This presentation premiered at WaterSmart Innovations

watersmartinnovations.com

Exploratory Data Analysis Supports Water Demand Model Development

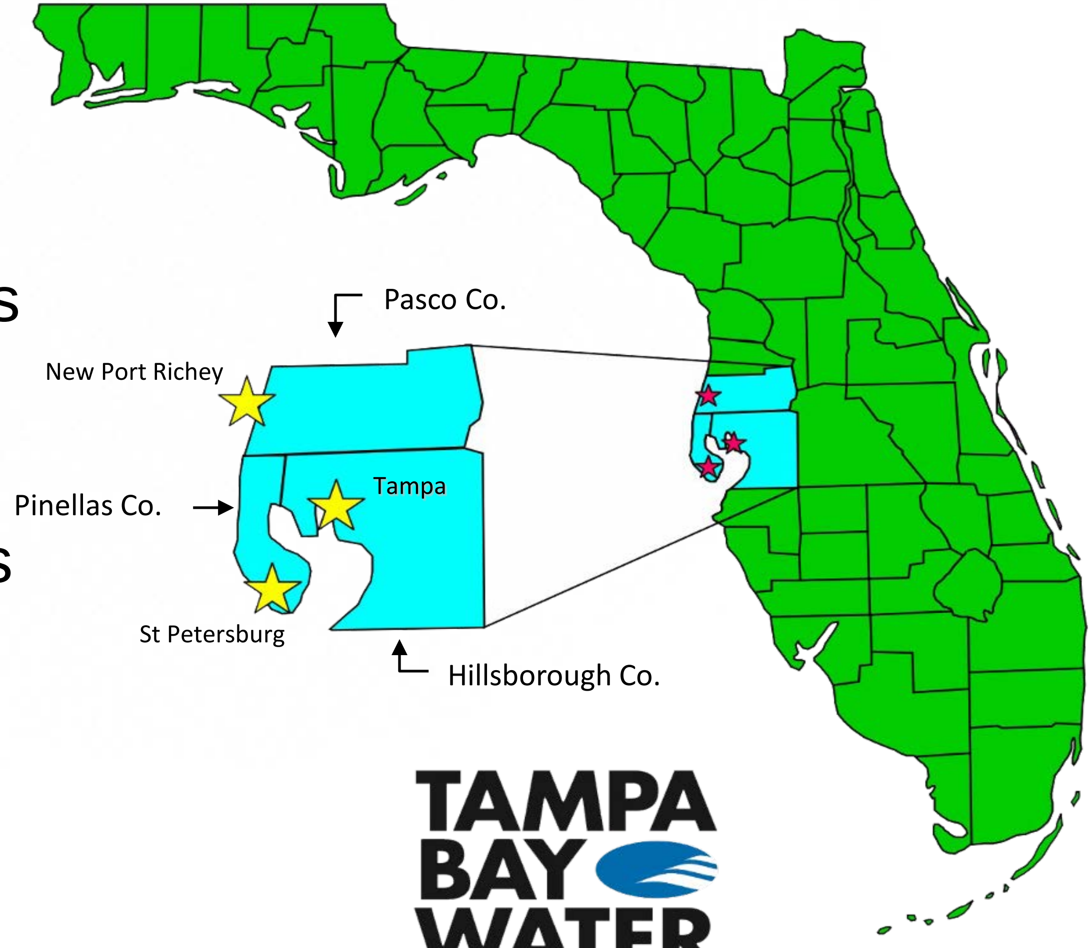
David Bracciano, Tampa Bay Water
Jack Kiefer, PhD, Hazen and Sawyer

Overview

- Background on Tampa Bay Water
- Role of demand forecasting in planning efforts
- Collaborative model development process
- Elements and examples of EDA
- Next steps

Agency Background

- Regional water supply authority serving over 2.4 million customers
- Six member governments, across three counties
- Member demands:
  - 2015: 227 MGD
  - 2035: 281 MGD



TAMPA BAY WATER
Supplying Water To The Region

Long-Term Demand Forecasting System (LTDFS)

- LTDFS designed to:
  - Track water consumption, socioeconomic, economic and policy conditions
  - Provide inputs for demand forecasting models
  - Prepare forecasts through implementation of models
  - Support water supply reliability ("just-in-time" supply development) efforts

2015 model redevelopment

- GOAL: Improve alignment of water use w/explanatory variables for econometric & end-use modeling
  - Better sectoral definitions
  - More uniform spatial designations
  - More direct demographic measurements
  - More flexibility for modeling/exploratory data analysis
  - Potentially stronger statistical relationships
  - Better support to other demand-oriented analyses and monitoring

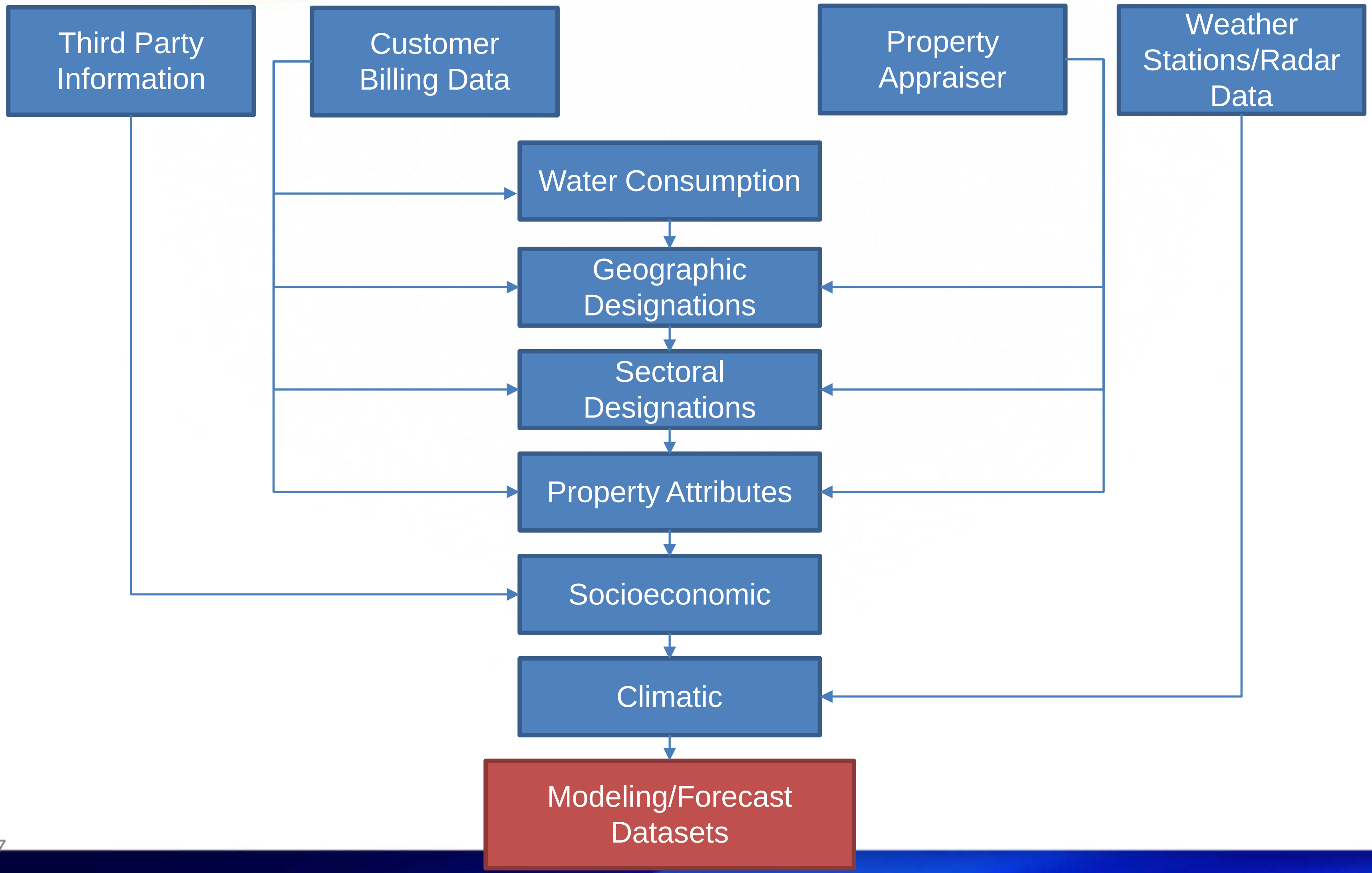
Database objectives

- Extensive database developed to support model redevelopment

1. Ensure acquired information can be maintained through time to support future evaluations

2. Standardize analytical routines so they can be replicated and updated efficiently through time

Source data dependencies



Collaborative Model Development Process

- Extension of collaborative database development
- Engage internal and external expertise to
  - Leverage technical specialties
  - Enhance institutional knowledge
  - Generate ideas and alternative modeling options
  - Cross-validate and replicate findings
  - Identify preferred forecasting model

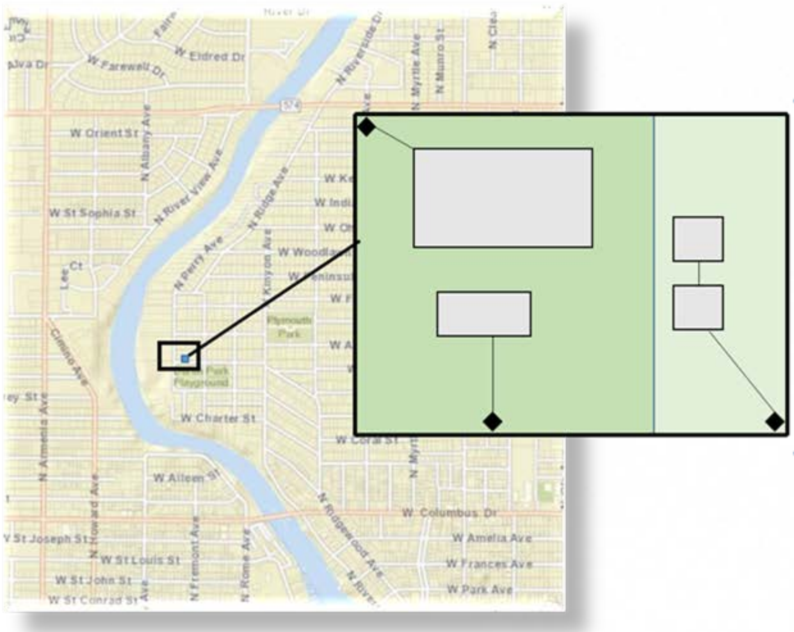
Exploratory Data Analysis

- How does water use vary?
  - Across individual users
  - Across geographic areas
  - Over time
- How can we best distinguish between signals and “noise”?
  - What variables are available to explain variability?
  - What variables can we generate forecasts or scenarios for?
- Iterative modeling, visualization, and diagnostics

Dimensions of the Data

- 696,466 parcels associated with 543,054 water using locations
  - SF: 500,126 locations
  - MF: 17,050 locations
  - NR: 25,878 locations

Unique Water Using Locations



- ❑ 1 water use **location** associated with:
  - ❑ 1 business or function associated with:
    - ❑ 4 buildings
    - ❑ 2 tax parcels
    - ❑ 3 water meters

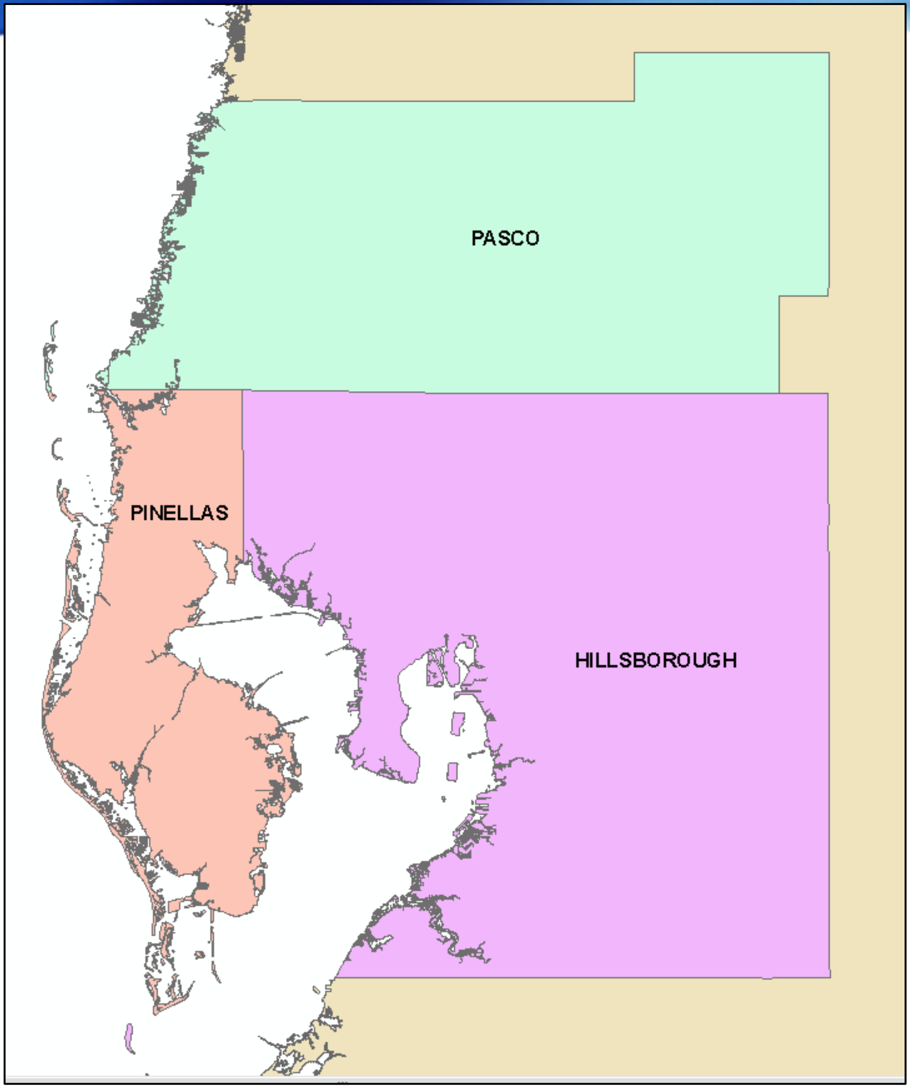
Source: Kiefer, J.C. and L.R. Krentz. 2016. *Evaluation of Customer Information and Data Processing Needs for Water Demand Analysis, Planning, and Management*. Denver: Water Research Foundation.

Dimensions of the Data

- Alternative geographic (cross-sectional) dimensions
  - Water using location
  - Block groups
  - Traffic analysis zones
  - Census tracts
  - Water demand planning areas (WDPAs)

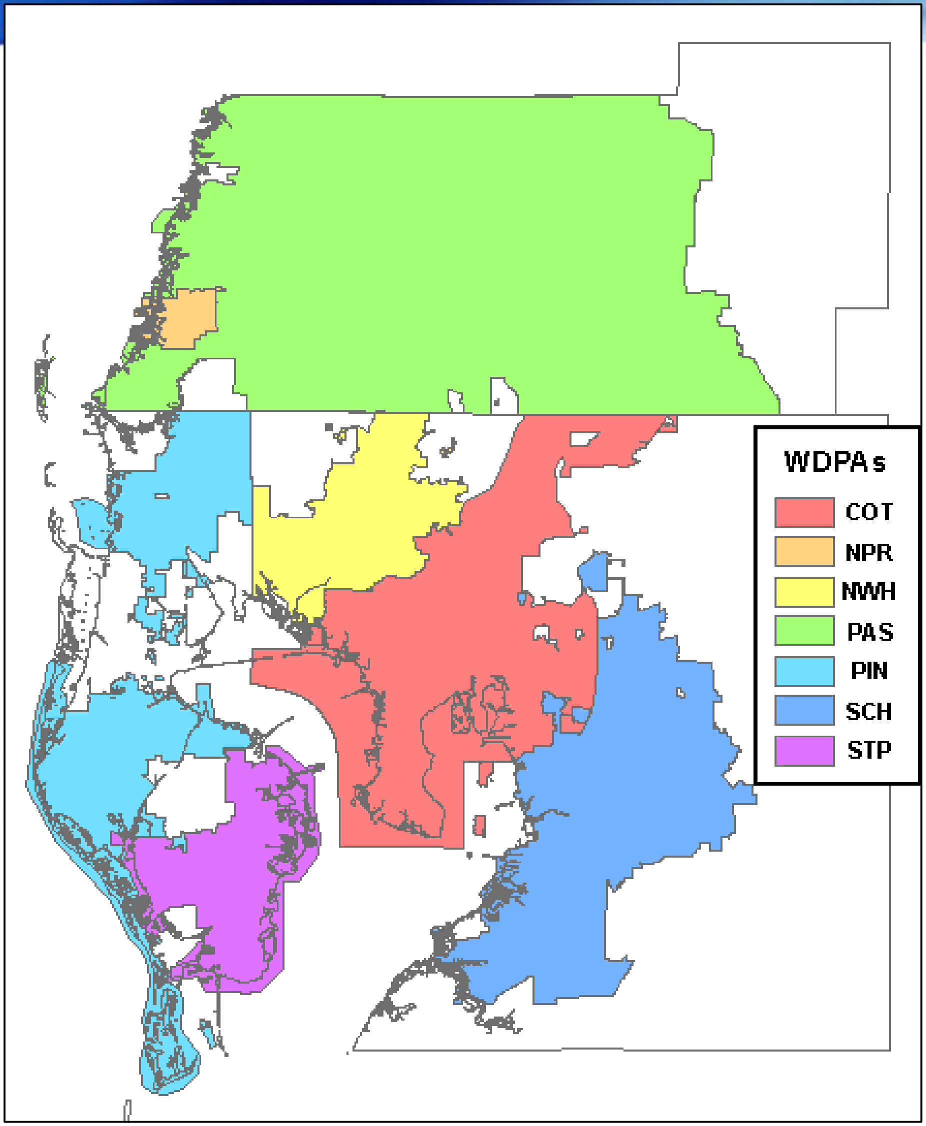
Counties

- Hillsborough
- Pasco
- Pinellas



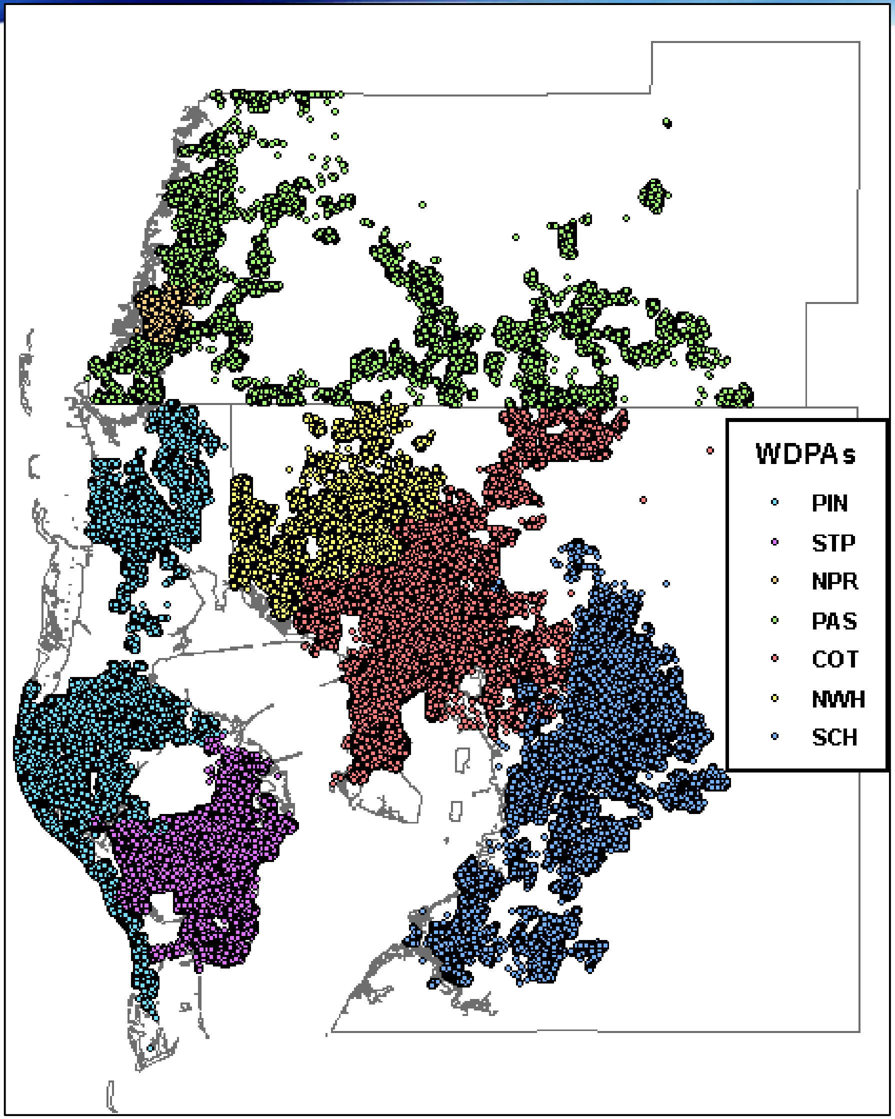
WDPAs

- City of Tampa (COT)
- New Port Richey (NPR)
- Northwest Hillsborough (NWH)
- Pasco County (PAS)
- Pinellas County (PIN)
- South Central Hillsborough (SCH)
- City of St. Petersburg (STP)



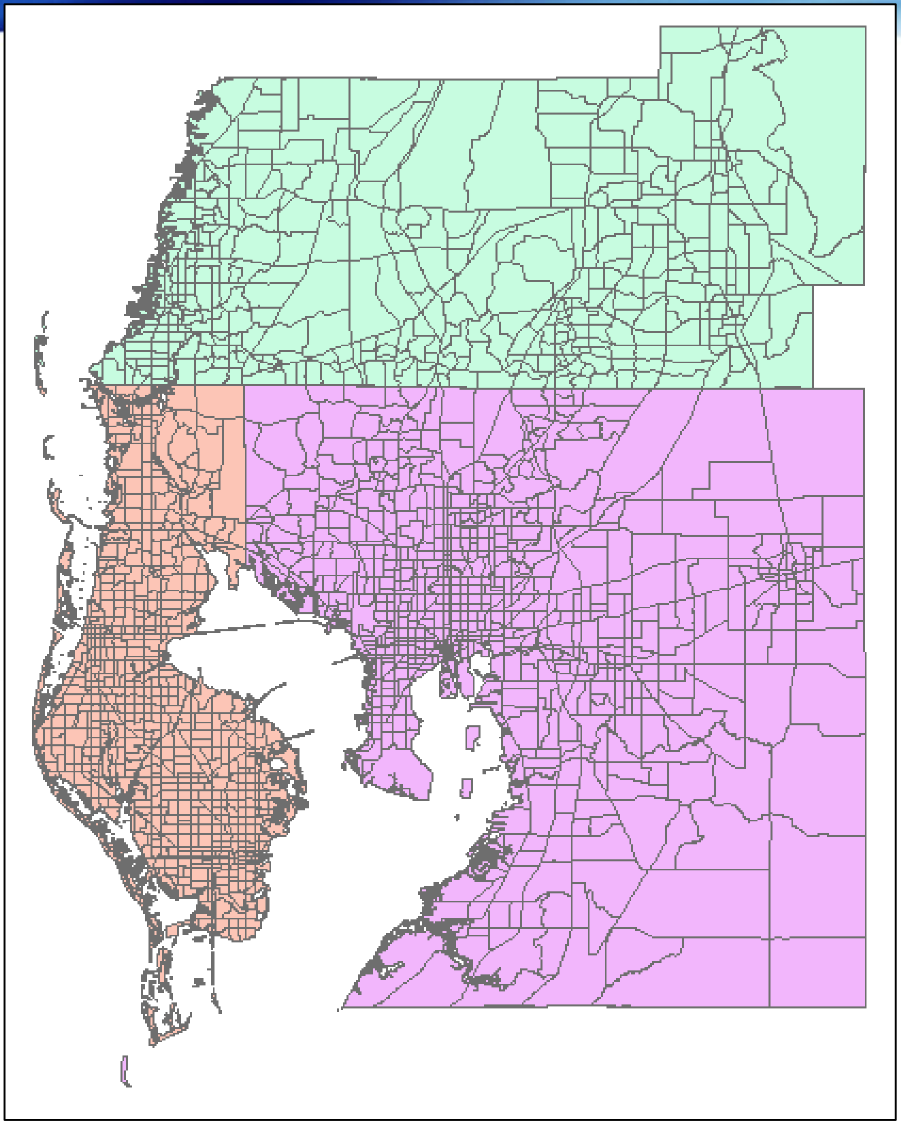
Water-using locations

WDPA	# of Locations
Pasco County	88,292
New Port Richey	7,590
Northwest Hillsborough	46,584
South Central Hillsborough	96,340
City of Tampa	121,805
Pinellas County	96,761
City of St. Petersburg	85,682



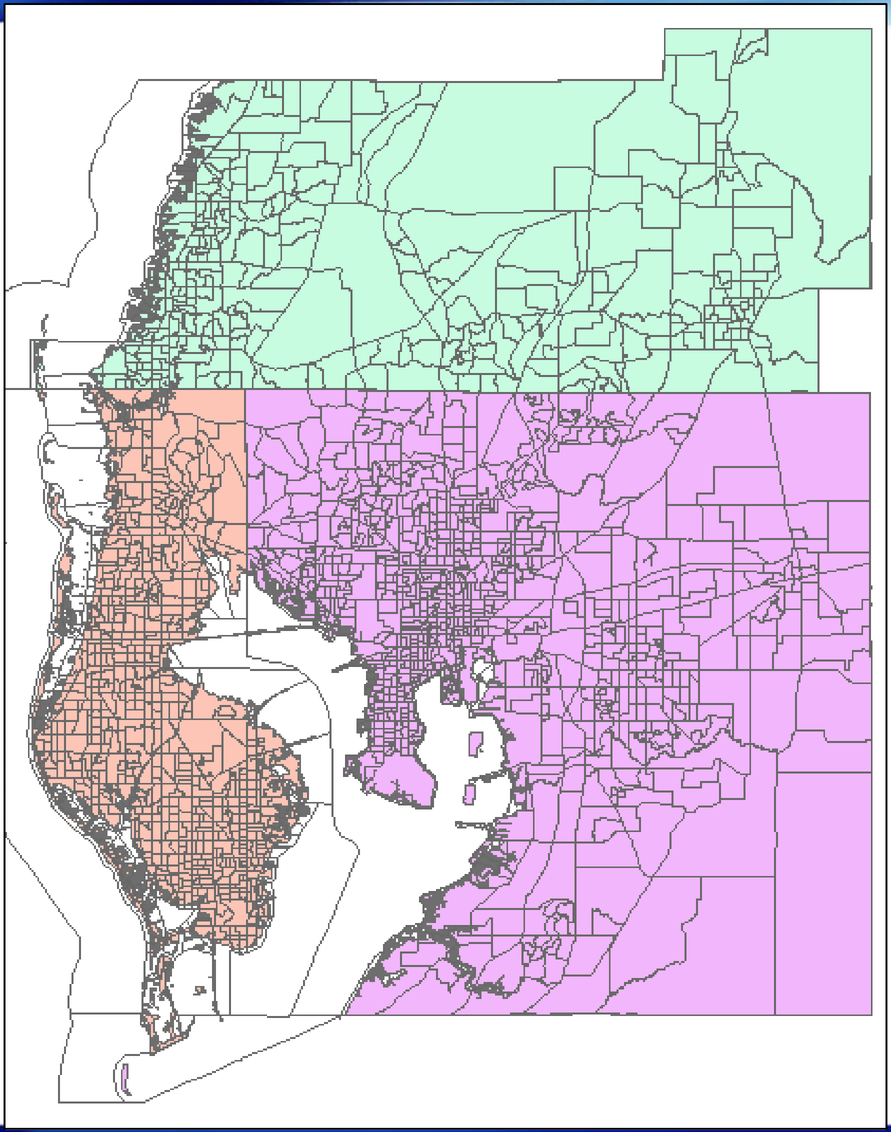
Traffic analysis zones (TAZs)

County	# TAZs	# TAZs having Water Use Locations
Hillsborough	794	689
Pasco	415	286
Pinellas	780	608
TOTAL	1,989	1,583



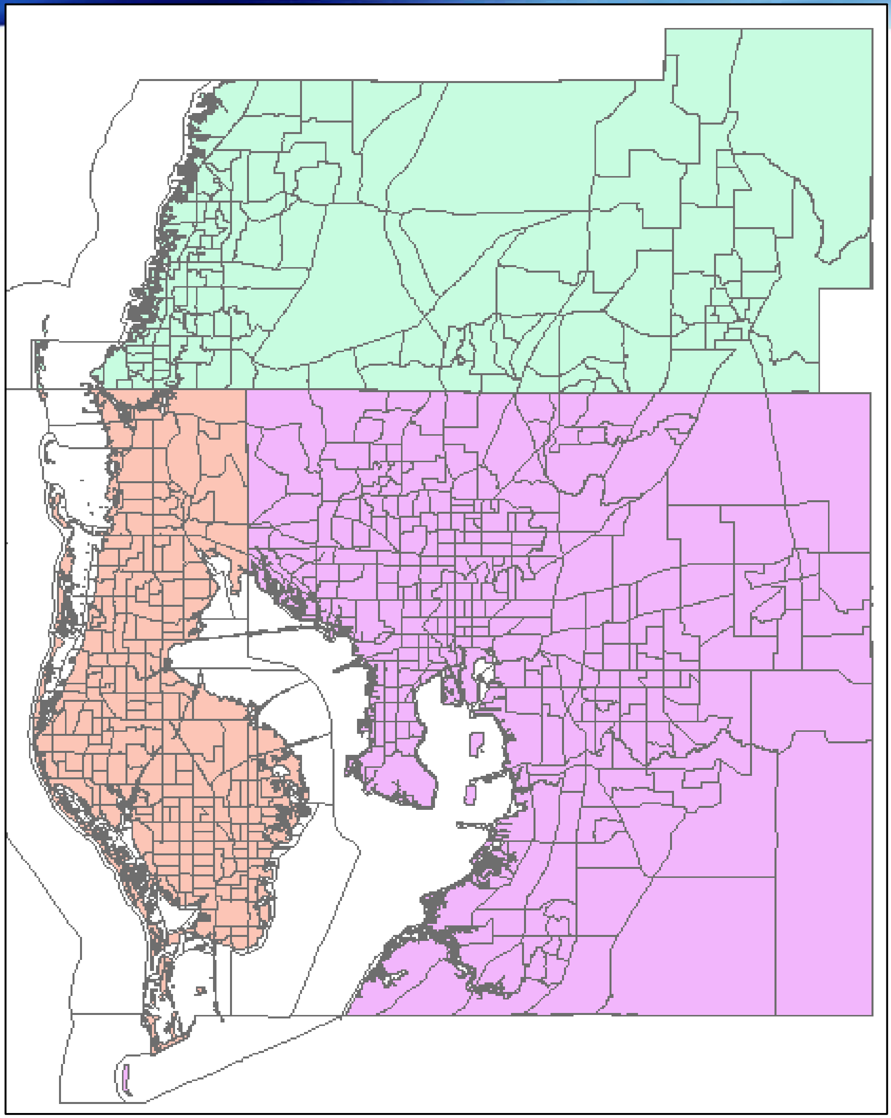
2010 Census Block Groups

County	# Block Groups	# Block Groups having Locations
Hillsborough	880	775
Pasco	308	264
Pinellas	721	572
TOTAL	1,909	1,611



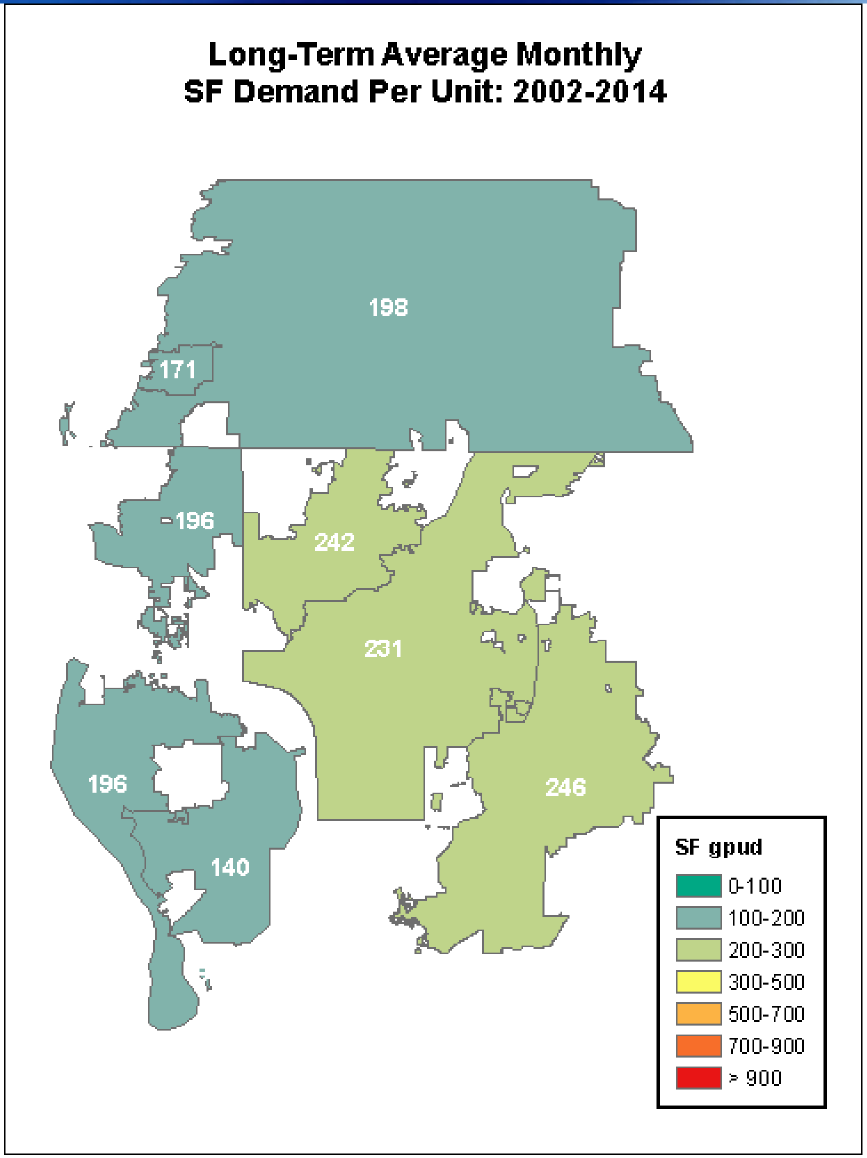
Census Tracts

County	# Tracts	# Tracts having Locations
Hillsborough	321	285
Pasco	134	124
Pinellas	246	220
TOTAL	701	629



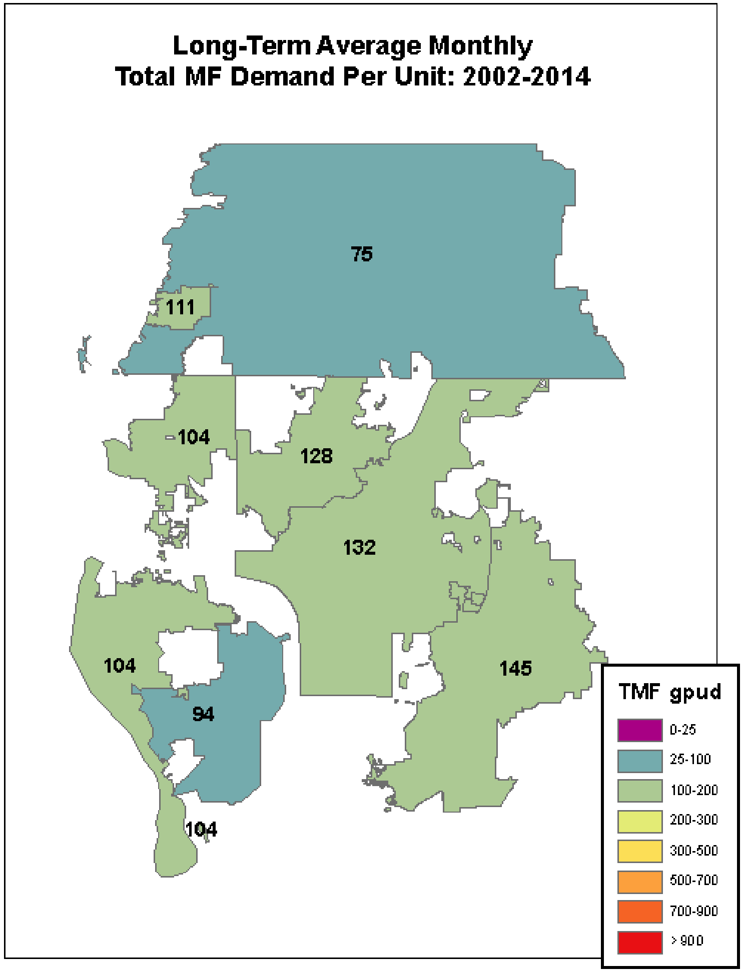
Spatial variability in water use per SF unit

- WDPA-level
  - PAS: 198 gpud
  - NPR: 171 gpud
  - NWH: 242 gpud
  - **SCH: 246 gpud**
  - COT: 231 gpud
  - PIN: 196 gpud
  - **STP: 140 gpud**



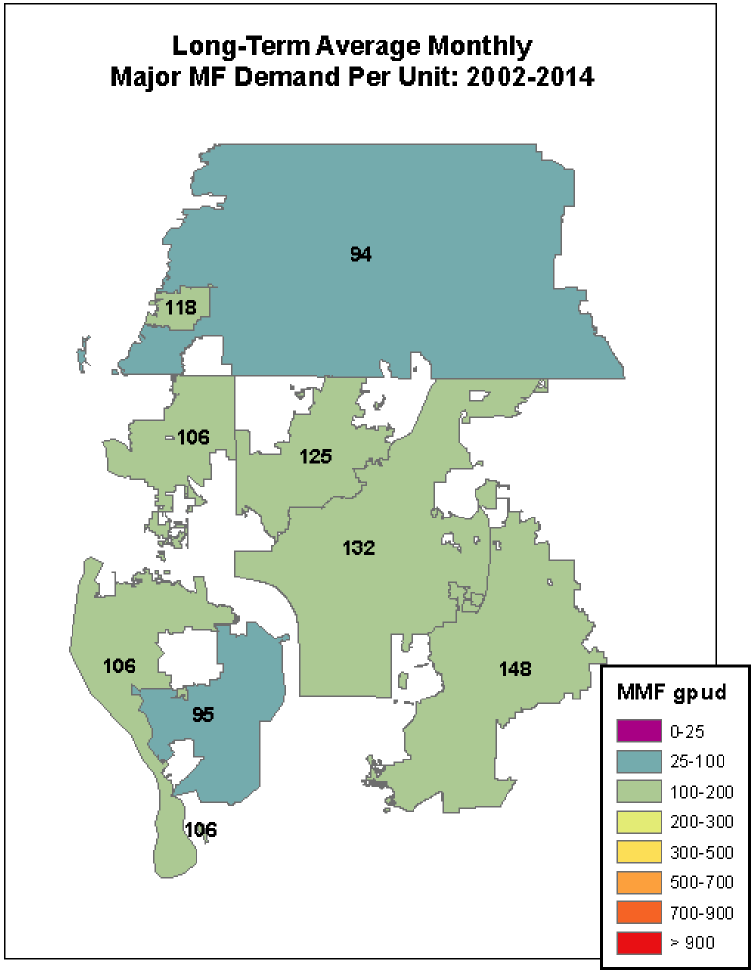
Spatial variability in water use per MF unit

- **PAS: 75 gpud**
- NPR: 111 gpud
- NWH: 128 gpud
- **SCH: 145 gpud**
- COT: 132 gpud
- PIN: 104 gpud
- STP: 94 gpud



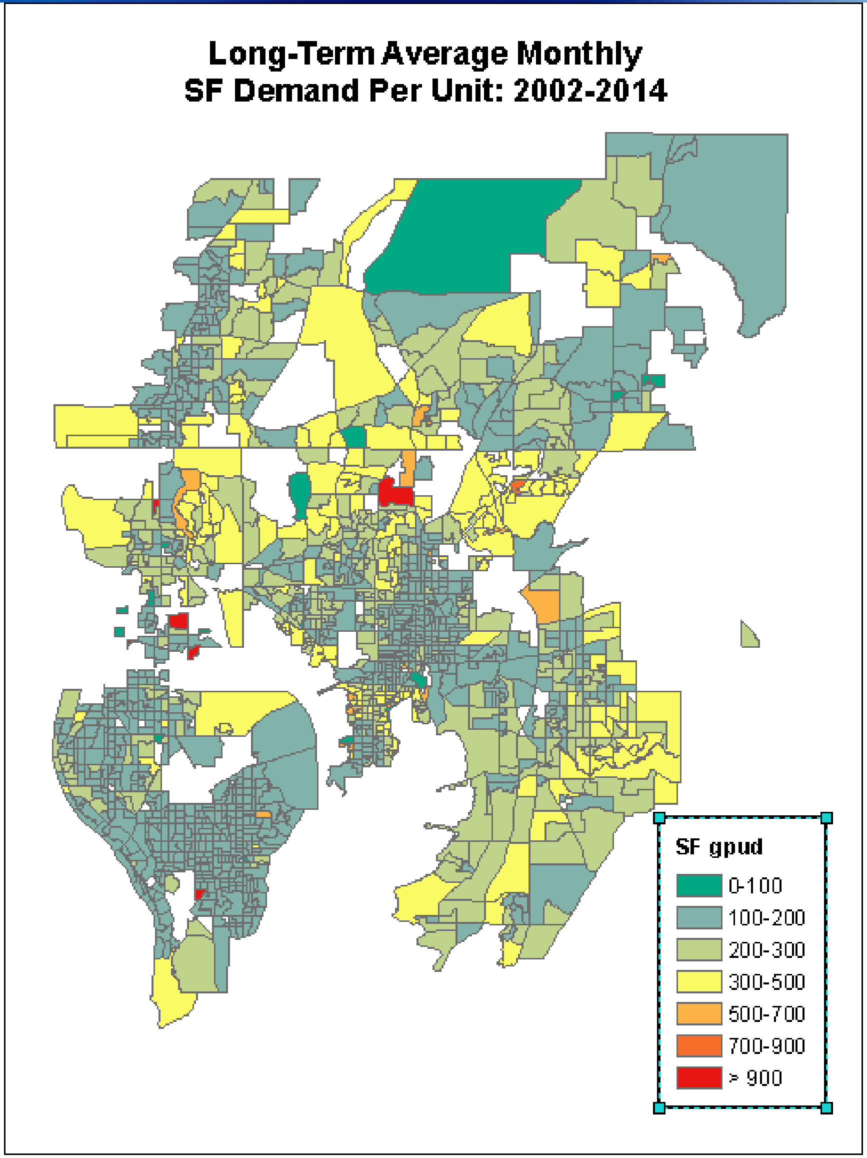
Spatial variability in water use per MF unit (drop Mobile Homes)

- **PAS: 94 gpud**
- NPR: 118 gpud
- NWH: 125 gpud
- **SCH: 148 gpud**
- COT: 132 gpud
- PIN: 106 gpud
- STP: 95 gpud



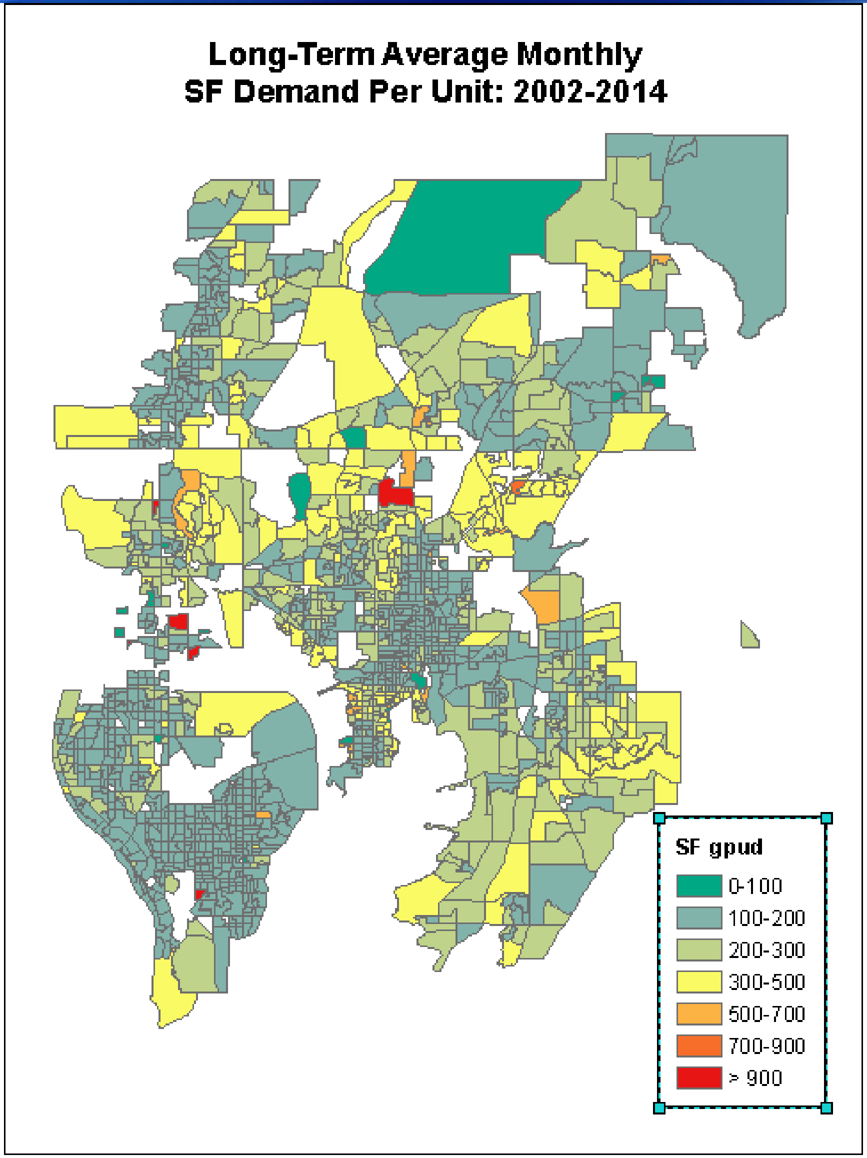
Spatial variability in water use per SF unit

- Block-group level
  - Variation much greater than WDPA level
  - Differences in demographics over small geographic scales

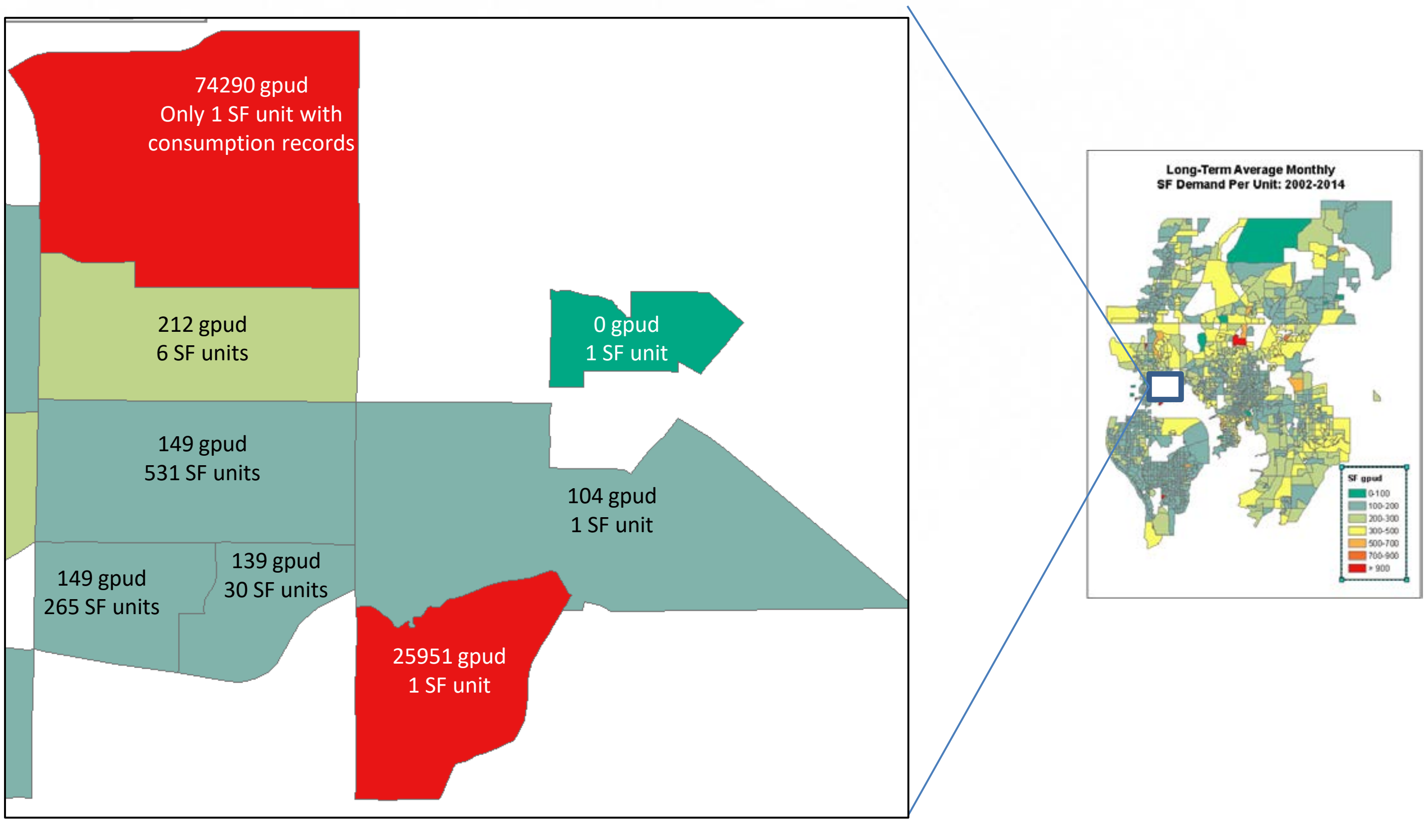


Spatial variability in water use per SF unit

- Hotspots and Coldspots
  - Extremely high and low demands
  - Investigate for possible data errors/instability



Spatial variability in water use per SF unit

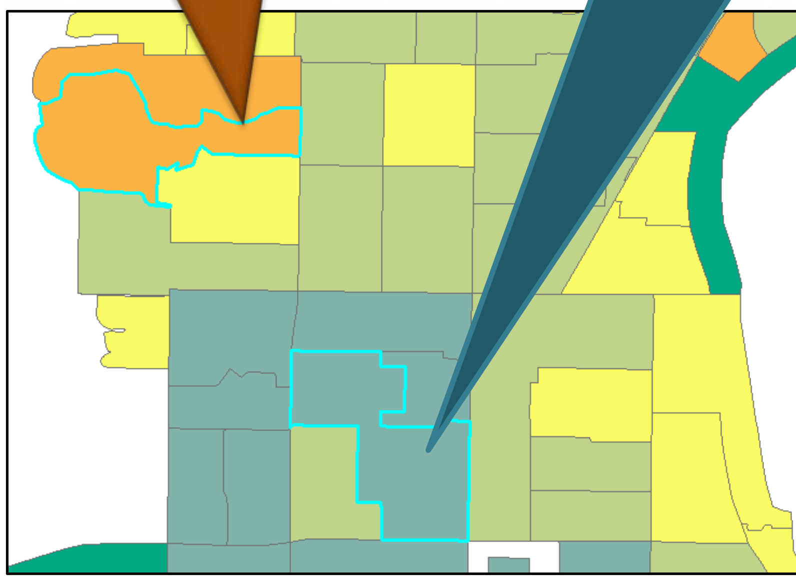


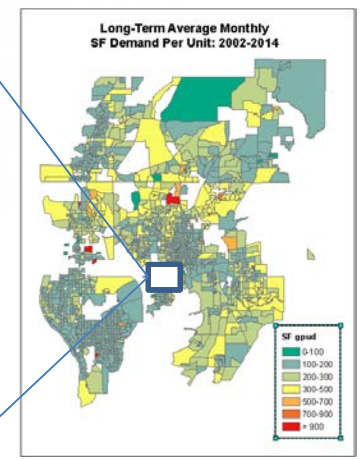
Block Group 120570064002

- Avg **SF** demand 522 gpud
- Median household income $155K (2015 $)
- Persons per household: 2.74
- Housing density: 3.39 units/acre

Block Group 120570066002

- Avg **SF** demand 183 gpud
- Median household income $58K (2015 $)
- Persons per household: 2.34
- Housing density: 5.32 units/acre

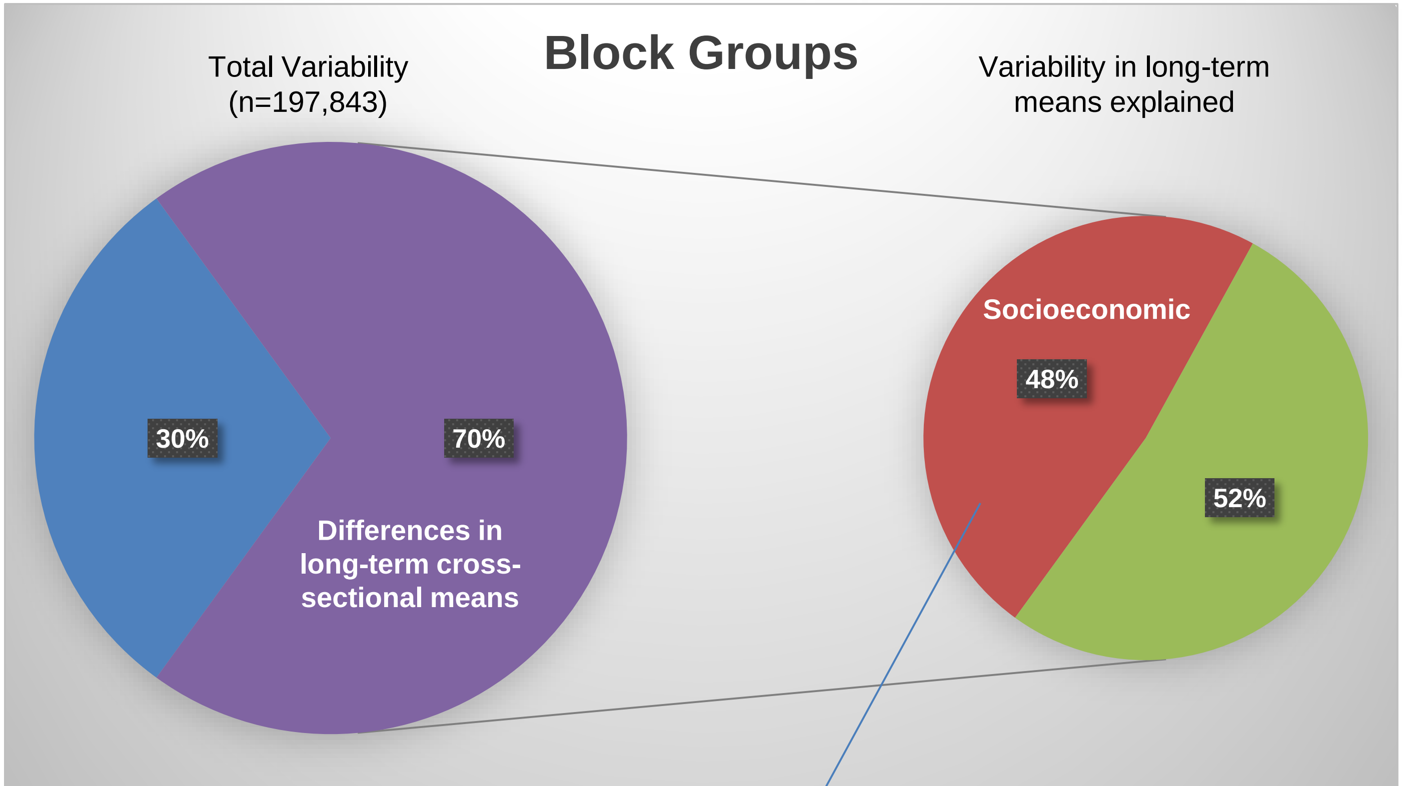




Econometric analysis

- Econometric models developed to explain variability
- Pooled time-series cross sectional data
  - Evaluate cross-sectional (geographic) component
  - Evaluate time series component
  - Combine
- Models by sector to include
  - Socioeconomics
  - Demographics
  - Land use
  - Access to reclaimed water
  - Weather and climate
  - Prices
  - Passive efficiency (trend)

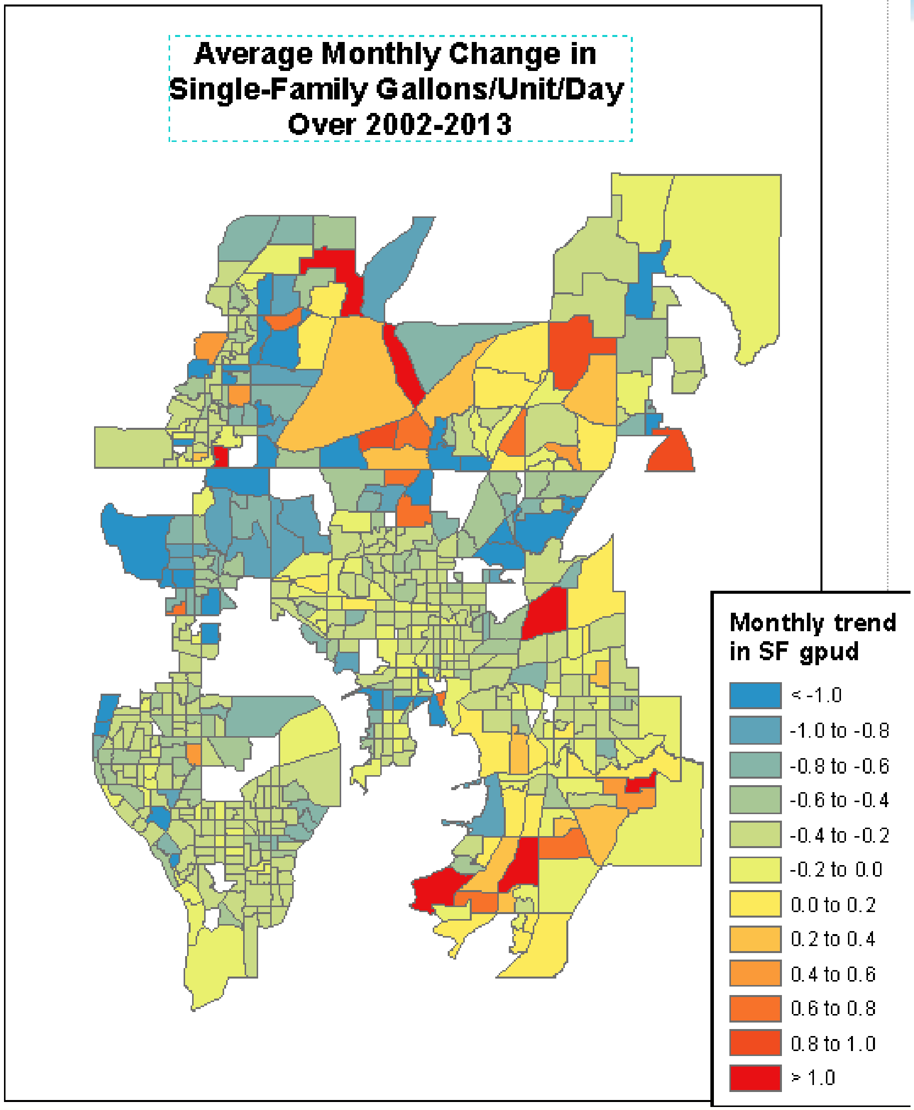
Example Models of Cross-Sectional Variability



$$SFq_i = k * INC_i^{0.38} * PRICE_i^{-0.51} * UNITS_{ACRE_i}^{-0.25} * HH_{SIZE_i}^{0.22} * RECL\%_i^{-0.05}$$

Trends in SF water use per unit

- Map of directional change
- What's behind trends?
  - Increases in housing density (-)
  - Increases in access to reclaimed water (-)
  - Increases in fixture efficiency (-)
  - More outdoor use in outer growth areas (+)

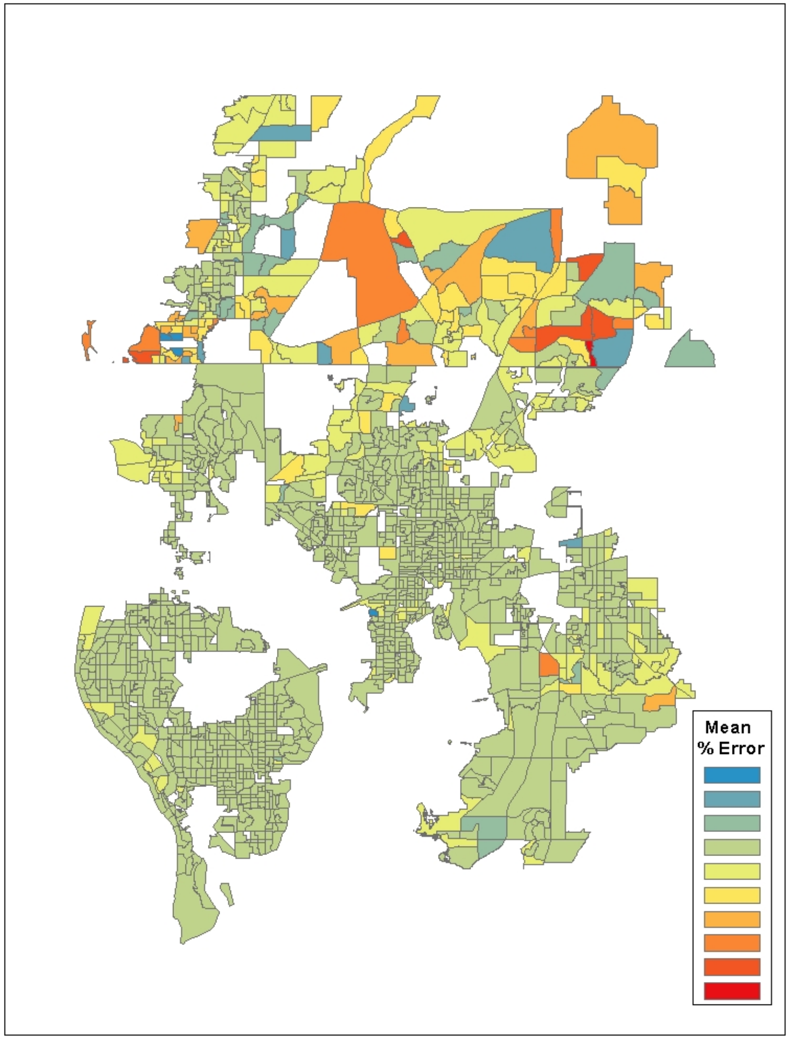


Modeling cross-sectional AND time-series data

- What have we learned thus far?
  - Results using Block Group, TAZ, and Tract averages are very similar
  - "Averaging up" to WDPAs increases the magnitude of estimated socioeconomic effects
    - Cross-sectional differences much greater than time variability
    - Parameter estimates compensate
  - Climatic model components are consistent across all geographic levels
  - 7% to 10% reduction in use over 12 years independent of specified price and socioeconomic effects

Spatial model error analysis

- Map of average SF model prediction errors by block group
- Large majority within +/- 3% of observed
- Higher errors in "transitional" areas



Conclusions

Major benefits of collaborative database and EDA

- Model development becomes transparent to the Agency
- A big box but not a black box
- Knowledge gain and transfer (2-way)
- Enhanced understanding of demand variability and uncertainty
- Permits explicit integration of demand forecast into supply planning process

Conclusions

Lessons learned

- Lot of data means lot of options
- Database development at granular spatial level requires QAQC
- Exposes staffing barriers and experience deficiencies that can be corrected

Next Steps

- Finalize recommended models for residential sectors
- Nonresidential EDA and model development
  - Enhanced classification capabilities
  - Additional uses outside forecasting
- Collaborative probabilistic forecast development
- Integration into supply reliability planning and demand management processes

Thank you! Questions?

David Bracciano, Demand Management Coordinator
dbracciano@tampabaywater.org
727-791-2313

Jack Kiefer, PhD, Senior Associate, Hazen and Sawyer
jkiefer@hazenandsawyer.com
618-889-0498