Task A.4. Final Initial Recommendations

Technical Memorandum

Prepared By 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

May 1, 2017

TABLE OF CONTENTS

LIST OF FIGURES

1.0 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 task (Task A) of its scope of work. This effort has consisted of reviewing applicable documents and background materials (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 this technical memorandum (Task A.4), which contains the panel's final initial recommendations for changes and modifications to MODFLOW and HSPF. The panel's recommendations are grouped into recommendations for changes to Version 1.1 of the NFSEG model that can be completed by July 1, 2017 (Phase 1) and changes that can be considered later for Phase 2 or for future updates.

2.0 SUMMARY OF KEY FINDINGS AND RECOMMENDATIONS

2.1 HSPF Changes (Brian Bicknell)

2.1.1 Phase 1: Changes to NFSEG Version 1.1 Groundwater Model That Can Be Completed by July 1, 2017

- All gages should be calibrated if sufficient data are available
- The objective function should increase the weighting of the total overall flow and the flow frequency relative to other measures
- Areal recharge should be less discontinuous at watershed boundaries; provide
- Provide details of the overall recharge computation from the land use category recharge
- In the documentation, include more detail of the PEST calibration and objective function, including components and their initial and revised weights

2.1.2 Phase 2: Changes to NFSEG Version 1.1 Groundwater Model To Be Considered for Phase 2 or Future Updates

- Parameters should be constrained to be similar in adjacent watersheds for the same land use category
- Water balance should be reviewed to ensure it is reasonable for all land uses; includes inputs (rainfall, irrigation), components of runoff, recharge, and components of ET
- Include tables of the hydrologic parameter values for all watersheds and land use categories in an appendix
- Compute and include tables of additional statistics, such as percent bias and % differences at high and low flows

2.2 MODFLOW Changes (J. Hal Davis)

There were several particularity good choices made when developing the NFSEG model, which were:

1. Employing a technical team and steering committee in the early stages of the model development process.
2. The lateral model boundary conditions were (whenever possible) extended to the groundwater flow system boundaries or to reasonable no-flow boundaries. Because of this, for the intended use of this model, the boundaries should never need to be changed. Thus in future revisions the only upgrading will probably need to be to internal processes.
3. The grid size appears to be as small as feasible given the limits of computer power. Although this results in longer run times it does give greater resolution to the hydrologic processes.

4. Using the HSPF program and methodology allowed baseflows to be determined in a way that incorporated site-specific recharge and ET across the model domain.

Items that still need to be addressed are:

1. Cross sections should be added showing geologic layering and model layering. There should be at least one cross section orientated down gradient and one cross gradient.
2. A discussion of the final model water budget should be added.
3. Potentiometric surface maps for the UFA for years 2001 and 2009, using the heads collected for calibration, should be added.
4. A map should be added showing the MOFFLOW recharge rates to the UFA (the rate crossing the top of the cells that represent the UFA).
5. The recharge rates shown in figure 2-29 seem counter intuitive. The recharge rates are highest in the northwest portion of the model domain where there is a dense stream network. In contrast, in the part of the model domain centered around the Suwannee River, there are very few streams due to the high infiltration rates, but the assigned recharge rates are some of the lowest. This needs further discussion in the report.
6. A map showing all the streamflow and spring flow calibration points, with estimated baseflows is needed to show the data that the model is calibrating to.
7. The discussion of the PEST calibration needs to be expanded. Since PEST requires the modeler to make decisions on how the objective function is formulated and what additional data is used to modify the objective function, it would be helpful to understand the logic of how PEST was applied.

2.3 MODFLOW Changes (Louis Motz)

2.3.1 Comments for Version 1.1 of the NFSEG Groundwater Model That Should Be Considered for Phase 1

- Model layers do not coincide everywhere with hydrogeologic units, particularly in layers 1, 2, and 3: this concept is confusing and whether it needs to be revised needs to be discussed further;

- A brief discussion needs to be provided in the model documentation report (NFSEG V1.0) concerning how the regional-scale NFSEG groundwater model will be used to provide boundary conditions for the sub-regional scale Keystone Heights transient groundwater model, which is currently under development by SJRWMD.

- Suggested editorial changes in the model documentation report are indicated, including the need for additional documentation and referencing.

2.3.2 Additions to Version 1.1 of the NFSEG Groundwater Model That Should Be Considered for Phase 2 or Future Updates

- The North Florida Unstructured Grid (NFUSG) Model project should be continued in order to represent explicitly selected conduit systems for which adequate data including cave maps are available.

- A verification scenario for the NFSEG groundwater model for hydrologic conditions different from 2001 and 2009 should be considered.

- "Pumps-off" scenarios for the NFSEG groundwater model should be investigated to assist in the assessment of Minimum Flows and Levels (MFL's) in the model area.

- Consideration should be given to making sure that model layers coincide with hydrogeologic and lithologic units.

2.4 MODFLOW Changes (James Rumbaugh, Environmental Simulations, Inc.)

2.4.1 Phase 1: Changes to NFSEG Version 1.1 Groundwater Model That Can Be Completed by July 1, 2017

- Make a PEST run where ET is an adjustable parameter to see how this affects estimation of Kz values and location of flooded cells.
- Make one MODFLOW run with pumps off. Present results to peer review panel but do not include in version 1.1 documentation.
- Review observation weights for consistency and document in v1.1 report. Report calibration statistics with and without weights in v1.1 report.

2.4.2 Phase 2: Changes to NFSEG Version 1.1 Groundwater Model To Be Considered for Phase 2 or Future Updates

- Evaluate the use of MODFLOW-USG using v1.1 calibration.
- Create verification run using 2010 data or other time period if 2010 is not significantly different from the current calibration periods.

2.5 Summary of Key Findings (Dann Yobbi, P.G.)

- Supporting documentation for the model, NFSEG v1.0 is inadequate and requires revision;

- The report does not provide the data and discussion (statistics, uncertainty, etc.) needed to confidently assess the ability of the model to adequately simulate the hydrologic system;

- Additional figures, tables, and text are needed to adequately assess model results, strengthen technical defensibility, and provide needed documentation.

3.0 PEER REVIEW REPORTS BY PANEL MEMBERS

3.1 HSPF Changes (Brian Bicknell)

3.1.1 Phase 1: Changes to NFSEG Version 1.1 MODFLOW Model That Can Be Completed by July 1, 2017

In the model documentation, it was implied that one of the reasons for not calibrating at some USGS stations was that they were not located in the watershed delineation correctly. There was no further information on the number of stations that were in this category. We recommend that the segmentation be adjusted slightly to accommodate any stations that have sufficient data available, and that the calibration be extended to these stations.

The model documentation includes a brief description of the use of PEST to calibrate the HSPF models, and it includes a list of eleven comparisons/statistics that are optimized in the objective function. Since this is a critical part of the model calibration process, we recommend that the documentation include more detail of the PEST calibration and the objective function, including more detailed descriptions of the components and their initial and revised weights.

There is an implication that the eleven listed components of the objective function were weighted equally. We recommend that an investigation be made of the effect of increasing the weighting of the total overall flow and the frequency distribution curve. It has been our experience that these two metrics are very effective in achieving a good quality hydrologic calibration. Since many of the Percent Differences in mean monthly flow shown in Table 15 of the model documentation are high, this might improve the calibration at some of these stations.

The computation and averaging of simulated recharge from the HSPF output and linkage with the groundwater model grid is a relatively tedious and complex procedure. It involves: 1) outputting the recharge components from the HSPF results for multiple land use categories and multiple watershed segments, 2) overlaying the land use category and HSPF segmentation map with the groundwater model grid, and 3) computing the weighted average over each grid cell. There was no detailed description of this process in the HSPF model documentation. We recommend that details of the recharge computation and linkage of the HSPF model to the groundwater model be described and presented graphically.

Ideally, the separate HSPF models would generate consistent recharge estimates; however, some evidence indicates that there are significant discontinuities in the computed recharge at watershed boundaries. We recommend that the calibration be adjusted to minimize these discontinuities, or at a minimum describe the methodology that is used to smooth the recharge for input to the groundwater model. We also recommend that the documentation include areal recharge displays (maps) for each land use category and the overall recharge to verify that it is relatively continuous.

3.1.2 Phase 2: Changes to NFSEG Version 1.1 MODFLOW Model To Be Considered for Phase 2 or Future Updates

Separate calibrations of the watershed model are performed at each gage, and this can often lead to very different parameter values for land areas that are very close to each other and are the same land cover/use category. The model would be more defensible if these types of discrepancies were minimized. In the model documentation, it was stated that: *"Regularization of parameters between watersheds using PEST is not planned, but a manual review and adjustment of parameter ranges was made to ensure that adjacent watersheds have similar parameter values."* Since the documentation does not contain sufficient information to judge the effectiveness of this regularization, we recommend that either an exhaustive listing of the parameter values be generated, or the regularization should be included in PEST, if possible. In any event, we recommend the development of an appendix to the final model report that contains tables of the hydrologic parameter values for all watersheds and land use categories.

A detailed watershed hydrologic model is an attempt to represent a conceptual model of the region. Therefore, the water balance components (input and simulated) should be reviewed for reasonableness. This involves computing model results for individual land uses, model segments, and for each HUC8 watershed, for the following water balance components:

- Precipitation (and irrigation)
- Total Runoff (sum of following three components)
- Overland flow runoff
- Interflow runoff
- Baseflow runoff
- Potential Evapotranspiration
- Total Actual Evapotranspiration (sum of following five components)
- Interception ET
- Upper zone ET
- Lower zone ET
- Baseflow ET
- Active groundwater ET
- Deep Groundwater Recharge/Losses

Although observed values are not available for most of these water balance components, the average annual values must be consistent with expected values for the region, as impacted by the individual land use categories. This is a separate consistency, or reality, check with data independent of the modeling (except for precipitation) to ensure that land use categories and the overall water balance reflect local conditions. One of the most important benefits of this task is that it will allow a separate check to ensure that the rainfall (and water use/application data), which are two of the primary driving force inputs to the model are correct – or at least reasonable. The software package HSPEXP+ contains a water balance summary output that would facilitate this task.

There are many model calibrations in the watershed model, and judging the calibration quality requires review of many graphical and statistical summaries for these models that indicate agreement at various flow regimes and over various time intervals (e.g., monthly, annual, daily). The standard HSPF calibration procedure, known as a "weight of evidence" approach embodies the concept that multiple statistics (and graphical) comparisons should be made to assess the performance of a model or quality of a calibration. We recommend that a larger set of calibration statistics be generated and compiled in an appendix. They will generally include some or all of the following for each calibration station:

Statistical tests:

- 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 (e.g., correlation coefficient, Nash Sutcliffe efficiency coefficient)

3.2 MODFLOW Changes (J. Hal Davis)

<u>3.2.1 Phase 1: Changes to NFSEG Version 1.0 Groundwater Model That Can Be Completed by July 1, 2017</u>

3.2.1.1 Miscellaneous Comments

I don't think that any single item listed would be hard to complete by July 1, but cumulatively they probably would be. Thus any items not completed by July 1 could be carried over.

8. Figure 2-1: If possible, the model grid should be added (tick marks along both axis would be fine).
9. Page 3; p1: The Model Code Selection section needs to include more discussion of the pros and cons of the selected model code. The effect of dissolution caves and high conductivity zones in the limestones needs to be discussed. The presence of the Falmouth cave system maps indicates extensive karst development.
10. Cross sections should be added showing geologic layering and model layering. There should be at least one cross section orientated down gradient and one cross gradient.
11. Page 38: p3: The vertical accuracy of the 3DEP data should be given.
12. Page 65: Potentiometric surface maps for the UFA for years 2001 and 2009, using the heads collected for calibration, should be added.
13. A discussion of the final model water budget should be added.
14. Simulated potentiometric maps for model layers 5, 6, and 7 should be added.

15. Figures 3-46, 3-47, 3-48, and 3-49: Any known aquifer test data should be added to the maps.

3.2.1.2 Comments on Recharge

Recharge to the NFSEG model is probably the most important part of the flow system to estimate accurately and is also probably one of hardest to determine accurately. As stated above the process using HFSP code and technique appears thorough and well thought out, but some additional discussion is needed to clarify the results of the recharge estimation and application to MODFLOW. Since the recharge is set (not a PEST estimation parameter) any difference in the estimated recharge rate from the actual recharge rate will be "made up for" in the PEST estimated conductivities making it especially important to understand.

Comments are listed as follows:

16. The report does a good job of walking the reader through the processes of how the recharge and ET data sets were created but does not show some of the important results of the creation process, or how they were incorporated into the MODFLOW model. The maps showing the assigned recharge rates (Figures 2-28 and 2-29), the assigned maximum saturated ET rates (Figures 2-30 and 2-31), and the assigned ET extinction depths (Figure 2-32) are helpful. But the result of this, the recharge rate to layer 1 of
17. MODFLOW is not given, so it is not clear what the relative recharge rates across the model domain are. A map needs to be added showing the recharge rates that are applied to layer 1 of MODFLOW.
18. A map also needs to be added showing the MOFFLOW recharge rates to the UFA (the rate crossing the top of the cells that represent the UFA).
19. The recharge rates shown in figure 2-29 seem counter intuitive. The recharge rates are highest in the northwest portion of the model domain where there is a dense stream network (figure shown below); it seems probable that this area should have high runoff rates and low recharge rates. In contrast, in the part of the model domain centered around the Suwannee River, there are very few streams due to the high infiltration rates, but the assigned recharge rates are some of the lowest. This needs further discussion in the report.
20. Since ET is used in the HFSP model and in MOFLOW, an explanation should be given to show that is this accomplished without double counting ET.

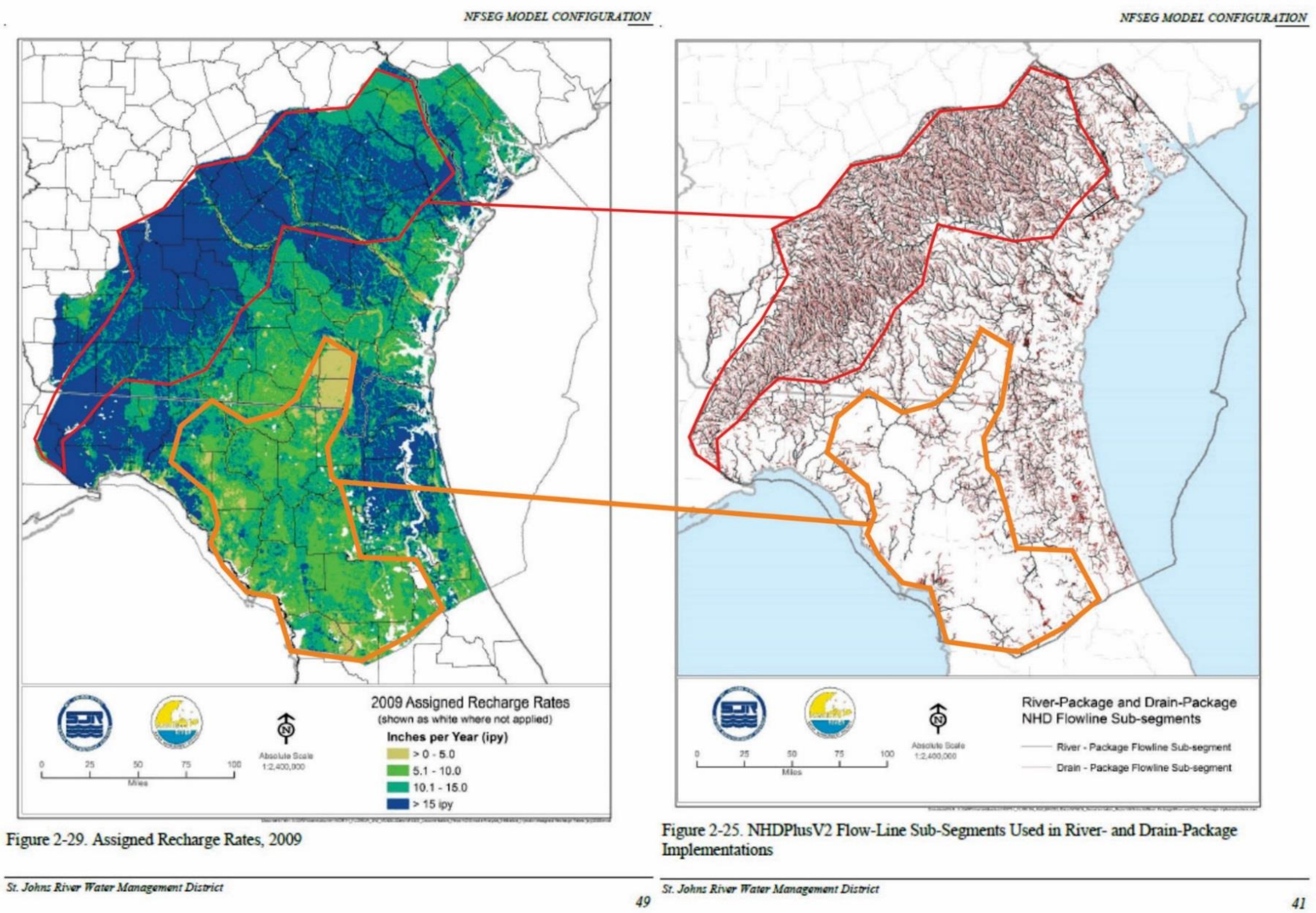

Figure 1. Assigned Recharge Rates, 2009

Figure 2. NHDPlus V2 Flow-Line Sub-Segments Used in River- and Drain-Package Implementation

3.2.1.3 Comments on Streamflow Calibration Points

The baseflows and spring flows used for MODFLOW calibration are discussed in several parts of the report but I still do not have a clear idea of how may calibration points there are, spatially where they are located, and the flow rates determined.

21. A map showing all the streamflow and spring flow calibration points, with estimated baseflows is needed to show the data that the model is calibrating to. If there are too many flow spring flow calibration points to easily put on a map, then only use the higher flow springs.

22. Page 66; p4: The method of determining baseflow by "equating it to the observed total stream flow rate at a given gage or the change in total stream flow between gages" needs to explained. And the method described as "adjusting the observed, total stream flows with the ratio of HSPF-simulated baseflows and total flows" also needs to be explained.

3.2.1.4 Comments on Sensitivity Analysis

23. The sensitivity analysis described in the report is difficult to translates into how heads and flows will be affect across the model. An overview discussion of what the sensitivity analysis is expected to show would be helpful.
24. Appendix J, while thorough, was difficult to understand. An overview in the report would help clarify the meaning of the Appendix.

3.2.1.5 Comments on PEST Calibration

1. The discussion of the PEST calibration needs to be expanded. Since PEST requires the modeler to make decisions on how the objective function is formulated and what additional data is used to modify the objective function, it would be helpful to understand the logic of how PEST was applied.

<u>3.2.2 Phase 2: Changes to NFSEG Version 1.1 Groundwater Model To Be Considered for Phase 2 or Future Updates</u>

1. Consider making layer 1 the surficial only, layer 2 the Miocene only, thus the UFA to start at layer 3 (having model layers follow lithologic layers). The reason is that when looking at the recharge, permeability, etc., maps it is hard to know where the UFA is present or where the other lithologic layers occur since model layers can cross several lithology boundaries.

<u>3.2.3 Verification for Year 2010</u>

I looked at the hydrographs in the conceptual model report and seemed to me that the hydrological conditions occurring in year 2010 were close to those occurring in year 2009, at least for groundwater levels. Since NFSEG model is steady state, then, when used for regulatory purposes it will need to be set up using some set of hydrologic assumptions, which I assume will be low-flow conditions to cover a worst-case scenario. It might be better to find a year in which those conditions occurred and verify to that year.

3.3 Review of NFSEG Groundwater Model Version 1.1 (Louis Motz)

3.3.1 Review Comments for the Groundwater Model Documentation Report (NFSEG V1.0) That Should Be Considered for Completion by July 1, 2017 (Phase 1)

1. p. 2, Table 2-1:
Is any pumping from Layer 2 included in the model (e.g., in the Keystone Heights area)?

2. Principal Representation of NFSEG Model Layers (pp. 2 and 5-7 and Tables 2-1 and 2-2) and Modeled Distributions of Horizontal Hydraulic Conductivities, Transmissivities, Vertical Hydraulic Conductivities, and Leakances (Figures 3-46 – 3-56):
The representation of the NFSEG layers described in pp. 2 and 5-7 and Tables 2-1 and 2-2 is somewhat complicated and has the potential to be very confusing. The three uppermost model layers are said to correspond to hydrogeologic units (surficial aquifer system, intermediate aquifer system, and Upper Floridan Aquifer), except where they do not. How extensive areally in the model do these exceptions occur, i.e., how much of layer 1 is considered to be the surficial aquifer and how much of layer 1 is considered to be part of the intermediate confining unit, Floridan aquifer system, or a combination thereof? How much (areally) of model layer 2 is considered to be part of the intermediate confining unit and how much of model layer 2 is considered to part of the Floridan aquifer system and/or surficial aquifer system?

If model layers 1, 2, and 3 do not always correspond to hydrogeologic units but instead, cross hydrogeologic boundaries, how is this accounted for during calibration? Are different limits in ranges for horizontal and vertical hydraulic conductivities set for different areas of the model for layers 1, 2, and 3? When layer 1 is part of the intermediate confining unit, Floridan aquifer system, or a combination thereof, will it be possible to make sure it has the same hydraulic properties as model layers 2 and 3 that also represent the intermediate confining unit and/or the Floridan aquifer? Will it be possible to present calibration results in terms of transmissivities for the Upper Florida Aquifer in areas where model layers 1 and 2 along with model layer 3 comprise the Upper Floridan Aquifer?

The calibration results in Figures 3-46 – 3-56 are difficult to interpret. For example, in the horizontal hydraulic conductivity map shown in Figure 3-46, how much of model layer 1 is the surficial aquifer and how of layer 1 is part of the relatively low permeability intermediate aquifer system and how much is part of the higher permeability Floridan Aquifer System? In the horizontal hydraulic conductivity map shown in Figure 3-47, how much of layer 3 is part of the Floridan Aquifer System and how much is part of the lower permeability surficial and intermediate aquifer systems? In the transmissivity map in Figure 3-50, do these transmissivities represent the total transmissivity of the Upper Floridan Aquifer or is it necessary to add transmissivities from layer 1 in areas where layer 1 is considered to be the uppermost part of the Floridan Aquifer to the values shown in Figure

3-50? In the leakance map in Figure 3-55, how much of layer 2 is the relatively low permeability confining unit and how much of layer 2 is part of the much more permeable Floridan Aquifer System (for which leakance should not be calculated)?

3. Better documentation is needed for terms used in the report. For example, these terms need to be explained better in the text and/or referenced to previous reports:

p. 4: "Fall Line" and "Gulf Trough"

p. 7: "Florida-Hatteras Slope"

4. p. 28, lines 24-25: Along the coasts of the Atlantic Ocean and Gulf of Mexico, the active domain of model layer 1 is bounded by constant-head grid cells used to represent the equivalent freshwater head of the ocean.
How were the equivalent freshwater heads calculated? Please provide an equation and/or reference to the calculations. Also, 'specified-head" should be used instead of "constant-head".

5. p. 38, lines 11-13: River-Package Implementation
Discharge between the groundwater flow system and perennial streams and lakes are *[is]* **represented in the NFSEG groundwater model by implementation of the MODFLOW river package.**
Will the NFSEG regional model provide sufficient detail to assess MFL's for streams and lakes? Please provide a brief description of the Keystone Heights transient groundwater model currently under development by SJRWMD, including how the NFSEG model will be used to provide boundary conditions for the transient model, which will be used to assess MFL's for lakes Geneva and Brooklyn.

6. p. 54, Figure 2-34:
Is Brunswick, Georgia a major pumping source? Should it be located on Figure 2-34? Does saltwater upconing into the Upper Floridan Aquifer still occur at Brunswick (see Maslia and Prowell 1990)?

7. pp. 75-76, Figures 3-1 and 3-2:
1st magnitude springs should be labeled on these figures.

8. p. 67 and pp. 83-84, Figures 3-9 and 3-10:
Horizontal head differences…are used as a calibration metric. Consideration of the distance between two wells at which head differences are measured should also be included, and it is recommended that the horizontal head gradient (head difference divided by distance) be used as a calibration metric instead of just the horizontal head difference between two wells.

9. p. 93, Table 3-2. Calibration Statistics: Graphical representation such as bar graphs for the calibration statistics would improve the reader's understanding of the results. This applies to estimated water use and groundwater model water budgets (in inches/year) for 2001 and 2009 as well.

10. p. 94, Table 4-1: Typo? Should this be Table 3-3?

11. pp. 96-97, Model Capabilities and Limitations:

Model limitations in terms of unconfined/confined conditions in the transition zone in Alachua County and other areas should be discussed in this section.

p. 96, lines 11-12: ….[the] model may be capable of simulating changes in flows and heads with an accuracy that is comparable to or better than models currently used for planning or regulatory purposes.
To which models is the current model being compared and what calibration parameters are being considered? This statement needs to be documented, which requires specifying the models being compared and making a detailed comparison of residual means of errors, root mean squares of errors, and other calibration statistics. Is it necessary to include this statement? If so, then it needs to be documented with comparisons to specific models.

p. 97, lines 8-10: D. Enhanced calibration rigor obtained by matching water levels and flows to two calibration periods (calendar years 2001 and 2009) that represent significantly different hydrologic conditions;
What is significantly different between 2001 and 2009? Are the differences between 2001 and 2009 explicitly explained anywhere in the report?

p. 97, lines 14-16: G. Expanded availability of water-level data…through implementation of sophisticated statistical estimation techniques;
What "sophisticated statistical estimation techniques" were implemented? Is it necessary to make this claim? The estimation techniques need to listed explicitly; otherwise, this sentence should be eliminated.

p. 97, lines 19-23: I. Inclusion of additional calibration constraints not used in the development of many of the models currently used in Florida, including: vertical head difference differences adjacent aquifers,…and a more extensive set of stream baseflow 'pickup' and spring-group observations.
To which models is the current model being compared? This statement needs to be documented, which requires specifying the models being compared and making a detailed comparison of calibration constraints included in the current model with calibration constraints in "many of the models currently used in Florida". Is it necessary to include this statement? If so, then it needs to be documented with comparisons to specific models.

12. Pp. 132 – 135, 4.0 References: The entire list of references needs to be alphabetized. Currently, there are two lists (pp. 132-133 and pp. 134-135).

3.3.2 Additions to Version 1.1 of the NFSEG Groundwater Model That Should Be Considered for Phase 2 or Future Updates

1. The North Florida Unstructured Grid (NFUSG) Model project should be continued in order to represent explicitly selected conduit systems for which adequate data including cave maps are available.

2. A verification scenario for the NFSEG groundwater model for hydrologic conditions different from 2001 and 2009 should be considered.

3. "Pumps-off" scenarios for the NFSEG groundwater model should be investigated to assist in the assessment of Minimum Flows and Levels (MFL's) in the model area.

4. Consideration should be given to making sure that model layers coincide with hydrogeologic and lithologic units.

3.4 Review of NFSEG Groundwater Model Version 1.1
(James Rumbaugh, Environmental Simulations, Inc.)

3.4.1 Phase 1: Changes to NFSEG Version 1.1 MODFLOW Model That Can Be Completed by July 1, 2017

I suggest that the weighting scheme on observation groups be reviewed prior to finalizing the version 1.1 calibration. In reviewing the latest PEST results, there seem to be about 23 different observation groups comprising about 11 major observation types. The dry and wet penalty observations are contributing almost nothing to the objective function, while the "qs_spring*" group contributes very little (~4,600). The remaining 8 major categories all have a contribution of between 17,000 to 103,000.

I would like to see an explanation of what each observation group in the PEST control file represents and how the weighting was determined. Most groups seem to have only a few different weights (e.g. 0.0, 0.1, 5.0, 1.0). The qr*, qspring*, and qs* groups are highly variable. Especially important is the reasoning behind removal of many of the observations (weight = 0.0). When reporting the calibration statistics in the version 1.1 report, I would like to see them with and without weights.

It would be useful to see one additional PEST calibration with ET as an adjustable parameter (evtrmul* parameters). I would like to judge the use of ET in reducing flooded cells and on how Kz values are estimated. When presenting the results of this calibration run, please provide a comparison of flooded areas and the effect on vertical hydraulic conductivity.

I would like to see one model run with pumps off using the final calibration for version 1.1. We can discuss this topic endlessly but until we see the results of an actual run, it's hard to know where to go with this issue. I would recommend this not be in the version 1.1 report but simply presented to the panel. Focus on a qualitative comparison to our best guess as to the predevelopment potentiometric surface in the UFA and also the effect on flooding in layer 1.

3.4.2 Phase 2: Changes to NFSEG Version 1.1 MODFLOW Model To Be Considered for Phase 2 or Future Updates

I suggest saving the verification run until the version 2 process. Use of 2010 for verification is fine if it is not just a duplication of 2001 or 2009. If the recharge is about the same as 2001 or 2009 then I do not think it would be a good candidate for verification.

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.

3.5 Preliminary Peer Review of the Draft Report "Development and Calibration of the North Florida Southeast Georgia Groundwater Model (NFSEGv1.0) (Dann Yobbi, P.G.)

3.5.1 Background

The North Florida Southeast Georgia (NFSEG V1.0) is a numerical groundwater flow model that simulates groundwater flow in portions of Florida, Georgia, and South Carolina. NFSEG v1.0 is a planning level version used for support of the North Florida Regional Water Supply

Plan. The St. Johns River (SJRWMD) and Suwannee River (SRWMD) water management districts are completing the NFSEG model Version 1.1 that will be used for the establishment/assessment of minimum flows and levels (MFLs) and for regulatory use. Finalization of the NFSEG and associated HSPF models mandates convening a panel of modeling experts to peer review the final phase of the Version 1.1 model interactively during model development. This provides an opportunity for the modelers to incorporate the peer panel's review suggested changes prior to model completion.

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

3.5.2 Introduction

The document is my (Dann Yobbi) comments regarding TASK A and is primarily based on information presented in the NFSEG 1.0 draft model report "Development and Calibration of the North Florida Southeast Georgia Groundwater Model (NFSEG V1.0)" by Fatih Gorden, Douglas Durden, and Trey Grubbs, 2016. As background, I reviewed supporting appendices A-J and three other NFSEG documents including:

- North Florida Southeast Georgia (NFSEG) Groundwater Flow Model Conceptualization by Douglas Durden, Tim Cera, and Nathan Johnson
- Data availability for development of the North Florida Southeast Georgia (NFSEF) Regional Groundwater Flow Model in its Potential Dorman by Douglas Durden
- NFSEG V1.1 Improvements

3.5.3 Review Comments

Supporting documentation for the model, NFSEG v1.0 is inadequate and requires revision. The report does not provide the data and discussion (statistics, uncertainty, etc.) needed to confidently assess the ability of the model to adequately simulate the hydrologic system. Additional figures, tables, and text are needed to adequately assess model results, strengthen technical defensibility, and provide needed documentation. Specified items and potential remedies are divided into a phase 1 and phase 2 completion targets and are listed below:

3.5.3.1 Phase 1: Supporting Documentation To Be Completed by July 1, 2017

- **Statistics and Model Results—** The evaluation of model results is incomplete and requires more thorough discussion and analyses. For example, the report does not have any discussion on model error by layer, model-wide trends in the spatial distribution of head residuals for each aquifer, errors/residuals in baseflows or spring flows, or if residuals are randomly distributed. Do simulated values seem reasonable? How well does the simulated water budget match budget values determined independently? The following is a short list of information that is needed:

  1. *Water Budget*-- Water budgets including sources of input, output, and relevant hydrologic components for the years 2001 and 2009; and comparisons to the model simulated water budgets are needed (tables and figures). Good examples of water budget presentation are found in (1) USGS WRIR 02-4207 (See fig. 23 and explanations on pp. 39 and 42); and (2) SJ2006-4 (pp. 127-133).
  2. *River Baseflows*--Table, graphs, figures showing simulated and estimated baseflows (including % differences) for all stream gages used for calibration of the river baseflows in MODFLOW.
  3. *Spring Flows*--Tables, graphs, figures showing simulated and estimated flow (including % differences) for all springs. Also, summary statistics (POR, mean, accuracy, etc.).
  4. *Head Residual Grouped by Layer*--Maps of simulated WT/POT surfaces and residuals by layer.
  5. *Leakage*--Maps showing simulated leakage through confining layers, lakes, streams, drains, etc.

Additionally, the calibration statistics requested by Liquid Solutions Group should be provided to the peer review panel prior to the Phase 1 Review Meeting (TASK B.2) I agree with their suggestion and as a reminder of LSG's request I have attached their appendix S1 to this review comment memo.

- **Hydraulic Properties—**A discussion of the range of observed values from APTs, choice of initial input values, and comparison to final values (calibrated) is needed. Additionally, APT values need to be plotted on figures 3-46 through 3-56.
- **Parameter Uncertainty—**The relative uncertainty of parameter values is needed to assess model reliability and accuracy. Good examples of this material are found in Sepulveda and others (2012) on fig. 97 and pages 118 and 122.
- **Assignment of Weights—**A discussion and justification of the assignment of weights is needed. Please describe the method used to assign weights. Was a trail-in-error

approach or some other method used to assign weights? Good example of presentation of this type of information is found in Sepulveda and others (2012) pages 74-75 and table 12.

- **Sensitivity Analysis—**The sensitivity analysis discussion is inadequate for determining the relative importance of input parameters and boundary conditions on calibration results. This discussion is necessary as sensitivity analyses are instrumental when determining model accuracy and guide the modeler in estimating appropriate weights for calibration targets. A good example is provided in Sepulveda and others (2012). They describe a useful approach for assessing many of the same model parameters utilized by this model including lumping of sets of parameters using multipliers to calculate composite scaled sensitivities. See pages 82-83 and figure 56 for a good example. A similar procedure and discussion is recommended for this report.

3.5.3.2 Phase 2: Final NFSEF V1.1 Model and Documentation

- **Conceptual Model**—Clear connection (nexus) between model layers and conceptualized (simplified) hydrostratigraphic units is needed. Good presentations of this information is found in Durdin (2012) and in Williams (2006). The hydrogeologic cross sections (figs. 32-34) in the cited above Data Availability Report need to be extended into Florida both in the north-south and east-west directions. A similar conceptual model discussion and visual presentation shown in figs. 25 and 26 in Williams (2006) would benefit this report.
- **Model Limitations—** This section of the report should clearly state all limitations and assumptions of the model. However, this section of the report is incomplete. For example, the report does not provide a discussion of the model's limitation regarding simulation of a conduit network nor the appropriateness of using a uniform orthogonal grid. Nor does the report provide a discussion on the limitations on the grid spacing for accurately simulating small scale hydrologic features such as spring flow, river discharge, infiltration or discharge at stream cells, or near large discharge wells. Limitations on the model's predictive and computational capabilities are controlled by the scale of the sub-divisions and this needs to be thoroughly discussed. Omitting these discussions may mislead the users of this model to believe that this approach is the best or only approach.
- **Verification ("History Match") –**Model verification or "history matching" is an essential aspect of the model development process. Without this analysis, defensibility of the model regarding its future use supporting management decisions by the District is limited at best. I therefore recommend that the calibrated NFSEG v1.1 model be verified using a different set of historical hydrologic data than used during the calibration process. The District has identified the need for verification in Task 7 of the Technical Team Charter (10/27/2011; appendix B). Comparison between field

(measured) and simulated water level and flow data highlight strength and weakness in the model construction and provide greater confidence in the calibration and predictive capabilities of the model.

3.5.3.3 Phase 2 or Future Updates

- **Proposed "Pumps Off" Use of the Model—**A "pumps off" application of the final NFSEG model has been discussed during technical meetings to evaluate ground water withdrawals on heads, spring flows, and surface water flows. One suggestion is to eliminate all pumping (set to zero) and run the model to steady state. This scenario likely is not an appropriate application of the calibrated NFSEG V1.1 model based on similar model applications used to evaluate the effects of net ground water withdrawals on the flow at Silver Springs. Results of a peer review of the NDM version 5 model (Anderson and Stewart, 2016) reported the following: "*The method of setting all pumping to zero and running the model to steady state is the least reliable modeling method for estimating the effect of pumping on the flow of Silver Springs as the water table and potentiometric surface rise well above land surface at steady state*". Anderson and Stewart (2016) also discuss two other scenarios that may be viable for determining baseline (pre-development) conditions for comparison to future pumping scenarios: (1a) vary pumping rates by a specified percentage of plus or minus 50 percent or (1b) vary by actual pumping rates rather than percentages--then plot predicted flows against total pumpage and determine intercept; and (2) setting model transient run time (simulation time) under a "pumps off" condition to one or more years. These applications and possibly a longer transient model run should be investigated to determine the most reliable modeling application for evaluating the effect of ground water withdrawals on heads and surface water flows.

3.5.3.4 Miscellaneous Comments on the NFSEG V1.0 Report

1. Table of Contents is incomplete. Several 3rd and 5th order headings are omitted.
2. Add Acronyms and Additional Abbreviations
3. Table 3-3--no discussion of the content or reference to this table in report.
4. Add tables 4-1 and 4-2 to List of Tables.
5. Acronyms need to be avoided for table and figure titles.
6. There is a duplicate listing of some of the references.
7. Introduction—Make sure to state that the NFSEG v1.1 model "*will be used as a <u>tool</u>*" to support development of the NFSEG. Also, suggest including a figure and/or table and/or discussion on the synthesis of the layering schemes from local or regional models into this single regional model. (See Sepulveda (2002), pgs. 46-54;

and fig 1—Model Domains –In Appendix B-- North Florida Southeast Georgia Groundwater Model Development Project Task and Time Line Work Plan)

8. Model grid (discretization) should be shown.
9. Pg.28 states "constant-head" grid cells but figure 2-42 indicates "specified heads". Specified is the preferred terminology.
10. Pg. 38-39—why are metric units reported?
11. Pg. 64 p.1—Why are no sensitivity analyses reported for the preliminary testing? What are the reasonable ranges for the parameters?
12. Pg.64 p. 2—No results were presented so what statistical measure is meant by a "***high level*** *of consistency"* between simulated and observed values?
13. Pg. 64 PEST Calibration—Description of PEST calibration is inadequate and requires additional discussion and information not currently included. A good example of a well written and organized discussion of PEST calibration is presented in Sepulveda and others (2012) pgs. 55-118.
14. Pg.70 p1—Figures 11, 22, 13, and 14 are not found in the report.
15. Pg.70 p2—Figure 16 is not found in report.
16. Pg.70 p3—Figure 15 should be 3-15.

3.5.3.5 Miscellaneous Comments on NFSEG V1.0 Figures

1. Figures 2-2 through 2-20-- Very pretty, but illegible/unreadable. Need better contrast among colors. Delete color for offshore areas. Model boundary is not labeled. Use consistent figure labels. Also, why are these figures in landscape while remaining figures in portrait layout?
2. Figure 2-25—Not very useful at this scale.
3. Figures 2-28 through 2-31—need better contrasting colors. Delete (ipy)
4. Figure 2-36—Avoid acronyms from figure title. Spell out DSS.
5. Figure 3-2—Not very useful at this scale. Need to focus on Florida and use colors for different magnitudes—not circle size.
6. Figures 3-3 and 3-4—Can't distinguish between river/drain sub-segment. Need better color contrast.
7. Figure 3-7—Need consistent legend titles—Should read "Vertical Head Difference" not "Values VHB".
8. Figure 3-9—Use the term gradient instead of horizontal head differences.
9. Figures 3-19 through 3-21—Suggest contouring the residuals to clearly identify any patterns in values.
10. Figures 3-25 through 3-29—Need to label green cone.
11. Figure 3-30—This is not the vertical head residual plot—it is the same as figure 3-5.
12. Figure 3-31-- This is not the vertical head residual plot—it is the same as figure 3-7.

13. Figure 3-32-- This is not the vertical head residual plot—it is the same as figure 3-6
14. Figure 3-33-- This is not the vertical head residual plot—it is the same as figure 3-8.
15. Figure 3-35—Label green cone and blue solid and dashed lines.
16. Figures 3-36 and 3-37—See comment # 8.
17. Figures 3-39 through 3-41—See comment # 15.
18. Figure 3-43--Difficult to read. Data should be in a table rather than on this figure. Unclear why 2 bars shown for each spring.
19. Figure 3-44 and 3-45—See comment #15
20. Figure 3-46—Interval 0 – 10 should be <10.
21. Figures 3-47 through 3-56—Hard to read. Poor color contrast. Delete offshore blue color. Smallest interval should be <50, <1,000, <0.01, etc. Also, need to explain areas with no color.
22. Figure 3-57—Potentiometric contour around Keystone Heights is 90.

3.5.3.6 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.

Durden, Douglas, 2012, Data availability for development of the north Florida southeast Georgia (NFSEG) regional groundwater flow model, in the area of its potential domain: http://northfloridawater.com/pdfs/NFSEG/NFSEG_Data_Availability_Report.pdf

Knowles, L., Jr., O'Reilly, A.M., and Adamski, J.C., 2002, Hydrogeology and simulated effects of ground-water withdrawals from the Floridan aquifer system in Lake County and in the Ocala National Forest and vicinity, north-central Florida: U.S. Geological Survey Water-Resources Investigations Report 02–4207.

Sepúlveda, N., 2002, Simulation of ground-water flow in the intermediate and Floridan aquifer systems in peninsular Florida: U.S. Geological Survey Water-Resources Investigations Report 02–4009.

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.

Williams, S.A., 2006, Simulation and effects of groundwater withdrawals from the Floridan aquifer system in Volusia County and vicinity: SJRWMD Technical Publication SJ2006-4.

Appendix S1: Requested Statistics and Plots from Liquid Solutions Group

We understand the Districts are currently making modifications to the MODFLOW and HSPF models based on internal and external comments received. Therefore, we propose that the information/data below only be produced for the next version of the NFSEG Model released to the Technical Team, and not the versions that have been published to date.

1) **Nash-Sutcliffe Coefficients**: Summary table of the Nash-Sutcliffe coefficients for stream gages used for the calibrations of the HSPF models.
2) **Average Daily Flows**: Simulated and observed average daily flow for the stream gages used for calibration of the HSPF models (tabular and graphical) for the period-of-record, 2001, 2009, and 2010.
3) **Average Yearly Flows**: Simulated and observed average yearly flow (including percent difference for each year) for the stream gages used for calibration of the HSPF models (tabular and graphical) for each year in the period-of-record.
4) **Daily Cumulative Frequency Distributions**: Simulated and observed daily cumulative frequency distributions for the stream gages used for calibration of the HSPF models (tabular and graphical).
5) **River Baseflows**: A summary of the average yearly modeled baseflows for 2001 and 2009 for all stream gages used for calibration of the river baseflows in the MODFLOW model. An example table is provided below.

Stream Gage Number and Name	2001 Observed/ Calculated Baseflow (MGD)	2001 Simulated HSPF Baseflow (MGD)	2001 Simulated MODFLOW Baseflow (MGD)	2009 Observed/ Calculated Baseflow (MGD)	2009 Simulated HSPF Baseflow (MGD)	2009 Simulated MODFLOW Baseflow (MGD)
Gage 1						
Gage 2						
Etc.						

6) **Closed Basin Recharge**: A summary of simulated and observed (or estimated target) average yearly recharge to the Upper Floridan aquifer associated with closed basin recharge, sinks, or injection wells represented with a Special Action in HSPF for the period of record, 2001, 2009, and 2010 as summarized in the example table below. Where a calibration target (qualitative or quantitative) was available for this parameter (e.g., Orange Lake sink), please provide the data used to develop the calibration target.

Closed Basin Recharge, Sink or Injection ID and/or Name	POR Historical Recharge (MGD)	POR Simulated Recharge (MGD)	2001 Historical Recharge (MGD)	2001 HSPF Simulated Recharge (MGD)	2009 Historical Recharge (MGD)	2009 HSPF Simulated Recharge (MGD)	2010 Historical Recharge (MGD)	2010 HSPF Simulated Recharge (MGD)
Orange Lake Sink								
Sink 2								
Etc.								