27	19.33	21.13	23.22	25.40
70%	16.64	17.37	19.45	21.28	23.45	25.67
75%	16.72	17.47	19.57	21.46	23.67	25.98
80%	16.81	17.57	19.73	21.66	23.93	26.30
85%	16.92	17.68	19.89	21.88	24.23	26.70
90%	17.04	17.84	20.09	22.16	24.59	27.16
95%	17.21	18.07	20.42	22.59	25.14	27.82

Table D.7
SC Hillsborough WDPA Probabilistic Forecast: Annual Average Water Demand

Water Year	2003	2005	2010	2015	2020	2025
Calibrated Point Forecast	25.09	26.40	30.25	33.66	37.62	41.63
Calibrated Probabilistic Forecast: Summary Statistics						
Mean	25.00	26.40	30.24	33.54	37.31	41.13
Standard Deviation	0.72	0.92	1.37	1.96	2.78	3.73
Median	24.99	26.39	30.25	33.65	37.56	41.55
Calibrated Probabilistic Forecast: Demand Distributions (Percentiles)						
5%	23.85	24.90	27.94	30.22	32.48	34.45
10%	24.09	25.23	28.46	31.01	33.82	36.28
15%	24.25	25.44	28.81	31.54	34.61	37.57
20%	24.38	25.63	29.11	31.98	35.17	38.46
25%	24.50	25.78	29.35	32.34	35.70	39.11
30%	24.61	25.90	29.55	32.62	36.14	39.67
35%	24.71	26.03	29.75	32.91	36.50	40.21
40%	24.80	26.15	29.93	33.16	36.88	40.69
45%	24.90	26.28	30.09	33.40	37.24	41.13
50%	24.99	26.39	30.25	33.65	37.56	41.55
55%	25.08	26.50	30.43	33.88	37.87	41.95
60%	25.17	26.62	30.59	34.12	38.18	42.37
65%	25.27	26.74	30.76	34.35	38.52	42.79
70%	25.37	26.88	30.95	34.59	38.84	43.24
75%	25.48	27.02	31.16	34.86	39.21	43.69
80%	25.61	27.18	31.39	35.16	39.62	44.19
85%	25.75	27.36	31.66	35.54	40.06	44.76
90%	25.93	27.61	31.97	35.96	40.60	45.41
95%	26.20	27.96	32.42	36.60	41.40	46.37

Appendix E
Derivation of Explanatory and Driver Projections for Point Forecasting

Projected values of explanatory and driver variables for the point forecast were derived using data from the modeling database as well as MPOs and BEBR. This section describes the mathematical basis for deriving these projections.

E.1 SF Households and MF Dwelling Units

The ratio of annual average single-family households to multi-family units was derived from number of SF and MF accounts in each WDPA for Water Year 2001.

1. Number of SF households was estimated as equal to number of SF accounts. For each WDPA, average number of SF households for WY 2001 was determined by totaling the number of accounts in TAZs within that WDPA for each month, then averaging across the water year.

$$N_{SF,2001,WDPA} = \frac{\sum_{m=\text{Oct }2000}^{\text{Sep }2001}\left(\sum_{TAZ \in WDPA} SFaccounts_{TAZ,m}\right)}{12} \qquad \text{(E.1)}$$

2. Number of MF dwelling units was estimated as equal to number of MF accounts times a unit-per-account factor. These factors were provided by member governments for a previous project (Ayres Associates, 1997) and are listed in Table E.1. For each WDPA, average number of MF dwelling units for WY 2001 was determined by totaling the number of accounts in TAZs within that WDPA for each month, averaging across the water year, then multiplying by the units per account factor for the WDPA.

$$N_{MF,2001,WDPA} = \frac{\sum_{m=\text{Oct }2000}^{\text{Sep }2001}\left(\sum_{TAZ \in WDPA} MFaccounts_{TAZ,m}\right)}{12} \times UnitsPerAccout_{WDPA} \qquad \text{(E.2)}$$

Table E.1
Number of MF Units Per Account by WDPA[1]

WDPA	# MF Units/account
Tampa	53
Pinellas	16
St Petersburg	11
New Port Richey	11
Pasco	2.5
NW Hillsborough	55
SC Hillsborough	55

3. For each forecast year, BEBR county-level projections of number of single-family housing starts was divided by BEBR county-level projections of total number of housing starts (BEBR, 2001a). These ratios were taken to represent the fraction of newly-constructed units in each WDPA and forecast year that were single-family units.

$$f_{SF,y,WDPA} = \frac{\#\,SF\;\; starts_{y,county}\;\; from\;\; BEBR}{\#\,total\;\; starts_{y,county}\;\; from\;\; BEBR} \qquad \text{(E.3)}$$

4. Number of new SF units for each TAZ and forecast year were determined by multiplying change in total number of units for each TAZ from the previous year (MPO data) by estimated fraction of new units for that year that are SF units (fSF,y,WDPA from Equation E.3). Number of new MF units for each TAZ was determined as the difference between total number of new units and number of new SF units.

$$\Delta N_{SF,y,TAZ} = f_{SF,y,WDPA} \times \left(total\;\#\;\; of\;\; new\;\; units_{y,TAZ}\;\; from\;\; MPO\right) \qquad \text{(E.4)}$$

$$\Delta N_{MF,y,TAZ} = total\;\#\;\; of\;\; new\;\; units_{y,TAZ}\;\; from\;\; MPO - \Delta N_{SF,y,TAZ} \qquad \text{(E.5)}$$

5. For each forecast year, number of SF and MF units were determined by adding the number of new SF and MF units to the previous year's SF and MF units. Average number of SF and MF units in 2001 (from Equations E.1 and E.2) were used as base year data.

$$N_{SF,y,TAZ} = N_{SF,y-1,TAZ} + \Delta N_{SF,y,TAZ} \qquad \text{(E.6)}$$

1 *Ayres Associates, 1997.*

$$N_{MF,y,TAZ} = N_{MF,y-1,TAZ} + \Delta N_{MF,y,TAZ} \tag{E.7}$$

6. Projections of NSF,y,TAZ and NMF,y,TAZ were summed by WDPA to produce WDPA projections of single-family and multi-family units.

$$N_{SF,y,WDPA} = \sum_{TAZ \in WDPA} N_{SF,y,TAZ} \tag{E.8}$$

$$N_{MF,y,WDPA} = \sum_{TAZ \in WDPA} N_{MF,y,TAZ} \tag{E.9}$$

E.2 Total Employees and Fraction Employment in Service, Industrial, and Commercial Categories

Projections of total employment and fraction employment in Service, Commercial, and Industrial categories were developed from MPO data.

Pinellas and Pasco County MPOs provided observed number of employees by TAZ and category for 1999 and projected number of employees by TAZ and category for 2005, 2010, 2015, 2020, and 2025. For each of these years, total employment and number of employees in each category was determined for Pinellas, St. Petersburg, New Port Richey, and Pasco WDPAs by summing employment data across TAZs within those WDPAs:

$$N_{NR,y,WDPA}^{category} = \sum_{TAZ \in WDPA} \#\ employees_{TAZ,category,y} \quad from\ MPO \tag{E.10}$$

$$N_{NR,y,WDPA} = \sum_{category} N_{NR,y,WDPA}^{category} \tag{E.11}$$

for WDPA = Pinellas, St Pete, Pasco, or New Port Richey, *category* = Service, Commercial, or Industrial, and *y* = 2005, 2010, 2015, 2020, and 2025. Yearly projections of total employment between 2002 and 2005 were interpolated between 2002 observed total employment in the modeling database and 2005 projected total employment. Total employment projections between 2005 and 2010, 2010 and 2015, 2015 and 2020, and 2020 and 2025 were also interpolated.

Hillsborough County MPO provided observed number of employees by TAZ and category for 1999 and projected number of employees by TAZ and category 2015 and 2025. Service and Commercial categories were subdivided by the Hillsborough MPO into “local” and “regional” employment, where local employment refers to employees living and working in the same TAZ and regional employment refers to employees living in a different TAZ from where they work. For each MPO-projected TAZ and year, Commercial employment was taken as the sum of local Commercial and regional Commercial employment, and Service employment was taken as the sum of local Service and regional Service employment. Projections of total employment for Northwest Hillsborough, South Central Hillsborough, and City of Tampa WDPAs were then formed in the same fashion and for St. Pete and Pinellas.

Projected values of fraction employment in each WDPA were determined by dividing total number of employees in a category and year by total number of employees in a year:

$$COM_{y,WDPA} = \frac{N^{Commercial}_{NR,y,WDPA}}{N_{NR,y,WDPA}} \qquad \text{(E.12)}$$

$$IND_{y,WDPA} = \frac{N^{Industrial}_{NR,y,WDPA}}{N_{NR,y,WDPA}} \qquad \text{(E.13)}$$

$$SER_{y,WDPA} = \frac{N^{Service}_{NR,y,WDPA}}{N_{NR,y,WDPA}} \qquad \text{(E.14)}$$

E.3 Per-Household Income

Experian provided observed yearly average income per household for a sampling of street addresses for 1999-2002. These data were geocoded to parcels through street address, allowing grouping of income data by TAZ. Median income per household was then determined as the median of income samples within each TAZ. These TAZ-level results were used in the modeling database.

Projections of per-household income were determined using 2002 values from the modeling database and projected real per-capita growth rates from BEBR (BEBR, 2001a). Base-year income was calculated for each WDPA by averaging median household incomes of TAZs within that WDPA. Projected household income was then derived by applying BEBR county-level year-to-year growth rates in real per capita income:

$$INC_{2003,WDPA} = INC_{2002,WDPA} \times \left(1 + \frac{g_{county,2003}}{100\%}\right) \qquad \text{(E.15)}$$

$$INC_{2004,WDPA} = INC_{2003,WDPA} \times \left(1 + \frac{g_{county,2004}}{100\%}\right) \text{ etc.,} \qquad \text{(E.16)}$$

where $g_{county,year}$ is the projected percent growth in real per capita income in a given year for the county corresponding to the WDPA.

E.4 Single-Family Persons per Household

Experian provided yearly observed persons per household data for a sampling of street addresses for 1999-2002. These data were geocoded to parcels through street address, allowing grouping of persons per household data by TAZ. Parcels were classified as single-family or multi-family by way of geocoded billing data, allowing Experian persons per household sample observations to be classified by sector. Single-family persons per household for each TAZ was then determined as the average of single-family persons per household samples within each TAZ. Likewise, multi-family persons per household for each TAZ was estimated by the averaging multi-family samples within each TAZ. These TAZ-level results were used in the modeling database.

Projections of single-family persons per household were determined using 2002 single-family and multi-family persons per household from the modeling database and projected population and total number of households from BEBR (BEBR, 2001a).

- The ratio of single-family to multi-family persons per household for 2002 was determined from the modeling database:

$$R_{2002,WDPA} = \frac{PPH_{SF,2002,WDPA}}{PPH_{MF,2002,WDPA}} \qquad \text{(E.17)}$$

- Projected total persons per household for each WDPA was derived by dividing BEBR-projected permanent residents by BEBR-projected households (not housing stock) for the associated county:

$$PPH_{TOT,y,WDPA} = \frac{Pop_{TOT,y,County}}{Households_{TOT,y,County}} \qquad \text{(E.18)}$$

- Single-family persons per household for each WDPA was then determined using projected SF and MF units and the SF-to-MF persons per household ratio, assuming the ratio would remain the same as in 2002 for all forecast years:

$$PPH_{TOT,y,WDPA} = \frac{\left(PPH_{SF,y,WDPA} \times N_{SF,y,WDPA}\right) + \left(PPH_{MF,y,WDPA} \times N_{MF,y,WDPA}\right)}{N_{SF,y,WDPA} + N_{MF,y,WDPA}} \qquad \text{(E.19)}$$

$$PPH_{MF,y,WDPA} = \frac{PPH_{SF,y,WDPA}}{R_{2002,WDPA}} \tag{E.20}$$

- Substituting Equation E.20 into E.19, single-family persons per household can be calculated:

$$PPH_{SF,y,WDPA} = PPH_{TOT,y,WDPA}\left(\frac{N_{SF,y,WDPA} + N_{MF,y,WDPA}}{N_{SF,y,WDPA} + {N_{MF,y,WDPA}}\big/{R_{2002,WDPA}}}\right) \tag{E.21}$$

E.5 Single-Family and Multi-Family Housing Density

The modeling database contained estimates of number of SF households and MF dwelling units by TAZ (see Section 1.2.2 and Appendix A.2.1 and A.2.2). In addition, parcels were geocoded to TAZs/WDPAs and classified as single-family, multi-family, or non-residential by way of geocoded billing data. Total single-family and multi-family parcel acreage for each TAZ was thus determined by summing parcel areas. This data was used as part of the modeling database.

Base-year weighted average of TAZ-level single-family housing density was determined for each WDPA using Water Year 2002 data in the modeling database. Weights were formed for each TAZ and month of Water Year 2002 by dividing number of SF households in that TAZ and month by number of SF households in the corresponding WDPA and month:

$$w_{SF,m,TAZ} = \frac{N_{SF,n,TAZ}}{\sum_{TAZ \in WDPA}\left(N_{SF,n,TAZ}\right)} \tag{E.22}$$

Average single-family housing density in each WDPA and month in WY 2002 was then determined by multiplying housing density for each component TAZ and month by the corresponding weight, then summing these products. Monthly WDPA single-family housing densities were then averaged across WY 2002:

$$HOU_{SF,2002,WDPA} = \frac{\sum_{m=Oct2001}^{Sep2002}\left(\sum_{TAZ \in WDPA} w_{SF,m,TAZ} HOU_{SF,m,TAZ}\right)}{12} \tag{E.23}$$

where $HOU_{SF,m,TAZ}$ is from the modeling database. Likewise, average multi-family housing density for 2002 was determined using monthly TAZ weighted-average multi-family housing densities:

$$w_{MF,m,TAZ} = \frac{N_{MF,n,TAZ}}{\sum_{TAZ \in WDPA} \left(N_{MF,n,TAZ} \right)} \tag{E.24}$$

$$HOU_{MF,2002,WDPA} = \frac{\sum_{m=Oct2001}^{Sep2002} \left(\sum_{TAZ \in WDPA} w_{MF,m,TAZ} HOU_{MF,m,TAZ} \right)}{12} \tag{E.25}$$

where $HOU_{MF,m,TAZ}$ is from the modeling database.

All projected values of single-family and multi-family housing density were assumed to be equal to 2002 values; that is, housing density was assumed to be constant over the forecast period, such that new housing would imply new developed residential acreage.

E.6 Temperature, Rainfall, and 0.01" and 1" Rainy Days

All weather projections were taken as long-term (1971-2000) normal values for the given calendar month and WDPA. Historical weather by WDPA and month for was determined from 1971-2000 weather station data using inverse-squared distance-weighted average, much as was done for TAZ weather data in the modeling database (Section 1.2.4, Figure 1.5). Weather stations used for this determination are listed in Table 1.1.

For a WDPA, the following procedure was performed to estimate 1971-2000 historical data for each weather variable.

- Distances were measured between the WDPA centroid and each weather station: $d_{WDPA,1}$ = distance between WDPA centroid and station 1, $d_{WDPA,2}$ = distance between WDPA centroid and station 2 etc.

- Raw weights were defined as the inverse square of the each distance, e.g.:

$$w_{WDPA,1} = 1 / \left(d_{WDPA,1} \right)^2 , \quad w_{WDPA,2} = 1 / \left(d_{WDPA,2} \right)^2 \text{, etc.} \tag{E.26}$$

- Normalized weights were defined by dividing each raw weight by the sum of raw weights:

$$\tilde{w}_{WDPA,1} = {w_{WDPA,1}} \Big/ {\sum w_i}, \quad \tilde{w}_{WDPA,2} = {w_{WDPA,2}} \Big/ {\sum w_i}, \text{ etc.} \qquad \text{(E.27)}$$

- Historical values of WDPA weather variables were determined for each month in the 1971-2000 period by multiplying station observations by corresponding normalized weights. For example, consider 1" rainy days in January 1976:

$$RD1_{Jan,1976,Pasco} = \tilde{w}_{Pasco,1} \times RD1_{Station1,Jan,1976} + \tilde{w}_{Pasco,2} \times RD1_{Station2,Jan,1976} + \ldots, \qquad \text{(E.28)}$$

In this manner, weighted average values of monthly mean maximum daily temperature, monthly total rainfall, and monthly total number of 0.01" and 1" rainy days were determined for each WDPA and month between 1971 and 2000. Long-term normal weather values were then determined by pooling WDPA weather values by calendar month and determining an average value for each calendar month: e.g.:

$$RD1NORM_{Jan,Pasco} = \frac{RD1_{Jan,1971,Pasco} + RD1_{Jan,1972,Pasco} + \ldots + RD1_{Jan,2000,Pasco}}{30}, \qquad \text{(E.29)}$$

Projected weather variables in each forecast month were assigned long-term normal values for the corresponding calendar month.

References

Ayres Associates, 1997. *Regional Water Supply Demand Management Plan.* Prepared for The West Coast Regional Water Supply Authority (now Tampa Bay Water), June 1997.

BEBR, 2001a. *Florida Long-Term Economic Forecast 2001, Volume 2 – State & Counties. Florida Bureau of Economic and Business Research (BEBR)*, Warrington College of Business Administration, University of Florida, September 2001.

BEBR, 2001b. *Florida Population Studies: Population Projections by Age, Sex, and Race for Florida and Its Counties, 2002-2020*. Florida Bureau of Economic and Business Research (BEBR), Warrington College of Business Administration, University of Florida, Vol 34, No. 3, Bulletin 130, October 2001.

BEBR, 2001c. *Florida Population Studies: Projections of Florida Population by County, 2000-2030.* Florida Bureau of Economic and Business Research (BEBR), Warrington College of Business Administration, University of Florida, Vol 34, No. 1, Bulletin 128, May 2001.

BEBR, 2002. *Florida Statistical Abstract 2002.* Florida Bureau of Economic and Business Research (BEBR), Warrington College of Business Administration, University of Florida.

Black & Veatch, 2002. *Consulting System Engineer's Report.* Black and Veatch Corp., May 2002.

Hillsborough County FL, 2002. *Hillsborough County MPO Socioeconomic Data Estimates and Projections 1999 2015 2025.*

Kiefer, J. and Porter, G., 2000. *Development of Probabilistic Water Demand Forecast for the San Diego County Water Authority.* PMCL.

Neter, J., et al., 1996. *Applied Linear Statistical Models, 4th Edition.* Chicago: McGraw-Hill/Irwin, 1996.

Pinellas County FL, 2001. *Pinellas County MPO Forecast 2025 Employment Socioeconomic Data.* June 2001.

Pasco County FL, 2001. *Pasco County MPO Forecast 2025 Socioeconomic Data Development.* Prepared by Tindale/Oliver and Associates, Inc., July 2001.

WCRWSA, 1998. *Northern Tampa Bay New Water Supply and Ground Water Withdrawal Reduction Agreement between West Coast Regional Water Supply Authority, Hillsborough County, Pasco County, Pinellas County, City of Tampa, City of St. Petersburg, City of New Port Richey, and Southwest Florida Water Management District.* April 1998.