**4.A. CONCLUSIONS AND RECOMMENDATIONS FOR MODFLOW**

**1. Model Objectives, Conceptualization, and Design**

The model objectives, as stated in the Purpose and Scope (p. 1), i.e., the “…primary purpose of the NFSEG model is to enable improved evaluations of inter-District (e.g., SJRWMD/SRWMD) and inter-state (e.g., Florida/Georgia) water-level changes in the surficial and Floridan aquifer systems from groundwater use over the model domain.”, are addressed in the report. The previously prepared conceptualization report (Durden et al. 2013) details the plan for construction of the NFSEG groundwater model, including model extent, configuration, and lateral and internal boundary conditions; an analysis and interpretation of data needed for determination of the model calibration years; a plan for determination of groundwater recharge and maximum saturated evapotranspiration rates; and proposed NFSEG model calibration objectives.

The design structure of the MODFLOW model is considered excellent for the following reasons:

* Model boundaries extend to natural flow boundaries;
* Model grid cell size was as small as possible;
* Boundary conditions are no-flow where possible;
* Surficial aquifer was modeled as an active layer;
* ICU was modeled as an active layer;
* Geology was well researched;
* Aquifers were well researched and described;
* Representation of the aquifer layering in the model was generally good (having the UFA cross layers 1 and 2 did add some complexity); and
* Recharge was heavily researched using HSPF.

**2. Assumptions and Limitations of Input Data**

The hydrology of the area (Section 2. Hydrology of the Area) is described in terms of the surface-water and groundwater systems including rivers, lakes, swamps and wetlands, the Atlantic Ocean and Gulf of Mexico, the Surficial and Floridan Aquifer Systems, and the Intermediate Confining Unit (ICU). Sources of data for the groundwater systems include previously published reports, e.g., Miller (1986) and Bush and Johnston (1988), published aquifer test results, measured groundwater levels from water management district files and reports, and reported spring flows. Baseflows, which are an important calibration metric along with groundwater levels and spring flows, were estimated by averaging the results of four different hydrograph separation techniques and a fifth approach that utilizes flow duration curves. In addition, an alternative approach was utilized for determining baseflow pickups between adjacent gages that bound river reaches with contributing basin areas in which the ICU is absent completely, and which lack a well-developed, channelized, surface drainage network. In this approach, baseflow pickup was determined by taking the difference in the total observed flows of the upstream and downstream bounding gages, based on the assumption that overland runoff is negligible in areas in which the Upper Floridan aquifer is unconfined. Cumulative baseflow estimates and baseflow pickup estimates were mostly derived using the same averaging technique used to estimate baseflows. Concentrated groundwater inflows, e.g., rapid infiltration basins (RIBs) and drainage wells, and injection wells, are briefly described, along with groundwater withdrawals for public/commercial/industrial/institutional supply; agricultural irrigation supply; recreational irrigational supply (e.g., golf courses); and domestic self-supply. Input for recharge and maximum saturated ET are obtained from the results of the HSPF model simulations.

The description of the surface-water system is acknowledged to be brief (p. 5), and expanding the discussion of baseflows should be considered. The relative accuracy of the available data for groundwater heads and groundwater flows needs to be acknowledged, i.e., groundwater heads would be expected to be accurate to within a few tenths of a foot, but errors in estimates of groundwater flows (spring discharges and baseflows) would likely be much larger, e.g., the baseflow estimates may be accurate only to within an order of magnitude (ASTM 2018). Also, the discussion of groundwater inflows and withdrawals (pp. 22-23 and Figures 2-44 – 2-47) and the representation of the inflows and outflows in the MODFLOW well package (p. 41 and Figures 3-41 – 3-44) needs additional explanation and detail that could be provided in an appendix. Such detail would include well locations, pumping rates, and water-use categories for 2001, 2009, and 2010.

The assumption that groundwater flow in the Floridan Aquifer System can be approximated as laminar flow and represented as a porous medium in MODFLOW is applicable at the scale of the NFSEG grid spacing (2,500 feet x 2,500 feet discretization). Within the model domain, localized areas of non-linear laminar or turbulent spring flow may occur, and the accuracy of model results would be affected in such sub-regional areas. However, based on a comparison of the application of the MODFLOW Conduit Flow Package and a standard MODFLOW application at Wakulla Springs by Kuniansky (2016), the assumption that the standard MODFLOW porous medium approach is applicable throughout the NFSEG model domain is reasonable and defensible.

**3. Model Calibration and Sensitivity**

The NFSEG model was calibrated for 2001 (a relatively dry year) and 2009 (a relatively wet year) using hydrologic data (observations) including water levels, spring discharges, and estimated base flows. The year 2010 was selected for the NFSEG model verification. An initial steady-state manual calibration was performed, and the results from the initial calibration were used to guide the PEST- (Doherty 2010) facilitated process to achieve the calibration criteria of minimizing differences between observed and simulated water levels, head gradients, baseflows, and stream flow gains and losses. Numerous realizations (model runs) were performed in PEST by automatically varying various model parameters. Pilot points were used to estimate the spatial distribution of particular hydraulic properties within a model layer. Hydraulic parameters (model input), listed below, were varied spatially to achieve acceptable water levels and flows at hydrologic features of interest:

1. Kh and Kv multipliers of each layer;
2. Anisotropic ratio of each layer;
3. GHB conductance for springs;
4. River-bed conductance multipliers for stream baseflows;
5. Drain conductance multipliers for ephemeral streams;
6. River-bed conductance multipliers for lakes; and
7. Lake-zone multipliers for ICU (layer 2) Kv beneath lakes.

Recharge and ET, obtained from the HSPF results, were treated as constant inputs and not adjusted during the PEST runs. The calibration results were evaluated for heads, flows, and input parameters. The results between observed and simulated values for heads and flows (spring discharges and baseflows) were evaluated using standard statistical comparisons that include mean error (ME), mean of absolute error (MAE), standard deviation of error (SD), and correlation coefficient (R2). The calibration results for heads are compared to calibration targets for which 80% of the groundwater heads residuals should be within ±5 feet and for which 50% of the groundwater head residuals should be within ±2.5 feet. In the simulation results for groundwater heads, 72% of the residuals are within ±5 feet and 42% of the residuals are within ±2.5 feet for 2001, and 74% of the residuals for heads are within ±5 feet and 48% of the residuals are within ±2.5 feet for 2009. Thus, the simulation results do not achieve the calibration targets, but it is concluded that “…the percentage of the groundwater level residuals within 2.5 feet and 5 feet, generally indicate a very good match between observed and corresponding simulated values.” (pp. 58-58). Calibration targets were not established for spring discharges and baseflows. As described in Section 7 (Sensitivity and Uncertainty Analysis), the sensitivity of NFSEG model outputs to individual parameters was evaluated thoroughly to understand the importance of the various model input parameters to the behavior of simulated flows and levels using two methods, i.e., by calculating “traditional” parameter sensitivities as well as calculating composite-scaled sensitivities.

The justifications for treating evapotranspiration and recharge as constants in the PEST calibration in NFSEG Version 1.1 need to be discussed further in this report. Allowing evapotranspiration and recharge to be adjusted during PEST runs should be evaluated further in any future revision of Version 1.1 of the NFSEG model. Also, it is recommended that the calibration targets for groundwater heads be re-examined to determine if a broader range of statistical analyses such as criteria for mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) (e.g., Anderson and Woessner 1992) would provide a better set of metrics to judge the results for 2001, 2009, and 2010. Similarly, calibration targets should be established for spring discharges and baseflows, keeping in mind that the observed (or estimated) values may not be nearly as accurate as measured groundwater heads. Also, the residual statistics in Sections 4 and 5 (Model Calibration and Model Simulations) and results of other statistical analyses should be compared to residual statistics that have been obtained for other comparable regional groundwater flow models, e.g., SWFWMD’S District Wide Regulation Model (DWRM) version 3 and Northern District Model (NDM) version 5 and steady-state results in Sepúlveda et al. (2012).

**4. Model Documentation (explanation of model, data sources, and assumptions)**

In general, supporting documentation for the NFSEG model is adequate to assess the model results. However, additional statistical metrics and tests of random and normal distribution of residuals on heads, spring flows, and base flow residuals are needed to strengthen technical assessment of the calibration. Also, the “brief description of the surface-water system” (p. 5) needs to be expanded to include more descriptive material and details about baseflows. A weakness of the report is the use of qualitative statements such as “*good match, good agreement, generally good match overall, very good agreement, generally poor to fair comparison, generally poor comparison, and aspirational values*” to assess the goodness of fit between simulated and observed groundwater heads, spring flows, and base flows. Such qualitative descriptors are not easily evaluated because one’s person view of what represents “good” agreement between the model and observations can vary from another, and, thus, the use of these descriptors should be avoided.

**5. Suitability of MODFLOW and Related HSPF Models for Intended Applications**

MODFLOW-NWT is a recent version of MODFLOW from the USGS and is appropriate given the stated objectives of the model. In this application, the use of MODFLOW is suitable for water-resource assessment, determining minimum flows and levels, and evaluating water-use permit applications on a regional scale.

**Note: items 6, 7, and 8 are combined into one response:**

**6. Appropriateness, Defensibility, and Validity of Model/Relationships;**

**7. Validity and Appropriateness of All Assumptions Used in Development of Model/Relationships; and**

**8. Deficiencies, Errors, or Source of Uncertainty in Model/Relationship Development, Calibration, and Application**

The model objectives, as stated in the purpose and scope, are addressed in this report, and the model conceptualization is described in the report by Durden et al. (2013), which presents the geology, physical structure of the aquifer, and other background information needed to begin assembling the model. In the report reviewed here, the range of errors in the determination of the baseflows is not reported as recommended by ASTM (2018), and additional documentation and discussion of spring flows and baseflows is needed. Calibration targets for spring flows and baseflows also need to be established, and consideration needs to be given to adjusting recharge and/or ET during the PEST calibration in any subsequent revision of version 1.1 of the NFSEG model. Additionally, there is some indication that head, spring flow, and base flow residuals are not randomly distributed in the model domain. A non-random, spatial distribution in residuals often indicates model bias and possible model error. To determine the validity of spatial randomness, the “run statistics” (Hill 1998) calculated by the MODFLOW Observation Process or similar code should be used as an independent measure of randomness.

These additional considerations potentially will have important impacts on the applicability of the NFSEG groundwater model. The chapter-by-chapter comments in the Detailed Comments section of this draft peer review report should also be considered. Most aspects of the NFSEG model are valid and appropriate; if the concerns of the reviewers recommending additional discussion and explanation are addressed, then the NFSEG model should be quite appropriate and defensible.