TECHNICAL MEMORANDUM

TO: NFSEGv1.1 Technical Peer Review Panel: Louis H. Motz, Ph.D., P.E., D.WRE, Chair, Brian R. Bicknell, J. Hal Davis, P.G., James Rumbaugh, P.G., Dann Yobbi, P.G.

FROM: St. Johns River Water Management District, Palatka, Florida And Suwannee River Water Management District, Live Oak, Florida

RE: **Reconciliation** responses to Independent Technical Peer Review of the North Florida Southeast Georgia Groundwater Model (NFSEG v1.1) August 22, 2018

DATE: January 25, 2019

BACKGROUND

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.

Version 1.0 of the NFSEG groundwater model and the HSPF-derived surface-water models was 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 Florida and south Georgia for their use and review. Version 1.1 of the NFSEG groundwater model and the HSPF-derived surface-water models has been developed to address changes and improvements recommended for Version 1.0. In this version, the model has been calibrated to steady-state hydrologic conditions representing 2001 and 2009. Surface-water models have been developed for all surface-water basins within the groundwater model boundaries using the Hydrological Simulation Program-FORTRAN (HSPF) software to improve initial estimates of recharge and maximum saturated evapotranspiration for input.

Preliminary calibration results for Version 1.1 of the NFSEG groundwater model and the HSPF-derived surface-water models were completed in May 2017, and documentation and model files of Version 1.1 of the NFSEG and HSPF models were completed for final peer review in April 2018.

In March 2017, SJRWMD and SRWMD requested an independent technical peer review of the NFSEG groundwater model and the HSPF models as the final phase of Version 1.1 of the model was being developed. Responsibilities of the Peer Review Panel included conducting a thorough review of the groundwater and surface-water models and model documentation reports and assessing the following topics:

- Model objectives, conceptualization, and design;
- Assumptions and limitations of input data;

- Model calibration and sensitivity;
- Model documentation (explanation of model, data sources, and assumptions);
- 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.

The majority of the peer review comments from the April17, 2018 summary review report were positive and indicated that the model development and results were well conceived and executed. This response memorandum presents a list of reviewer's requests for more detail and cites the location of each reviewer comment following the comment number. This is followed by a table in blue with the description of a response that the District determined would meet the team's request (Requested Action Item). The location of each response added to the final report is provided in the second column of each table (District Response). All typographical errors and issues with Figures, Tables and Appendices will be corrected in the final report.

REVIEWER COMMENT 1 (Chapter 2.1, Section 2A)

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

> "… this report [does not] sufficiently discuss .. the springs and the baseflows in the rivers. The baseflow discussion is too short, and for the intended use of this model, a thorough documentation and understanding of the baseflows is very important. In addition, ASTM (2018) (also see response to Question 3.J.5) recommends that the error range associated with each calibration target be identified, in addition to the value to be used for calibration. This was not done for the baseflows."

Requested Action Item	***District Response***
Discuss baseflow and spring-flow contributions to river flows in greater detail.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")* *Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Identify error ranges associated with baseflow targets.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*

REVIEWER COMMENT 2 (Chapter 2.1, Section 2C)

Was the numerical [MODFLOW] model constructed in a manner that is consistent with the underlying conceptual model, using appropriate data and methods of analysis?

> "The description of baseflows requires further discussion in this report. Also, it appears that spring flows were given much larger weights than river baseflows in the calibration, causing PEST to produce closer matches to the springs and poorer matches to the rivers; this point needs to be discussed in this report as well. Not having recharge or evapotranspiration as PEST parameters requires further discussion in this report and further consideration as possible calibration parameters in any future revision of version 1.1 of the NFSEG model. Further discussion in this report should include an estimate of the accuracy of the recharge and evapotranspiration values calculated in HSPF, an explanation of why springs were simulated in

layer 3 and rivers were simulated in layer 1, and whether manually adjusting recharge and evapotranspiration would result in better matches for the river baseflows."

Requested Action Item	***District Response***
Provide additional discussion of baseflows. Describe reasons for assigning larger weights to spring flows than baseflows in the PEST calibration process.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches") Chapter 4 (Under section entitled "Individual spring flows (2001 and 2009)"*
Describe reasons for not treating recharge and maximum saturated ET as PEST calibration parameters.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*
Discuss the accuracies of the recharge and maximum saturated ET distributions calculated by HSPF.	*Chapter 3 (under section entitled "Recharge and Evapotranspiration")*
Explain why springs were simulated primarily in layer 3 while rivers were simulated primarily in layer 1.	*Chapter 3 (Under section entitled "River Boundaries") Chapter 3 (Under section entitled "General Head Boundary Conditions")*
Discuss as to whether manually adjusting recharge and maximum saturated ET would result in better matches for river baseflows.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*

REVIEWER COMMENT 3 (Chapter 2.1, Section 3A)

Is the parameterization scheme used in the PEST calibration appropriate?

"… not making recharge a PEST parameter needs an explanation (or inclusion as a parameter).

"The justifications for treating evapotranspiration and recharge as constants in the PEST calibration in NFSEG Version 1.1 need to be discussed further in this report. Allowing evapotranspiration and recharge to be adjusted during PEST runs should be evaluated further in any future revision of Version 1.1 of the NFSEG model."

Requested Action Item	***District Response***
Explain why recharge is not a PEST calibration parameter.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*
Justify treating maximum saturated ET and recharge as constants in the PEST calibration process.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*

REVIEWER COMMENT 4 (Chapter 2.1, Section 3B)

Were the types of observations and their implementation in the PEST calibration appropriate, given the intended use of the model?

" A better description of the reasoning for the weights assigned to each observation group should be provided, however. The report does a good job of documenting the weights that were used but does not really get into the logic behind the choice of weights. Also, the river baseflow determination needs more discussion and documentation (as mentioned earlier). "

Requested Action Item	*District Response*
Provide a better description of the reasoning for the weights assigned to each observation group.	*Chapter 4 (starting at second paragraph under section entitled "Weighting Scheme")*
Provide more discussion and documentation of river baseflow determination.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*

REVIEWER COMMENT 5 (Chapter 2.1, Section 3C)

Have the 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?

"First, since the report goes into considerable detail on parameter and observation groups, it would be consistent to add a table of the contributions of each observation group to the objective function. The objective function is described in general in the report, but the actual results from the PEST run are not documented. A table is provided of head statistics but not for spring flows and base flows. Spring data and baseflow pick-up estimates in Appendices E and F should also show the percent error in spring flow and base flow values to give the reader a better indication of the degree of fit with the flow observations. In addition, the match for important springs is provided in table form for the 2010 verification simulation (Table 5-2), but spring flow matches should also be tabulated and evaluated for the 2001 and 2009 calibration periods. Also, in the no-pumping simulation, estimates for historical heads and spring flows were used to evaluate the no-pumping simulation results, but estimates for baseflows were not made. A number of rivers in the model domain have gages that date back to the 1930s; if possible, these data should be used to estimate historical baseflows, which could also be used to evaluate the no-pumping simulation. In addition, however, return flow should not be included in the no-pumping simulation."

Requested Action Item	*District Response*
Add a table of contributions of each observation group to the objective function.	*In Appendix "M", a new appendix that we created for this purpose, referenced in section called "Calibration Results"*
Add percent errors of simulated vs. estimated spring flows and baseflow pickups in Appendices E and F.	*In Appendix "N", another new appendix, referenced in section called "Observed versus Simulated Spring Flows"* *In Appendix "O", another new appendix, referenced in section called "Quality of Baseflow Matches"*
Provide tables of the matches for important springs for 2001 and 2009, as was done for 2010.	*Inserted in chapter 4 as Tables 4-7 and 4-8 under section entitled "First Magnitude Springs and Spring Groups"*
Estimate historical baseflows for gages having hydrographs that extend back to the 1930s or earlier for further evaluation of the pumps-off simulation.	*Inserted in Chapter 5 under section entitled "Baseflow Estimates from 1933 through 1942"*

REVIEWER COMMENT 6 (Chapter 2.1, Section 3I)

Have the limitations of the final version of the NFSEG groundwater flow model been adequately described in the model documentation?

> " Section 8 Model Limitations concisely explains the limitations of the model with the following exception: in the calibration section (Section 4 Model Calibration, p. 55), it is stated that structural errors typically are the largest source of errors in a model. This should be repeated in Section 8."

Requested Action Item	*District Response*
Repeat in Section 8 of the NFSEG v1.1 final report that structural errors are typically the largest sources of errors in a model.	*Additional discussion of structural error inserted in Chapter 7 under section entitled "Parameter Uncertainty Analysis Results", second paragraph.*

REVIEWER COMMENT 7 (Chapter 2.1, Section 3J - Obj. 5)

Do the model calibration statistics meet industry standards in ASTM Standard Guide for Calibrating a Ground-Water Flow Model Application, Designation D 5981-96 (2008).

> "In the MODFLOW simulation, calibration targets for heads were established prior to the calibration process but not for spring flows and baseflows, which should be established. Based on ASTM (2018), one criterion for accepting a calibration is that the residual for heads should be a small fraction of the difference between the highest and lowest heads across the model area. This criterion should be checked in addition to the calibration results for heads and residuals described in Section 4 Model Calibration of the draft model report. In addition, targets for spring flows and baseflows should be established based on the accuracy of the observed (or estimated) values for these parameters. ASTM (2018) recognizes that errors in the estimates for groundwater flow rates will usually be larger than errors in the estimates of heads and, in particular, that baseflow estimates are generally accurate only to within an order of magnitude. In such cases, the upper and lower bounds on the acceptable modeled value of baseflow can be equal to the upper and lower bounds on the estimate. This limit should be recognized when establishing calibration targets and evaluating the calibration for baseflows in the NFSEG groundwater model."

Requested Action Item	*District Response*
Determine the ratios of head residuals to the difference between the maximum and minimum heads within the model domain.	*Figures 4-11b and 4-12b, section entitled "Groundwater-Level Residuals of Layer 1"* *Figures 4-21b and 4-22b, section entitled "Groundwater Level Residuals of Layer 3"*
Determine acceptable upper and lower bounds for simulated spring flows and baseflows as the upper and lower bounds of estimated values. For baseflows, suggests using a range of +/- of an order of magnitude of respective estimated values.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")* *Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Spring-flow error bounds should be established, in addition to baseflow error bounds.	*Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Provide comparison of simulated spring flows and baseflows to corresponding estimated values and	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*

provide indications of degree to which simulated values fall within established error bounds (i.e., "…targets for springs and baseflows should be established based on the accuracy of the observed (or estimated) values").	*Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*

REVIEWER COMMENT 8 (Chapter 2.2, Section 1A)

Does the documentation provide a clear and appropriate description of the NFSEG groundwater flow model and supporting HSPF surface-water models?

> "The areas that are lacking are: 1) the presentation of the calibrated model parameters and 2) calibrated model water balances. HSPF models should include documentation of the key hydrologic parameters that are used to calibrate a model. The NFSEG HSPF model is quite complex due to the large geographic area and the large number of unique HSPF models that are included. In order for a reviewer to determine whether the various parameters are within reasonable/valid ranges, the documentation should include an appendix that summarizes the parameter values with tables and maps."

Requested Action Item	***District Response***
Document the key hydrologic parameters used in HSPF calibration in an appendix in tables and maps.	*Appendix R* *Appendix T*
Provide calibrated HSPF model water balances.	*Appendix T*

REVIEWER COMMENT 9 (Chapter 2.2, Section 1B)

Are the purposes and scope of the HSPF documentation clearly stated and sufficient to document the models? Is the content of the documentation consistent with the stated purpose and scope of the document?

> "The calibration approach should include discussion of the effects of calibration of flows affected by tides and significant man-made influences on the predicted recharge. The calibration results (shown in the 55 watershed-specific appendix sections) should include a brief discussion of man-made influences and other causes of poor calibration for poorly calibrated gauges. The calibration section should also include documentation of the key hydrologic parameter values obtained or reproduced from a nearby watershed during calibration (see question #1D), and also the simulated water balance summaries described under question #3G. "

Requested Action Item	***District Response***
Include a discussion of the effects of tides and manmade structures on predicted recharge and the calibration in general.	*Chapter 9, Table 9-17 in section "Calibration Results", Appendix R and T*
Include documentation of the key hydrologic parameter values obtained or reproduced from a nearby watershed and the simulated water balance summaries described under question #3G.	*Appendix R and T*

REVIEWER COMMENT 9 (Chapter 2.2, Section 1D)

After reading the documentation, are the purposes, scope, strengths/weaknesses, intended use, and limitations of the NFSEG model understandable?

“The calibration of watersheds with tidal and man-made influences on measured flows should be discussed in the calibration approach, and the possible effects on computed recharge should be evaluated. Since PEST is used for the automated calibration, the effects of specific objective function components on calibration should be discussed in this section.”

Requested Action Item	***District Response***
Discuss the calibration of watersheds with tidal and manmade influences on measured flows in the discussion of the calibration approach. Evaluate the effects on computed recharge.	*Chapter 9, Table 9-17 in section “Calibration Results”, Appendix R and T*
Discuss the effects of specific objective function components on the calibration.	*Chapter 9, Tables 9-14 and 9-15 in section “Parameter Estimation with PEST”*

REVIEWER COMMENT 10 (Chapter 2.2, Section 2F)

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?

“ … this could be backed up more clearly in the documentation, including a description of the feature(s) that are non-standard, and citation of a document that confirms the District’s prior validation of the non-standard version. The primary feature that is not in the publicly-available version is an optional method for computing surface runoff from a standard pervious land area (PERLND). This feature is utilized to improve the simulation of surface runoff from the land areas categorized as wetlands and water in the NFSEG model. (LSG—p. 15).”

Requested Action Item	***District Response***
Describe non-standard feature(s) in the documentation. Include a citation of a document that confirms the District’s prior validation of the non-standard version.	*Chapter 9, in section “Surface FTABLES”*

REVIEWER COMMENT 11 (Chapter 2.2, Section 2Fa)

Unique aspects of these systems were represented with Special Actions or with other features of HSPF. Are these conceptually sound and implemented appropriately:

“This simulated spring outflow was calibrated to measured spring flows, which is very innovative. However, this aspect of the spring feature does not seem to be included in the documentation.”

“Review of a HSPF model input file with a closed basin indicates that it is implemented correctly; however, the values of the reach-specific parameters used to represent the invert, the maximum flow and depth above invert where maximum flow begins could not be verified as part of this peer review.”

Requested Action Item	***District Response***
Include a description of the spring-flow simulation feature in the documentation.	*Chapter 9, in section “Representation of Springs to Improve HSPF Calibration”*
Include a description of the invert, maximum flow and depth above	*Chapter 9, in section “Closed*

invert where maximum flow begins for each closed-basin reach in the discussion of closed-basin representation.	*Basin Representation"*

REVIEWER COMMENT 12 (Chapter 2.2, Section 2Fc)

Were the types of observations and their implementation in the PEST calibration appropriate, given the intended use of the model?

> **"**Because PEST is not yet in common usage by HSPF modelers, it is recommended that the objective function components be more completely described, especially the effects of adjusting the relative weights."

Requested Action Item	***District Response***
Describe PEST objective-function components more completely, especially regarding adjustments of relative weights.	*Chapter 9, in section "Parameter Estimation with PEST"*

REVIEWER COMMENT 13 (Chapter 2.2, Section 3C)

Have the differences between observations and their simulated equivalents (model residuals) been described sufficiently?

> **"**The main recommendation is to include a very brief discussion of the modelers' conclusions and evaluation of the reasons for poor agreement in the calibration results at gauges that are poorly calibrated. These reasons can be a combination of poor observed data, tidal effects, man-made influences in the watershed, unmodeled groundwater gains/losses, and uncertainty in a key input."

Requested Action Item	***District Response***
Include a brief discussion of the reasons for poor agreement in the calibration results for gauges that are poorly calibrated.	*Chapter 9, Table 9-17 in section "Calibration Results" and Appendix T*

REVIEWER COMMENT 14 (Chapter 2.2, Section 3D)

Have the values of calibrated parameters been described appropriately?

> **"No,** the HSPF documentation does not include appropriate description of the primary hydrologic parameter values obtained during calibration. The minimum set of calibrated parameters that should be documented in an appendix (tables and maps) are listed below.

AGWRC - Base groundwater recession
BASETP - Fraction of remaining ET from baseflow
CEPSC - Interception storage capacity
DEEPFR - Fraction of groundwater inflow to deep recharge
INFILT - Index to infiltration capacity
INTFW - Interflow inflow parameter (omit due to low value)
IRC - Interflow recession parameter (omit due to low value of INTFW)
KVARY - Variable groundwater recession
LZETP - Lower zone ET parameter
LZSN - Lower zone nominal soil moisture storage
UZSN - Upper zone nominal soil moisture

The main purposes of this recommendation are to: 1) ensure that the parameters have reasonable values, i.e., they are within valid ranges for the respective process formulations and for the specific land cover and climate; and 2) ensure that the variation over the model domain and within specific watersheds is reasonable. The standard requirement for any HSPF model documentation includes summaries of the key calibrated (and assumed) hydrologic parameters listed above."

Requested Action Item	***District Response***
Document the values of the key calibration parameters listed above as resulting from the calibration process.	*Appendices R and T*
Enable judgment of reasonableness of ranges in these values in view of land covers and climate.	*Appendices R and T*
Ensure that variations are reasonable within specific watersheds and across the model domain.	*Appendices R and T*

REVIEWER COMMENT 15 (Chapter 2.2, Section 3E)

Does the final version of the model appear to be adequately calibrated given the available data for calibration and the state of knowledge (and lack thereof) of the hydrologic system prior to development of the model?

"It is noted that at several gauges, there are large, virtually constant differences between the simulated and observed flows that are caused by either an error in the model or a significant man-made influence. These should have been investigated and either documented, if it is man-made; or corrected, if a model error was the cause. Examples are gauges 02197500 and 02198500, both in the Savannah River. It is assumed in these cases that the calibration criteria used by PEST were affected by objective function components other than the total flow, (e.g., total actual ET)."

"The main questions or concerns with the calibration are related to the effects on recharge of calibration to observed flows that are affected by tidal and (especially) man-made influences. The discussion should include an analysis of this impact. Possibly, the effect is small for the same reason that the calibration did not adjust the simulated flow to match observed in the examples of large, constant differences in the two Savannah River gauges noted above. It is assumed that other criteria in the objective function prevented the large changes that would be needed to bring the flows into better agreement."

Requested Action Item	***District Response***
At several gauges, there are large, virtually constant differences between simulated and observed flows. This should be investigated for determination of causes and documented. If model error is the cause, then the problem should be corrected.	*Chapter 9, Table 9-17 in section "Calibration Results", Appendix R and T.* *Identified as improvement for future NFSEG versions.*
Include an analysis of the effects of tidal and manmade influences on calibration results.	*Chapter 9, Table 9-17 in section "Calibration Results", Appendix R and T.* *Identified as improvement for future NFSEG versions.*

REVIEWER COMMENT 16 (Chapter 2.2, Section 3G)

Has the complete model water balance, accounting for all water sources and sinks, been assessed and found reasonable?

"Not completely. This question seems to be addressed primarily to the MODFLOW model. However, it is also applicable to the HSPF model. The Districts should generate and document (in an appendix) summaries of the average annual HSPF water balance results for the individual land areas (PERLND and IMPLND). This 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 in addition to averages over all land covers. The primary purpose for this output is to determine the reasonableness of the amounts. It allows the modeler to identify errors in the input data such as rainfall, PET, and irrigation; and unreasonable water balance quantities caused by the automated calibration. In addition, the calibration of total actual ET to expected annual amounts can be verified.

Based on a review of preliminary water balance data that the District recently produced for individual years (2001, 2009, and 2010), it is recommended that the water balance should be computed for the full period of calibration instead of individual years, and it should be included in an appendix so that model reviewers can compare the results with input data (rainfall, irrigation, etc.) and the calibrated of total ET, in addition to verifying that the other components are reasonable."

District Response

Requested Action Item	***District Response***
Document summaries of average annual HSPF water-balance results for the individual land areas (PERLND and IMPLND). Include a summary of the inputs, ET losses, runoff losses to streams by soil layer, and groundwater recharge. Weighted average summaries should be generated for each land cover in addition to averages for all land covers.	*Appendix T*
Produce the water balances for the full period of calibration instead of for individual years.	*Appendix T*
Place water-balance results in an appendix to enable comparisons of results with input data and the calibrated total ET, in addition to other components.	*Appendix T*

REVIEWER COMMENT 17 (Chapter 3.1, Section 2)

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

Requested Action Item	*District Response*
Expand discussion of baseflows.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*
Acknowledge relative accuracies of groundwater heads and groundwater flows.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")* *Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Provide more detailed description of groundwater inflows and withdrawals, possibly in an appendix.	*Chapter 4 We can share the data on which the NFSEG well-input files are based; however, the files are quite large. Therefore, some discussion of how to transfer the data will need to occur. In the meantime, please note that the NFSEG well input files have been posted on the NFSEG website.*

REVIEWER COMMENT 18 (Chapter 3.1, Section 3)

Model Calibration and Sensitivity

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

District Response

Requested Action Item	*District Response*
Discuss justification for treating recharge and maximum saturated ET as constants in the PEST calibration process.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*
Re-examine calibration targets for heads to determine if a broader range of statistical analyses such as criteria for mean error (ME), mean absolute error (MAE), and root mean square (RMSE) would provide a	*Chapter 4 (Under section entitled "Comparisons of Groundwater-Level Residual Statistics to Other Models"); also Chapter 4 (Under section entitled*

better set of metrics to judge the simulation results for 2001, 2009, and 2010.	*"Comparison of NFSEG v1.1 Calibration Statistics in Portions of Model Domain that corresponds to the North Florida Water-supply Planning areas versus Overall Model Domain""*
Establish calibration targets for springs and baseflows, keeping in mind that spring-flow and baseflow estimates are much less accurate than head measurements.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")* *Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Compare residuals/residual statistics of NFSEG to those of other comparable regional groundwater flow models.	*Chapter 4 (Under section entitled "Comparisons of Groundwater-Level Residual Statistics to Other Models")*

REVIEWER COMMENT 19 (Chapter 3.1, Section 4)

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

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

District Response

Requested Action Item	***District Response***
Apply tests of random and normal distribution of residuals on heads, spring flows, and baseflow residuals.	*Chapter 4 (Under section entitled "Statistical and Spatial Trends in NFSEG v1.1 Groundwater-Level Residuals and Comparison of Trends to Other Groundwater Models")*
Provide additional descriptive material and details about baseflows.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*
Remove generalized descriptions of calibration results.	*The "Calibration Results" section was revised to accomplish this.*

REVIEWER COMMENT 20 (Chapter 3.1, Section 8)

Deficiencies, Errors, or Sources of Uncertainty in Model/Relationship Development, Calibration, and Application

> " In the report reviewed here, the range of errors in the determination of the baseflows is not reported as recommended by ASTM (2018), and additional documentation and discussion of spring flows and baseflows is needed. Calibration targets for spring flows and baseflows also need to be established, and consideration needs to be given to adjusting recharge and/or ET during the PEST calibration in any subsequent revision of version 1.1 of the NFSEG model. Additionally, there is some indication that head, spring flow, and base flow residuals are not randomly distributed in the model domain. A non-random, spatial distribution in residuals often indicates model bias and possible model error. To determine the validity of spatial randomness, the "run statistics" (Hill 1998) calculated by the MODFLOW Observation Process or similar code should be used as an independent measure of randomness. "

District Response

Requested Action Item	***District Response***
Report range of errors in baseflows.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")*
Establish calibration targets for spring flows and baseflows.	*Chapter 4 (Under section entitled "Quality of Baseflow Matches")* *Chapter 4 (Under section entitled "Spring-Flow Target Uncertainty")*
Consider making recharge/maximum saturated ET adjustable calibration parameters.	*Chapter 4 (Under section entitled "Recharge and Maximum Saturated ET as Calibration Parameters")*
Perform testing to determine if non-randomness of residuals is a problem. Suggests use of MODFLOW Observation Process (Hill 1998) or similar code as an independent test of spatial randomness.	*Chapter 4 (Under section entitled "Statistical and Spatial Trends in NFSEG v1.1 Groundwater-Level Residuals and Comparison of Trends to Other Groundwater Models")*

REVIEWER COMMENT 21 (Chapter 3.2, Section 3)

HSPF Model Calibration and Sensitivity

> " Some stream gauges where the data are uncertain (i.e., poor quality as judged by USGS) have unsatisfactory calibration statistics. These are generally locations that are influenced by tidal flows, man-made structures and flow modifications, and unusually flat or areas of strong groundwater interaction with surface flows. These poorly-calibrated gauges should be discussed briefly in the calibration summaries for each HUC8 watershed. "
>
> "Sensitivity is not addressed in the documentation for the NFSEG HSPF model. … A possible future enhancement would include sensitivity analysis of these parameters in selected watersheds."

District Response

Requested Action Item	***District Response***
Discuss reasons for poor calibration results of	*Chapter 9, Table 9-17 in section "Calibration Results", Appendix R and T*

specific gauges in the calibration summaries.	
Include sensitivity analysis in selected watersheds.	*Identified as improvement for future NFSEG versions.*

REVIEWER COMMENT 22 (Chapter 3.2, Section 4)

HSPF Model Documentation (explanation of model, data sources, and assumptions)

> "In the calibration approach section, watersheds with tidal and man-made influences on measured flows should be discussed, and the possible effects on computed recharge should be evaluated. Since PEST is used for the automated calibration, the effects of specific objective function components on calibration should be discussed in the section on PEST. In the calibration section, the final parameter values of selected HSPF parameters should be compiled and summarized, and HSPF water balance summaries should be compiled and summarized to verify their reasonableness and verify that the total actual ET calibration to expected/literature values is adequate. "

District Response

Requested Action Item	***District Response***
Discuss possible effects on computed recharge of watersheds with tidal and manmade influences on measured flows.	*Chapter 9, Table 9-17 in section "Calibration Results", Appendix R and T*
Discuss the effects of specific PEST objective function components on calibration.	*Chapter 9, in section "Parameter Estimation with PEST".* *Identified as improvement for future NFSEG versions.*
Compile and summarize calibrated HSPF parameters and HSPF water balances.	*Appendix T*

REVIEWER COMMENT 23 (Chapter 3.2, Section 8)

HSPF Deficiencies, Errors, or Sources of Uncertainty in Model/Relationship Development, Calibration, and Application

> "Some of the watersheds are affected by processes that are not included in these models due to the limitations imposed by the large area and large number of models. These include man-made modifications, tidal effects, and large groundwater influence on surface water flows. The modelers made a decision to not include man-made changes in the models, and HSPF is generally not capable of representing significant groundwater or tidal effects without additional conceptualization and use of special features. Therefore, it is fair to say that the underlying HSPF process relationships are somewhat limited for accurately calibrating watersheds with these conditions unless they are explicitly included by the modeler. This is illustrated in many of the poorly calibrated gauges in the model. However, some of the poorly calibrated watersheds are likely resulting in reasonable and appropriate recharge, since many of the objective function criteria are being satisfied. In those watersheds where the percent bias is extremely high (and therefore the recharge is more likely to be invalid), it is recommended (in future calibrations of the model) that the model be modified to represent the man-made influences, or alternatively those watersheds should be assigned parameter values from a nearby watershed that is well

calibrated. This recommendation of using calibrated parameters from another watershed should also be applied to gauges that have strong tidal influences."

District Response

Requested Action Item	***District Response***
Include man-made modifications to the hydrology within the HSPF models.	*Identified as improvement for future NFSEG versions.*
Use parameters from nearby subwatersheds where the calibration is not affected by man-made modifications to the hydrology of the system.	*Identified as improvement for future NFSEG versions.*