HENRY
Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

Conference Paper, Published Version

Cera, Tim

Calibratin Of St.Johns River Watershed Hydrology Models for the Water Supply Impact Study

Zur Verfügung gestellt in Kooperation mit/Provided in Cooperation with:
Kuratorium für Forschung im Küsteningenieurwesen (KFKI)

Verfügbar unter/Available at: https://hdl.handle.net/20.500.11970/109712

Vorgeschlagene Zitierweise/Suggested citation:
Cera, Tim (2012): Calibratin Of St.Johns River Watershed Hydrology Models for the Water Supply Impact Study. In: Hagen, S.; Chopra, M.; Madani, K.; Medeiros, S.; Wang, D. (Hg.): ICHE 2012. Proceedings of the 10th International Conference on Hydroscience & Engineering, November 4-8, 2012, Orlando, USA.

Standardnutzungsbedingungen/Terms of Use:

Die Dokumente in HENRY stehen unter der Creative Commons Lizenz CC BY 4.0, sofern keine abweichenden Nutzungsbedingungen getroffen wurden. Damit ist sowohl die kommerzielle Nutzung als auch das Teilen, die Weiterbearbeitung und Speicherung erlaubt. Das Verwenden und das Bearbeiten stehen unter der Bedingung der Namensnennung. Im Einzelfall kann eine restriktivere Lizenz gelten; dann gelten abweichend von den obigen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Documents in HENRY are made available under the Creative Commons License CC BY 4.0, if no other license is applicable. Under CC BY 4.0 commercial use and sharing, remixing, transforming, and building upon the material of the work is permitted. In some cases a different, more restrictive license may apply; if applicable the terms of the restrictive license will be binding.

Proceedings of the 10th Intl. Conf.on Hydroscience & Engineering, Nov. 4-7, 2012, Orlando, Florida, U.S.A.

CALIBRATION OF ST. JOHNS RIVER WATERSHED HYDROLOGY MODELS FOR THE WATER SUPPLY IMPACT STUDY

Tim Cera[1]

ABSTRACT

This paper describes the calibration process for the Water Supply Impact Study (WSIS) watershed hydrology models. Within the St. Johns River watershed there were 47 flow or stage data collection sites during the calibration period (1995-2005) covering 260 of the approximately 900 subwatersheds.

Hydrological Simulation Program - FORTRAN (HSPF) was used for the watershed simulations. HSPF is a water balance model, with a mix of physical and empirical process representations, and is highly parameterized. The calibration process started with a series of workshops for the modeling team to come to uniform understanding of model parameter ranges, estimates of parameter sensitivity, model representation of processes, input data, refining an existing 'common logic' for HSPF modeling of Florida hydrology. The common logic was then used to configure the Parameter ESTimation (PEST) suite of utilities for model parameter optimization.

The Nash-Sutcliffe efficiency (NSE) statistic was used to indicate with one number the calibration performance. A NSE of 1 is a perfect simulation of the observed data, and 0 would mean that the average of the observed data would be better than the simulation. Of the calibrated locations, NSE of 41 calibrations were greater than 0.5 and seventeen were greater than 0.75 which corresponds to ratings of 'satisfactory' and 'very good' respectively.

1. INTRODUCTION

In 2006, the St. Johns River Water Management District (SJRWMD), South Florida Water Management District (SFWMD), and the Southwest Florida Water Management District (SWFWMD) recommended capping groundwater use at the 2013 demand level in order to prevent harm to ground water resources and natural systems of the region. The three districts agreed that alternative water supply (AWS) sources would need to be developed to meet water demands above the 2013 level. The SJRWMD's Water Supply Plan (St. Johns River Water Management District 2006) identified several potential AWS sources. Among these sources were surface water from the Ocklawaha River and surface water from the St. Johns River.

The Governing Board of the SJRWMD subsequently determined that a water supply impact analysis was necessary to ensure that the withdrawal of surface water from the Ocklawaha and St.

---

[1] Senior Professional Engineer, Bureau of Engineering, St. Johns River Water Management District, Palatka, FL, 32178-1429, USA (tcera@sjrwmd.com)

Johns rivers would not cause unacceptable environmental effects in the St. Johns River. Consequently, the board called for the St. Johns River Water Supply Impact Study (WSIS). The goal of the St. Johns River Water Supply Impact Study was to provide an analysis of the potential environmental effects to the St. Johns River of surface water withdrawals for consumptive use.

As part of this study, the District determined that the development of basin-scale framework computer models would best meet the current and future needs to assist the District in managing water resources in a cost and time efficient manner. A framework model is a large-scale computer model that simulates the hydrologic and water quality processes in a basin with adequate detail to be meaningful. The simulation environment must address relevant issues related to the computer simulation of hydrologic, hydrodynamic, and water quality processes in selected District watersheds and District receiving water bodies. For watershed modeling, the District chose the Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) interface and the Hydrologic Simulation Program – Fortran (HSPF) as the framework model. The Environmental Protection Agency (EPA) has sponsored the BASINS and the HSPF projects for many years for hydrologic and water quality simulations. Combined they are used by the EPA and stakeholders across the country to assist in the development of Total Maximum Daily Loads (TMDL) and they are part of the EPA's TMDL Toolkit.

The SJRWMD recently completed the WSIS. See reference (St. Johns River Water Management District, 2012) for the location of the published report.

This paper describes the calibration process for the WSIS watershed hydrology models fully detailed in Chapter 3 of the WSIS report (St. Johns River Water Management District, 2012). Within the St. Johns River watershed there were 47 flow or stage data collection points during the calibration period (1995-2005) covering 260 of the approximately 900 subwatersheds.

As mentioned earlier, HSPF was used for the watershed hydrology simulations. HSPF is a water balance model, with a mix of physical and empirical process representations, and is highly parameterized. The calibration process started with a series of workshops for the modeling team to establish a ‘common logic’ to come to uniform understanding of model parameter ranges, estimates of parameter sensitivity, model representation of processes, input data, and Florida hydrology. The common logic was then used to configure the Parameter ESTimation (PEST) suite of utilities for model parameter optimization.

PEST sees a model as anything that accepts text files as input and produces text files as output. Scripts or batch files can act as ‘composite models’ to pre-process model input files, run the actual model, and then post-process the model output files before used by PEST.

For the HSPF models PEST developed an optimized parameter set by minimizing the difference between the model output hydrograph and five objective functions; the observed frequency distribution curve and the observed daily, monthly, annual, and calibration period average flows. Since PEST replaces the manual adjustment and testing of possibly hundreds of parameters, each modeler had time to evaluate their calibration in light of their knowledge of hydrology and the subwatershed they were simulating.

The PEST optimization process could require thousands of runs depending on the number of PEST adjustable parameters to be optimized. To make these model runs as efficient as possible Parallel PEST, and later BeoPEST was used to run the models in parallel on the SJRWMD Linux cluster.

The Time Series PROCessor (TSPROC), part of the PEST Surface Water Utilities package, was used in two different ways (called ‘contexts’ within TSPROC) in the calibration process. The first use was for development of the initial PEST control and instruction files, which can be very complicated if setup by hand. The second use was for post-processing of the results of the HSPF model in the ‘composite model’ seen by PEST.

The primary purpose of the watershed hydrology models for the WSIS project was to establish flow boundary conditions to the mainstem hydrodynamic model. Only the watershed hydrology models for the Upper St. Johns River Basin (USJRB) were used directly for the evaluation of the WSIS surface water withdrawals since the USJRB is outside of the domain of the mainstem hydrodynamic model.

The general form and character of the calibration process presented herein is shown in Figure 1.

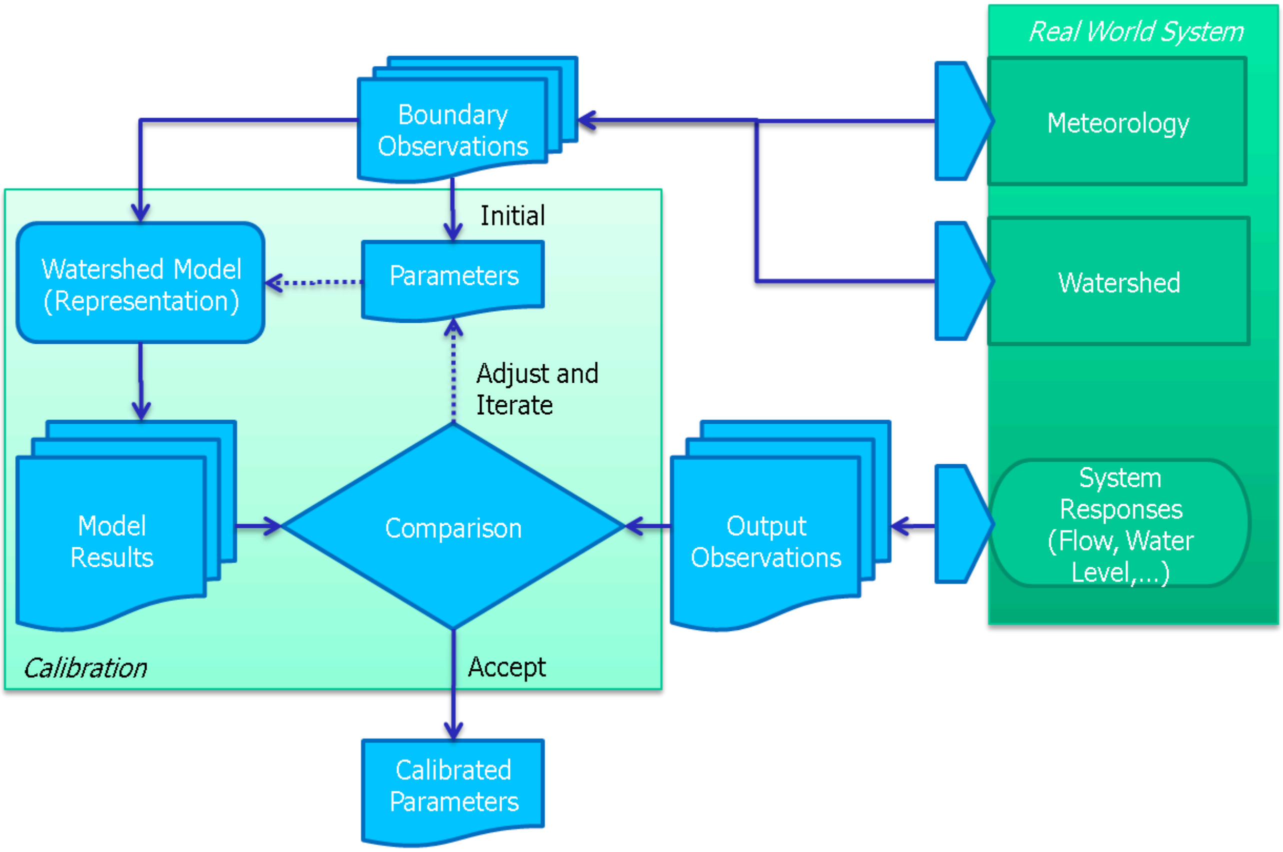

Figure 1 Conceptual diagram of the calibration process.

2. OBSERVED DATA

2.1. Observed Precipitation

The SJRWMD maintains both point rain gauge and Doppler radar rainfall data sets. A contractor creates for SJRWMD a daily Doppler radar rainfall data set on a 2-km grid adjusted to a network of rain gauges. This adjustment forces the Doppler total rainfall over long periods to match the total from the coincident rain gauges. The Doppler radar rainfall data starts in 1995 and continues to the present. The SJRWMD also acquired National Weather Service (NWS) data for the simulation period (1974 to 2008) from 25 separate daily and hourly point rain gauges throughout the St. Johns River watershed. The Doppler and point rain gauges form fundamentally different data sets and cannot be intermingled. The primary difference is that Doppler averages rainfall over a relatively large area (2 × 2 km) while gauges provide data at a specific point (0.2 m circle). Because of the difference in spatial scale, Doppler records rainfall that the rain gauges miss and averages intense rainfall over the grid cell.

Although many of the watershed models that formed the foundation of this project were already calibrated using the Doppler radar data set, long-term statistics and analyses were needed for WSIS that would cover at least 30 years. The Doppler radar data set only provides 13 years of rainfall data, whereas some NWS stations have data back to the early 1900s. This long-term simulation requirement forced the use of the NWS point rain gauges for the scenario simulations. Because the Doppler and rain gauge data sets are fundamentally different and should not be intermingled, the watershed models were recalibrated and all scenarios run with the long-term NWS rain gauge data. The period chosen for the model scenario simulations ran from 1975 through 2008. The weather and climate was variable during this time and can be considered a good representation of long-term rainfall and evaporation patterns.

A Thiessen polygon network was developed to establish the area of influence for the NWS rain gauges used in this study, but was not used to weight the rainfall amounts (Figure 2). Even though more evenly distributed rain volumes can be obtained by area weighting of the multiple rain gauges that cover each watershed, this process can also reduce rainfall intensities. Rainfall intensity is a major factor in determining surface runoff. A reduction in intensity by area weighting can arbitrarily shift the model parameters to increase infiltration and reduce simulated runoff. Therefore, a single rain gauge was selected for each subwatershed based on the Thiessen polygon area that covers the majority of the subwatershed.

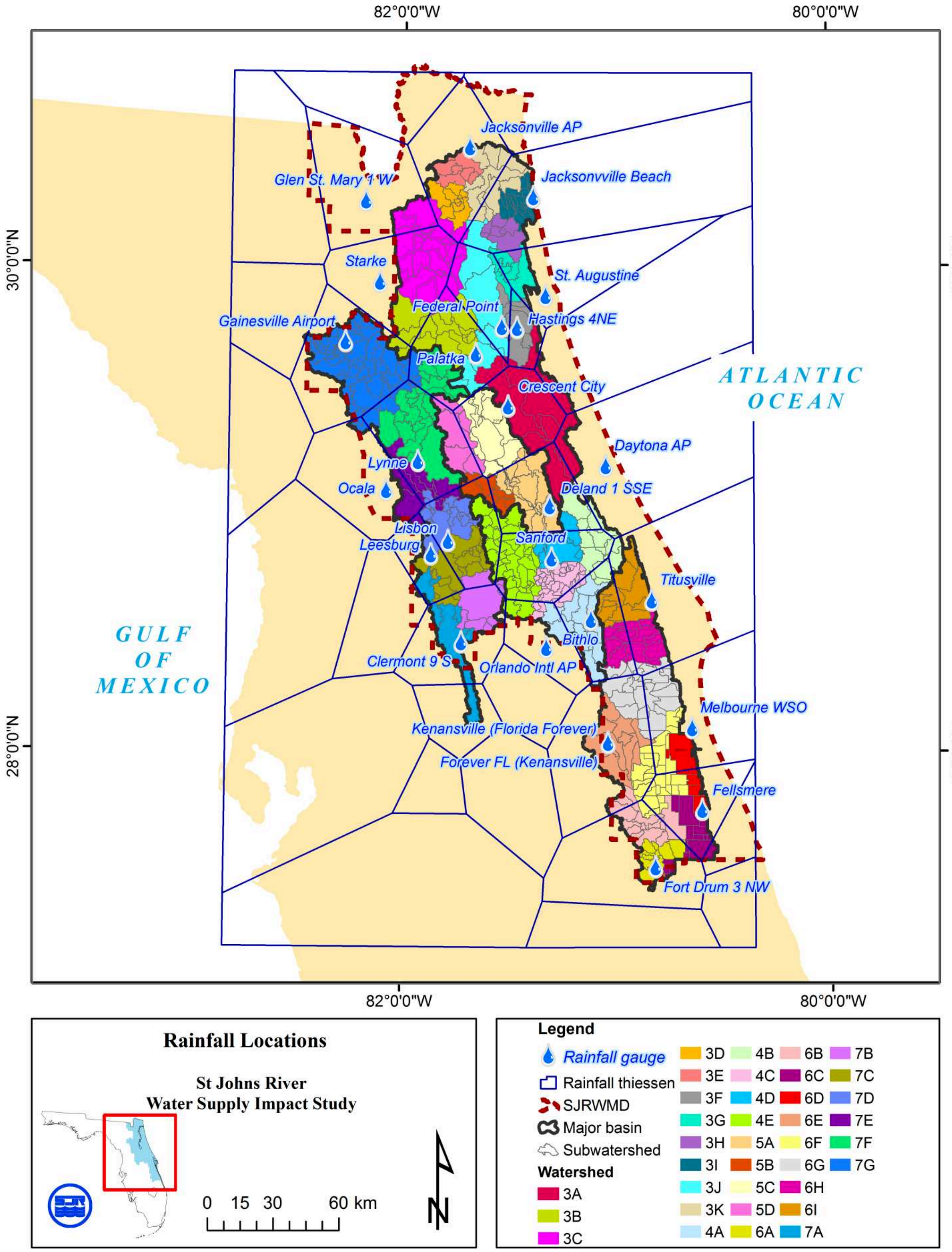

Figure 2 Rain gauge station locations and Thiessen polygons indicating area of influence.

Average annual rainfall varies from 46 to 57 in. across SJRWMD (Figure 3). Note that in most of SJRWMD, the 1995 through 2006 average rainfall is slightly higher than the longer-term average from 1960 through 2006. The primary cause of this difference is increased hurricane and tropical storm landfall within SJRWMD. Rainfall amounts vary greatly on an annual basis. Figure 4 shows the annual variation of SJRWMD annual averages for Orlando and Jacksonville International Airport. There is no easily identifiable trend either spatially or in time between these two stations.

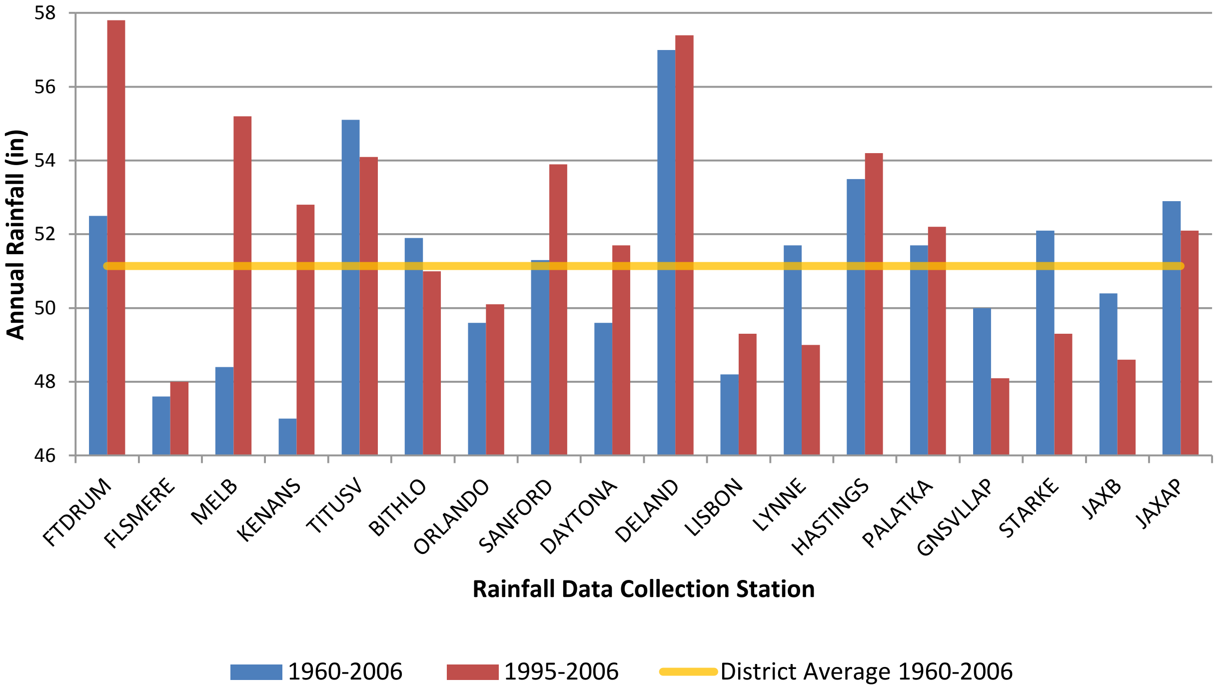

Figure 3 Average annual precipitation at rainfall gauge stations, arranged approximately south to north.

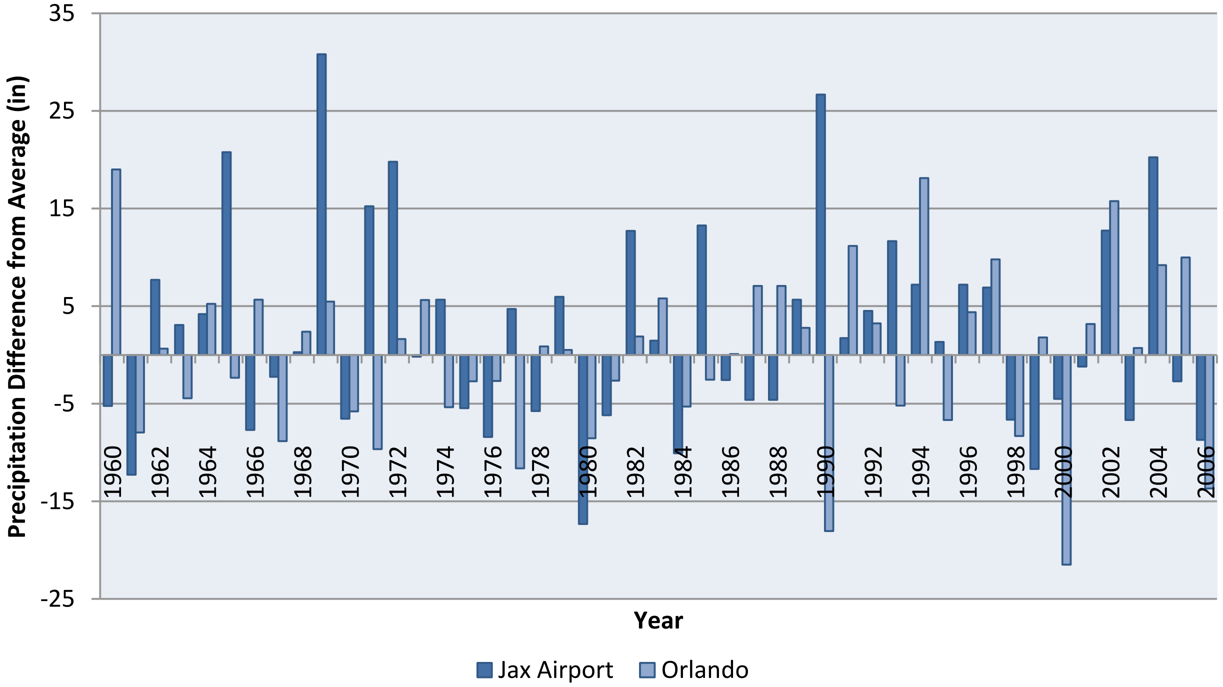

Figure 4 Yearly differences from station average in precipitation at Jacksonville International Airport and Orlando rain gauges.

The spatial distribution of rainfall varies widely across the St. Johns River watershed. The entire river watershed receives rain on the same day less than 0.5% of the time and receives no rain 14% of the time. It rains an average of 104 days yr^{-1} with a range throughout the St. Johns River watershed of 75 to 120 days as shown on Figure 5.

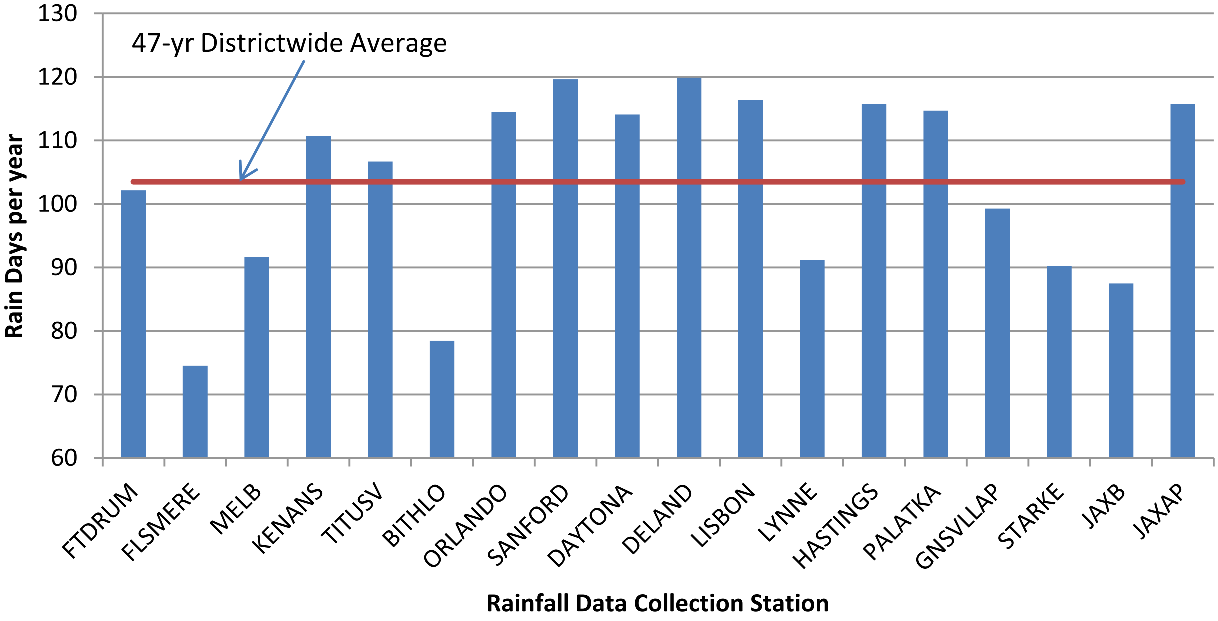

Figure 5 Days per year with precipitation at rainfall gauge stations, arranged approximately south to north

The NWS rain data are not processed by the NWS to fill in missing values or address other data issues. NWS uses flags when a value is good, missing, or accumulated (indicating a total value from several previous time intervals). To be useful for modeling and compilation of statistics, all missing data must be filled in with estimated data and all data marked as accumulated must be disaggregated into the appropriate previous time steps.

The processing of the rainfall data involved the following three steps:

1. Disaggregating accumulated data and assigning values to previous time intervals so that the total over those previous intervals equals the accumulated value
2. Filling in missing data
3. If daily, then disaggregating daily data to hourly data

The precipitation records from the NWS stations sometimes have flagged records indicating an accumulation since the last recorded value. For example, an hourly gauge may have a value for 1:00 a.m. and an accumulated value at 5:00 a.m. There is no information about the period between the data points, so the 5:00 a.m. point is an accumulation of rain since 1:00 a.m. Distributing accumulated values into the hourly values within the aggregation interval involves the following process:

- Distribute accumulated rainfall across the period using the closest volumetric rainfall from nearby hourly stations. The nearest station gets priority as a reference; however, if a secondary station is significantly closer to the total accumulated rainfall, the secondary station is used as a reference for distributing accumulated data.
- If none of the nearby stations have a reasonable distribution, then a triangular distribution is used to estimate data. Small events (under 0.5 in.) may generally fall in 1 hr in the afternoon or evening. Larger events should follow a triangular distribution over three to five hrs.
- All hourly NWS stations have a separate daily recorder, except Lynne (Table 1). Missing hourly rainfall data are estimated by using the daily values from the same location. If the

rainfall volumes are consistent, the missing hourly data are estimated from the daily data according to the process described above for aggregated data. The hourly data are compared against monthly and annual totals from the daily rain data set to make sure a significant rainfall event was not missed in the hourly rainfall data.

Table 1 Rain accumulation interval at rain gauges sorted from upstream to downstream (south to north)

Rain Gauge	Daily Recorder	Hourly Recorder
Fort Drum	X	
Fellsmere	X	
Melbourne	X	X
Kenansville	X	
Titusville	X	
Bithlo	X	
Orlando	X	X
Sanford	X	
Daytona	X	X
DeLand	X	
Lisbon	X	X
Lynne		X
Ocala	X	
Crescent City	X	
Hastings	X	
Palatka	X	
Gainesville Airport	X	X
Starke	X	
Jacksonville Beach	X	
Jacksonville Airport	X	X

The software package Watershed Data Management Utility (WDMUtil) was used to disaggregate daily rainfall data to hourly rainfall data. WDMUtil is a powerful tool for hydrologic data visualization, statistics, editing, and management of Watershed Data Management (WDM) files (Lumb, Carsel, & Kittle, 1988). WDMUtil is now part of the BASINS project funded by the EPA.

The rainfall data were loaded into a WDM file where WDMUtil was used to estimate an hourly rainfall distribution for each site, when necessary. As stated above, the two closest hourly rainfall stations to each daily rainfall station were used for this estimation. Only long-term National Oceanic and Atmospheric Administration (NOAA) hourly stations were used to disaggregate the long-term NWS rainfall data.

2.2. Evaporation

Potential evaporation is defined as the evaporation from a shallow body of water. Traditional potential evaporation data are estimated by measuring the water level in a shallow pan of water called a “Class A” pan. A factor of around 0.75 is applied to pan evaporation data to account for all of the unknowns (such as heating of the pan itself) that would tend to overestimate potential evaporation. The pan factor is dependent on local conditions and varies among “Class A” pans, but in all cases it is determined by professional judgment. Rarely are all of the site requirements satisfied for a “Class A” pan. Pan evaporation data within SJRWMD are sparse, problematic, inaccurate, and highly variable among the few data collection sites available.

Because of the problems with the pan evaporation estimates, potential evaporation estimates were developed for this project using the Hargreaves method scaled with a factor to a detailed evaporation estimate using the Priestly-Taylor method. The Priestly-Taylor method was applied by the USGS in a cooperative project with SJRWMD and others to use satellite measurements of radiation for the evaporation estimate. This method provides a consistent evaporation estimate both spatially (2 × 2 km) and temporally across SJRWMD. Unfortunately, the period of record only runs from 1995 to 2007 and simulation of the WSIS scenarios required input data from 1975 through 2008. As part of the plan to standardize long-term input data to the HSPF hydrologic models, an estimate of potential evaporation was developed based on the Hargreaves equation. The Hargreaves method requires only measured maximum and minimum air temperature data and seems to be less sensitive than other methods (including Priestly-Taylor) by the condition of the data collection site, such as arid or semiarid climate and vegetated or non-vegetated land cover. Other than temperature, the Hargreaves method requires solar information including extraterrestrial radiation and sunlight hours, which is calculated from the time of year and the latitude of the station. Various studies have compared Hargreaves method against measured and estimated potential evaporation for 11 locations and Hargreaves method ranked the most accurate of all methods that require only temperature (Hargreaves & Allen, 2003).

A Thiessen network was developed to assign evaporation data from the 20 available meteorological stations used to calculate potential evaporation (Figure 6). The selection of the meteorological stations to use in each watershed was based on the station that covers the most area in the watershed based on the Thiessen network.

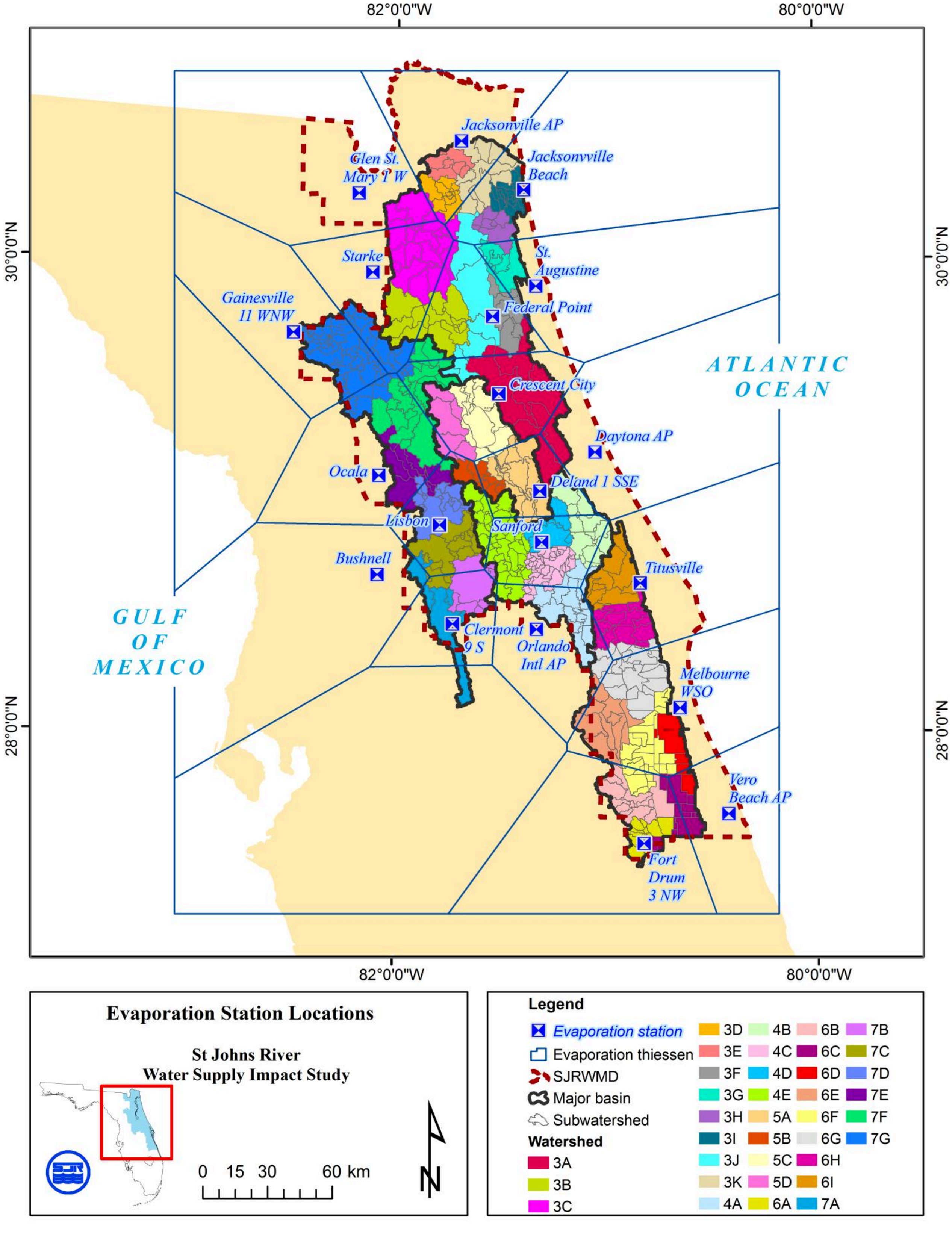

Figure 6 Evaporation station locations, where potential evaporation is computed, and Thiessen polygons defining each area of influence.

Temperature data obtained from 20 meteorological stations in four basins for the period of 1975 to 2006 were used to compute Hargreaves potential evaporation. The Hargreaves estimates are then adjusted to the Priestly-Taylor estimates using an adjustment coefficient for each of the sites.

The Hargreaves method that was used is summarized as follows. Extraterrestrial radiation is computed as a function of the declination of the sun and latitude using Equations 3–1 and 3–2. Equation 3–1 calculates declination of the sun (dec) where declination is in radians and J is the Julian day within the year.

$$dec = 0.4101 \sin\left[\frac{2\pi(J-80)}{365}\right] \tag{1}$$

Extraterrestrial radiation (R_a) is estimated using solar declination (dec) computed in Equation 1 and the longitude in radians ($lrad$).

$$R_a = \frac{118}{\pi}\{\cos^{-1}(-\tan(dec)\tan(lrad))\sin(lrad)\sin(dec) + \cos lrad \cos dec \sin \cos -1 - \tan dec \tan lrad \tag{2}$$

where R_a units are in $MJ\ m^2/day$.

Hargreaves potential evaporation (mm day^{-1}) is computed using Equation 3–4. In Equation 3–4, K is an adjustment constant obtained by regressing Hargreaves potential evaporation against what is considered a more robust Priestly-Taylor potential evaporation.

$$pot.\ evap. = K\left\{0.408 * 0.0023 * R_a * (T_{mean} + 17.8) * \sqrt{T_{max} - T_{min}}\right\} \tag{3}$$

An example of the process to determine the K adjustment factor for the Hargreaves estimates is presented in Figure 7 and Figure 8. The K coefficient for each evaporation station is presented in Table 2.

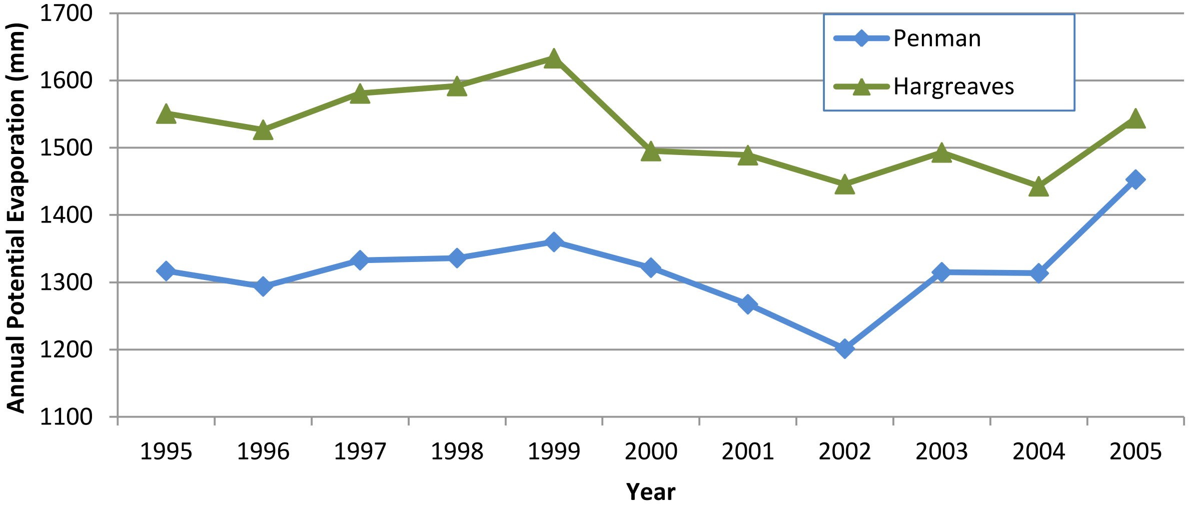

Figure 7 Annual potential evaporation comparison between Priestly-Taylor (with satellite radiation measurements) and Hargreaves for Gainesville.

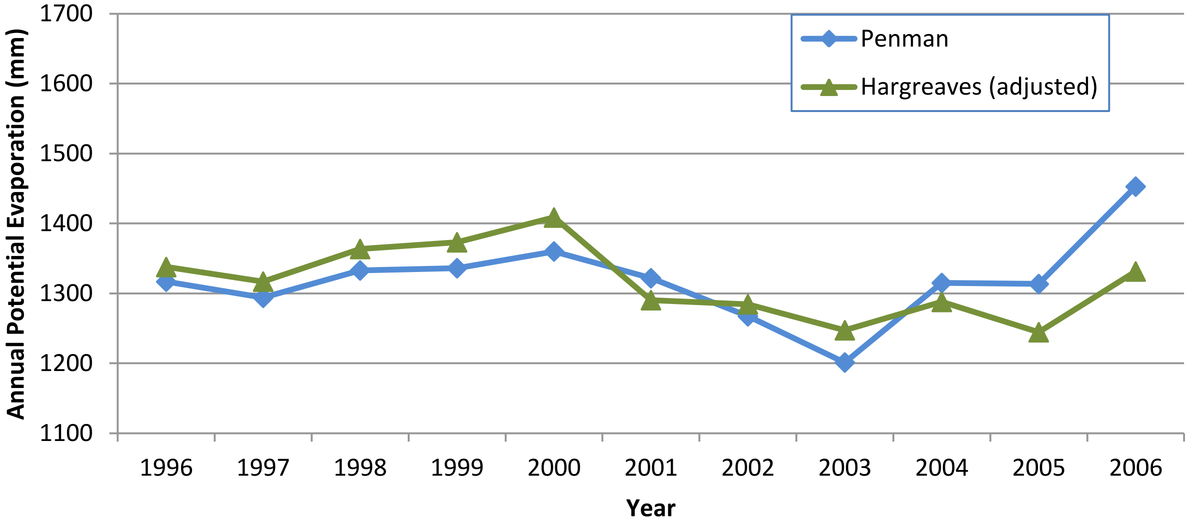

Figure 8 Annual potential evaporation comparisons after Hargreaves method was adjusted for Gainesville.

Table 2 Coefficients used to adjust Hargreaves potential evaporation estimate to Priestly-Taylor

Evaporation Stations		Hargreaves Adjustment Coefficient
Name	Abbreviation	
Bushnell	BUSHNELL	0.8425
Clermont	CLERMONT	0.8714
Crescent City	CRESCENT	0.9056
Daytona Beach	DAYTONA	0.9342
DeLand	DELAND	0.8726
Federal Point	FEDPT	0.9057
Ft Drum	FTDRUM	0.8663
Gainesville	GNSVILLE	0.8431
Glen St. Mary	GLNSTMRY	0.8663
Jacksonville International Airport	JAXAP	0.9381
Jacksonville Beach	JAXB	1.1193
Lisbon	LISBON	0.9114
Melbourne	MELB	0.9264
Ocala	OCALA	0.8101
Orlando	ORLANDO	0.9109
Sanford	SANFORD	0.8888
St. Augustine	STAUG	0.9952
Starke	STARKE	0.8665
Titusville	TITUSV	0.9940
Vero Beach	VERO_BCH	0.9582

The estimated annual evaporation is higher in the southern area of the river near Vero Beach and lower in the northern area near Jacksonville. This difference is expected because the average temperature is higher in the southern area of the river. There is also a difference in evaporation between the eastern coastal areas and the western ridge areas, with higher evaporation in the inland areas and lower evaporation near the more humid coaSt. These differences in evaporation are summarized in Figure 9.

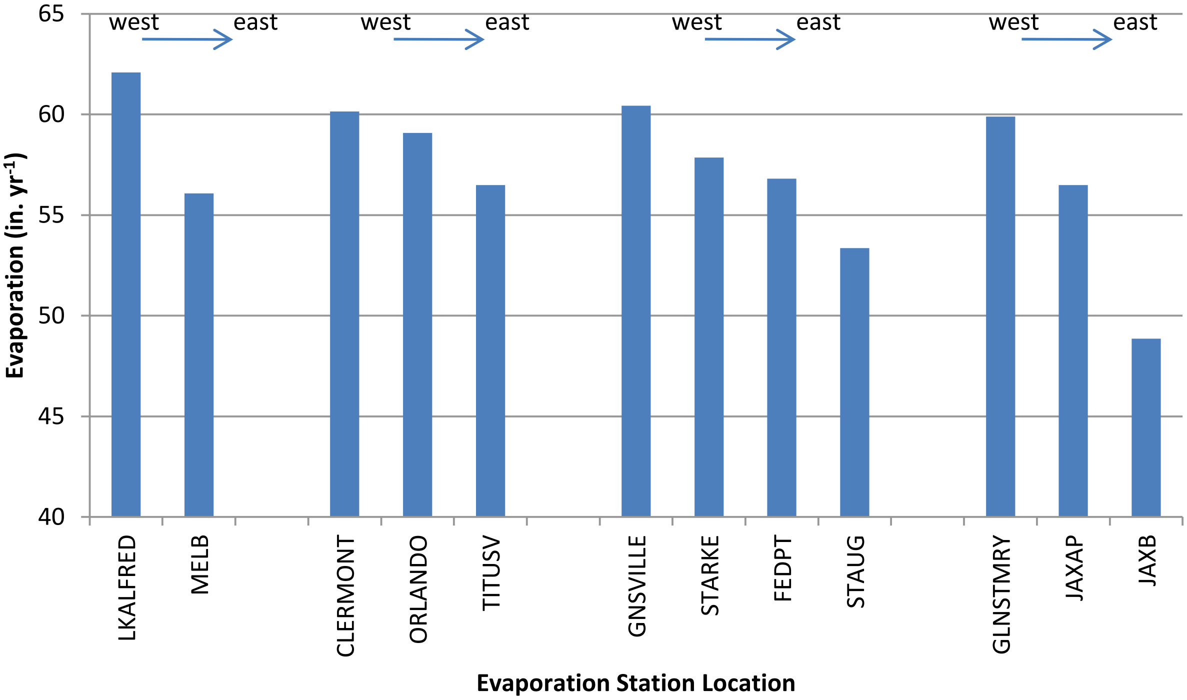

Figure 9 Average annual potential evaporation arranged in four approximate west to east cross-sections, with cross-sections arranged from south to north.

2.3. Observed Flow and Water Level Data

An important and underappreciated aspect of almost all published stream flow data is that stream flow is not measured directly but estimated from a stage-discharge relationship. A stage-discharge relationship serves as a model relating an easy to measure water level with very difficult, time-consuming, and expensive flow measurements. Even though stage-discharge relationships are well known and can be an efficient and effective method to estimate stream flows, only a small portion of large watersheds are gauged.

When developing a stage-discharge relationship, flow measurements are plotted with their corresponding stages, and a curve is approximately fitted through the points. That curve becomes the relationship for estimating flow, given stage within the range of flow measurements. For flows outside the range of the flow measurements, the curves are extended using logarithmic plotting, velocity-area studies, or using the results of indirect measurements.

The stage-discharge relationship is subject to change because of changes in the physical features that affect the gauge site. The stage-discharge relationship can be changed temporarily because of aquatic growth or debris, downstream flow obstructions may produce backwater effects that reach the gauge, and upstream obstructions may also change the cross sectional area of flow.

WSIS uses USGS stream flow data for calibration except for a few stations. The USGS rating curve model has errors associated with the estimated flow. Even though there are several ways to estimate the rating curve error (Dymond & Christian, 1982), the USGS has established a subjective estimate of annual flow data quality based on a review of measured data, datum shifts, and other characteristics of the flow measurement station. Table 3 describes the USGS system of data quality estimation (Kennedy, 1983). The USGS system provides a general site-specific estimate of error, and there may be significantly more error where there are few flow measurements in the rating curve (e.g., at high and low flows). USGS gives a single quality category for each year of data.

Table 3 USGS flow data quality categories. Source: (Kennedy, 1983).

Quality Category	Description
Excellent	95% of daily discharges within 5% of "true"
Good	95% of daily discharges within 10% of "true"
Fair	95% of daily discharges within 15% of "true"
Poor	Daily discharge have less than "fair" accuracy

There are other inherent difficulties in flow measurement in Florida due to the shallow slope, poorly defined cross sections, and tidal influences. Most USGS flow measurement stations in Florida are rated "Fair." Although it would be difficult to collate the data for each station and each year, only about 10% of Florida's stations rate a "Good" classification. An "Excellent" rating for a station in Florida is very rare.

The overall locations of the flow observation stations and their corresponding gauged watersheds are presented in Figure 10.

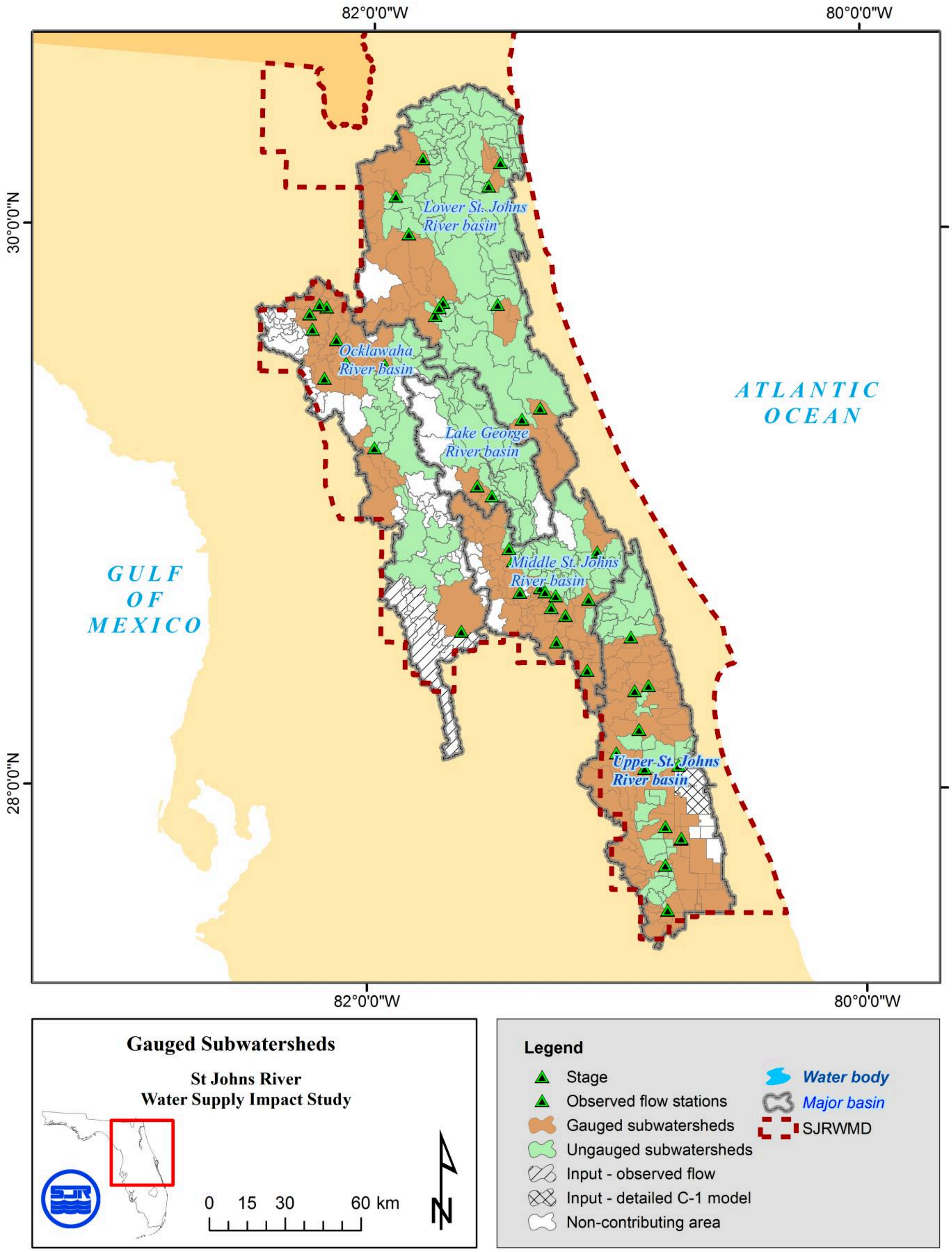

Figure 10 Map of flow and water level gauges, and gauged, ungauged, special input and noncontributing watersheds for St. Johns River watershed.

3. WATERSHED PHYSIOGRAPHY

3.1. Watershed, Major Basin, Planning Unit, and Subwatershed Boundaries

The terminology used in this paper to describe the hydrologic boundaries mostly follows Technical Publication SJ97-1 (Adamus, Clapp, & Brown, 1997).

- Watershed: A collection of major basins that contribute to a single water body. Five major basins numbered 3 through 7 comprise the St. Johns River watershed (see Table 4 and Figure 11).
- Major Basin: The SJRWMD is divided into ten major basins (Table 4) numbered one through ten.
- Planning Unit: The major basins are subdivided into planning units. Planning unit boundaries are based on tributary areas for larger rivers and streams or areas with similar characteristics. Each major basin has a varying number of planning units uniquely labeled with a capital letter starting with "A."
- Subwatershed: In Technical Publication SJ97-1 this is analagous to "Planning Unit ID," which is also described as "7.5-Minute Quad Basin." Aside from minor edits, the subwatersheds boundaries used in the modeling for this study matches the boundaries of the in Technical Publication SJ97-1. The "Planning Unit ID" in SJ97-1 and subwatersheds in this study are uniquely numbered within each planning unit starting with "1". The subwatershed numbers used for the WSIS were assigned to make the hydrologic connection apparent and do not match the Planning Unit ID numbers in Technical Publication SJ97-1.

Table 4 Area estimates for all major basins within SJRWMD's jurisdiction from TP SJ97-1 (Adamus, Clapp, & Brown, 1997)

Number	Major Basin Name	Area (mi^2)	
		St. Johns River Watershed (WSIS Project Area)	Other SJRWMD Major Basins
1	Nassau River Basin		432
2	St. Mary's River Basin		951
3	Lower St. Johns River Basin	2,755	
4	Middle St. Johns River Basin	1,205	
5	Lake George Basin	817	
6	Upper St. Johns River Basin	1,748	
7	Ocklawaha River Basin	2,116	
8	Florida Ridge Basin		692
9	Northern Coastal Basin		681
10	Indian River Lagoon Basin *		1,163
Total area of the St. Johns River Watershed (WSIS project area)		8,641	
Total area of SJRWMD major basins outside of the WSIS project area			2,536

*Includes 134-mi^2 Interbasin Diversion Planning Unit (6D) which historically was part of the upper St. Johns River, and depending upon the time period, completion of restoration projects, and management of operations can be hydrologically split between the Indian River Lagoon and upper St. Johns River major basins.

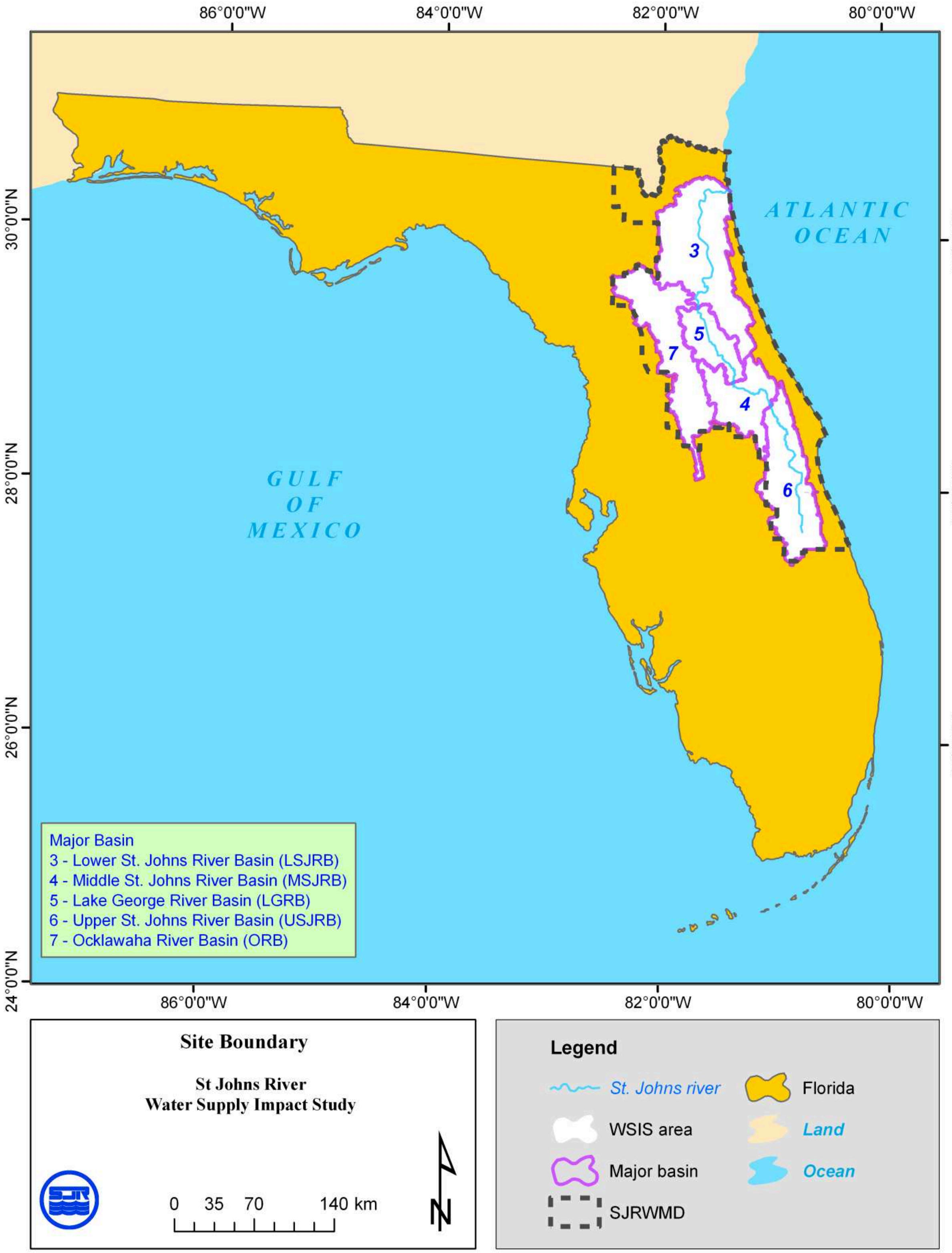

Figure 11 Water Supply Impact Study (WSIS) project boundary

The SJRWMD has jurisdiction over all or parts of eighteen counties in northeast Florida encompassing approximately 11,177 mi^2. In general, SJRWMD boundary follows major hydrologic boundaries as they were estimated when the five water management districts were created by the

State of Florida in 1976. The St. Johns River is one of the few northward flowing rivers in the United States. It is about 300 mi long from its headwaters near Florida's Turnpike in Okeechobee and Indian River counties to its mouth at Jacksonville, Florida. The river has a total fall of approximately 25 ft over its length, thus having an average gradient of less than 0.1 ft mi^{-1}. The major basins that make up the St. Johns River watershed account for approximately 65% of SJRWMD's jurisdictional land area.

The best estimate for the total area of the St. Johns River watershed is found in Technical Publication SJ97-1 (TP SJ97-1) (Adamus, Clapp, & Brown, 1997) and is approximately 8,641 mi^2. There is a difference in total watershed area and the area used in the HSPF hydrologic models. These differences come from four sources: updates in watershed boundaries; subtraction of the river surface area not modeled in the HSPF hydrologic model (but included in the hydrodynamic model of the river); subwatersheds that do not contribute surface water runoff; and shift of watershed boundaries due to flood, water quality, and environmental restoration projects. This modeled area of the watershed is called the contributing area. The contributing area to the St. Johns River is approximately 7,466 mi^2. Unless specified explicitly, watershed areas in this paper are the contributing areas. Table 5 illustrates the differences between total and contributing areas.

Table 5 Comparison of area estimates for the major basins within the St. Johns River Watershed, arranged approximately upstream to downstream.

Major Basin Number	Major Basin Name	TP SJ97-1 Major Basin Area (mi^2)*	Model Major Basin Area (mi^2)
6	Upper St. Johns River Basin	1,748	1,739
4	Middle St. Johns River Basin	1,205	1,020
5	Lake George Basin	817	512
7	Ocklawaha River Basin	2,116	1,590
3	Lower St. Johns River Basin	2,755	2,605
	Total	**8,641**	**7,466**

Source: TP SJ97-1 (Adamus, Clapp, & Brown, 1997)

* Included acreage that does not contribute to surface runoff

At the major basin level described in Table 5 there is likely little difference in watershed boundaries, but at the planning unit and subwatershed levels, the watershed boundaries in the current HSPF hydrologic models were reviewed and adjustments were made to remove gaps, eliminate overlaps, and make other modifications based on new hydrologic information.

3.2. 1995 Land Use

The 1995 aerial interpretation of land use for this study was developed under contract to Geonex, Inc. based on 1994 and 1995 color-infrared aerial photography of the entire SJRWMD. These data layers support many projects throughout SJRWMD as a snapshot of land use and land cover.

The aerial photography was produced by the National Aerial Photography Program (NAPP) from Jan 1994 through Dec 1995, with the bulk of photos taken in 1994.

A photo interpretation key (PI key) was developed to facilitate a uniform assessment across SJRWMD and establish other necessary interpretation standards. The minimum mapping unit areas from the PI Key are found in Table 6. The land uses mapped are defined by the Florida Land Use and Cover Classification System (FLUCCS) codes (State of Florida, 1999) resulting in approximately 140 distinct land use classifications in the SJRWMD.

Table 6 Minimum mapping size for aerial photography interpretation to establish land use

Land Use	Minimum Mapping Unit
Upland classes	2.0 ac
Water and wetland classes	0.5 ac
Rivers and canals	10 m or greater in width and continuous
Roads and railroads	All major transportation corridors
Utility corridors	30 m or greater in width

For this effort, the detailed FLUCCS coded land uses were grouped into categories according to similar hydrologic response. There are 15 HSPF land use groups. The wetland land use category is split into two parts depending on whether they are riparian (adjacent to the river or stream) or non-riparian (i.e., an upland wetland). This split of wetland areas improves the hydrologic representation of the watershed. Wetland areas listed in this paper are a summation of riparian and non-riparian wetland areas unless specified otherwise. In portions of the Ocklawaha basin, the forestland use group was divided into a forest (90% of the area) and a forest regeneration land use (10% of the area). The land use groups are listed in Table 7.

Table 7 Land use groups for HSPF hydrologic modeling.

HSPF Hydrologic Modeling Land Use Number		HSPF Hydrologic Modeling Land Use Group Name	Special Category	Note
1		Low-density residential	–	< 2 dwelling units per acre
2		Medium-density residential	–	2 to 5 dwelling units per acre
3		High-density residential	–	> 5 dwelling units per acre
4		Industrial and commercial	–	–
5		Mining	–	–
6		Open and barren land	–	–
7		Pasture	–	–
8		Agriculture general	–	–
9		Agriculture tree crops	–	–
10		Rangeland	–	–
		Forest	–	–
	11*		Forest	90% of Forest land use area: only used in portions of the Ocklawaha River Basin
	14*		Forest Regeneration	10% of Forest land use area: only used in portions of the Ocklawaha River Basin
12		Water	–	–
		Wetland	–	–
	13*		Riparian Wetlands	Wetland land use is split between riparian and non-riparian wetlands according to the drainage pattern within each subwatershed
	15*		Non-riparian Wetlands	

*In some cases calculated as part of another land use category

Table 8 Summary of the 1995 HSPF land use groupings in the St. Johns River Basin.

HSPF Hydrologic Modeling Land Use Number and Group	1995 Land Use (acres)		2030 Land Use (acres)	
1. Low-density residential	263,841	4.9%	787,264	14.7%
2. Medium-density residential	247,710	4.6%	540,955	10.1%
3. High-density residential	78,947	1.5%	169,659	3.2%
4. Industrial and commercial	140,282	2.6%	301,998	5.6%
5. Mining	20,515	0.4%	14,973	0.3%
6. Open and barren land	112,207	2.1%	58,512	1.1%
7. Pasture	505,701	9.4%	343,102	6.4%
8. Agriculture general	267,970	5.0%	148,201	2.8%
9. Agriculture tree crops	144,268	2.7%	72,500	1.3%
10. Rangeland	272,895	5.1%	136,985	2.5%
11. Forest	1,659,119	30.9%	1,139,307	21.2%
12. Water	286,016	5.3%	286,016	5.3%
13. Wetlands	1,374,656	25.6%	1,374,656	25.6%
Total	5,374,127	100.0%	5,374,127	100.0%

4. HYDROLOGICAL SIMULATION PROGRAM–FORTRAN (HSPF)

The Hydrological Simulation Program–FORTRAN (HSPF) is a comprehensive hydrology (water quantity) and water quality modeling system. Currently HSPF is part of the BASINS modeling environment. HSPF is highly regarded as a complete and defensible watershed model for the simulation of hydrology and water quality for both conventional and toxic pollutants. The simulation results of the HSPF model consist of a time history of the runoff flow rate and can include sediment load and nutrient and pesticide concentrations along with a time history of water quantity and quality at nearly any point in a watershed.

The model evolved from the 1960's Stanford Watershed Model. In the 1970s, water-quality processes were added. Development of a FORTRAN version, incorporating several related models using software engineering design and development concepts, was funded by the Athens, Georgia, EPA Research Lab in the late 1970s. In the 1980s, pre-processing and post-processing software, algorithm enhancements, and use of the USGS WDM system were developed jointly by the USGS and EPA. The HSPF model has been successfully applied in climatic conditions around the world. The HSPF model currently enjoys the joint sponsorship of both the EPA and the USGS and continues to undergo refinement and enhancement of its component simulation capabilities, along with user support and code development. (United States Geological Survey, 2010)

A watershed is conceptually represented in HSPF as a series of storage compartments (e.g., surface depressions, soil zones, groundwater zones, river segments). Based on the principal of mass conservation, HSPF performs continuous budget analysis of water quantity and quality for these storage compartments. Given the inputs of meteorological time series and the parameter values related to watershed characteristics, HSPF generates time series of runoff, stream flow, loading rates, and concentrations of various water quality constituents.

Although most parameters of HSPF can be specified by watershed spatial and physical data (e.g., land use, topography, stream characteristics, and soil properties), a few parameters, such as those related to infiltration, evaporation, and instream kinetics, need to be determined in the model calibration process. Model calibration is the process of adjusting values of model parameters to accurately reproduce the observed flow and water quality data for a given compartment. Once calibrated, the HSPF model is considered to accurately represent the hydrologic and water quality processes in a watershed and can be used for scenario analysis.

A watershed and its stream network are characterized in HSPF by various pervious land segments (PERLND), impervious land segments (IMPLND), and reaches/reservoirs (RCHRES) based on subwatershed delineation, land uses, and the impervious percentage for each land use. watersheds are grouped into 13 categories, with two additional special categories. The four urban categories are further divided into pervious and impervious fractions. The pervious portion of a land use category is represented as PERLND, and the impervious portion of a land use category is represented as IMPLND. For modeling purposes, the stream network in a subwatershed is grouped together and represented as RCHRES. The geometric and hydraulic properties of a RCHRES are represented in HSPF by FTABLEs, which describe the relationships among stage, surface area, volume, and discharge for the reach segment. Detailed description of these submodules can be found in Bicknell, et al. (2001).

A series of model simulation graphics are provided to illustrate the HSPF model (Figure 12, Figure 13, Figure 14, and Figure 15). Hydrologic simulation for PERLND and IMPLND is carried out in the PWATER submodule (Figure 13) and the IWATER submodule (Figure 14). The simulated hydrologic processes for PERLND include interception, infiltration, evapotranspiration, runoff, and deep percolation. The simulated processes for IMPLND are similar to those for PERLND except there are no infiltration and subsequent subsurface processes. Hydraulic behaviors in RCHRES are simulated in the HYDR submodule (Figure 15).

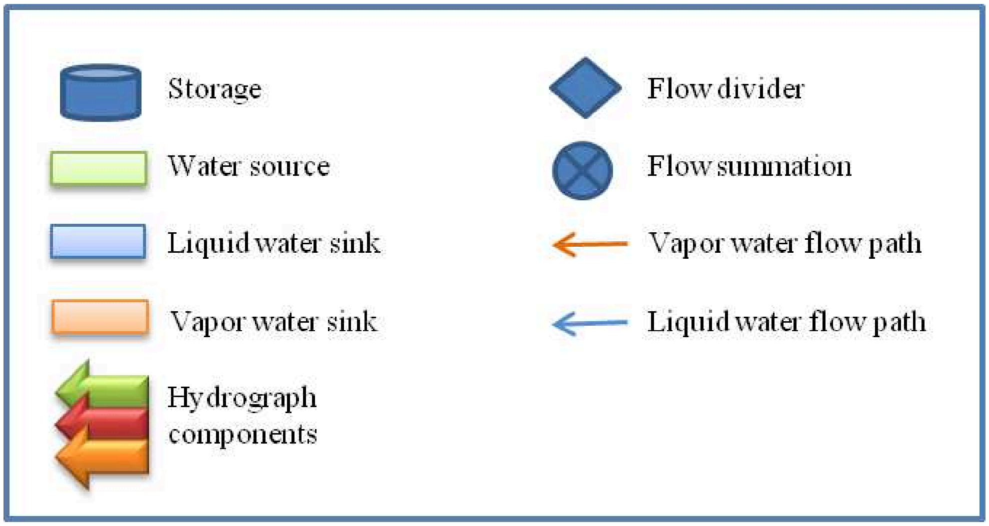

Figure 12 Legend for HSPF model simulation graphics in Figure 13, Figure 14, and Figure 15.

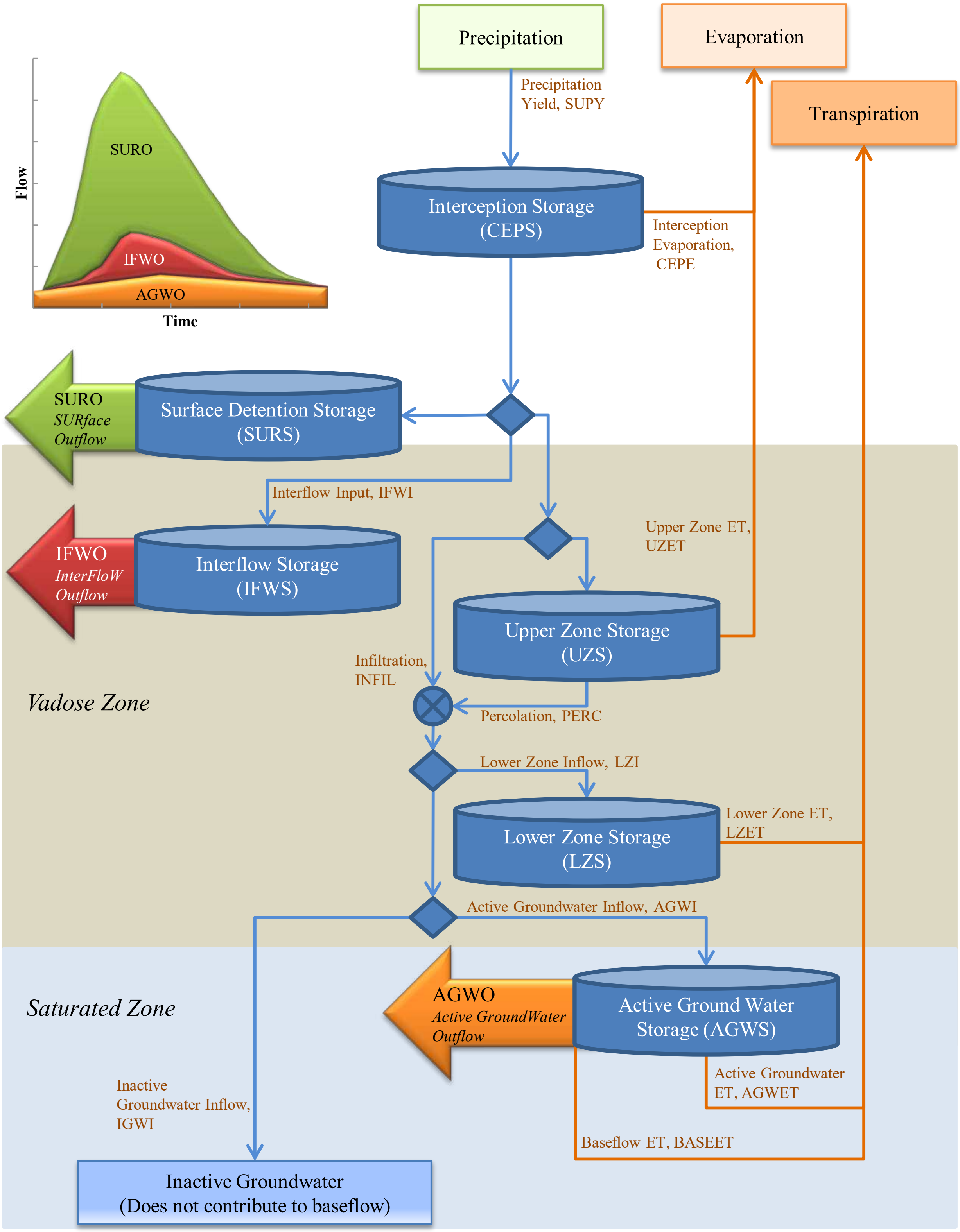

Figure 13 Illustration of water storage and movement in the HSPF model pervious land element (PERLND).

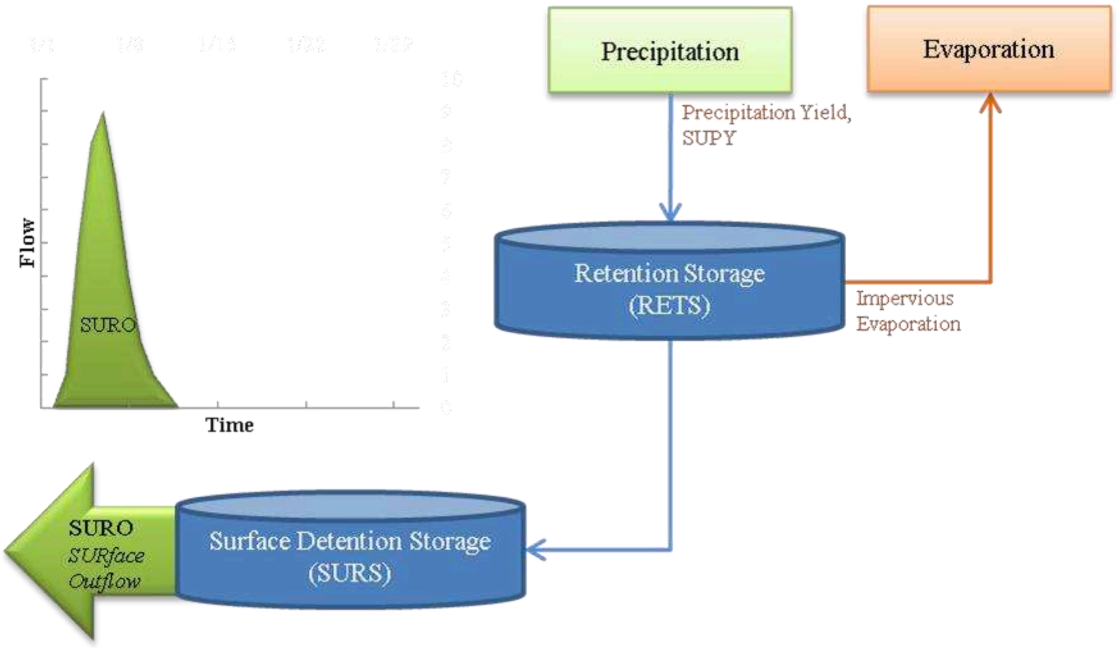

Figure 14 Illustration of water storage and movement in the HSPF model impervious land element (IMPLND).

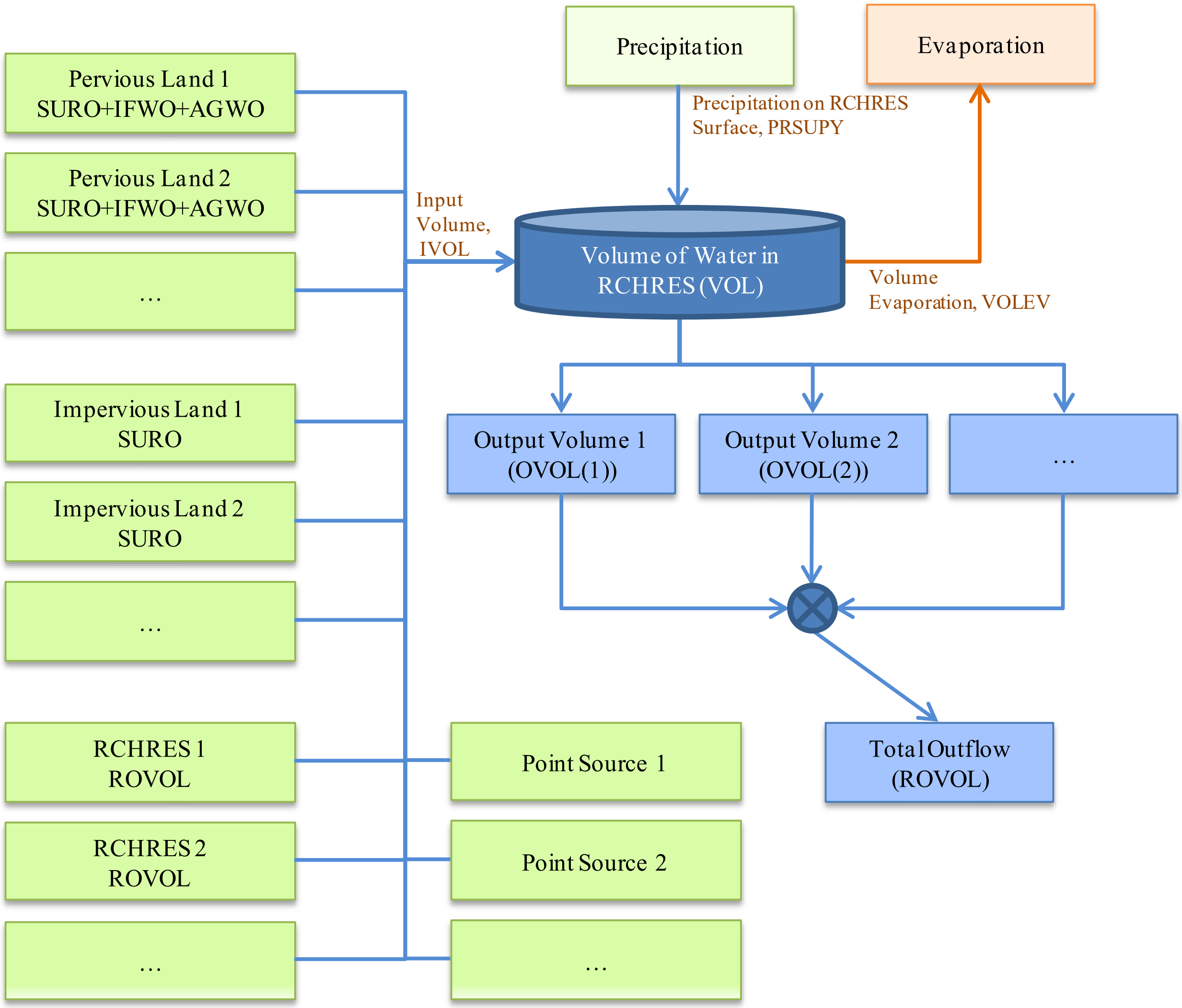

Figure 15 Illustration of water collection and movement in the HSPF model reach/reservoir element (RCHRES).

5. CALIBRATION PROCESS

The default calibration period selected for the WSIS HSPF hydrologic models is from 1995 to 2006, though the actual calibration period depended on available meteorologic and flow observation time series. This default period was selected for three reasons:

- The baseline for groundwater planning programs at SJRWMD is 1995.
- Due to extensive data and computer run time requirements, the EFDC hydrodynamic model of the main stem of the St. Johns River simulates from 1995 to 2006.
- A major portion of the USJRB project was completed by 1995 and project activities were relatively stable between 1995 and 2006.

5.1. Model Input Parameters–Common Logic

The changes to the model concerning land use, precipitation, and evaporation require a complete examination of the model parameters. Different SJRWMD engineers originally modeled watersheds with the HSPF hydrologic model for various purposes and developed model parameters that were characteristic of the individual basins. For the WSIS program, SJRWMD has developed a common logic (Table 9) describing reasonable parameter value ranges for all HSPF hydrologic models in SJRWMD. This HSPF common logic was derived from an evaluation of the possible range of model parameters for Florida's unique hydrology, extensive SJRWMD experience, and the parameter ranges common in other parts of the world (EPA, July 2000).

Table 9 Common logic/understanding for the range of a few HSPF parameters applied to watersheds within the SJRWMD.

Parameter	Description	SJRWMD	(EPA, July 2000) Tech Note 6
		Min/Max	Min/Max
AGWRC	Base groundwater recession	0.9/0.999	0.85/0.999
BASETP	Fraction of remaining evapotranspiration from baseflow	0.0/0.1 a little higher is OK	0.0/0.2
CEPSC	Interception storage capacity	0.03/0.20 in.	0.01/0.40 in.
DEEPFR	Fraction of groundwater inflow to deep recharge	0.0/0.6 1.0 is OK if ephemeral stream	0.0/0.5
INFEXP	Exponent in infiltration equation	2.0/2.0	1.0/3.0
INFILD	Ration of max/mean infiltration capacities	2.0/2.0	1.0/3.0
INFILT	Index to infiltration capacity	0.01/1.0 in./hr See table in notes	0.001/0.5 in./hr
INTFW	Interflow inflow parameter	0.0/3.0	1.0/10.0
IRC	Interflow recession parameter	0.50/0.70	0.30/0.85
KVARY	Variable groundwater recession	0.0/3.0 (1/in.)	0.0/5.0 (1/in.)
LSUR	Length of overland flow	200/500 ft	100/700 ft
LZETP	Lower zone evapotranspiration parameter	0.20/0.70	0.10/0.90
LZSN	Lower zone nominal soil moisture storage	2.0/10.0 in.	2.0/15.0 in.
NSUR	Manning's "n" for overland flow	0.15/0.35	0.05/0.50
PETMAX	Temperature below which evapotranspiration is reduced	35.0/45.0 °F	32.0/48.0 °F
PETMIN	Temperature below which evapotranspiration is zero	30.0/35.0 °F	30.0/40.0 °F
SLSUR	Slope of overland flow plane	0.001/0.15	0.001/0.30
UZSN	Upper zone nominal soil moisture	0.10/1.0 in. 4.0 for wetlands	0.05/2.0 in.

5.2. Directly Connected Impervious Area (DCIA)

Impervious areas include all surface areas that prevent water from infiltrating into the ground. Typical impervious areas are buildings/roofs, roads, and parking lots. These impervious areas can be classified into two categories: DCIA and nondirectly connected impervious area (NDCIA). DCIAs are the impervious areas that directly connect to the drainage network with no opportunity for infiltration (e.g., a parking lot that drains directly to a creek). NDCIAs are the impervious areas that drain to pervious areas (e.g., a rural home surrounded by a vegetated area). In this study, only DCIAs are modeled as IMPLND and NDCIAs are part of the PERLND land use element.

Among the HSPF 13 land use groups, the four urban land categories are assumed to have some DCIA. The four urban land groups are low-density residential, medium-density residential, high-density residential, and industrial and commercial. Estimation of the percent DCIA for WSIS in each urban land use category stems from observed flows of small storm events, because most runoff during small storms is generated from DCIA. Impacts of changing percentages of DCIA on total mass balance and seasonal flow distribution were also considered. The proportion of DCIA in each urban land use category is attributed to IMPLND for the HSPF hydrologic model (Table 10). The remaining nine land use categories are assumed to consist of pervious (PERLND) elements.

Table 10 Percentages of Directly Connected Impervious Area.

HSPF Hydrologic Modeling Land Use Group	% Imperviousness
Low-density residential	5
Medium-density residential	15
High-density residential	35
Industrial and commercial	50

5.3. Parameter Estimation

Calibration of a model is an iterative process of changing parameters, running simulations, checking results, and repeating until an acceptable match is made between the simulated and observed data (Figure 1). A calibrated model is one that most closely resembles the behavior of the systems in the real world. When manually performed, model calibration can be a time consuming endeavor. In addition, it can be difficult to maintain a consistent approach of parameter adjustments among a diverse group of engineers, such as the nine HSPF modelers for the WSIS project. To reduce the time burden on our modelers and have a consistent calibration framework, a parameter estimation model optimization tool called PEST (which stands for Parameter ESTimation) was used to assist in model calibration (Doherty, 2004).

PEST sees a model as a ‘black box’. The only requirements of that ‘black box’ is that it accepts text files as input and produces text files as output (Figure 16). Even if the model requires binary input or output, scripts or batch files can act as ‘composite models’ to pre-process model input files, run the actual model, and then post-process the model output files before use by PEST. PEST template files for creation of model input, and a parsing language to extract the required simulated values from model output, allows PEST to adapt to the model rather than changing the model itself.

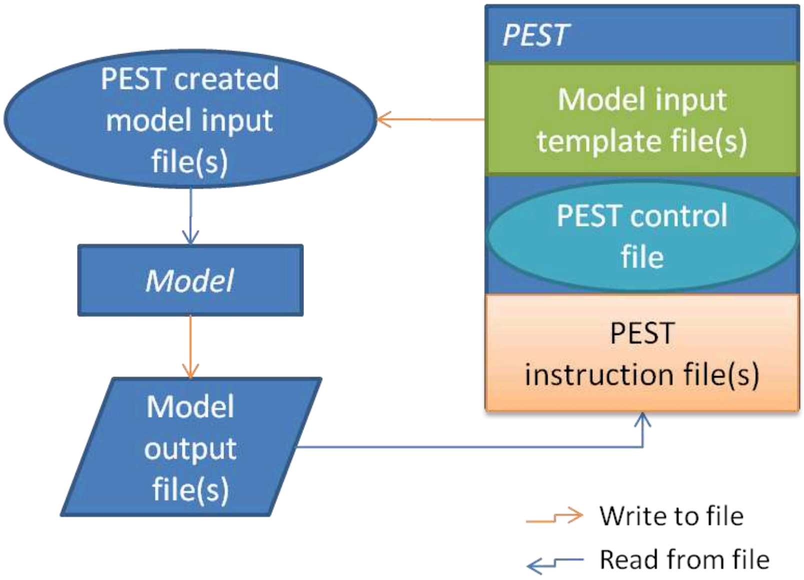

Figure 16 Conceptual model of PEST interaction with a model. Model template file(s) are used be PEST to create new data sets with changed parameters in the correct places. PEST instruction file(s) are use to parse the model output so PEST can compare modeled against the objective functions (observations).

PEST is a nonlinear parameter estimator that will adjust model parameters to minimize the discrepancies between simulated and corresponding real-world measurements using weighted, least-squares optimization. To accomplish this is must run the model many times, using different parameter sets, then analyzing the effect of changing each parameter on the difference between simulated and observed objective functions. Objective functions are observations or statistics of the observations. PEST evaluates parameter changes based on the improvement against the objective functions and decides whether to undertake repeated optimization until no further improvement can be achieved. The modeler must define the objective functions (based on observations), and select pertinent parameters and set the parameter's upper and lower bounds for adjustment.

Four objective functions were established for this project: daily flow, monthly flow, annual flow, and flow duration curves. Gauged and simulated flows are compared within these four objective functions to address daily flow variability, seasonal variability, annual discharge characteristics, and overall discharge characteristics. The modeler assigns weights to each objective function based on the importance of each discharge component that will obtain the best overall match between gauged and simulated discharge.

The PEST utility was used to optimize the parameters lower zone nominal soil moisture storage (LZSN), lower zone evapotranspiration (LZEPT), index to infiltration capacity (INFILT), upper zone nominal soil moisture (UZSN), base groundwater recession (AGWRC), interflow inflow (INTFW), interflow recession (IRC), fraction of groundwater inflow to deep recharge (DEEPFR) and the wetland surface runoff FTABLE storage-runoff relationship. Relative values of parameters were established by the modelers among land uses to produce expected relative runoff amounts. Urban land, including impervious area, produces the most runoff, agriculture produces the next largest runoff, open land and rangeland produce less, and forest and wetlands produce the least runoff. PEST allows parameters to be "tied" to a "parent" parameter. In this way, all of the tied parameters are adjusted equally among the various land uses. In general, LZSN, LZEPT, INFILT, and UZSN parameters are tied together among land uses. The exception to this is wetlands. Wetland parameters give emphasis to larger upper zone storage and lower infiltration rates. For this reason, wetland parameter sets are not comparable to other land uses and are adjusted independently. The parameters AGWRC and DEEPFR are applied to the entire watershed. In addition, PEST allows parameters to be "fixed" and not adjusted. For example, in many cases of INTFW and IRC, these parameters usually are given a restricted range close to zero or fixed to zero or a very small number (Table 9).

An absolute critical component of our ability to calibrate these models within the time-frame allotted was the use of an incredibly powerful utility, also written by PEST's John Doherty, called the Time Series PROCessor (TSPROC). TSPROC is the main component of the 'PEST Surface Water Utilities' suite of programs. TSPROC interprets a simple scripting language to process and analyze time-series. Illustrated in Figure 17, by running in a different context the same TSPROC script used later for the composite model, has the ability to create PEST control and instruction files (instruction files are PEST's model output parsing language). TSPROC development has been adopted by Steve Westenbroek of the Wisconsin USGS and is in the process of being updated and published as a part of the USGS technical report series (Westenbroek, Doherty, Walker, Kelson, Hunt, & Cera, in review).

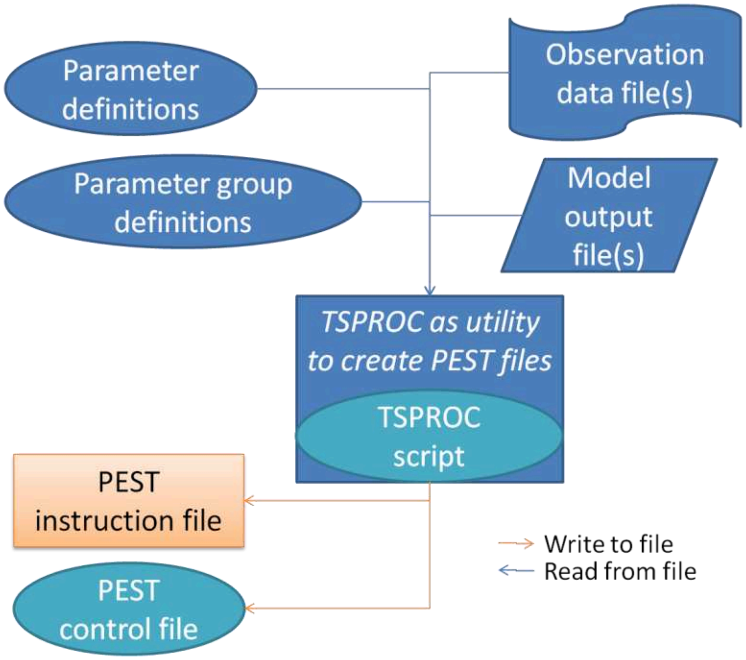

Figure 17 Conceptual data flow when using TSPROC to create PEST control and instruction files.

When used as a post-processing tool in a composite model, TSPROC will reformat the model output to a consistent TSPROC output format (Figure 18). The 'composite model' can be developed in just about any programming or scripting language, typically though it is a shell script on Linux and a batch file on Windows.

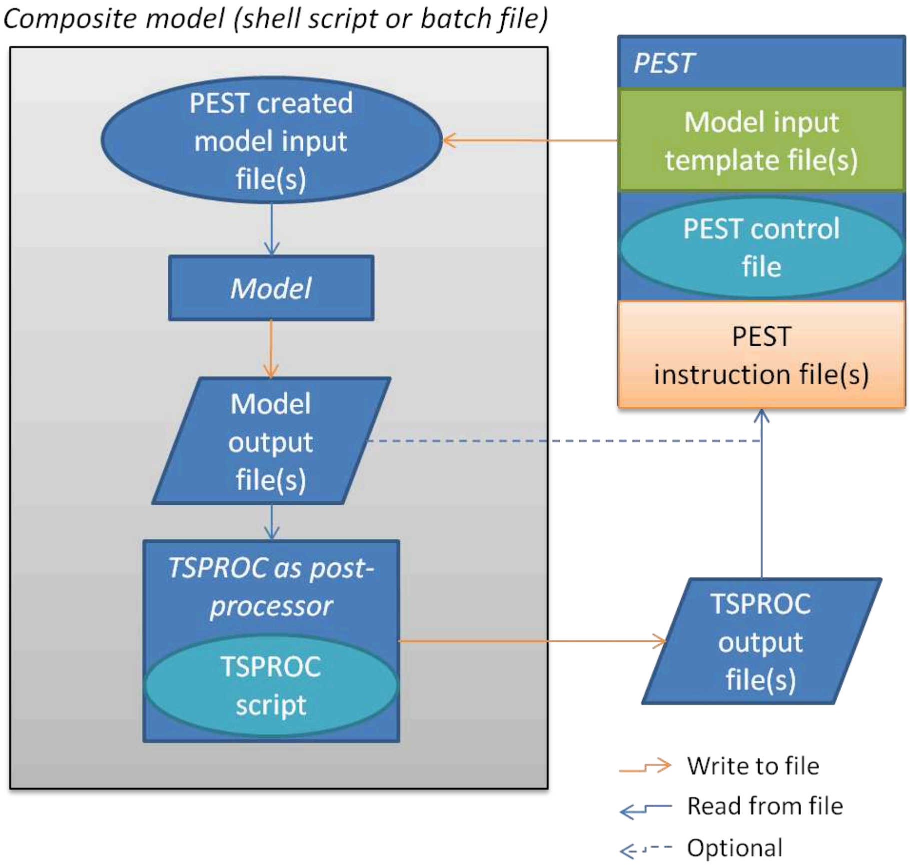

Figure 18 Conceptual data flow and setup to use PEST with a composite model and TSPROC as a post-processor.

One of the utilities that is part of the PEST suite, is called PAR2PAR. PAR2PAR is needed if you want to enforce a relationship between two parameters. For example, *parm1* may have a reasonable range of 1 to 10, whereas *parm2* has a range of 2 to 11. There is additional information though in that the modeler knows that *parm1* will always be less than 0.80 of *parm2*. PAR2PAR can enforce this kind of requirement. It does add a bit of complexity as seen in Figure 19.

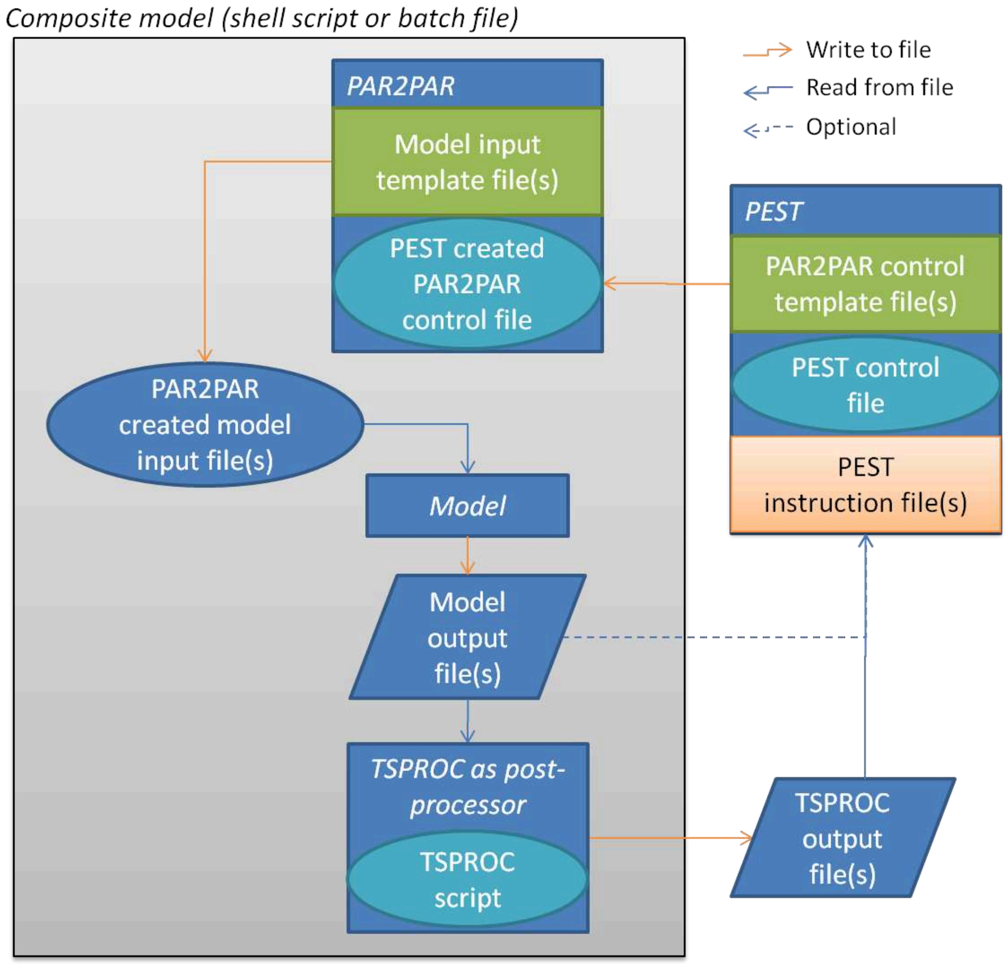

Figure 19 Conceptual data flow of PEST with PAR2PAR to define relationships between parameters, along with TSPROC as a model post-processor.

6. PEER REVIEW

A peer review of the all of the initial HSPF hydrologic model calibrations was performed by Intera Incorporated (Intera). Intera's detailed report was completed in September 2009 and contained recommendations for improvement of the models. Their report looked at the entire modeling effort, and identified the following four main recommendations to improve the model calibration:

- Re-examination of directly connected impervious area (DCIA) values: For some basins, the DCIA values, particularly for industrial and commercial use, seem higher than accepted values. Since DCIA is used to determine the areas of PERLND and IMPLND segments, the model is highly sensitive to changes in these values. It is also very important to remain systematic in the definition of the parameters. Justification is necessary to describe basin-to-basin differences in any model parameters based on the same mapping data.
- Consideration of changes to retention storage capacities (RETSC) for IMPLND segments. Since these segments are not routed to storage attenuation reaches, but rather directly to discharge reaches, RETSC should be increased in order to account for storage in conveyance systems and ponds that most impervious runoff undergoes prior to discharge.

- Re-examination of active groundwater evapotranspiration (AGWET). Currently, the majority of segments have no AGWET. This should be calibrated accordingly in the context of the depth of the water table and vegetation type.
- Implementation of storage attenuation for PERLND and IMPLND segments. This can be accomplished using storage attenuation reaches.

The SJRWMD reviewed these recommendations and implemented them in the following manner.

6.1. Re-examination of Directly Connected Impervious Area values

The DCIA values were generally too high and varied among the models. In many cases, DCIA values were simply adopted from predecessor models, which were not always focused on water supply issues. What had been conservative assumptions for other purposes (e.g., flooding or water quality) were not necessarily appropriate for the WSIS. In no case were DCIA values adjusted to calibrate the models. A very good estimate of DCIA can be found by analyzing the results from small storms that occur after a dry period, since surface water flows for these kind of storms would predominately be from DCIA. From this analysis, DCIA from older models, and Intera's recommended values the new DCIA values were established across the board as; 5% of lower density residential, 15% of medium-density residential, 35% of high-density residential, and 50% of industrial/commercial.

6.2. Consideration of changes to retention storage capacities (RETSC) for IMPLND segments

The RETSC value was too small and it was increased from various values to a standard of 0.1 in. Though increased to a larger value we did not adjust RETSC to represent detention storage. RETSC affects both peak and volume strongly, whereas detention storage affects peak strongly but affects volume only weakly.

6.3. Re-examination of active groundwater evapotranspiration (AGWET)

The use of AGWET parameter was not sufficient for many of the models that have shallow water tables. The AGWET parameter values were compared to the depth to water table map. The AGWET parameter was changed in all models to a range of values consistent with the depth to water table map for that subwatershed.

6.4. Implementation of storage attenuation for PERLND and IMPLND segments

Additional storage was necessary to have a better representation of the hydrology. Surface FTABLEs were used to implement this storage, which are part of the high water table algorithms in the HSPF hydrologic model. The surface FTABLEs are used to represent the storage in non-riparian wetlands.

In the original HSPF hydrologic model construction, the 13 land uses (in PERLNDs and IMPLNDs) were routed directly to their associated streams (RCHRES). The Intera peer review suggested that routing some of the flow from upland surface areas to upland wetlands would provide a better representation of the subwatershed. The initial model construction implicitly represented this storage by adjusting other model parameters in the calibration process.

Wetlands tend to slow movement of water because of surface storage. One result of this is that wetland areas have a larger potential for evapotranspiration. HSPF hydrologic modeling provides the option to define surface outflow as a function of surface detention depth. This feature allows improved representation of the surface storage and attenuated surface runoff typical of wetlands.

The first step in this process is the definition of the upland surface areas that would drain to these wetlands. The SJRWMD contains thousands of wetlands that range in size from less than one acre to thousands of acres. The wetlands were classified as either riparian (directly connected or adjacent to a reach) or non-riparian (not directly connected to a reach). An additional wetland land use classification was created for the non-riparian wetlands. Drainage areas for the non-riparian wetlands were determined by using SJRWMD Digital Elevation Model overlaid by the HSPF hydrologic modeling land use groups to determine the drainage area of each non-riparian wetland. The processing generated tables showing the portion of each land use that drained to the non-riparian wetlands for each subwatershed.

A surface FTABLE was developed for each upland wetland. The area used in the FTABLE matched the area of the wetland. Development of the storage-outflow relationship begins with the general function:

$$Q = ay^m \tag{4}$$

Where:

Q = fraction of storage that runs off per hour
y = normalized depth above the invert
a, m = PEST optimized coefficient and exponent

PEST was used to optimize the wetland storage-outflow relationship by adjusting the depth of incipient flow and equation parameters. The lower and upper bounds for the depth of incipient flow are 0.01 to 11.99 in. The lower and upper bounds for the equation coefficient are 0.00 to 0.10. The lower and upper bounds for the equation exponent are 1 to 10. The storage-outflow relationship is typically used to populate the FTABLE at depths of 12, 24, and 36 in.

Separation of the watershed into areas that drain to non-riparian wetlands and the reach is an easy Geographical Information System (GIS) exercise for the 1995 land use. The land use prediction to support the estimated 2030 population was based on shifting land uses across the entire subwatershed but did not otherwise have a spatial component so it could not be split with GIS into non-riparian and riparian drainage areas. The area percentage split in 1995 between non-riparian and reach was maintained for 2030 for each non-urban land use. The new urban land use was prorated between non-riparian and riparian areas based on its percentage of total urban lands.

7. CALIBRATION RESULTS

A review of the data for the tributary stream gauges can reveal differences in the hydrologic response of these streams. One simple measure of response is to represent the flow measurements as a flow rate per square mile of watershed. The discharge in cubic feet per second per square mile (cfs/mi^{-2}) allows a direct comparison of observed and simulated (calibrated) flows for the 50 gauged site locations used in the surface water models for the WSIS (Figure 20). These flows are averaged over the calibration period and reduced to cfs/mi^{-2}. Most of the flows are less than 2.0 cfs mi^{-2}, with the exceptions being those discharges that contain spring flows. Note that the lower discharge per unit area watersheds typically had non-contributing surface area in the basin.

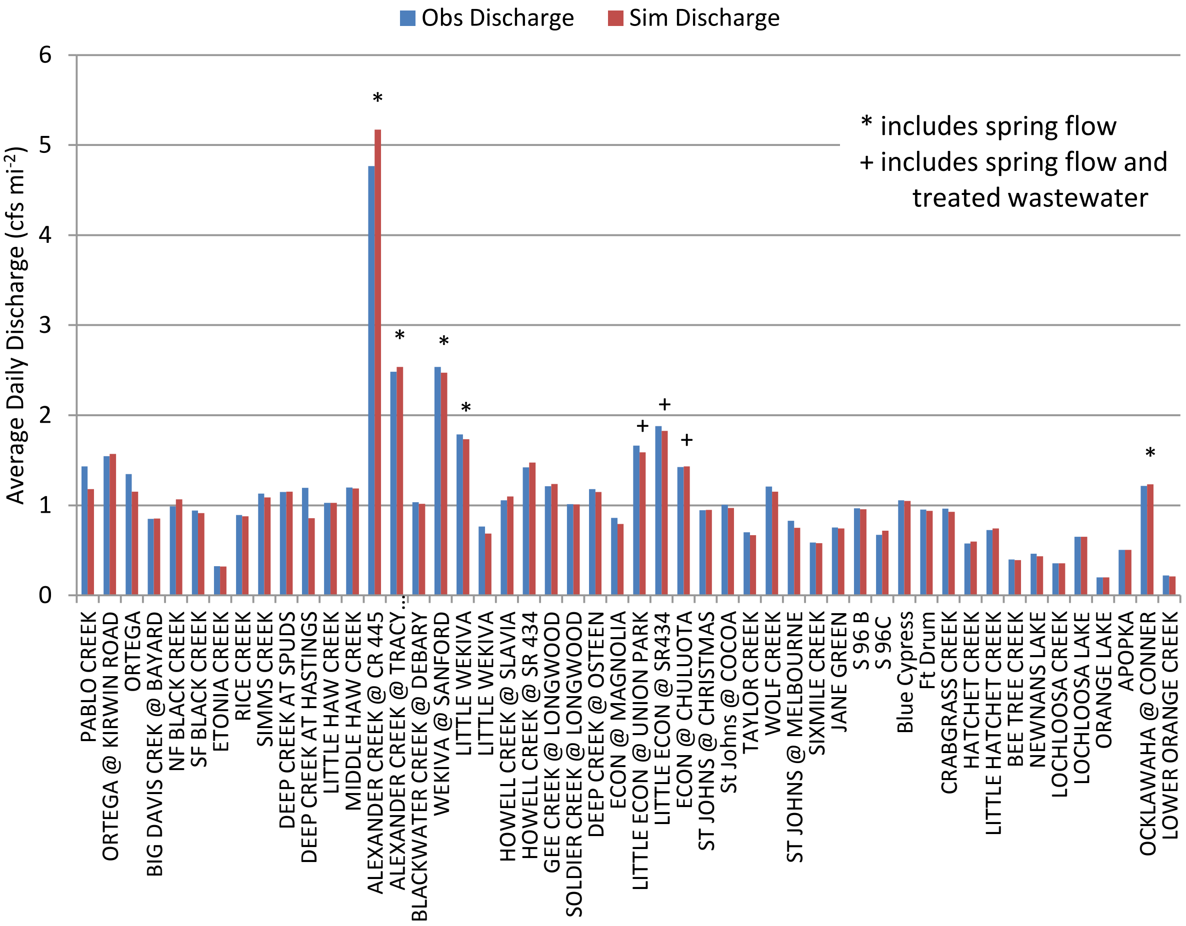

Figure 20 Comparison of observed (Obs) and simulated (Sim) discharge at calibration sites used in the WSIS watershed hydrology models.

The difference between the observed and simulated flows is small, which, along with the other model statistics, indicate than the models are a good representation of the watershed hydrology. The differences in reported versus simulated flows may be explained by a limited number of hydrologic factors, which all may affect runoff characteristics.

A very common measure of the performance of a hydrologic model is the Nash–Sutcliffe statistic (Moriasi, Arnold, Van Liew, Bigner, Harmel, & Veith, 2007). The Nash–Sutcliffe statistic ranges from zero to one, where zero would mean that the average of observations is a better model and one is a perfect match between simulated and observed data (Table 11). Negative Nash–Sutcliffe values are possible, although they do not have a particular meaning.

Table 11 Grading model calibration performance with the Nash–Sutcliffe statistic (NSE) and the Percent Error of the Mean (PEM). Adapted from (Moriasi, Arnold, Van Liew, Bigner, Harmel, & Veith, 2007).

Performance Rating	Nash–Sutcliffe (Monthly)	Percent Error of the Mean (Monthly)
Very good	0.75 < NSE < 1.00	< ±10
Good	0.65 < NSE < 0.75	±10 < PEM < ±15
Satisfactory	0.50 < NSE < 0.65	±15 < PEM < ±25
Unsatisfactory	< 0.50	> ±25

The calibration performance results for the WSIS watersheds show that the calibrated model is rated “good” or “very good” for 30 of 39 flows and “unsatisfactory” for only 2 of 39 flows (Calibration performance for water levels were rated “very good” for all four subwatersheds (Table 14).

Table 12). Performance of calibration results for dynamically managed structures were rated “good” for one of four structures and “unsatisfactory” for two of four structures (Table 13). This result is not unexpected, as human influence on the structures is extensive and is not readily reproduced in a model. For example, flows at S96C and S96B discharged into the same receiving water pool during the model period, causing backwater effects, which reduced cumulative flow from the structures. Decisions as to which structure would be closed to allow for design discharge from the other structure were based on a multitude of factors, including upstream agricultural pumping, distribution of regional rainfall, and anticipated atmospheric conditions. These ratings were not deemed critical, as the differences in parameters, such as flow and stage, from the model scenarios would drive the environmental evaluations, not their absolute values. In two cases, flow observations were not used for calibration because of poor (“unsatisfactory”) flow records or short periods of record. Four watersheds were calibrated using water level observations of nearby lakes instead of discharge. Calibration performance for water levels were rated “very good” for all four subwatersheds (Table 14).

Table 12 Calibration performance between simulated and observed flows at USGS and SJRWMD stations.

HSPF Hydrologic Model	Planning Unit	Subwatersheds	Flow/Water Level Station	Nash–Sutcliffe (Performance Rating)*	Percent Error of the Mean
Lower St. Johns River Basin (3)					
Little Haw Creek	3A: Crescent Lake	1	USGS 02244420	0.62 (Satisfactory)	1.94
Middle Haw Creek	3A: Crescent Lake	3	USGS 02244320	0.74 (Good)	1.72
Etonia Creek	3B: Etonia Creek	21, 26, 31, 41, 42, 43	USGS 02245050	0.69 (Good)	-1.77
Rice Creek	3B: Etonia Creek	2	USGS 02244473	0.80 (Very good)	3.08
Simms Creek	3B: Etonia Creek	44	USGS 02245140	0.66 (Good)	3.11
North Fork Black Creek	3C: Black Creek	6, 10, 11, 15, 16, 17	USGS 02246000	0.81 (Very good)	1.12
South Fork Black Creek	3C: Black Creek	1, 2, 3	USGS 02245500	0.75 (Very good)	2.78
Ortega at Jacksonville	3D: Ortega River	1	USGS 02246300	0.70 (Good)	13.19
Deep Creek at Spuds	3F: Deep Creek	2	USGS 02245260	0.68 (Good)	-0.34
Big Davis Creek at Bayard	3H: Julington Creek	3	USGS 02246150	0.65 (Good)	1.79
Pablo Creek at Jacksonville	3I: Intracoastal Waterway	2, 4, 5, 6, 3	USGS 02246828	0.65 (Good)	4.93

HSPF Hydrologic Model	Planning Unit	Subwatersheds	Flow/Water Level Station	Nash–Sutcliffe (Performance Rating)*	Percent Error of the Mean
Middle St. Johns River Basin (4)					
Econlockhatchee River at Magnolia Ranch	4A: Econlockhatchee River	1	USGS 02233001	0.22 (Unsatisfactory)	-17.88
Little Econlockhatchee River near Union Park	4A: Econlockhatchee River	11	USGS 02233200	0.61 (Satisfactory)	5.62
Little Econlockhatchee River at SR 434	4A: Econlockhatchee River	11, 12, 13	USGS 02233475	0.83 (Very good)	3.00
Econlockhatchee River near Chuluota	4A: Econlockhatchee River	1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13	USGS 02233500	0.72 (Good)	1.23
Deep Creek near Osteen	4B: Deep Creek	2, 4	USGS 02234100	0.79 (Very good)	-0.08
Howell Creek near Slavia	4C: Lake Jesup	1, 2, 4, 5, 6	USGS 02234324	0.70 (Good)	-1.31
Howell Creek at SR 434	4C: Lake Jesup	1, 2, 4, 5, 6, 7, 8, 9	USGS 02234344	0.73 (Good)	0.74
Gee Creek near Longwood	4C: Lake Jesup	10, 12, 13, 14, 15	USGS 02234400	0.60 (Satisfactory)	-0.77
Soldier Creek near Longwood	4C: Lake Jesup	17, 18, 19, 20, 21, 23	USGS 02234384	0.64 (Satisfactory)	0.39
Blackwater Creek near DeBary	4E: Wekiva River	8, 9, 10, 12	SJRWMD 30143084	0.80 (Very good)	1.35
Wekiva River near Sanford	4E: Wekiva River	19, 25, 26, 27, 28	USGS 02235000	0.68 (Good)	2.36
Little Wekiva River	4E: Wekiva River	20, 21, 22, 23, 24	SJRWMD 09502132	0.66 (Good)	1.56
Lake George Basin (5)					
Alexander Creek at CR 445	5B: Alexander Springs	1, 2	SJRWMD 18523784	0.79 (Very good)	-1.02
Alexander Creek at Tracy Canal	5B: Alexander Springs	3, 4, 5	SJRWMD 18553786	0.80 (Very good)	-2.57

HSPF Hydrologic Model	Planning Unit	Subwatersheds	Flow/Water Level Station	Nash–Sutcliffe (Performance Rating)*	Percent Error of the Mean
Upper St. Johns River Basin (6)					
Fort Drum Creek	6A: Fort Drum Creek	1, 2, 3, 4, 5, 6, 10	USGS 02231342	0.72 (Good)	1.31
Blue Cypress Creek	6B: Blue Cypress Creek	5, 6, 7, 8, 9, 10	USGS 02231396	0.52 (Satisfactory)	0.69
Crabgrass Creek	6E: Jane Green Creek	8	USGS 02231565	0.43 (Unsatisfactory)	3.55
Sixmile Creek	6F: St. Johns Marsh	2	USGS 02231454	0.60 (Satisfactory)	0.40
St Johns River near Melbourne	6F: St. Johns Marsh	7	USGS 02232000	0.88 (Very good)	6.22
Wolf Creek near Deer Park	6G: Lake Poinsett	4	USGS 02232200	0.61 (Satisfactory)	4.86
St Johns River near Cocoa	6G: Lake Poinsett	5, 6, 8, 10, 11, 12, 13	USGS 02232400	0.85 (Very good)	3.47
St Johns River near Christmas	6HI: Tosohatchee-Puzzle Lake	1, 2, 3, 4, 5, 6, 7, 8	USGS 02232500	0.88 (Very good)	-0.42
Ocklawaha River Basin (7)					
Ocklawaha River at Conner	7EF: Marshall Swamp-Rodman Reservoir	1, 2, 3, 4, 23, 5	USGS 02240000	0.98 (Very good)	0.01
Lower Orange Creek	7G: Newnans Lake-Orange Lake-Orange Creek	6, 7, 8, 9, 10	USGS 02243000	0.94 (Very good)	4.45
Bee Tree Creek	7G: Newnans Lake-Orange Lake-Orange Creek	6, 7	SJRWMD 02850235	0.86 (Very good)	1.68
Hatchet Creek	7G: Newnans Lake-Orange Lake-Orange Creek	1, 2, 3, 4, 5	SJRWMD 01950193	0.80 (Very good)	0.05
Little Hatchet Creek	7G: Newnans Lake-Orange Lake-Orange Creek	9	SJRWMD 02840233	0.78 (Very good)	0.01
Lochloosa Creek	7G: Newnans Lake-Orange Lake-Orange Creek	16, 17, 18, 19	SJRWMD 01930189	0.86 (Very good)	0.65

*See Table 11 for the performance rating scale.

Table 13 Calibration performance between simulated and observed flows at dynamically managed structures.

HSPF Hydrologic Model	Planning Unit	Subwatersheds	Flow/Water Level Station	Nash–Sutcliffe (Performance Rating)	Percent Error of the Mean
Upper St. Johns River Basin (6)					
S-96C	6B: Blue Cypress Creek	1, 2, 18	SJRWMD 0098	0.57 (Satisfactory)	-6.75
S-96B	6C: Fellsmere	3, 4, 5, 8, 10, 11, 12, 16, 17	SJRWMD 0096	0.11 (Unsatisfactory)	0.92
Jane Green Creek	6E: Jane Green Creek	1, 2, 3, 4, 5, 6, 7, 9	USGS 02231600	0.69 (Good)	0.09
Taylor Creek	6G: Lake Poinsett	15, 16, 17	USGS 02232415	0.29 (Unsatisfactory)	-5.74

*See Table 11 for the performance rating scale.

Table 14 Calibration performance between simulated and observed water level measurements at lakes.

HSPF Hydrologic Model	Planning Unit	Subwatersheds	Flow/Water Level Station	Nash–Sutcliffe (Performance Rating)	Percent Error of the Mean
Ocklawaha River Basin (7)					
Apopka	7BCD: Lake Apopka-Lake Harris-Lake Griffin	9	SJRWMD 30003000	0.85 (Very good)	-0.05
Newnans Lake	7G: Orange Creek	8, 10, 11, 12, 13	SJRWMD 04831007	0.86 (Very good)	0.05
Lochloosa Lake	7G: Orange Creek	20, 21, 22, 23, 24, 25, 26, 27, 31, 32, 33, 34, 35, 36, 37	SJRWMD 71481615	0.88 (Very good)	-0.05
Orange Lake	7G: Orange Creek	20, 21, 22, 23, 24, 25, 26, 27, 31, 32, 33, 34, 35, 36, 37	SJRWMD 02611465	0.97 (Very good)	0.03

*See Table 11 for the performance rating scale.

8. SUMMARY

The first step in the scientific foundation of WSIS was to evaluate surface water runoff of the St. Johns River watershed using the HSPF hydrologic model. Ninety-seven individual HSPF runoff models were developed, and 47 were directly calibrated against observed hydrologic data using

1995 land use conditions and 1995 through 2006 meteorologic data. This period was selected, in part, due to relatively constant physical conditions in the watershed. The calibration resulted in an acceptable fit for average daily discharges and a good fit for monthly and yearly results. Of the 47 calibrated watersheds, there were 21 'very good', 22 'good' and 'satisfactory', and only four 'unsatisfactory' values for the Nash–Sutcliffe statistic, which is a common rating of the performance of hydrologic models. The 'unsatisfactory' calibrations were of managed systems that are typically very difficult to match because of the vagaries of human behavior that drive the decisions on when and how much water is released or pumped. The parameters from the calibration models were then applied to the 50 watershed models that did not have gauges for calibration. This process allowed for development of a 1995 condition model for every subwatershed within the St. Johns River watershed. The results from over 900 subwatersheds were then aggregated as needed to develop flow input into the hydrodynamic model.

9. FUTURE WORK

9.1. Predictive uncertainty

The PEST optimization framework in parameter estimation mode will develop a parameter set where the simulated results have a minimum error variance compared to the observations. Within the PEST suite of utilities you can use the calibrated parameter set to develop a number of parameter sets that come close to the calibration results. For any of these new parameter sets, the model could be called calibrated. By analyzing these parameter sets, where each one 'calibrates' the model, you could determine the predictive uncertainty of the simulation results.

9.2. Quantitative analysis of observational error impact on the calibration

When comparing to observations in modeling, especially in the geosciences, the observations are used as if they have no error. We know this isn't true, but it is difficult issue to address. Establishment of confidence limits on observed flows can be done, but it is a difficult and time-consuming process.

9.3. Regularization

PEST run in regularization mode would normalize parameters between models, by balancing the parameters being nearly equal between models and the optimization of each model's parameter set.

9.4. Incorporation of non-contributing areas

There are several subwatersheds in the SJRWMD that have no observed surface flow. These subwatershed have been mis-named as 'non-contributing' when actually there could be baseflow into adjacent subwatersheds. Currently HSPF would implicitly handle significant baseflow from adjacent subwatersheds by adjusting the parameters to minimize the local baseflow contribution. A better representation of the system should be attainable by breaking the 'non-contributing' subwatersheds up and adding them to adjacent subwatersheds.

9.5. Modeling of springs

Currently spring flow is added to each HSPF subwatershed model as an external time-series. A regional model could be setup that would actually model the spring flow. As a point of interest, springs in SJRWMD are almost always around the 'non-contributing' areas mentioned in 9.4.

ACKNOWLEDGEMENTS

Several persons deserve acknowledgment for their contributions to this work. The project reviewers from Interra gave us helpful and constructive advice to improve the model calibration and enforce consistency across the project. John Doherty (author of PEST) was an invaluable guide helping us kick start the PEST/TSPROC/PAR2PAR driven calibration process, without which we would likely still be working on the project. And of course, the SJRWMD staff that poured time and effort into the WSIS project (Table 15).

Table 15 SJRWMD staff the worked on the WSIS watershed hydrology calibration and modeling.

SJRWMD Staff	Role
Timothy B. Cera, P.E. Florida Professional Engineer license: 52592	Editor, author, technical reviewer and manager; computer programming, lead for communication with National Academy of Sciences Review Committee, coordinated model development, and implemented training for HSPF, HSPF common logic, and PEST
Dale R. Smith, P.E. Florida Professional Engineer license: 32054	Working group leader, editor, author, technical reviewer and manager
Michael G. Cullum, P.E. Florida Professional Engineer license: 41869	Editor, project review and coordination
Marc Adkins, P.E. Florida Professional Engineer license: 42447	Author, hydrologic simulation of Black Creek, Deep Creek, Ortega, Trout, Little Wekiva; Helped to implement and use PEST with HSPF
Joseph Amoah, Ph.D., P.E. Florida Professional Engineer license: 73043	Author, hydrologic simulation of Crescent Lake and lower Orange Creek watersheds
David Clapp	Data management lead, author, hydrologic simulation of Orange Creek and Lake George watersheds
Robert Freeman, P.E. Florida Professional Engineer license: 17282	Author, hydrologic simulation of Etonia Creek watershed
Matt Hafner	Data management, author, hydrologic simulation of Lake George watersheds
Xiaoqing (Shaw) Huang, Ph.D.	Author, hydrologic simulation of Ocklawaha watersheds.
Yanbing Jia, Ph.D., P.E. Florida Professional Engineer license: 68422	Author, hydrologic simulation of Econlockatchee, Lake Monroe, and Lake Jesup watersheds; Water quality simulation of explicit modeling of BMP application to future land use
Tom Jobes	Author, computer programming, establishment of overall hydrologic framework, development of non-riparian/riparian methodology, hydrologic simulation of Upper Basin
Liang-Tsi (Maria) Mao, Ph.D., P.E. Florida Professional Engineer license: 46733	Author, hydrologic simulation of Six Mile, Big Davis, Pablo Creek watersheds and the north and south watersheds adjacent to the Lower St. Johns River

REFERENCES

Adamus, C., Clapp, D., & Brown, S. (1997). *Technical Publication SJ97-1: Surface Water Drainage Basin Boundaries [of the] St. Johns River Water Management District: A Reference Guide.* Palatka, FL: St. Johns River Water Management District.

Bicknell, B. R., Imhoff, J. C., Kittle, J. J., Jobes, T. H., & Donigian, J. A. (2001). *HYDROLOGICAL SIMULATION PROGRAM - FORTRAN, HSPF, Version 12, User's Manual.* Athens, GA: U.S. Environmental Protection Agency - National Exposure Research Laboratory.

Doherty, J. (2004). *PEST, Model-Independent Parameter Estimation, User Manual: 5th Edition.* Watermark Numerical Computing.

Dymond, J. R., & Christian, R. (1982). Accuracy of discharge determined from a rating curve. *Hydrological Sciences Journal, 27(4)* , 493-504.

EPA. (July 2000). *BASINS Technical Note 6: Estimating Hydrology and Hydraulic Parameters for HSPF.* EPA.

Hargreaves, G. H., & Allen, R. G. (2003). History and Evaluation of Hargreaves Evapotranspiratioin Equation. *Journal of Irrigation and Drainage Engineering* , 53-63.

Kennedy, E. J. (1983). Computation of Continuous Records of Streamflow. In *Techniques of Water Resource Investigactions, Book 3, Chapter A13.* U.S. Geological Survey.

Lumb, A. M., Carsel, R. F., & Kittle, J. L. (1988). Data management for water-quality modeling development and use. *Proceedings of the International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology.*

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bigner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of ASABE, American Society of Agricultural and Biological Engineers* , vol 50, #3, pp 885-900.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bigner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of ASABE, American Society of Agricultural and Biological Engineers* , vol 50, #3, 885-900.

St. Johns River Water Management District. (2006). *District Water Supply Plan, 2005, Technical Publication SJ2006-2.* Palatka, Florida: St. Johns River Water Management District.

St. Johns River Water Management District. (2012, February 14). *Final Report.* Retrieved August 5, 2012, from St. Johns River Water Management District web site - http://sjrwmd.com: http://www.sjrwmd.com/watersupplyimpactstudy/

State of Florida. (1999). *Florida Land Use, Cover and Forms Classfication System.* Geographic Mapping Section, Surveying and Mapping, Department of Transportation.

United States Geological Survey. (2010, January 15). *HSPF.* Retrieved January 18, 2012, from http://water.usgs.gov/software/HSPF

Westenbroek, S. M., Doherty, J., Walker, J. F., Kelson, V. A., Hunt, R. J., & Cera, T. (in review). *TSPROC - A general time-series processor to assist in model calibration and result summarization.* United States Geological Survey.