TETRA TECH

MEMO

To:	Tony Countryman, Northwest Florida Water Management District (NWFWMD)
Cc:	Jack "Trey" Grubbs and Kathleen Coates, NWFWMD
From:	Scott Simpson and Pete Andersen, Tetra Tech David Simon and Chuck Spalding, McDonald Morrissey Associates (MMA)
Date:	May 20, 2020
Subject:	Transient Calibration of the Region II Groundwater Flow Model

1.0 INTRODUCTION

The Northwest Florida Water Management District (NWFWMD, or "the District") is updating its existing groundwater flow and transport models to assess the need to develop and establish minimum aquifer levels to manage saltwater intrusion in the Upper Floridan aquifer along the coastal portion of Water Supply Planning Region II (Santa Rosa, Okaloosa, and Walton Counties; **Figures 1 and 2**). Modeling tasks are anticipated to include the following: (1) perform transient calibration of the Region II MODFLOW ("R2MF") groundwater flow model with updated pumpage and hydrologic data, (2) build and verify a sub-regional, Coastal Region II SEAWAT ("CR2SWT") model from two existing sub-regional DSTRAM models, (3) calibrate the CR2SWT model with updated pumpage, water level and water quality data, and (4) perform predictive simulations. This technical memorandum addresses Task 1: performing a transient calibration of the existing R2MF model with updated pumpage and hydrologic data.

An existing MODFLOW-based groundwater flow model was developed in 2000 for the Upper Floridan aquifer in planning Region II (HydroGeoLogic, 2000). The model was subsequently used to provide water level boundary conditions for the transient calibration of two sub-regional DSTRAM models extending from predevelopment to 2004 (HydroGeoLogic, 2005; HydroGeoLogic, 2007). Several years ago, the regional model was updated with additional pumpage for years 2005 through 2009 to run predictive simulations.

This technical memorandum describes modifications—including updated pumping, boundary conditions, and parameterization—that were made to the existing R2MF model to enable transient simulation from 1942 through 2015. Automated calibration using PEST (Doherty, 2019) was performed in favor of the manual approach used by HydroGeoLogic (2000), and additional types and number of calibration targets were included in the R2MF model.

2.0 MODEL CONFIGURATION AND UPDATES

2.1 MODEL CONFIGURATION

The three-dimensional finite difference grid from the previous model (HydroGeoLogic, 2000) was used without modification. This grid covers an area of approximately 10,600 square miles and encompasses the entirety of four Florida counties (Escambia, Santa Rosa, Okaloosa and Walton), portions of three other Florida counties (Holmes, Washington and Bay), and part of southern Alabama (**Figure 1**). The R2MF model grid consists of 5 layers, 114 rows, 175 columns. Row and column spacing is variable and grid cell sizes range from approximately 6 square miles (mi^2) in the Northwestern Highlands to approximately 0.2 mi^2 near the center of the domain (**Figures 1 and**

1165 Sanctuary Parkway, Suite 270, Alpharetta, GA 30009
Tel 770-619-9950 Fax 770-619-9903 tetratech.com

2). Each model layer contains a unique hydrostratigraphic unit including (from top to bottom; **Figures 3 and 4**): the surficial aquifer system (including the sand-and-gravel aquifer, where present), the intermediate aquifer system (IAS; a regional confining unit), the Upper Floridan aquifer (UFA), the Bucatunna clay confining unit (where present), and the Lower Floridan aquifer (LFA) (HydroGeoLogic, 2000). Where the Bucatunna clay confining unit is not present, layers 3, 4, and 5 represent collectively the undifferentiated Upper Floridan aquifer. The layer 1 (surficial aquifer system or sand-and-gravel aquifer) is specified as an unconfined aquifer whereas layers 2 through 5 are considered confined (HydroGeoLogic, 2000).

The boundary conditions within the model included:

- No-flow boundaries beneath Layer 5,
- Constant head boundaries representing bays and the Gulf of Mexico in Layer 1 (**Figure 5**),
- Head-dependent flux (“General Head”) boundaries along the entire eastern, western, and southern boundaries in layers 3, 4, and 5, and along the eastern boundary in layer 1,
- Rivers in layer 1 representing larger rivers and creeks (**Figure 5**),
- Rivers in layers 1 and 3 representing Holmes Creek and the Choctawhatchee River (**Figure 5**),
- Drains in all layer 1 cells not containing a constant head boundary (**Figure 5**) to prevent unrealistically high simulated layer 1 heads,
- Pumping and injection wells, and
- Domain-wide, time-constant groundwater recharge to layer 1 equal to approximately 20 inches/year.

The R2MF model’s original documentation (HydroGeoLogic, 2000) provides more detailed descriptions of model layering, boundary condition selection and parameterization, and the hydrogeologic conceptual model upon which the R2MF model is based.

2.2 MODEL UPDATES

A few updates were made to the R2MF model provided by the District (HydroGeoLogic, 2000). One such update was replacing the prior groundwater extraction and injection rates with updated rates provided by the District. All other revisions made to the R2MF model are described in greater detail below.

2.2.1 MODFLOW Version & Packages

The R2MF model provided by the District (HydroGeoLogic, 2000) was developed for MODFLOW-96 (Harbaugh and McDonald, 1996). The model was updated using Groundwater Vistas (Environmental Simulations, Inc., 2019) to be run using MODFLOW-2005 (Harbaugh, 2005) to take advantage of more recent refinements to MODFLOW. In the process of this model translation, the Discretization (DIS) Package was added and the Block-Centered-Flow (BCF) Package was replaced with the Layer Property Flow (LPF) Package.

2.2.2 Simulation Period Updates

The original R2MF model (HydroGeoLogic, 2000) was calibrated to 1990 steady-state conditions. A subsequent application used the model in transient mode to simulate groundwater flow from pre-development (1942) to 2009. One long (36,500 day) transient stress period representing pre-development conditions was followed by annual stress periods to 2009. For this transient calibration task, six additional annual stress periods were added to the model to represent years 2010 through 2015. Pumping rate information for years 2010 to 2015 provided by the District were added to the Well (WEL) Package. Additional pumping and injection well information was received from the District on two separate occasions, and new pumping and injection wells—as well as updated injection rates for existing injection wells—were added to the WEL Package. Previously calibrated model boundary conditions for the River (RIV), General-Head (GHB), and Drain (DRN) Packages were held constant for the annual stress periods representing pre-development to 2015.

2.2.3 Constant Head Boundaries

The layer 1 constant heads representing the Gulf of Mexico and adjoining bays were revised late in the calibration process to account for the effects of saltwater density. Revised heads were defined in terms of the equivalent freshwater head (h_f) based on the densities of seawater (ρ_s) and fresh water (ρ_f), point-water heads (h=0), and the cell-specific bottom elevations of layer 1 (Z) as in SEAWAT (Langevin et al., 2003):

$$h_f = \frac{\rho_s}{\rho_f}h - \frac{(\rho_s - \rho_f)}{\rho_f}Z$$

Prior to this revision, all constant heads were assigned a head of zero.

3.0 MODEL CALIBRATION

3.1 CALIBRATION TARGETS

Transient calibration of the updated R2MF model was originally intended to include the following targets: annual average groundwater heads, spatial and temporal differences in annual average heads, and spatial changes in river baseflow. However, during the calibration process additional targets—including prior estimates of aquifer transmissivities and pre-development groundwater heads—were added to the original target dataset.

3.1.1 Groundwater Heads

Mean annual groundwater head measurements were provided in a database by the District. From this database only groundwater heads from wells located in model layers 2, 3, and 5 were used as calibration head targets. The target dataset was refined further during the calibration process as some head targets were determined to be based on either unrepresentative head data (e.g. water levels from pumping wells) or erroneous measurements. After the removal of all erroneous and unrepresentative targets (as determined by District staff), the head target dataset used to inform and guide the calibration process included 139 well locations (**Figure 6**) and 2,769 head targets (**Appendix A**). For practical purposes, mean annual groundwater head measurements were assumed to represent conditions at the approximate midpoint of each year (June 30) during calibration.

3.1.2 Head Differences

Head difference target datasets were constructed from the District's groundwater head database. The individual wells used to develop head difference targets were chosen to capture the observation-based shape and timing of development of the extraction-induced cone of depression in the Upper Floridan aquifer.

Vertical head difference (VHD) targets were used to guide the automated calibration process toward accurately representing vertical gradients in groundwater heads between model layers. These targets were taken as the difference between head targets from wells located within the same model grid cell (i.e. same row and column) in the same year, but in different layers. A total of 7 vertical well pairs were identified (**Figure 7**) between model layers 2 and 3, 1 and 3, and 3 and 5 which provided total of 110 vertical head difference targets (**Appendix A**).

Horizontal head difference (HHD) targets were used to capture the lateral differences in groundwater heads within the model domain and were centered on the cone of depression. These targets were taken as concurrent differences between mean annual groundwater heads at a well located in the center of the cone of depression (NWF ID = 1894 in layer 3; NWF ID = 3210 in layer 5) and mean annual groundwater heads from wells in the same model layer surrounding the center well (**Figure 8**). A total of 12 well pairs were identified in the Upper and Lower Floridan aquifers which provided a total of 507 horizontal head difference targets (**Appendix A**).

Temporal head difference (THD) targets were created to ensure the calibrated model would be able to capture the timing of the cone of depression development at individual wells. Wells with longer records of measurements were chosen for THDs, and individual head difference targets were calculated as the difference between each head measurement and the latest (i.e. most recent) reliable head measurement. In total, 44 wells were chosen (**Figure 9**) and 1,504 temporal head difference targets were defined for these wells (**Appendix A**).

3.1.3 Baseflow Targets

Differences in baseflow between upstream and downstream gages (**Figures 2 and 5**), or "incremental baseflows," were used as qualitative targets during calibration. The District provided annual average baseflow estimates for two upstream gages—the Choctawhatchee River at Caryville, FL (USGS 02365500) and Holmes Creek at Vernon, FL (USGS 02366000) gages—and one downstream gage (Choctawhatchee River near Bruce, FL; USGS 02366500). These estimates were used to calculate a total of 39 annual average incremental baseflow targets spanning the 1950-1978 and 2006-2015 periods. Stream flow data were not available for the Holmes Creek at Vernon, FL gage for the years 1979 through 2005, so annual average baseflow estimates could not be made. Therefore, annual average incremental baseflow targets could not be created for this time period. These targets were compared to cumulative simulated River Package flows between the upstream and downstream gages (**Figure 5**). The incremental baseflow targets were assigned weights of zero, and therefore did not contribute to the objective function being optimized but were monitored throughout the calibration process to ensure approximate agreement between simulated and estimated values.

3.1.4 Target Groups and Weighting

A calibration plan was developed prior to calibration that included how the initial target datasets would be subdivided into groups and subsequently weighted during calibration. Groundwater heads, VHD, and HHD targets were each grouped into three time periods (1942-1965, 1966-1990, and 1991-2015) for a total of nine (9) target groups. The THD targets were divided into three target groups based on well locations: (1) wells outside the SEAWAT model domain, (2) the eastern half of wells (by well count) in the SEAWAT domain, and (3) the western half of wells (by well count) in the SEAWAT domain. The PWTADJ1 utility from the Parameter Estimation ("PEST") Suite (Doherty, 2019) was used to reassign weights to the resulting 12 head and head difference target groups such that each group contributed equally to the objective function being minimized. The ultimate objective of this reweighting approach was to reduce the likelihood of spatially or temporally biased model errors. The District approved the final calibration plan (**Appendix B**) following minor modifications.

3.1.5 Additional Targets

Two additional types of targets were added after the calibration process began: aquifer transmissivity targets and pre-development head targets. Each new type of target was assigned to a new target group during subsequent calibration runs.

Estimated transmissivities (T) from aquifer performance tests (APTs) conducted at 39 locations (**Figure 10**) were provided by the District and used as calibration targets. All APT-based T estimates were compared to the modeled transmissivity of the corresponding well open interval. Model transmissivities were calculated by the PEST utility program, PAR2PAR, based on well open-interval information, model layer elevations, model cell-specific kriging factors, and horizontal hydraulic conductivity pilot point values. Transmissivity targets were assigned to three different groups based on geographic location. One group included the nine APT locations near the center of the pumping-induced UFA cone of depression, a second group included 13 APT locations within approximately 15 miles of locations in the first group, and the third group included the remaining 17 APT locations (all of which were in the east-central portion of the domain). Within each group, multi-well APTs were assigned weights twice those of single-well APTs.

Pre-development head targets were developed using data provided by the District and the interpreted Floridan aquifer potentiometric surface developed by Bush and Johnston (1988). A total of 98 pre-development head targets were used in calibration (**Figure 11**): 16 were based on measurements provided by the District (all of which were made before 1940), and 82 additional targets were defined at unique, evenly-spaced points along the onshore portion of the SEAWAT model boundary based on interpolation from the Bush and Johnston (1988) map. Pre-development head measurement targets were assigned weights double those used for the SEAWAT model boundary head estimates.

3.2 CALIBRATION PARAMETERS

A total of 795 calibration parameters were adjusted during model calibration. Nearly all these parameters—783, or 98.5% of all calibration parameters—represented hydraulic conductivities at pilot points. The sand-and-gravel aquifer (Layer 1) was assumed to have a uniform value of horizontal hydraulic conductivity (Kh), however all other model layers were assigned spatially variable conductivities using pilot points. The 12 calibration parameters that were not conductivity pilot points were:

1. Layer-specific values of horizontal-to-vertical anisotropy (Kh/Kv) for layers 1, 2, 3, and 5 (4 calibration parameters),
2. Hydrostratigraphic unit (HSU)-specific value of Kh/Kv for the Bucatunna clay portion of layer 4 (1 calibration parameter),
3. HSU-specific values of storage coefficient for layers 1, 2, 3, 5, and the Bucatunna clay portion of layer 4 (5 calibration parameters),
4. Horizontal hydraulic conductivity (Kh) of layer 1 (as noted above), and
5. A domain-wide scaling factor (multiplier) applied to the prior model's River (RIV) Package conductance values.

Pilot points were used to spatially interpolate layer 4 storage coefficients, however the 215 pilot points used for this purpose (**Figure 12**) were assigned one of two values: pilot points in areas where the Bucatunna clay was present (i.e. red pilot point locations in **Figure 13**) were assigned the Bucatunna clay's storage coefficient, and pilot points in areas where the Bucatunna clay was absent (i.e. the yellow pilot point locations in **Figure 13**) were assigned the layer 5 storage coefficient.

The pilot point-based calibration parameters consisted of:

1. Unique vertical hydraulic conductivity (Kv) pilot points encompassing all of layer 2 (215 pilot points; **Figure 12**),
2. Unique horizontal hydraulic conductivity (Kh) pilot points encompassing all of layers 3 (215 pilot points) and 5 (215 pilot points; **Figure 12**), and
3. Unique vertical hydraulic conductivity (Kv) pilot points encompassing the western and southern portions of layer 4 that represent the Bucatunna clay and Bucatunna clay-to-undifferentiated Floridan Aquifer System (FAS) transition (138 pilot points; red and green locations in **Figure 13**). The 90 undifferentiated FAS pilot points in layer 4 (yellow locations in **Figure 13**) were assigned Kv values based on the Kh of the coinciding layer 5 pilot point and the layer 5 horizontal-to-vertical anisotropy ratio (Kh/Kv), which resulted in layer 4 Kh values that were nearly identical to those in layer 5 where the Bucatunna clay formation is absent.

3.2.1 Initial Values and Ranges

Initial calibration parameter estimates were based on the prior Region II MODFLOW model (HydroGeoLogic, 2000). The riverbed conductance multiplier was initially 1.0, and all layer- or HSU-specific conductivities, anisotropies, and storage parameters were assigned the same values as the prior model. Initial conductivities at pilot points were based on the prior model's conductivity (Kh or Kv, as appropriate) at each pilot point location.

The parameter bounds used during calibration were defined based on parameter type. Hydraulic conductivities, storage parameters, and the RIV conductance multiplier were allowed to vary between 10% and 1000% of the initial value (i.e. initial value +/- one order or magnitude), except for the layer 1 specific yield, which was allowed to vary between values of 0.025 (i.e. 10% of the initial value) and 0.30. Horizontal-to-vertical anisotropy ratios could vary between 50% and 150% of the initial values.

3.3 PROCESSING FRAMEWORK

Existing PEST-compatible pre- and postprocessing utility programs were used to automate all necessary model pre-processing and post-processing steps. The pre-processor programs employed to create the necessary model datasets served one of two purposes: translation of pilot point values to parameter fields, or translation of calibration parameters to model parameters. Post-processing utilities were used to extract (and convert, if necessary) model-generated outputs to simulated quantities directly comparable to calibration targets.

3.3.1 Pre-Processing

3.3.1.1 Hydraulic Property Kriging and Interpolation

The locations and spacing of pilot points used to represent the model's various hydraulic properties were initially selected in order to approximate the Bucatunna clay-to-undifferentiated FAS transition from the prior model (HydroGeoLogic, 2000). All pilot points were regularly spaced throughout most of the model calibration process (**Figure 12**). The PPK2FAC1 and FAC2REAL utilities (Doherty, 2019) were used to calculate kriging factors and interpolate hydraulic properties from kriging factors, respectively, for all properties based on the regularly spaced pilot points. During the course of model calibration the District and Tetra Tech modeling staff determined that introducing supplemental, irregularly-spaced Kv pilot points in layer 4 (**Figure 13**) could allow for greater refinement of the transition from the Bucatunna clay-confined portion of the LFA to the undifferentiated FAS than the original, regularly-spaced points (**Figure 12**). Therefore, the PLPROC utility (Doherty, 2016)—which, unlike PPK2FAC1, allows kriging parameters to vary spatially and produces smooth parameter fields regardless of pilot point spacing—was used to krige and interpolate layer 4 Kv in the final calibrated model. PPK2FAC1 and FAC2REAL perform these functions for all other properties.

3.3.1.2 Calibration Parameter-to-Model Parameter Translations

The PAR2PAR utility (Doherty, 2019) was used to perform the following functions:

1. Scale river boundary conductance values using the riverbed conductance multiplier and write the MODFLOW River Package,
2. Facilitate the conversion of undifferentiated FAS Kh pilot point values to layer 4 Kv pilot point values using the undifferentiated FAS Kh/Kv, and
3. Assign and populate the layer 4 storage pilot point file with Bucatunna clay or layer 5 storage values (as appropriate)[2].

[2] Properties at pilot points where the Bucatunna clay is absent in the eastern portion of model layer 4 were assigned the same values as layer 5.

The TWOARRAY utility (Doherty, 2019) was used to calculate the necessary MODFLOW Kh arrays for layers 2 and 4 from (1) the vertical hydraulic conductivity arrays written by FAC2REAL and PLPROC, respectively, and (2) each layer's respective Kh/Kv anisotropy.

PAR2PAR was also used to calculate model transmissivities at the locations of the 39 APT-based transmissivity estimates. Well screen elevations were used to determine the interval of each APT well in each model layer. The resulting open-interval lengths were used along with the appropriate layer- and cell-specific horizontal hydraulic conductivities to calculate model transmissivities for comparison to the transmissivity targets.

3.3.2 Post-Processing

After each R2MF model was executed using MODFLOW-2005, five post-processing utilities from the PEST suite were used to extract model results for comparison to calibration targets: MOD2SMP, SMP2SMP, MOD2SMPDIFF, SMPDIFF, and BUD2SMP1 (Doherty, 2019).

The MOD2SMP and SMP2SMP utilities were used to, respectively, spatially and temporally interpolate simulated groundwater heads. The MOD2SMPDIFF utility was used to calculate simulated vertical and horizontal head differences between well pairs (using output from MOD2SMP), and SMP2SMP was subsequently used to temporally interpolate the resulting head differences to the head difference target dates. Simulated temporal head differences were calculated from model results using the SMPDIFF utility based on simulated heads at a reference time that was unique to each well (due to varying water level target availability and quality). Simulated incremental changes in baseflows were calculated using the BUD2SMP1 utility and were temporally interpolated to estimate flows at target dates—June 30 of each year—using SMP2SMP.

3.4 CALIBRATION METHODS

Two different parameter estimation algorithms were used during calibration of the R2MF model: the global search Covariance Matrix Adaptation-Evolution Strategy (CMA-ES) algorithm (Doherty, 2019), and the commonly applied, gradient-based PEST algorithm as implemented in PEST++ (Welter et al., 2012). Both algorithms were used during the calibration process to vary parameter values in different, systematic ways, and were executed whenever the parameter values, bounds, model inputs, boundary conditions and/or target datasets were revised. Calibration runs executed during the calibration process are documented in **Appendix C**.

The final set of calibration runs that identified the model parameters documented herein consisted of the following steps:

1. Perform a global search of the parameter space using CMA-ES,
2. Adopt the best parameter set identified by CMA-ES,
3. Manually adjust selected parameter values from the previous step, as needed, and
4. Update the parameters, with a "local search" of the parameter space using PEST++, while holding constant any manually-adjusted parameters from Step 3.

During the final CMA-ES run, Step 3 became necessary to counteract obvious "overfitting" of several Kv pilot points in the Bucatunna clay-to-undifferentiated FAS transition zone. Step 4 was subsequently executed to improve the match between simulated and target values after calibration metrics were degraded as a result of Step 3.

4.0 CALIBRATION RESULTS

The calibration metric goals defined prior to beginning the calibration process and the corresponding calibration metric values calculated for the final calibrated model are presented in **Table 1**. Separate metric goals were defined for the full target dataset and for the subset of targets inside the CR2SWT model domain.

4.1 CALIBRATION PARAMETERS

Parameter bounds and the final (calibrated) values for all pilot point-based hydraulic conductivity parameters are provided in **Appendix D**. The parameter bounds and final values for all non-pilot point calibration parameters are provided in **Table 2**.

Calibrated horizontal hydraulic conductivities of layers 3 and 5 are illustrated in **Figures 14 and 15**, respectively. Corresponding transmissivities of layers 3 and 5 were subsequently calculated and are shown in **Figures 16 and 17**. Calibrated vertical hydraulic conductivities of layers 2 and 4 are illustrated in **Figures 18 and 20**, respectively. The resulting leakances of layers 2 and 4 are illustrated in **Figures 19 and 21**, respectively. The storage coefficient of cells in layer 4 were calculated by first assigning the value of the storage coefficient of the Bucatunna clay or that of the LFA to each of the 215 pilot points at the locations shown in **Figure 12**, and subsequently interpolating from these pilot points using ordinary kriging. The resulting distribution is shown in **Figure 22**.

4.2 CALIBRATION TARGETS

All calibration targets and comparable simulated quantities are provided in **Appendix A**. Observed and simulated well hydrographs at groundwater head target locations are shown in **Appendix E**. Mean and mean absolute head residuals for each target well during the model simulation period are illustrated in **Figures 23 and 24**, respectively. Observed and simulated head values for the entire target dataset are compared in **Figure 25**. Simulated head contours and observed contours (provided by the District) for model layer 3 are shown for two selected years (2000 and 2015) in **Figures 26 and 27**. The mean absolute errors for all VHD targets are shown in **Figure 28**, and the individual target and simulated VHD targets are compared in **Figure 29**. Mean absolute errors for HHD targets are similarly shown in **Figure 30** and a comparison of target and simulated HHD targets is shown **Figure 31**. THD mean absolute errors are shown in **Figure 32** and a comparison of the target and simulated THD targets is shown in **Figure 33**. Pre-development head contours from Bush and Johnston (1988) are compared with the 1942 layer 3 heads simulated by the calibrated R2MF model in **Figure 34**, and the target and simulated pre-development targets are compared in **Figure 35**. APT-estimated transmissivities are compared with model transmissivities in **Figure 36**.

4.3 WATER BUDGETS & FLUXES

Domain-wide water budget components from the end of the pre-development, 2000, and 2015 stress periods are summarized and contrasted in **Table 3**. Pre-development, 2000, and 2015 rates of leakage from layer 2 (IAS) to layer 3 (UFA) are shown in **Figures 37, 38, and 39**, respectively. Negative values in **Figures 37-39** indicate upward flow from the UFA to the IAS, and positive values indicate downward flow and UFA recharge.

5.0 DISCUSSION

5.1 CALIBRATION METRIC GOALS

The ability of any model to replicate field observations cannot be known with any degree of certainty prior to commencing the calibration process. Despite this uncertainty, however, all but one of the predefined target metric goals was achieved by the final calibrated model (**Table 1**). The one target metric goal that was not met—the mean absolute error (MAE) goal for groundwater heads inside the CR2SWT model subregion—was within 0.3 ft of the 5 ft MAE goal[3]. All other target types—including the three head difference target types (i.e. VHDs, HHDs, and THDs)—achieved much lower calibration metric statistics than the predefined calibration goals.

5.2 CALIBRATED PARAMETERS

The final values of calibration parameters not defined using pilot points (**Table 2**) reveal several patterns. The transient recalibration process resulted in refinements to the prior model's horizontal-to-vertical anisotropy estimates, which were 1.0 for layer 1 and 35.0 for all other layers (HydroGeoLogic, 2000). Model-wide anisotropies in the two confining units—the IAS (Kh/Kv = 32.0) and Bucatunna clay (Kh/Kv = 36.9)—are comparable to those from the original R2MF model. Anisotropies in the Floridan aquifer layers (3 and 5) were reduced substantially to 2.04 and 17.9 (or by approximately 94% and 49%), respectively, during calibration which suggests that the prior model's assumption of identical anisotropies in layers 2-5 may have been an oversimplification of the system's properties.

The storage coefficients of layers 2-5 all increased and the specific yield of layer 1 decreased relative to the original R2MF model's uncalibrated values (**Table 2**). The calibrated specific yield of the sand-and-gravel aquifer (0.0252) is near the lower limits of both the specified parameter range and the estimated range for sands (Anderson and Woessner, 1992). Due to the fact that (1) storage coefficients were included in the transient model calibration (unlike in the original steady-state R2MF model), and (2) the transient R2MF model matches field observations quite well, the calibrated storage coefficients derived during the transient calibration represent more thoroughly tested and proven—and therefore more defensible—estimates than those used in the prior model.

Hydraulic conductivities at pilot points (**Appendix D**) generally increased over the course of the calibration process. Of the 215 pilot points in layers 2, 3, and 5, Kh increased at 123 (57%) layer 3 pilot points and at 112 (52%) layer 5 pilot points, and Kv increased at 109 (51%) layer 2 pilot points. Horizontal conductivities and transmissivities in layers 3 and 5 vary spatially; however, the only clear spatial pattern in either layer is that both have higher conductivities and transmissivities in the east-central portion of the model domain (**Figures 14-17**), which is consistent with the available APT estimates. The Bucatunna clay-to-undifferentiated FAS transition in layer 4 is reflected by the relatively sharp transition from Kv values of approximately 1.0E-07 ft/d to Kv values more than one million times higher (i.e. greater than 0.1 ft/d) over horizontal distances of approximately 5-10 miles (**Figure 20**). This sharp transition from where the Bucatunna clay confines the LFA to the undifferentiated FAS is consistent with the conceptual model and was present in the original R2MF model (HydroGeoLogic, 2000).

5.3 WATER BUDGETS & FLUXES

The instantaneous domain-wide water budgets from the end of the pre-development, 2000, and 2015 stress periods (**Table 3**) reveal several noteworthy patterns. First, the first (36,500 day) stress period does approximate a steady-state, pre-development condition; the pre-development storage inflow and outflow rates at the end of this period are small relative to both the other concurrent pre-development fluxes and storage fluxes in the other selected periods. Recharge is the dominant source of water (i.e. inflow) to the model and the largest outflows during all three periods shown are to the River and Drain boundaries that represent rivers and evapotranspiration, respectively.

The effects of increased pumping between pre-development and 2000 conditions increased inflows from and reduced outflows to both GHB and CHD boundaries. The overall ("net") change in GHB flow (4.03E+05 ft^3/d; **Table**

[3] Note: The CR2SWT model subregion groundwater head MAE goal (5 ft) was achieved earlier in the calibration process. However, with the addition of the pre-development head targets to the calibration target dataset and subsequent calibration runs, the 5 ft MAE goal was not achieved.

3) accounted for approximately 7% of the net increase in pumping from pre-development to 2000 (5.75E+06 ft^3/d). The reductions in net CHD outflow (1.65E+06 ft^3/d) and net RIV outflow (2.24E+06 ft^3/d) from pre-development to 2000 respectively accounted for approximately 29% and 39% of the pumping increase during this period (**Table 3**). Cumulative net leakage from the IAS (layer 2) to the UFA (layer 3) increased by approximately 2.17E+06 ft^3/d from pre-development (**Figure 37**) to 2000 (**Figure 38**), which represents approximately 38% of the net difference in pumping between pre-development and 2000 (5.75E+06 ft^3/d).

The overall reduction in net pumping of approximately 17% from 2000 (5.75E+06 ft^3/d) to 2015 (4.79E+06 ft^3/d) reversed some of the pre-development to 2000 trends (**Table 3**). CHD outflows increased and GHB and CHD inflows decreased between 2000 and 2015. The reduction in inflows from CHDs during this period is evident when comparing the magnitude and direction of leakage between the IAS and UFA in 2000 (**Figure 38**) and 2015 (**Figure 39**): the red area beneath the Gulf of Mexico and adjoining bays (indicating upward flow) is larger in 2015 than in 2000, and the downward leakage rates in 2015 have generally lower magnitude than in 2000. Reduced pumping in coastal Bay County after 2000 led to the reestablishment of the upward hydraulic gradient—present during pre-development—by 2015.

5.4 CALIBRATION TARGETS

5.4.1 Groundwater Heads

The calibrated model shows minimal bias in simulated post-development groundwater heads domain-wide (mean error, "ME" = 0.40 ft; **Table 1**) and within the CR2SWT model subregion (ME = -0.08 ft). As with the ME statistics, post-development groundwater head MAE statistics indicate a better overall match to observations in the CR2SWT model subregion (MAE = 5.28 ft) than domain-wide (MAE = 5.93 ft). Examination of the ME and MAE results by well for the entire simulation period also indicate minimal spatial bias: positive and negative errors are interspersed (**Figure 19**) and wells with large MEs and MAEs are interspersed with wells with lower errors (**Figures 23 and 24**).

5.4.2 Head Differences

Observed vertical head differences are reproduced well by the calibrated R2MF model based on the VHD MAE/range metrics in **Table 1** (6.78% domain-wide and 6.20% in the SEAWAT domain). The overall directions of the vertical gradients are simulated accurately (**Figure 29**), although the model tends to underestimate the magnitude of both positive and negative head differences. There does appear to be some degree of spatial bias in the VHD targets; the well pairs with the greatest mean absolute errors are in the eastern portion of the domain and are from well pairs in layers 1 and 3 or layers 2 and 3 (**Figure 28**).

Horizontal head differences between the well pairs shown in **Figure 30** are simulated very well based on the observed versus simulated comparison (**Figure 31**) and the MAE/range metrics in **Table 1** (3.35% domain-wide and 5.28% in the SEAWAT domain). Inspection of mean absolute errors of the individual HHD target well pairs (**Figure 30**) reveals that all layer 3 HHD well pairs have MAEs below 8 ft and that three of the four layer 5 well pairs have MAEs below 5 ft.

Among the three types of head difference targets, THDs are generally replicated the best by the calibrated R2MF model. The MAE/range metrics (2.07% in the SEAWAT domain and 1.88% domain-wide; **Table 1**) indicate that simulated temporal differences in heads match the observed differences extraordinarily well. The greatest mean absolute THD errors are near the center of the pumping-induced cone of depression near Fort Walton Beach, FL (**Figure 32**). This apparent spatial bias is not unexpected, or indicative of poor model performance, given that the temporal head differences of the greatest magnitudes were observed in this area (see well hydrographs in **Appendix E**).

5.4.3 Baseflow Targets

The qualitative incremental baseflow targets for the 1950-1978 and 2006-2015 periods are not matched as well as the other calibration target groups. However, this result was to be expected considering that weights of zero were given to all baseflow targets during calibration. Although the MAE/range is approximately 15% (**Table 1**) the temporal variability in incremental baseflows is not captured by the model (**Appendix A**). Simulated baseflows generally decrease over time (in response to the general increase in groundwater pumping) and range from 738-753 ft^3/s (average = 748 ft^3/s). The estimated baseflow targets have a much wider range (338-1764 ft^3/s) and higher average (861 ft^3/s). However, it is noteworthy that baseflow separation analyses and the resulting baseflow estimates typically have large uncertainties associated with them.

5.4.4 Pre-Development Heads

The pre-development head targets were added late in the calibration process; therefore, these targets were not assigned pre-calibration target metrics (**Table 1**). However, the pre-development heads have similar mean absolute errors (i.e. less than 6 ft) to the post-development, measurement-based head targets. The pre-development targets are, on average, underestimated slightly with mean errors of approximately 2 ft in the SEAWAT domain and domain-wide. This bias, which is relatively small compared to the approximate 110 ft range in pre-development head target values, is apparent in the comparison of observed and simulated target values (**Figure 35**) since many more targets fall below the perfect fit line.

Pre-development heads are generally reproduced well by the calibrated R2MF model (**Figure 35**), particularly along the CR2SWT model boundary (**Figures 34 and 35**).

5.4.5 Transmissivity Targets

Pre-calibration metric goals were not defined for transmissivity targets because, like the pre-development head targets, these targets were added after the calibration process began. The 39 APT-based aquifer transmissivity targets are generally reproduced by the calibrated R2MF model horizontal hydraulic conductivity distributions (**Figure 36**); all modeled transmissivities are within one order of magnitude of the APT estimates. The overall replication of these targets indicates that the calibrated R2MF model has properties that are consistent with the actual hydraulic properties near all available APTs performed in the model domain.

6.0 CONCLUSION

The R2MF model was calibrated to replicate mean annual groundwater levels, spatiotemporal differences in mean annual groundwater levels, pre-development groundwater levels, and APT-derived transmissivity estimates. All targets are replicated reasonably well, and the three spatiotemporal difference target groups—vertical, horizontal, and temporal head differences—are reproduced very well. The subset of groundwater heads in the area of greatest current interest to the District—the area within the CR2SWT model domain—are generally reproduced as well as or better than the overall head target dataset by the calibrated model. Qualitative comparisons of the simulated Upper Floridan aquifer potentiometric surface and Choctawhatchee/Holmes baseflows under various pumping conditions with observed interpolations and estimates are also generally agreeable.

The calibrated R2MF model's ability to accurately simulate heads and head differences through space and time demonstrates it is a suitable tool for defining boundary conditions for the sub regional CR2SWT model and providing initial estimates of aquifer parameters for the sub-regional model. The R2MF model may also be a suitable tool for assessing the effects of new withdrawals, especially within the existing UFA cone of depression where the model has demonstrated the ability to simulate–within acceptable limits—the effects of historical pumping.

7.0 REFERENCES

Anderson, M.P., and Woessner, W.W., 1992. Applied Groundwater Modeling – Simulation of Flow and Advective Transport. Academic Press, Inc.

Bush, P.W., and Johnston, R.H., 1988, Ground-Water Hydraulics, Regional Flow, and Ground-Water Development of the Floridan Aquifer System in Florida and in Parts of Georgia, South Carolina, and Alabama: U.S. Geological Survey Professional Paper 1403-C, 80 p., 17 plates.

Doherty, J., 2016. PLPROC: A Parameter List Processor, User Manual. Watermark Numerical Computing and National Centre for Groundwater Research and Training (Australia), 244 p.

Doherty, J., 2019. PEST: Model-Independent Parameter Estimation, User Manual, 7th Edition. Watermark Numerical Computing, 393 p.

Environmental Simulations, Inc., 2017. Groundwater Vistas, Version 7.

Langevin, C.D., Shoemaker, W.B., and Guo, W., 2003, MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model—Documentation of the SEAWAT-2000 Version with the Variable-Density Flow Process (VDF) and the Integrated MT3DMS Transport Process (IMT). U.S. Geological Survey Open-File Report 03-426. 52 p.

Harbaugh, A.W., 2005, MODFLOW-2005, the U.S. Geological Survey modular ground-water model -- the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16.

Harbaugh, A.W., and McDonald, M.G., 1996, User's Documentation for MODFLOW-96, an update to the U.S. Geological Survey Modular Finite-Difference Ground-Water Flow Model: U.S. Geological Survey Open-File Report 96-485. 63 p.

HydroGeoLogic, Inc. 2000. Modeling of Ground Water Flow in Walton, Okaloosa and Santa Rosa Counties, Florida. Technical report prepared for Northwest Florida Water Management District. 117 p.

HydroGeoLogic, 2005. *Saltwater Intrusion in the Floridan Aquifer in Walton, Okaloosa, and Santa Rosa Counties, Florida: Western Model Domain Final Report*, HydroGeoLogic, Inc., Herndon, Virginia.

HydroGeoLogic, 2007. *Saltwater Intrusion in the Floridan Aquifer in Walton, Okaloosa, and Santa Rosa Counties, Florida: Eastern Model Domain Final Report*, HydroGeoLogic, Inc., Herndon, Virginia.

Welter D. E., Doherty J. E., Hunt R. J., Muffels C. T., Tonkin M. J., Schreuder W. A., 2012, Approaches in Highly Parameterized Inversion: PEST ++, a Parameter ESTimation Code Optimized for Large Environmental Models, U.S. Department of the Interior, U.S. Geological Survey. Techniques and Methods, Book 7, Section C5