NORTH FLORIDA SOUTHEAST GEORGIA GROUNDWATER MODEL
INDEPENDENT TECHNICAL PEER REVIEW

Task B.2. Phase 1 Draft NFSEG Version 1.1 Model

Prepared by

NFSEGv1.1 Technical Peer Review Panel

Louis Motz, Chair
Brian Bicknell
J. Hal Davis
James Rumbaugh
Dann Yobbi, P.G.

Prepared for

St. Johns River Water Management District
Palatka, Florida

and

Suwannee River Water Management District
Live Oak, Florida

June 30, 2017

CONTENTS

INTRODUCTION 3

SUMMARY OF KEY FINDINGS 5
1. HSPF Changes (Brain Bicknell) 5
2. MODFLOW Changes – (J. Hal Davis).. 5
3. MODFLOW Changes (Louis Motz).. 5
4. MODFLOW Changes (James Rumbaugh) 6
5. MODFLOW Changes (Dann Yobbi, P.G.)... 6

INDIVIDUAL PEER REVIEWER REPORTS..7
1. HSPF Changes (Brain Bicknell) 7
2. MODFLOW Changes – (J. Hal Davis).. 10
3. MODFLOW Changes (Louis Motz).. 18
4. MODFLOW Changes (James Rumbaugh) 24
5. MODFLOW Changes (Dann Yobbi, P.G.)... 31

INTRODUCTION

The North Florida Southeast Georgia (NFSEG) groundwater model is being developed by the St. Johns River Water Management District (SJRWMD) and the Suwannee River Water Management District (SRWMD) to provide a shared tool that can be used by both water management districts to assess the impacts of current and future groundwater withdrawals on water resources in north Florida. The model encompasses parts of Florida, Georgia, and South Carolina covering an area of approximately 60,000 square miles. The model is fully three-dimensional and utilizes seven layers to represent the surficial aquifer system, the intermediate confining unit, the Upper Floridan aquifer, the middle semiconfining unit, the upper zone of the Lower Floridan aquifer, the lower semiconfining unit, and the Fernandina Permeable zone of the lower Floridan aquifer where these hydrogeologic units are present. In its present form, the model has been calibrated to steady-state hydrologic conditions representing 2001 and 2009. To improve initial estimates of recharge and maximum saturated evapotranspiration for input to the NFSEG groundwater model, surface-water models have been developed for all surface-water basins within the groundwater model boundaries using the Hydrological Simulation Program-FORTRAN (HSPF) software. Version 1.0 of the NFSEG groundwater model and the HSPF-derived surface-water models were completed in 2016 and distributed in August 2016 to stakeholder groups that consisted of government organizations, water utilities, private industry, and environmental organizations and other interested parties throughout north Floridan and south Georgia for their use and review. Version 1.1 of the NFSEG model, which incorporates changes and improvements to Version 1.0, currently is under development by SJRWMD and SRWMD.

A panel of modeling experts was convened by SJRWMD and SRWMD in March 2017 to provide independent technical peer review of the NFSEG groundwater model and the HSPF models as the final phase of Version 1.1 of the model is being developed. This is intended to provide opportunities for the SJRWMD and SRWMD modeling team to incorporate peer review suggested changes into the model as it is being completed. Responsibilities of the Peer Review Panel include conducting a thorough review of the groundwater model and model documentation report and assessing the following topics:

- Model objectives, conceptualization, and design;
- Assumptions and limitations of input data;
- Model calibration and sensitivity;
- Model documentation;
- Suitability of MODFLOW and related HSPF models for the intended applications;
- Appropriateness, defensibility, and validity of the model/relationships;
- Validity and appropriateness of all assumptions used in the development of the model/relationships; and
- Deficiencies, errors, or sources of uncertainty in model/relationship development, calibration, and application.

To date, the Peer Review Panel has completed the first two tasks (Tasks A and B) of its scope of work. The effort for Task A consisted of reviewing applicable documents and background materials prepared for Version 1.0 of the NFSEG model and proposed improvements for Version 1.1 (Task A.1), attending a kick-off meeting at the SJRWMD in Palatka on March 29, 2017 (Task A.2), preparing draft initial recommendations that were presented at a teleconference to SJRWMD, SRWMD, and stakeholders on April 13, 2017 (Task A.3), and preparing and submitting a technical memorandum on May 1, 2017 (Task A.4), which contained the panel's final initial recommendations for changes and modifications to MODFLOW and HSPF. The panel's recommendations were grouped into recommendations for changes to Version 1.1 of the NFSEG model that could be completed by July 1, 2017 (Phase 1) and changes that could be considered later for Phase 2 or for future updates.

The effort for Task B has consisted of reviewing Phase 1 results for Version 1.1 of the NFSEG model. This has included reviewing preliminary model calibration results presented by SJRWMD and SRWMD at a teleconference on May 5, 2017 and making suggestions to facilitate the model improvements proposed by SJRWMD and SRWMD (Task B.1) and reviewing Phase 1 model files, draft figures and tables, and calibration statistics and attending a technical review meeting in Palatka on June 21, 2017 (Task B.2). This technical memorandum presents a summary of key findings as well as specific suggestions from each peer review panel member for completing outstanding tasks during the remainder of the NFSEG Version 1.1 development period (Phase 2 period) so that the NFSEG Version 1.1 model can be finalized. The specific suggestions include consideration of a no-pumping/pre-development scenario, an uncertainty analysis, and a verification run for the model, and editorial suggestions for figures and tables. Preliminary answers to Task D.2 questions regarding questions #2A-F Model Implementation and #3A-D, G, and H Model Calibration and Application also are included.

SUMMARY OF KEY FINDINGS

1. HSPF Changes (Brain Bicknell)

- The HSPF models are difficult to review because of the lack of a detailed summary of the model parameters by watershed and land use/cover category. Also, the absence of a complete set of calibration statistics for all the calibration sites has hindered our review. In our experience developing and reviewing HSPF-based models, knowledge of the model parameter values along with the calibration statistics covering various flow regimes are critical. The following statistics should be generated:

    - Overall error statistics (e.g., mean error, percent bias, mean absolute error, RMS error)
    - Error in high, medium and low flows (e.g., 10% high, 25% high, 50% high, 50% low, 25% low, 10% low)
    - Correlation tests (correlation coefficient, Nash Sutcliffe efficiency coefficient)

- The objective function has apparently been adjusted for this model phase, but it still seems to be unable to achieve reasonable calibration for several of the sub-watersheds within a HUC8 (based on review of updated graphics from the St. Johns and Suwannee River basins). The details of the objective function should be documented and should include evaluation of the effects of the individual components. In the previous recommendations, we suggested that the total overall volume and the flow frequency comparison should be given increased weight. If this advice was followed, what effect did the changes have on the calibration? Also, some of the poorly-calibrated sites that contain “good quality” observed flow data should be evaluated by the Districts to determine the likely cause of the poor calibration statistics. This would improve the understanding of model limitations and effects of the objective function.

2. MODFLOW Changes – (J. Hal Davis)

- The method of calculating baseflows at specific sites needs better documentation.
- If feasible, baseflows should be determined using different methods and the reason for choosing a specific method given.
- For model documentation purposes, a figure should be added showing calculated and simulated baseflows at significant gaging sites (sites with publicly available records).

3. MODFLOW Changes (Louis Motz)

- Overall, several significant improvements have been made in the NFSEG Version 1.1 Model. However, additional improvements should be considered. The presentation of calibration results for heads and spring flows could be improved by graphically presenting the results in terms of the distribution of residuals in histograms in addition to the x-y plots already included. Improving the calibration results for heads and spring flows and particularly for vertical head differences should be continued.

- The no-pumping/predevelopment simulation and the 2010 verification run that are to be considered in the next step of model development and review will certainly need to be discussed further. In particular, for the no-pumping/predevelopment simulation, just setting groundwater discharge to zero may not be sufficient, but deciding how to represent predevelopment rainfall, evapotranspiration, land use, and heads and flows at the boundaries will be very challenging.

4. MODFLOW CHANGES (JAMES RUMBAUGH)

- Establish calibration goals for the flux observations (river base flow and spring flows)
- Consider removing temporal head difference observations
- Consider removing the lake observations
- Change ET to an estimated parameter and activate the wetting penalty observation type to determine if the large flooded areas can be alleviated.
- Evaluate parameters that hit an upper or lower bound in the latest PEST run
- Review APTs in layer 3 where transmissivity exceeds 1million cubic feet per day to determine if they are appropriate for comparison to the model.
- Adjust areas in layer 4 where vertical hydraulic conductivity significantly exceeds horizontal hydraulic conductivity.

5. MODFLOW CHANGES (DANN YOBBI, P.G.)

- Base flows need to be reevaluated/reviewed prior to recommencement of numerical model simulations.

INDIVIDUAL PEER REVIEWER REPORTS

1. HSPF CHANGES (BRAIN BICKNELL)

Specific Suggestions for Model Improvements:

1. The calibration of HUC8 watersheds without regularization of the parameters in adjacent watersheds can lead to parameter sets that appear very different. There should be an effort to improve the consistency or similarity of parameter sets in adjacent basins. The lack of a detailed parameter summary makes it difficult to evaluate whether this is a problem.
2. An important part of the HSPF models' response is determined by the amount of storage and the shape of the discharge function in the stream/river reaches. This information is embodied in the reach FTABLEs. In flat areas where there are large amounts of storage in wetlands that are part of (and adjacent to) the reaches, adjustment of the storage and discharge relationship in the FTABLEs has been found to improve the flow agreement and shape of hydrographs. Since many of the existing FTABLEs are the default tables generated by the BASINS tool, which is probably not appropriate for typical Florida watersheds, improvement of these FTABLEs (and the ones generated from other sources) by calibration of storage and discharge should be considered. This would be an additional objective function component.
3. The Districts should attempt to determine the impact of a poor-quality calibration of streamflow at some of the calibration sites in a HUC8 watershed on the predicted recharge in that HUC8. Assuming that an area experiences 18 inches of annual runoff, and assuming that the measured flow and the other water balance components are "correct", then a 20% error in the simulated runoff would translate to a 3.6-inch error in annual recharge. This is likely overstating the effect, since many calibration sites have smaller errors, and errors in both directions (over and under predictions) will tend to balance out the problem. However, we recommend that the Districts attempt to determine the uncertainty in the recharge prediction that is due to poorly calibrated total annual flow.
4. The Districts should generate summaries of the average annual HSPF water balance results for the land areas (PERLND and IMPLND). The water balance provides a summary of the: 1) inputs (rainfall, irrigation), 2) evapotranspiration losses, 3) runoff losses to streams (by soil layer), and 4) groundwater recharge. Weighted average summaries can be generated for each land cover in a watershed and also averages over all land covers. The HSPEXP+ software package provides a tool for generating these water balance results from HSPF output.

Specific suggestions for Next Steps:

a. No-pumping / predevelopment simulations
   Will the recharge from the existing HSPF models be used for these simulations? Will an attempt be made to change the HSPF models to a pre-development scenario and re-run to

generate recharge? If so, all land use would be converted to the natural categories, and irrigation and related withdrawals turned off.

b. Uncertainty analysis approach
   Will the uncertainty analysis be extended to the HSPF models?
c. 2010 verification run

Preliminary Answers to Task D Questions:

1. Model Documentation from WMDs
   A. The documentation is clear, but incomplete. It should be more detailed about model parameter values, calibration statistics, and the calibration objective function. It should also include more details (e.g., tables of annual averages) of the rainfall, PET, and irrigation-related inputs. These could be included in appendices.
   B. See item A.
   C. The documentation is readable and figures are clear. The format is satisfactory. As listed above, additional detail of at least three types of documentation should be included. These are parameter values, additional calibration statistics, and calibration objective function details and impacts.
   D. Most of the model is understandable. Additional documentation describing the limitations is needed.
2. Model Implementation
   A. The HSPF models are based on a good conceptual model and seem to have incorporated good decisions about data and the approach to HSPF modeling. (See below in item F. b.)
   B. The model code is appropriate.
   C. Yes.
   D. Yes.
   E. Yes.
   F. HSPF Questions
      a. I am satisfied that the version of HSPF used is defensible and appropriate. However, this should be described more clearly in the documentation, including a description of the features that are non-standard, and citation of a document that confirms the District's prior validation of the non-standard version.
      b. In general, I am satisfied that best available information was used to develop the HSPF models:
      - Input rainfall and potential evapotranspiration data, which are the primary driving forces of the model utilize the best available, consistent data source (NLDAS) that covers the entire model area; and the PET data have been adjusted to represent the correct quantity required by HSPF, i.e., lake evaporation.
      - The effects of irrigation are being represented using standard methodologies. The quantities are estimated using an appropriate model of soil moisture and available rainfall, and the amounts are made consistent with known records of local pumping and withdrawal data.

- The model area has been segmented (delineated) into watersheds based on the USGS HUC8 watersheds, with sub-watersheds based on elevation (DEM) data. The hydrology of separate land covers/uses is represented by segmenting the sub-watersheds into individual hydrology computational units (pervious and impervious land segments) using NLCD 2001 coverage.
- Closed basins, i.e., those that do not contribute surface flow to the major streams and rivers are modeled in a creative and effective manner that allows the hydrology parameters for those areas to be consistent with the adjacent contributing land areas. This leads to a consistent method for estimating recharge for these basins.
- Spring discharge is represented appropriately, and a creative method for predicting spring discharge quantities for some major springs using model-generated flows from the springsheds has been conceptualized and incorporated into the model. The comparison of predicted and measured spring discharges suggests the method is valid, and can be applied to periods (and springs) that are unmeasured.
- The calibration of the modeled streamflows to measured USGS gage flows using the PEST program is appropriate to ensure a defensible and consistent approach. The calibration objective function includes a series of metrics that are primarily related to the streamflow (i.e., at various seasons and time intervals and hydrologic regimes), but also attempts to constrain the simulated evapotranspiration to a consistent set of literature values for the different vegetative land covers. This is designed to improve the recharge prediction by constraining a significant water balance component that is difficult to measure, to a consistent and reasonable range.
- Note: I have not closely reviewed the Special Actions code for the springs and closed basins to see if it is correct. Also, I have not looked closely at the irrigation implementation in the model files. That will be done.

3. Model Calibration and Application
   A. Is this question intended to apply to HSPF? If so: The PEST scheme used for HSPF appears reasonable; however, there is insufficient description of the objective function and the effects of its various components on the calibration quality. (Refer to Item 2 under Specific Suggestions for Model Improvements.)
   B. Unknown
   C. The model residuals have not been described adequately. A more complete list of typically used statistics should be generated for each HSPF calibration site. (Refer to Item 1 under Specific Suggestions for Model Improvements.)
   D. Values of calibrated parameters have not been adequately documented. (Refer to Item 1 under Specific Suggestions for Model Improvements.)
   E. It is not clear that the final version of the model is adequately calibrated, since a complete set of model statistics has not been produced.
   F-J Appear to be specific to the MODFLOW GW Model.

2. MODFLOW CHANGES – (J. HAL DAVIS)

Specific Suggestions for Model Improvements:

There needs to be better documentation of the baseflows used for calibration. This documentation should include:

1) For each of the gages on the Withlacoochee, Alapaha, Suwannee, and Santa Fe Rivers there needs to be a plot covering the period from 2001 to 2009 (plus a few years before and after). In the text, there should be a discussion describing what the plotted flows indicate about the groundwater flow system.
2) On each plot, there should be an indication of the calibration baseflow determined (if one was determined) and the model simulated values. They should all have a simulated value.
3) I would also recommend doing this at any other gages deemed critical for the calibration.

These are my most important suggestions and I have included them for the following reasons. The accurate determination of baseflows is probably the most crucial factor in calibrating the groundwater model. The water levels used for calibration are usually accurately measured and straight forward. In contrast baseflows are difficult to determine and usually have a large uncertainty (at least compared to the water levels).

Calibrating to the baseflows determines the volume of water moving through the aquifer, and the volume of water will determine (or strongly influence) the recharge rates, magnitude of the impact of pumping, and the hydraulic conductivities determined.

I reviewed the river flows at the upstream gages located on the Withlacoochee River (at Quitman), Alapaha River (at Statenville), and Suwannee River (at Fargo) to see what they showed (figure 1). The flows in the Withlacoochee River (at Quitman) seem to recede back to about 30 cfs during extended periods of low rainfall, the flows in the Alapaha River (at Statenville) seem to recede back to about 50 cfs, and the flows in the Suwannee River (at Fargo) seem to recede back to about 20 cfs. The combined flows in these gages recedes back to about 100 cfs. I have canoed these rivers in periods of low water over the years and have never seen anything that would indicate that they are connected to the UFA upstream of the gages. So, I would attribute these flows to the surficial aquifer system based on my personal observations.

In addition, the flow in the Suwannee River at White Springs (figure 2) is similar that at Fargo (figure 1), indicating that the Suwannee River may not have significant interaction with the UFA until below this gage.

The baseflows used for model calibration for 2001 were 650.7, 587.6, and 155.6 cfs at Quitman, Statenville, and Fargo respectively; and 1,372.3 and 1,144.9 cfs at Quitman and Statenville (no value was given for Fargo). These seem too high.

At the Ellaville gage on the Suwannee (the next one down from White Springs) the baseflow seems to recede back to about 900 cfs indicting a strong connection with the Floridan aquifer. The increasing baseflows downstream at Luraville (1100 cfs), Branford (1500 cfs) and Wilcox (3000 cfs) indicate an increasing influence of the UFA.

It is interesting that the all the gages on the Suwannee that interact with the UFA seem to recede back to a constant value over time, even after periods heavy rainfall year. This could be used as evidence that a steady-state model is appropriate for simulating flow conditions over periods of many years. I had hoped that this low baseflow would be useful in understanding the model calibration but it obviously does not include the higher baseflows during the period of recession thus limiting its usefulness.

One possible line of evidence for baseflows would be to subtract the combined flows of the Withlacoochee (at Quitman), Alapaha (at Statenville), and Suwannee (at Fargo) from flows in Suwannee at White Springs, Ellaville, Luraville, and Branford (for the gage at Wilcox the flow in the Santa Fe [at Worthington] would also need to be subtracted). Not that this would be the final word on baseflow but would be another line of evidence, in addition to the baseflow method the modelers have already used. In the final analysis, it is up to the modelers to decide how they determine the calibration baseflows since they need to justify them.

As a final note, I do like that the model is being calibrated using many baseflows. It adds to the workload but will improve the overall quality of the model.

Specific suggestions for Next Steps:

Verification Run – 2010

To judge the overall quality of the 2010 verification run I would use the same criteria as for the 2001 and 2009 simulations, which is a comparison of the measured and simulated heads and flows.

Uncertainty Analysis Approach

In the report, there should more discussion on what the uncertainty analysis indicates about the model, such as which parameters are the best defined (and least well defined) and how the uncertainty in these would affect the model predictions.

No-Pumping / Predevelopment Simulations

I plotted the long-term flows in the Suwannee River at Branford to help determine if there were definite trends in flow (figure 3). I also estimated baseflow (I am not sure that the baseflow method used is best, but it does at least remove the higher flows). Both plots do show a downward trend in flows from 1998 to present. From 1970 to 1998 there appears to be no definite trend and has historically low variability.

It is possible that the lower flows from 1998 to present are a result of higher temperatures causing higher ET. Fortunately, the model includes ET so this effect (or non-effect) may possibly be determined. Based on just this gage (it is the only one I used) it appears that it may be difficult to determine smaller changes in Suwannee River flow due to the natural variability occurring. If this turns out to be the case, then the effect of predevelopment pumping will need to be determine based on water levels alone.

One thing that may be helpful would be a discussion on what constitutes success for this task.

Preliminary Answers to Task D Questions:

2. Model Implementation:

   A. Is the conceptual model appropriate for the intended use of the model? For example, are critical physical and hydrologic processes represented appropriately?

   Mostly I think so. The lateral boundaries are excellent. The inclusion of all the aquifers in the model (and not just the UFA) is good.

   I am not very knowledge on HSPF method so I cannot independently verify the output. But the detailed approach to determining ET and recharge in individual basins appears sound. As described above, I do think that there needs to be more discussion on baseflows.

   B. Is the model code appropriate, given the intended use of the model?
   I think that MODFLOW can give useful results for this hydrologic system. In the report, the modelers need to explain their reasoning in choosing the code and the limitations.

   C. Was numerical model constructed in a manner that is consistent with the underlying conceptual model, using appropriate data and methods of analysis?
   Again, I think so. Assuming baseflows are addressed further.

   D. Was the hydrologic model code selected appropriate for its intended use?
   Same answer as above. I think MODFLOW can give useful results for this hydrologic system. In the report, the modelers need to explain their reasoning in choosing the code and the limitations.

   E. Was the use of HSPF as a method to develop recharge and maximum saturated ET that is assigned to the MODFLOW groundwater flow model a valid and defensible method?
   Seems so, but I am not familiar enough with HSPF to made a judgement call on the details.

   F. Questions specific to HSPF model:

      a. The version of HSPF utilized for the hydrologic models is a non-standard version of HSPF that is not publicly available. Is the version of HSPF utilized appropriate and defensible?

      b. Was best available information utilized to develop the HSPF hydrologic models?

         i. Unique aspects of these systems were represented with Special Actions or with another feature of HSPF and are these conceptually sound and implemented appropriately:

1. RCHRES representation of Inactive Groundwater Storage to represent spring discharges
2. Closed basins
3. Drainage wells and swallets
4. Implementation of water use
    a. Agricultural irrigation
    b. Urban
        i. Septic
        ii. Irrigation
    c. Golf Courses
    d. Reuse spray fields

3. Model Calibration and Application:

A. Is the parameterization scheme used in the PEST calibration appropriate?
I would defer to Jim on the specifics of the PEST calibration.

B. Were the types of observations and their implementation in the PEST calibration appropriate, given the intended use of the model?
I would defer to Jim on the specifics of the PEST calibration.

C. Have differences between observations and their simulated equivalents (model residuals) been described sufficiently. For example, have an appropriate set of summary statistics, plots, and maps been presented that allow for evaluation of model limitations, (such as model bias and uncertainty) in a manner that meets or exceeds existing professional practices.
As discussed earlier, I would like to see more comparisons of the simulated baseflows to the actual baseflows. The measured heads are directly compared to the simulated heads (in map view) in the report. This provides an excellent and easily understood way to evaluate the model. I think that he same comparison should be given for the flows.

D. Have the values of calibrated parameters been described appropriately, using (for example) maps illustrating the range and spatial distribution of parameter values.
For the measured heads, yes, but not for the flows.

E. Has the complete model water balance, accounting for all water sources and sinks, been assessed and found reasonable?
I think so.

F. Have the uncertainty of key model parameters and predictions been assessed using methods that are appropriate and that meet or exceed typical practice for developing groundwater flow models? Has a detailed statistical assessment of uncertainty in modeled groundwater level and spring flow estimates been provided?
For the groundwater levels, yes.

Review of document: NFSEG_Proposed_Improvements_Status-June2017

Each item is addressed:

1. Update river and drain package

Comment: Excellent. The overall components of the MODFLOW water budget look reasonable. I went through the MODFLOW output budget (for the 2001 simulation). Now recharge is 85% of inflow (very good), constant heads are 2.2% of the inflow and 4.1% of the outflow (very good, this indicates that boundary conditions are only accounting for a small amount of water), and all the other flow components look reasonable.

2. Update and recalibrate HSPF models

No comment.

3. Improve simulated SAS water levels

Comment: Seems like a reasonable approach.

4. Reassess the use of MNW2 package for modeling multi-aquifer wells

Comment: Seems like a reasonable approach.

5. Improve simulated spring flows

Comment: Seems reasonable.

6. Improve baseflow simulations in the groundwater model in critical areas

Comment: Seems reasonable.

7. Improve point-source recharge distribution

Comment: Seems reasonable.

8. Improve aquifer parameter estimates in the model

Comment: Seems reasonable.

9. Null Space Monte Carlo Uncertainty Analysis

Comment: Seems reasonable, especially the concern about accurate recharge rates.

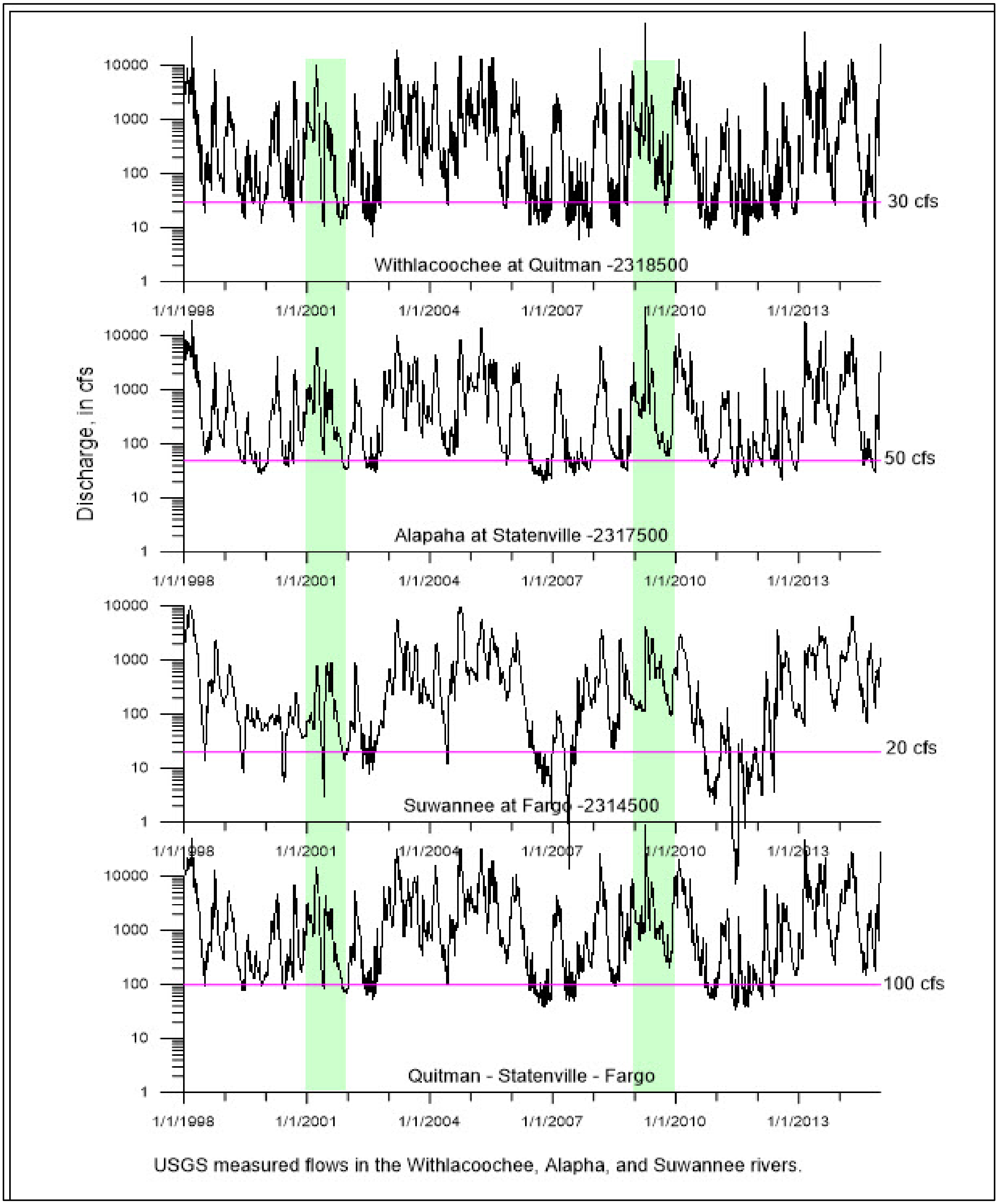

Figure 1. Measured flows in the Withlacoochee, Alapaha, and Suwannee Rivers

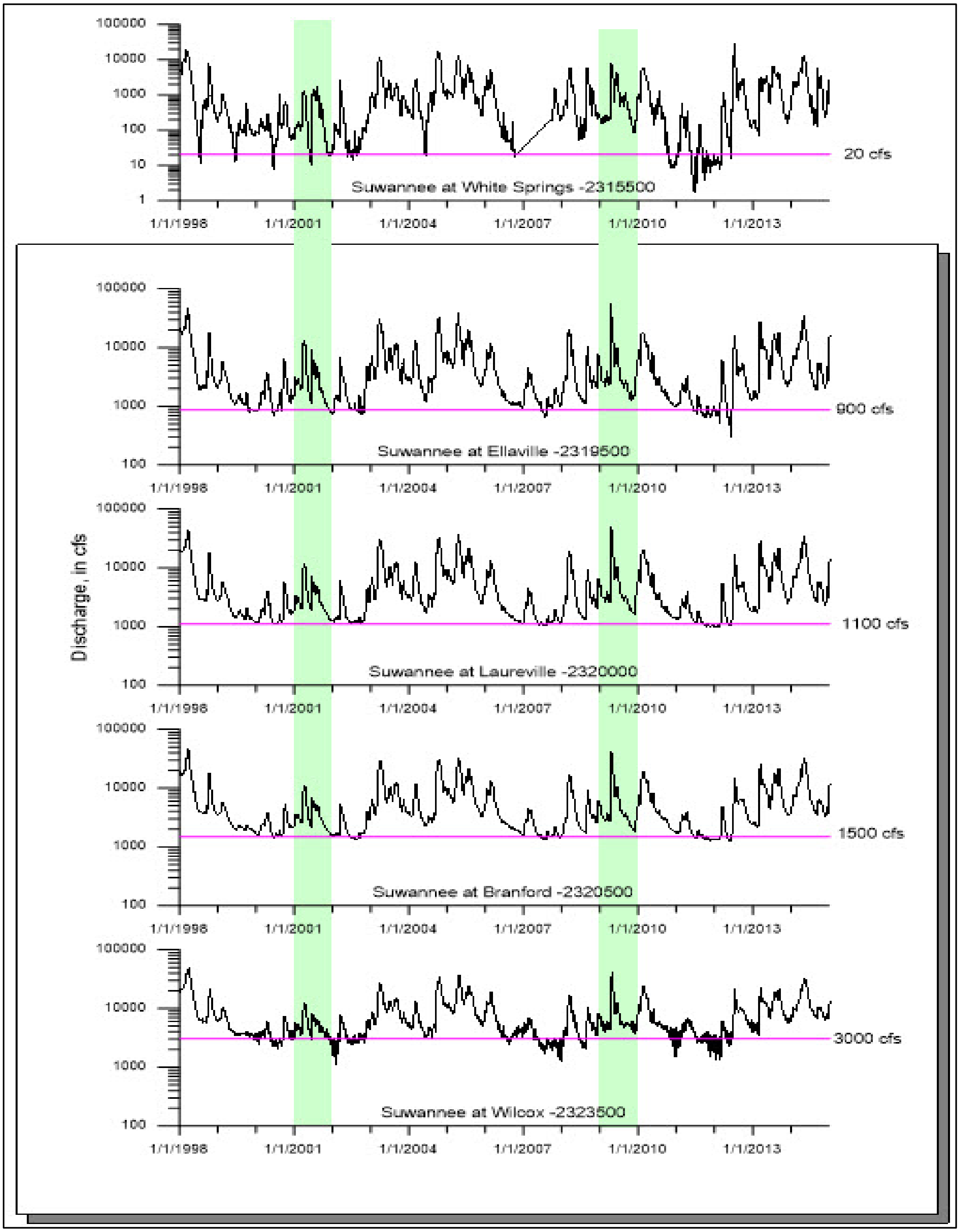

Figure 2. Measured flows in the Suwannee River at various gages.

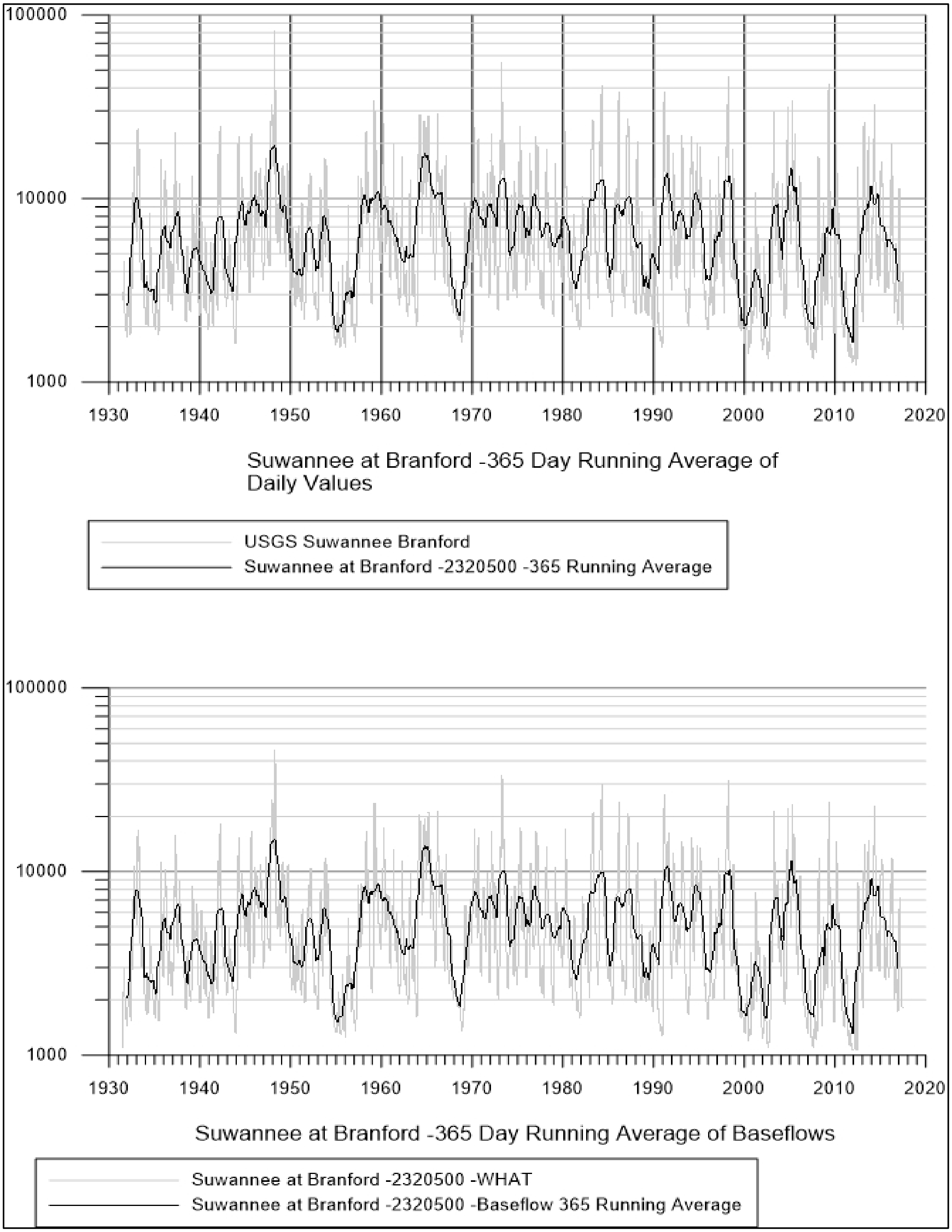

Figure 3. Suwannee at Branford for years 1930 to 2017.

3. MODFLOW CHANGES (LOUIS MOTZ)

Specific Suggestions for Model Improvements:

1. In Figures 84-89, it is not very easy to get a sense of how well the model is calibrated by just looking at these residual maps. Adding histograms that show the distributions of the residuals in terms of the means and standard deviations (in addition to Figures 90-96) would improve the presentation of these results. The same comments apply to the Figures 97-100, but there may not be enough residual values to plot histograms for these results.
2. The results plotted in Figures 103-104, particularly Figure 104, indicate that vertical head differences between layers 3 and 5 in 2001 and 2009 are not well represented in the calibrated model. Can these results be improved?
3. In Figures 105-108, it is very difficult to understand what these results represent. In Figure 105, values of ± 5 feet that are identified as residuals are shown over a large part of the model area in Florida and significant areas of Georgia. Does this mean that the simulated heads for 2001 in a large part of model layer 3 are in error by ± 5 feet? Similarly, in Figure 106, values of -14.9 to 0 feet and 0.1 to 15.0 feet that are identified as residuals are plotted over most of the model area. Does this mean that simulated heads for 2009 over most of model layer 3 are in error by as much as ± 15 feet? Are the values plotted in Figures 105-106 residuals or measured head differences? Simulated versus observed values of heads (not residuals) appear to be plotted in Figures 107 and 108. How were these values obtained from Figures 105 and 106?
4. In Figures 111-112, the simulated and observed spring discharge values match reasonably well except for one value in each figure. Is this the same spring for both 2001 and 2009? Can this apparent error be fixed?
5. Tables 2, 3, 4, and 8 are comprised of a large number of pages in text files, i.e., as many as 57 pages (Table 2). The information in these lengthy tables needs to be made available in spreadsheet format.
6. It is never made clear in the text exactly why years 2001 and 2009 were selected for the calibration years, i.e., what do these years represent? The obvious answer is that they represent a dry year and a wet year, respectively. However, are these the driest and wettest years in the recent period of year or were they selected because they are considered “representative” or were they selected for other reasons such as best available data sets? Providing annual rainfall totals for the model area in the form of a table and bar graph for the period of record at least from 2001 to 2010 and explaining the logic of selecting years 2001 and 2009 for the model calibration would be very helpful. Note: Please see Suggestions for Further Discussion 3. 2010 Verification Run (below) for similar discussion.

Specific Suggestions for Next Steps:

No Pumping/Predevelopment Simulation

No-pumping scenarios in which groundwater withdrawals are set to zero in the NFSEG model could be employed to estimate drawdowns due to pumping that have occurred in the surficial and Floridan aquifer systems. However, there likely will be different results for heads for the drier year (2001) versus the wetter year (2009) that will have to be accounted for. Differences in recharge, evapotranspiration, and specified heads and flows at the model boundaries for 2001 and 2009 also will have to be considered. Matching the pre-development potentiometric surface of Johnston et al. (1980) will require reviewing how the pre-development map was prepared and estimating recharge, evapotranspiration, and specified heads and flows at the model boundaries for pre-development conditions. Determining how significant land-use changes from pre-development time to 2001 and 2009 have been and whether they should be considered also needs to be discussed.

Uncertainty Analysis

An uncertainty analysis that could provide a probabilistic description of parameters of interest such as heads and spring flows and impacts due to future pumping should be considered. Sepúlveda et al. (2012) describe using a PEST application PREDUNC6 (Doherty 2010a and b) to calculate uncertainties in horizontal and vertical hydraulic conductivities in layers 1, 3, 5, and 7 in a groundwater model of the surficial and Floridan aquifer systems in east-central Florida that is similar in complexity to the NFSEG groundwater model. Also for the east-central Florida groundwater model, Sepúlveda and Doherty (2015) developed probabilistic descriptions of heads, spring flows, and drawdowns due to changes in groundwater withdrawals from 1999 to 2035. The Null Space Monte Carlo method, a stochastic probabilistic approach, was used in this application. Similar analyses should be considered for the NFSEG model.

2010 Verification Run

Selecting a representative year that was not used in the model calibration for a verification or model validation run is often done as part of developing a numerical model. However, several points need to be clarified to justify this selection. The first step is justifying the original selection of 2001 and 2009 for the calibration. It is never made clear in the text exactly why these years were selected, i.e., what do these years represent? The obvious answer is that they represent a dry year and a wet year, respectively. However, are these the driest and wettest years in the recent period of year or were they selected because they are considered “representative” or were they were they selected for other reasons such as best available data sets? The next step should involve justification for considering 2010 for the verification run. Are rainfall totals for 2010 somewhere between those for 2001 and 2009, or was 2010 more extreme (drier or wetter) than 2001 and 2009? Providing annual rainfall totals for the model area in the form of a table and bar graph for the period of record at least from 2001 to 2010 would go a long way toward explaining the logic of using these years in the model calibration and verification. Finally, it should be pointed out that the current view among many modelers is that a model cannot be verified or validated but only invalidated and that verification and validation have been replaced with other types of model performance evaluation such as parameter estimation and forecast uncertainty analysis (Anderson et al. 2015). Note: Please see **Suggestions for Further Discussion 2. Uncertainty Analysis** (above) for related discussion.

Editorial Suggestions for Figures and Tables

- Figure 1: Fall Line and Gulf Trough need to be identified and referenced in the text.
- Figures 22–25: Where is the 2001 potentiometric surface?
- Figures 45-46: Major pumping centers occur in the vicinity of Brunswick, so this city should be added to these maps (and to the base map).
- Figures 45-46: In addition to these figures, please plot total pumpage for model area for 2001 and 2009 by category for agricultural, public supply, domestic self-supplied, and landscape/ recreational/ aesthetic categories (bar graphs).
- Figures 76-77 and 80-81: make sure contours are adequately labeled.
- Figures 113-120: The reader is referred to Tables 10 and 11 for the USGS gage ID. The tables where the extremely large number of discharge values and residuals plotted in Figures 113-120 are located should also be indicated on the appropriate figure.
- Figures 121-122: The model-wide average recharge rates (inches/year) for 2001 and 2009 should be listed on these figures.
- Figures 129-130: Please check the figure titles shown in the list of figures. The water budgets plotted in Figures 129 and 130 appear to be summaries for the groundwater model, not for layer 3 as indicated in the list of figures. Water budgets summarized for the entire model are more meaningful than water budgets for just layer 3 in these figures.
- Figure 135: How well do the estimated transmissivity values for layer 3 match the APT results plotted on the same figure?
- Table 5: Values for proposed targets for mean error, root mean square error, and mean absolute error need to be included in the table. Values should be based on and referenced to generally recognized standards such as those of ASTM and/or USGS.
- Tables 2, 3, 4, and 8: These tables are comprised of a large number of pages in text files, i.e., as many as 57 pages (Table 2). In their present format, the tables are not of much value for review and checking results. The information in these lengthy tables needs to be made available in spreadsheet format.
- Table 9: The proposed title in the list of tables does not match the content of this table.
- Table 10: Table 10 apparently contains the content indicated in the list of tables as being in Tables 9 and 10.
- Table 11: Table 11 apparently contains the content indicated in the list of tables as being in Tables 11 and 12.
- Table 12: Table 12 apparently contains the content indicated in the list of tables as being in Table 13.

Preliminary Answers to Task D.2 Questions #2A-FD and #3A-D, G, and H

2. Model Implementation:

A. **Is the conceptual model appropriate for the intended use of the model? For example, are critical physical and hydrologic processes represented appropriately?**
   Yes, it appears that the physical, hydrologic, and hydrogeologic features in the model area are appropriately conceptualized.

B. **Is the model code appropriate, given the intended use of the model?**
   Yes, MODFLOW is an appropriate code to use for a large regional model.

C. **Was the numerical model constructed in a manner that is consistent with the underlying conceptual model, using appropriate data and methods of analysis?**
   Yes, the numerical model appropriately represents the underlying conceptual model.

D. **Was the hydrologic model code selected appropriate for its intended use?**
   Yes, HSPF is an appropriate model code to use to simulate the unsaturated zone and its interaction with the underlying zone of saturation.

E. **Was the use of HSPF as a method to develop recharge and maximum saturated ET that is assigned to the MODFLOW groundwater flow model a valid and defensible method?** Brian Bicknell needs to answer this question.

F. **Questions specific to HSPF model:** Brian Bicknell needs to answer this question.
   a. **The version of HSPF utilized for the hydrologic models is a non-standard version of HSPF that is not publicly available. Is the version of HSPF utilized appropriate and defensible?**
   b. **Was best available information utilized to develop the HSPF hydrologic models?**
   i. **Unique aspects of these systems were represented with Special Actions or with other feature of HSPF and are these conceptually sound and implemented appropriately:**
      1. **RCHRES representation of Inactive Groundwater Storage to represent spring discharges**
      2. **Closed basins**
      3. **Drainage wells and swallets**
      4. **Implementation of water use**
         a. **Agricultural irrigation**
         b. **Urban**
            i. **Septic**
            ii. **Irrigation**
         c. **Golf Courses**
         d. **Reuse spray fields**

3. Model Calibration and Application:

A. Is the parameterization scheme used in the PEST calibration appropriate?
Jim Rumbaugh needs to answer this question.

B. Were the types of observations and their implementation in the PEST calibration appropriate, given the intended use of the model?
Jim Rumbaugh needs to answer this question.

C. Have differences between observations and their simulated equivalents (model residuals) been described sufficiently? For example, have an appropriate set of summary statistics, plots, and maps been presented that allow for evaluation of model limitations, (such as model bias and uncertainty) in a manner that meets or exceeds existing professional practices.
Residuals and simulated vs. observed heads (Figures 84-108), spring flowrate residuals and simulated vs. observed spring discharges (Figures 109-112), and estimated baseflow pickup residuals (Figures 113-120) are presented. Residuals of heads appear to be randomly distributed in the model layers in which they occur. However, an uncertainty analysis has not been conducted yet, and the model documentation text for NFSEG Version 1.0 has not been updated.

D. Have the values of calibrated parameters been described appropriately, using (for example) maps illustrating the range and spatial distribution of parameter values?
A large number of maps (Figures 131-141) appropriately illustrate the range and spatial distribution of parameter values. However, there is still a question about how well the spatial distribution of transmissivity in model layer 3 matches the APT results shown in the same figure (Figure 135).

G. Has the complete model water balance, accounting for all water sources and sinks, been assessed and found reasonable?
Yes, the water budgets for the groundwater model for 2001 and 2009 (Figures 129 and 130) appear to be reasonable.

H. Have the uncertainty of key model parameters and predictions been assessed using methods that are appropriate and that meet or exceed typical practice for developing groundwater flow models? Has a detailed statistical assessment of uncertainty in modeled groundwater level and spring flow estimates been provided?
Not yet. According to SJRWMD, an uncertainty analysis will be performed in Phase 2 as the model is completed.

References Cited

Anderson, M.P., W.W. Woessner, and R. J. Hunt. 2015. Applied Groundwater Modeling: Simulation of Flow and Advective Transport. Second Edition. Elsevier, 564 pp.

Doherty, J. 2013a. *PEST, Model-Independent Parameter Estimation- User Manual*, 5th Edition, Brisbane, Australia. Watermark Numerical Computing.

Doherty, J. 2013b. *Addendum to the PEST Manual*, Brisbane, Australia. Watermark Numerical Computing.

Johnston, R.H., R.E. Krause, F.W. Meyer, P.D. Ryder, C.H. Tibbals, and J.D. Hunn. 1980. *Estimated potentiometric surface for the tertiary limestone aquifer system, Southeastern United States prior to development*. U.S. Geological Survey Open-File Report 80-406. Tallahassee, Florida, 1 sheet.

Sepúlveda, N. and J. Doherty. 2015. Uncertainty Analysis of a Groundwater Flow Model in East-Central Florida. Groundwater, 53(3), 464-474, May-June.

Sepúlveda, N., C.R. Tiedeman, A.M. O'Reilly, J.B. Davis, and P. Burger. 2012. *Groundwater flow and water budget in the surficial and Floridan aquifer systems in east-central Florida.* U.S. Geological Survey Scientific Investigations Report 2012-5161, 214 pp.

4. MODFLOW CHANGES (JAMES RUMBAUGH)

Specific Suggestions for Model Improvements

The calibration has established goals for heads but not fluxes, such as spring flows and river baseflows. I believe such goals should be established and that the tables comparing these fluxes to the model calibration should include a percent error column. I created a simple Table R-1 that shows springs and spring groups over 200 cfs, along with percent error in the 004b PEST run. There are several springs and spring groups that have significant error, which is more visible when expressed as a percent. I believe the next run should address the largest of these errors, which may require some hand tuning of the model prior to the PEST run. This same comment applies to river base flows although I have not presented a table for those.

Two other observation categories should be addressed in the report if they are kept in the analysis. These include the temporal head difference observations (td_lay1, td_lay2, etc.) and the lake (qlake01, qlake09) observations. Both are not matched very well in the calibration as shown in Figures R-1 and R-2. I recommend that the temporal head observations be removed from the calibration. PEST is not really estimating anything currently that would affect these observations. Recharge and ET are fixed and not estimated and the model is steady state so that storage is not estimated. I do not believe any other parameters in the PEST run can really do much to improve the temporal head match. One exception would be river conductance. I noted that 692 river cells have different conductances between 2001 and 2009. Were these estimated by PEST or where these assumptions imposed on the model?

I have been a bit confused on the purpose of the qlake observation types. Given the lack of fit, I am not sure that these are doing anything useful in the PEST run. I would appreciate hearing the District staff views on this.

I suggested in the last round of comments that a PEST run be made where ET is allowed to vary. I reiterate this suggestion and add that the wetting penalty observations should be give non-zero weights in this run. There are many wetting penalty observations that have very large residuals in the 004b PEST results. Currently ET is treated as a given in the model and I believe there is enough uncertainty in the calculation of ET by HSPF and the regional PET estimates that some adjustment is warranted. This should help reduce some of the large flooded cell issues in the model.

In looking at the parameters estimated by PEST, I noticed that 265 hit their maximum bound and 192 hit their lower bound. Most of these seemed to be vertical K pilot points. It may be useful to review the location of such points and see if the bounds could be adjusted further to allow more flexibility. In some areas, though, I also saw that vertical K is higher than horizontal K, primarily in layer 4. Figure R-3 shows these locations. I realize that layer 4 is an aquitard but having Kz > Kx is unusual.

Figure 135 shows the transmissivity of the UFA with posted APT results. There seems to be a disconnect, though, between APTs and T values in the highest transmissivity areas. All APTs in the zones with transmissivity greater than 1 million ft2/d are less than 380,000 ft2/d. Most APTs in this area are less than 100,000 ft2/d. Is this because the APTs were only partially penetrating the UFA? I think some explanation is needed for this discrepancy. If the APTs are not representative, then I suggest they be removed from the figures. There is also one APT over 1 million that is in an area with calibrated values in the 1,000 to 50,000 ft2/d range. Again, some explanation is required on this.

Specific Suggestions for Next Steps:

No-Pumping/Predevelopment Simulations

I tend to view this scenario as a qualitative type of simulation. We really do not know much that is quantitative about the predevelopment potentiometric surfaces and spring/river flows. Therefore, in my view you simply turn off the pumps and do a qualitative comparison between the "best guess" potentiometric surface from USGS to see any obvious discrepancies. Flooded cells have already been checked and do not seem to be an issue. The final check is on simulated spring flows and base flows. If those also show no obvious problems, then I would say the no-pumping/predevelopment scenario is done. I know that I may be in a minority on this but I do not believe that, given the lack of hard data, a huge effort is required.

Uncertainty Analysis Approach

I prefer the Null Space Monte Carlo style of uncertainty analysis. This is the most general type and will give the District the ability to assess the uncertainty in any prediction made with the model. It is a complex and time-consuming analysis so I would suggest leaving this until the model calibration is finalized so it only has to be performed once. Other types of uncertainty analysis are often either specific to a particular prediction or only give a measure of parameter sensitivity. Neither would be of much use from a practical standpoint.

2010 Verification Run

I would treat it as a calibration run, where the analysis of residuals for the various observation groups are presented in the same way as for the 2001 and 2009 steady state calibrations. Also, present the aspects of the model that change, such as recharge and ET distributions, and an analysis of changes in pumping.

Editorial Suggestions for Figures

- Graphs of observed vs simulated should have the X and Y axes be the same length. Having unequal axis length, makes it more difficult to interpret the degree of scatter around the ideal line.
- Graphs of observed vs simulated should be plotted for all types of observations used in the PEST run. Currently the temporal head difference targets and lake targets are not shown.
- Spring and river flux tables should show a percent error in addition to residual.

<u>Deficiencies Related to Task D2 Questions</u>

- There should be goals established for spring and river baseflow target types, just as in the head targets. For example, a certain percentage would be within +/- 10% error.

- I think some further analysis needs to be performed to see whether the qlake01 and qlake09 observation types are really providing useful information to the calibration. The matches achieved currently seem to indicate that these observations are not matched very well and may not be achieving their goals.

- Likewise, I think the same sort of assessment needs to be made for the td_lay1 through td_lay7 observations groups for temporal changes in head. Model parameters that are most likely to influence these observations are recharge, ET, and storage/specific yield. None of these parameters are being estimated. I think these observations, which account for a significant proportion of the measurement objective function, cannot be matched in the current calibration strategy.

It is not clear how item 3H in Task D2 is going to be addressed if Null Space Monte Carlo (NSMC) is not undertaken. From the presentation on June 21, staff indicated that NSMC may not be performed.

Table R-1. Percent Error of Spring Flux Targets for Springs > 200 cfs

Observation	Measured	Calculated	Residual	Percent Error
qs09_2320500	-6,367.2	-4,607.4	-1,759.8	27.6%
qs09_2323500	-6,014.3	-6,219.6	205.3	-3.4%
qs01_2323500	-4,296.2	-5,186.3	890.1	-20.7%
qs01_2320500	-3,568.0	-3,797.0	229.0	-6.4%
qs09_2330000	-1,477.2	-1,963.2	486.0	-32.9%
qs09_2228000	-1,298.1	-973.4	-324.7	25.0%
qs09_2317620	-1,260.3	-1,180.3	-80.1	6.4%
qs09_2319000	-1,145.6	-1,087.0	-58.6	5.1%
qs01_2330000	-957.9	-1,292.5	334.6	-34.9%
qs09_2315500	-924.9	-261.4	-663.5	71.7%
qs09_2322500	-869.2	-764.4	-104.8	12.1%
qs01_2319000	-861.5	-862.1	0.6	-0.1%
qr09_lsf_sprgrp	-794.0	-740.6	-53.4	6.7%
qs01_2317620	-752.3	-748.9	-3.4	0.5%
qspring09_587	-712.0	-709.2	-2.8	0.4%
qr09_2322500s	-630.0	-594.2	-35.8	5.7%
qr01_lsf_sprgrp	-611.0	-665.6	54.6	-8.9%
qs01_2322500	-607.9	-609.8	1.8	-0.3%
qs01_2228000	-587.4	-699.6	112.2	-19.1%
qr09_2313100	-558.0	-562.0	4.0	-0.7%
qr09_wacissa_sprgrp	-529.0	-524.2	-4.8	0.9%
qr01_2313100	-517.0	-513.3	-3.7	0.7%
qr09_silver_sprgrp	-501.0	-511.3	10.3	-2.0%
qr01_2322500s	-482.0	-523.6	41.6	-8.6%
qr09_crystal_sprgrp	-467.0	-461.0	-6.0	1.3%
qspring09_12341	-451.0	-437.1	-13.9	3.1%
qr01_silver_sprgrp	-445.0	-427.0	-18.0	4.0%
qspring09_n011235002	-435.5	-417.2	-18.3	4.2%
qspring01_n011235002	-418.8	-431.9	13.1	-3.1%
qr01_crystal_sprgrp	-409.0	-414.3	5.3	-1.3%
qspring09_7943	-404.0	-122.4	-281.6	69.7%
qspring01_7943	-386.0	-103.8	-282.2	73.1%
qs01_2315500	-302.3	-237.9	-64.5	21.3%
qr01_wacissa_sprgrp	-272.0	-288.3	16.3	-6.0%
qs09_2231000	-261.3	-192.6	-68.7	26.3%
qr09_iche_sprgrp	-254.0	-250.5	-3.5	1.4%
qr01_iche_sprgrp	-207.0	-195.3	-11.7	5.6%

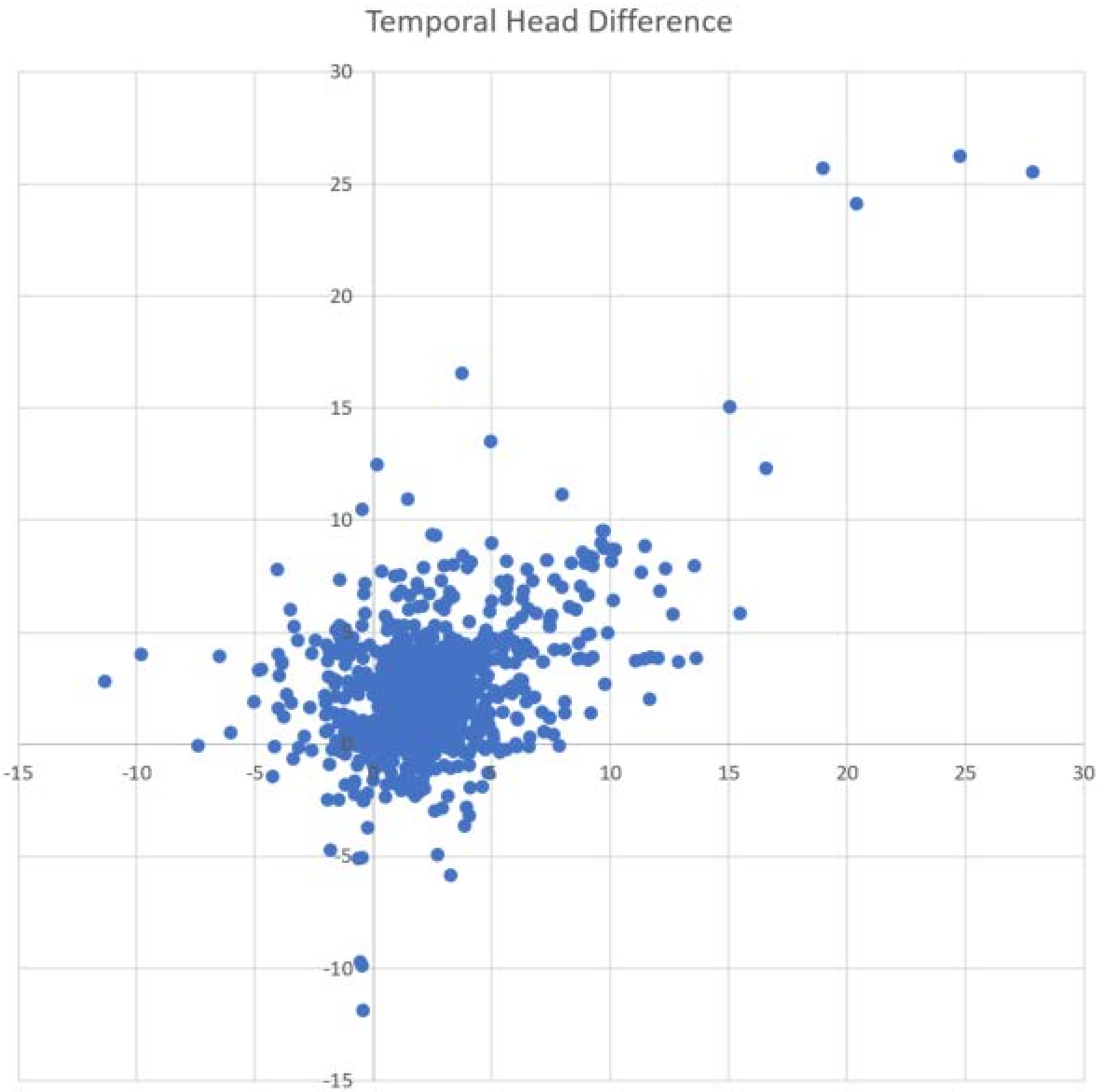

Figure R-1. Observed vs. Simulated Temporal Head Differences (004b).

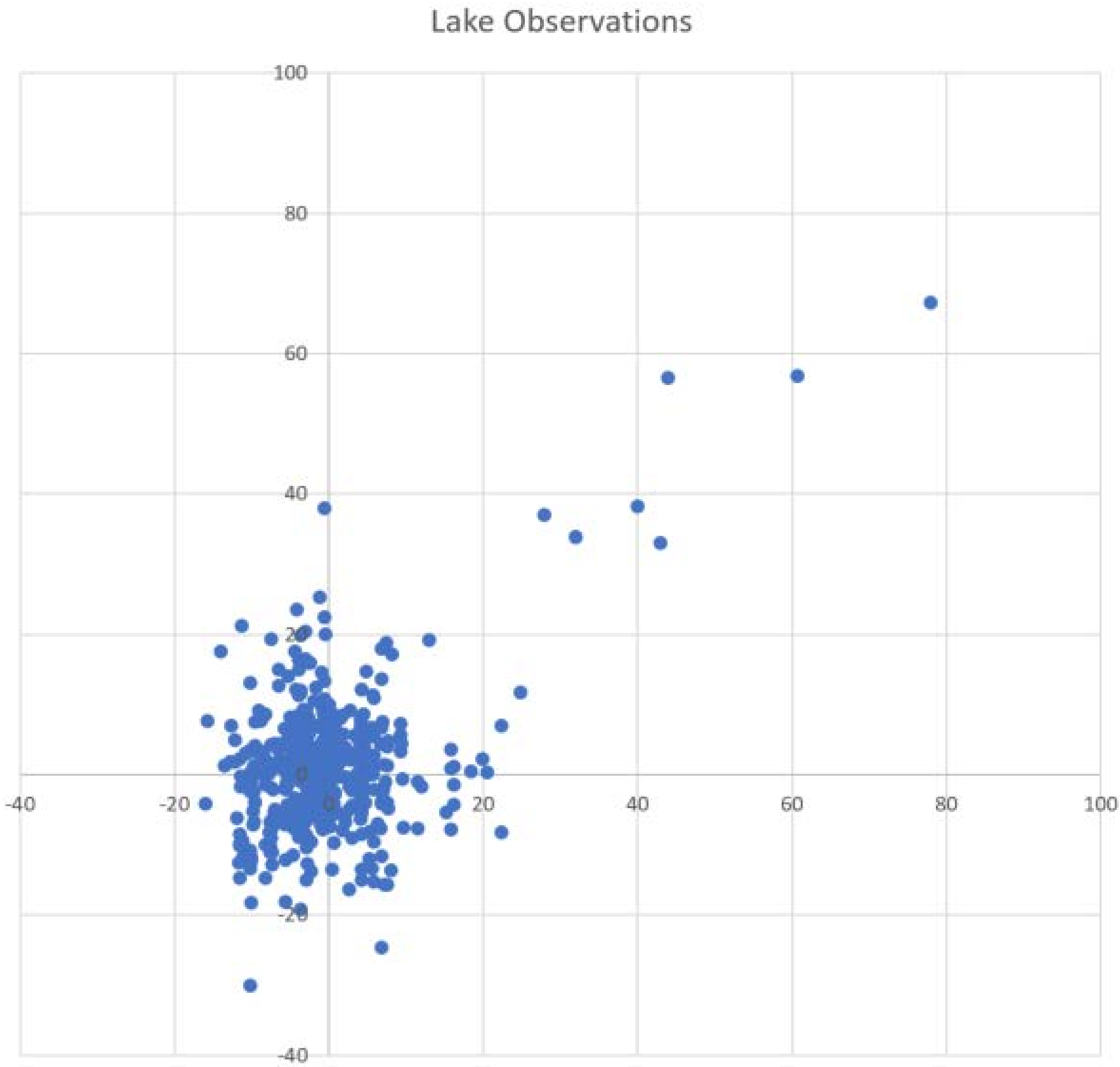

Figure R-2. Observed vs. Simulated Lake (qlake01/09) Observations.

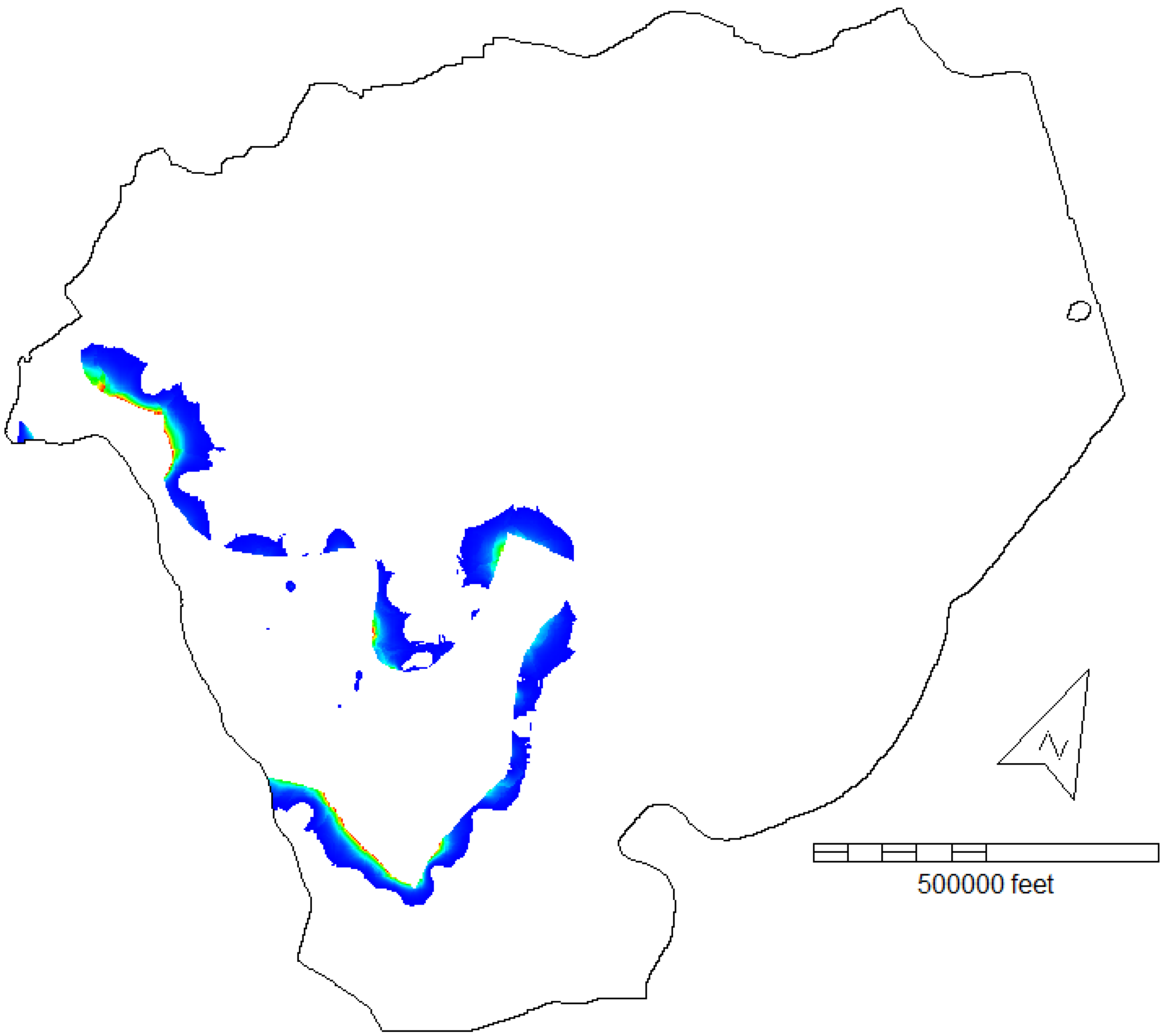

Figure R-3. Areas in Layer 4 where Kz is Greater than Kx

5. MODFLOW CHANGES (DANN YOBBI, P.G.)

TheCase004b model was presented to the peer review panel at a meeting on June 21, 2017 in Palatka, Florida. Relevant comments/questions provided below are numerically organized into the 4 categories as requested by the District.

Four major tasks comprise the peer review process. Each task includes specified deliverables and stakeholder, District, and reviewer participation meetings.

Task A. NFSEG V1.0 Review
Task B. Phase 1 Draft NFSEG v1.1 Model
Task C. Phase 2 Review
Task D. Final NFSEG v1.1 Model and Documentation

Specific Suggestions for Model Improvement

a. Base flow

Reasonable base-flow flux targets are critical for the success of this model as a useful water management tool. The data and methodology reviewed by Hal Davis and me, determined that current targets need to be reevaluated/reviewed prior to recommencement of numerical model simulations.

While it is unclear as to the methodology used to estimate base flow by the District; base flow estimates at 4 selected sites were reviewed by Hal Davis, peer reviewer. The base flows at Alapaha at Statenville (02317500), Suwannee at Brandford (02320500), Withlacoochee at Pinetta (02319000), and Withlacoochee at Quitman (02318500) gaging stations were significantly higher than field observations. This finding may imply that base flows assigned in the model are too high. Additionally, table 10 shows assigned model weights for these sites as 0.00 or 0.01 suggesting extremely limited influence on the objective function. Causation for the low weight assignment is unclear and should be discussed.

The scientific defensibility of base flow estimates presented is unproven due to (1) the inconsistent methodology used to estimate base-flow among the simulated sites; and (2) selected methodology to estimate base flow at each site is unidentified.

The draft report states that stream base-flow estimates, assigned in the NFSEG model, were quantified using 1 of 3 possible methods. Method 1 assigns base flow equivalent to stream flow at a gaging station or the change in flow between gaging stations. Method 2 assigns base flow using the U.S. Geological Survey public domain stream flow separation program, PART (Rutledge, 1993). Stream flow separation techniques partition base flow from total stream flow rate at a given gage. Method 3 adjusts field observations of stream flow at gaging stations using HSPF model derived ratios of simulated base flows to total stream flows.

Improving defensibility of the base flow estimates requires inclusion of the following discussions: (1) the methodology for each site needs to be identified; (2) why a single consistent methodology was not or could not be applied to all sites; and (3) document the calibration criterion.

To enhance my review, I consulted several USGS base flow experts and reviewed their study results. Halford and Mayer (2000) question the accuracy of recession-curve (separation) methods ‘‘because the major assumptions of the methods are commonly and grossly violated.’’ It is suggested when separation methods are employed, multiple methods of estimation are recommended because the accuracy of the results from a single method is difficult to assess. An alternative to separation methods are constructing exceedance plots (figure 1). These plots likely will provide a more reasonable estimate than hydrograph separation methods to estimate ground-water discharge to streams in the NFSEG model domain. A window of exceedance plot between 60 and 80 percent is suggested, where residuals are 0 in the window (see table 5 and figure 50 in Spechler and Halford, 2001). Kuniansky (1989) used the 60-percent duration of flow for creeks and rivers around Baton Rouge. Based on the findings of these studies in the southeastern US; I suggest that a comparison of base flow be evaluated using the NFSEG 2001 and 2009 base-flow values with base flows calculated using the exceedance methods.

To further strengthen support for the base flows being simulated, the NFSEG base-flow values should be compared to the current USGS Floridan Aquifer Ground Water Availability model (https://fl.water.usgs.gov/floridan/)? Base flows in the USGS model were estimated using the USGS GW Toolbox (https://water.usgs.gov/ogw/gwtoolbox/) programs. The GW Toolbox outputs base-flow estimates using a variety of methods and the mean of all methods is their base-flow flux target.

b. MODFLOW-USG

Will the District use MODFLOW-USG? Below is Jim Rumbaugh’s recommendation copied from the May 1, 2017 technical memo.

“I would like to see the District evaluate the use of MODFLOW-USG for NFSEG instead of MODFLOW-NWT. MODFLOW-USG has the same basic capabilities of NWT but also has some enhanced capabilities that make future use of NFSEG much easier and more appropriate. These new features include:

- *Use of nested grids within the parent regional model allow local refinement while retaining the regional nature of the simulation. This significantly reduces the issue of boundary effects on a local predictive simulation.*
- *Nested grids can have more layers than the parent regional model in areas where vertical gradients or local lithologic changes are important.*
- *MODFLOW-USG can eliminate inactive and pinched out cells. This reduces the memory requirements of the simulation and speeds up the simulation.*
- *MODFLOW-USG (beta version) can simulate turbulent flow in circular conduits using the Connected Linear Network (CLN) Package. Turbulent flow approximations include Manning, Darcy-Weisbach, and Hazen-Williams. CLNs can be used to simulate surface streams, subsurface conduits, and discrete fracture networks. In areas of known karst features, CLNs can be embedded in the regional model to account for enhanced flow in these areas.”*

Specific Suggestions for Next Steps:

No-pumping/predevelopment simulation

The District Staffs and stakeholders should take the lead for this task followed by an independent review of recommendations by the peer panel. An encouraging outcome presented at the meeting was the absence of widespread land-surface flooding due to simulated rejected recharge from the SAS during the “pumps off” simulation.

Uncertainty analysis approach

Good examples for presenting parameter and prediction uncertainty analyses are found in reports by Sepulveda and others (2012 p. 118, 122) and Sepulveda and Dougherty (2014). These publications document uncertainty for the east-central Florida ground-water flow model. Similar analyses would benefit the NFSEG model.

2010 verification run-what to show

A good example for showing a model’s ability to represent water-level change for the 1995 calibration and 2010 verification models is presented for the NDM5 in Anderson and Stewart (2016, section 4.0). A similar analysis is recommended for NFSEG including but not limited to:

- Comparison of the water balance data for the steady state and verification models.
- Plots of observed steady state and verification model heads.
- Plots of steady state calibration and verification base flow error.
- Plots of steady state calibration and verification spring flow error.

Editorial suggestions on figures

a. Is there anything missing?

- Water budgets for ground-water basins.
- Scatter plots of estimated transmissivity and simulated transmissivity.
- Representation of model layer/stratigraphic units.

b. Suggestions for improving

- A well-defined statement of the “intended application of the numerical model” and clear justification of assigned weights for flows, heads, parameters, etc. are needed.
- See comments on specific figures, tables, and statistics previously provided to the District in my May review write-up.

Task D2 Questions

3C--An appropriate set of summary statistics, plots, and maps have not been presented.

3D--Calibrated parameters have not been appropriately described.

3H--Uncertainty of key model parameters and predictions have not been assessed.

References

Anderson, P. and Stewart. M., 2016, Peer Review of the Northern District Model Version 5 and Predictive Simulations October 10, 2016--Final Report. In files of SJRWMD and SWFWMD, Palatka and Brooksville, FL.

Halford, K.J. and Mayer, G.C., 2000, Problems associated with estimating ground water discharge and recharge from stream-discharge records: Ground Water, v. 38, no. 3, p. 331-342.

Kuniansky, E. L., 1989, Geohydrology and simulation of ground-water flow in the "400-foot", "600-foot"' and adjacent aquifers, Baton Rouge area, Louisiana: State of Louisiana Department of Transportation TR 49.

Rutledge, A.T., 1993, Computer programs for describing the recession of ground-water discharge and for estimating mean ground-water recharge and discharge from streamflow records: USGS WRIR 93-4121.

Sepúlveda, N. and Doherty, J. (2015), Uncertainty Analysis of a Groundwater Flow Model in East-Central Florida. Groundwater, 53: 464–474. doi:10.1111/gwat.12232

Sepúlveda, Nicasio, Tiedeman, C.R., O'Reilly, A.M., Davis, J.B., and Burger, Patrick, 2012, Groundwater flow and water budget in the surficial and Floridan aquifer systems in east-central Florida: U.S. Geological Survey Scientific Investigations Report 2012–5161

Spechler, R.M. and Halford, K.J., 2001, Hydrogeology, water quality, and simulated effects of ground-water withdrawals from the Floridan aquifer system, Seminole County and vicinity, Florda: USGS WRIR 01-4182

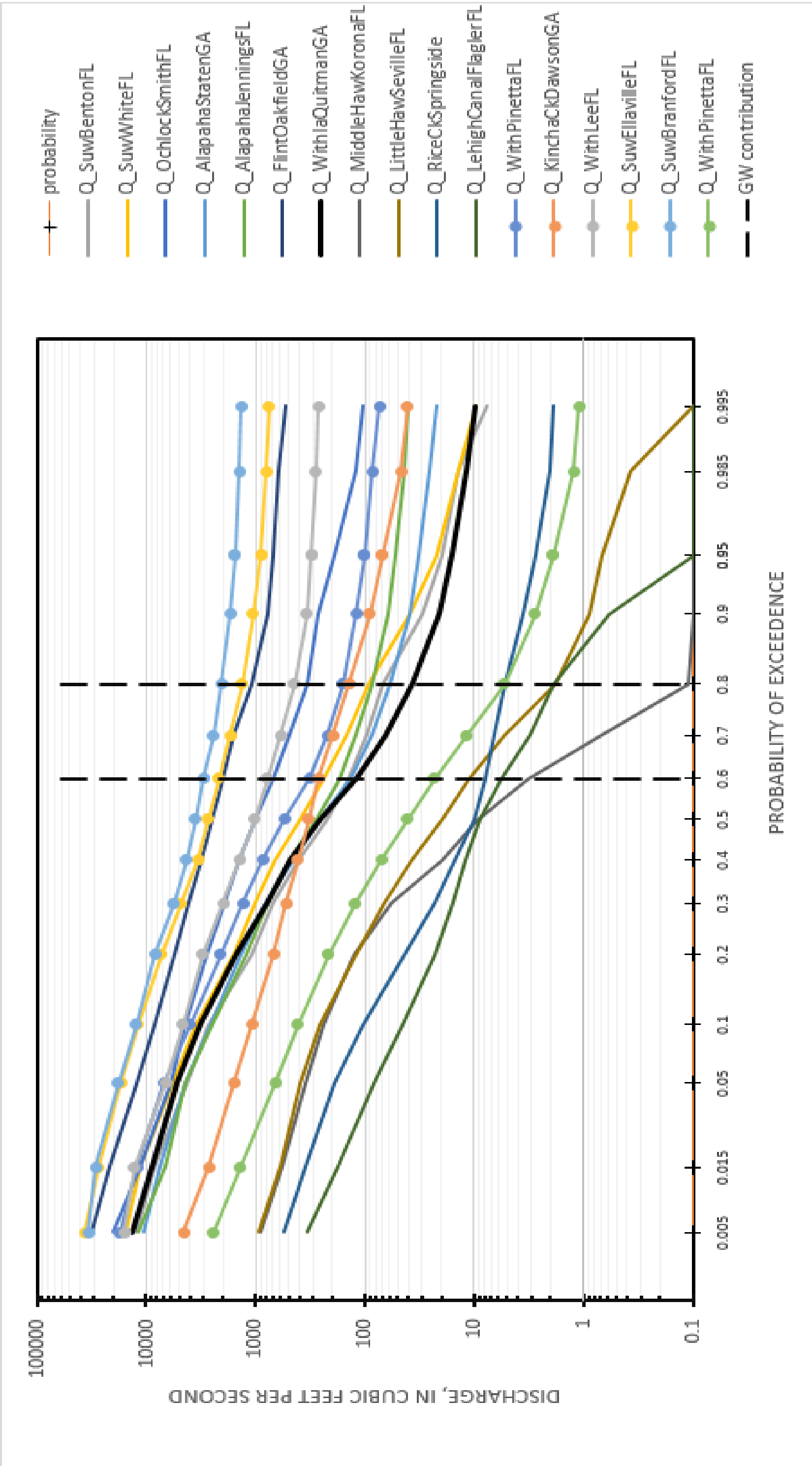

Figure 1. Duration of daily flow 1995-2010 at selected USGS stations in the NFSEG model domain.