

EXECUTIVE SUMMARY

INTRODUCTION

Tampa Bay Water is Florida's largest wholesale water supplier. Tampa Bay Water's member governments – Hillsborough, Pasco and Pinellas Counties and the cities of New Port Richey, St. Petersburg and Tampa – serve more than 2 million residents in the Tampa Bay region. Tampa Bay Water currently operates twelve wellfields. Eleven of these wellfields are operated as an integrated system under the Optimized Regional Operations Plan (OROP). The coordinated operation of these wellfields is guided by an operational optimization model that determines the pumping rates required to satisfy forecasted demands at eleven points of connection while emphasizing that environmental impacts are minimized.

A key component of the OROP system is the integrated hydrologic simulation model of the Central Northern Tampa Bay area (CNTB) developed by SDI Environmental Services, Inc. (SDI). The CNTB model employs SDI's ISGW (Integrated Surface water Ground Water) software which couples surface water hydrology, as simulated by the HSPF model (Hydrologic Simulation Program – FORTRAN; Johanson et al., 1984), and groundwater hydrology, as simulated by MODFLOW (McDonald and Harbaugh, 1988). Given its important role in Tampa Bay Water's wellfield management system, understanding and quantifying the uncertainty of predictions made by the CNTB model is a high priority to ensure long-term sustainability of the water resource.

ISGW Model Application

The CNTB model covers an area of 3,600 square miles and incorporates all eleven Tampa Bay Water wellfields comprising the OROP as well as the SC Hillsborough wellfield. HSPF simulates the surface water system using surface topography of the region as well as the region's hydrography represented as reaches of the numerous surface water bodies represented in the model. MODFLOW simulates water levels and groundwater flow directions and velocities in two principal hydrostratigraphic units, the Surficial Aquifer System (SAS) and the Upper Floridan Aquifer System (UFAS). The UFAS is confined

over much of the model area by the Intermediate Confining Unit (ICU), which is not explicitly discretized within the MODFLOW model. Hundreds of ICU 'holes' are probably present over the model domain. These holes are formed by collapses or sinkholes and often result in direct communication between surface water bodies and the UFAS hydrologic systems.

The hydrologic system within the CNTB model domain was simulated from 1971 through 1999 using one-week time intervals as described above. The surface water system was simulated using 119 sub-basins and 75 reaches. The groundwater system was simulated using two layers (the SAS and the UFAS), 153 rows and 152 columns with cell dimensions ranging from one-quarter mile to one-mile.

Study Objective

The two objectives of this study are to:

1. Conduct a comprehensive review of the spatial and temporal characteristics of hydrologic inputs and physical properties used to develop the CNTB model as well as provide an error analysis of predicted stream discharge, groundwater hydraulic head, and drawdown.
2. Develop recommendations and a scope of work to enhance the model calibration, improve the model predictive capabilities, and if possible, quantify the uncertainty in model predictions.

Scope of Work

The approach used in this study focused upon the analysis of model prediction errors and the sensitivity and correlation of model parameters to the calibration statistics to ascertain if the model suffers from over-parameterization, a concern presented in an earlier uncertainty review of the ISGW model (Jacobs and Gelhar (2000)). The approach accomplishes the following objectives:

- Resolves the issue of potential model over-parameterization so that any limitations with the present model will be clearly identified and understood,

- Provides an estimation of the uncertainty in the most important model parameters so that the reasonableness of the current model can be understood,
- Provides all the necessary background information and guidelines for Phase III tasks.

The approach tests the predictive power of the model by examining its basis (underlying data) and adequacy as an approximation of reality and its performance with both calibration and a validation data set of surface water stage height and groundwater levels from the SAS and UFAS over the period of 1997, 1998 and 1999. Since the CNTB model only used data through 1996 as its calibration target set, the measured data after 1996 provides an excellent record from which the model's predictive capability could be evaluated.

This scope of work was therefore organized in the following general sequence:

- Spatial and temporal analysis of model prediction errors for the surface water and groundwater systems from 1997 through 1999,
- Model parameters' evaluation and analysis,
 - Evaluation of precipitation data and ET quantification
 - HSPF input-parameter evaluation and development of parameter ranges
 - MODFLOW input-parameter evaluation and development of parameter ranges
 - Spatial and temporal range of water-level hydrographs
- Sensitivity analysis
- Scope of work for subsequent study

SPATIAL AND TEMPORAL ANALYSIS OF MODEL PREDICTION ERRORS

Calibration and Validation Accuracy

Jacobs and Gelhar (2000) point out that the current CNTB121 model may suffer from over-parameterization. This can lead to the condition where the model appears to be well calibrated, but actually has limited predictive power. This hypothesis was tested by

comparing the model predictive accuracy over the calibration period (1971-96) and the validation period (1997-99).

During calibration, the CNTB121 model used hydrologic observations that spanned a period of about 26 years starting in January 1971 and ending in December 1996. Observations between January 1997 and December 1999 were used for model validation. Over-parameterization is evident when the model accuracy for the calibration period is much better than the model predictive accuracy for the validation period. In order to conduct this comparison the following questions were posed:

- Are these two periods hydrologically similar?
- Is model performance consistent over the two periods?
- Are there any temporal characteristics of model prediction errors?
- Are there any spatial patterns of model prediction errors?

The first question aims to evaluate whether the hydrologic conditions during the validation period are similar to those of the calibration period. For example, if groundwater levels observed during the validation period are much lower than those observed during the calibration period, any discrepancy in model predictions accuracy might be attributed to the fact that the evaluations do not consider similar conditions. Therefore, it is important to provide quantitative evaluations of the similarities (and differences) between these two periods. Once the calibration and validation periods are compared, the second question aims to provide a quantitative assessment of whether over-parameterization is a problem and the extent of over-parameterization effects.

Each question was answered in the context of the surface water system (by evaluating stream discharge rates) and the groundwater system (by evaluating groundwater hydraulic heads). The general finding was that the CNTB-121 model appears to provide reasonable accuracy in terms of predicting long-term average hydrologic system evolution. Furthermore, the model validation results clearly show that the model predictive accuracy is reflective of the model's calibration accuracy. While the model does not display symptoms of over-parameterization, the model's predictions are highly questionable both for the short term (weeks) and for dry conditions. The model's predicted short-term water-level changes are highly questionable which makes it very unreliable to use within the OROP program.

MODEL PARAMETERS' EVALUATION AND ANALYSIS

The analysis of model's prediction errors raised the question of whether model input parameters are properly derived from the physical system. An evaluation of the basic data that was used to generate model input parameters and their relationships to supporting data and the physical hydrologic system was conducted. For the surface water model, the issues involved evaluation of (1) whether the parameter values are reasonable and within expected ranges for the CNTB region, and (2) whether the parameter values reflect land use conditions on the model segments (sub-basins). Two major issues were identified for the surface water model parameters. First, some parameters did not appear to be directly derived from the physical system. Parameters such as overland flow slope and hydraulic length should not be adjusted during calibration. These parameters should be calculated using the most detailed topography information available. Then, they should be kept constant throughout the calibration process. Other parameters, such as Manning's roughness coefficient, should be assigned a range of reasonable values based on land use information and reported values in the literature as well as any local information. During calibration, such parameters should only be allowed to vary within these established ranges.

The other identified model parameterization issue was that some parameters had values that were well outside their reasonable ranges. For example, the upper zone nominal storage parameter (UZSN) and infiltration coefficient (INFILT) had seasonal values that were much higher than typical values, even for similar hydrologic conditions. The need for such values during calibration typically indicates either other model parameterization problems or other conceptualization problems. For example, model parameter values could be adjusted to account for the lack of impervious land segments or interflow. It is clear that model re-calibration is necessary to adjust such parameters.

For evapotranspiration (ET) and precipitation, it was found that more data is needed for both precipitation and ET to guide the model parameterization. For precipitation, short-term needs include re-evaluation of the rainfall disaggregation methodology currently used to distribute precipitation across the model domain spatially and temporally. In the long term, more hourly rain gages are needed throughout the model domain. Calibration

of NEXRAD data using rain gages can greatly enhance the temporal and spatial rainfall representation within the model.

For ET, it was determined that the Holdridge method of ET estimation needs to be replaced with another ET estimation technique such as FAO-56 (Allen et al., 1998). In the long term, more satellite data should be processed to estimate ET over different seasons and for a reasonable amount of time using the SEBAL method. We recommend that at least two estimates are obtained each season for about 2 to 3 years. Once these ET estimates are available, a thorough evaluation of the spatial ET variability must be carried out in order to define the relative contributions of different systems to ET. This will greatly enhance the spatial and temporal ET representation within the model.

The groundwater component of HSPF is not used to simulate the groundwater regime because its simplistic representation is not suitable for Tampa Bay Water's water-supply management needs. Instead, MODFLOW (McDonald and Harbaugh, 1988), a distributed parameter finite-difference flow model, is used. In general, distributed models require extensive amounts of data that is never practically obtainable which results in prediction uncertainty. An evaluation of the groundwater model input parameters addressed the following issues:

- Are the aquifer data (estimates of SAS hydraulic conductivity, ICU leakance, UFAS transmissivity, UFAS anisotropy, and SAS and UFAS specific yield and storativity) representative of the aquifer?
- Has the model reasonably incorporated the field data?
- What flow model uncertainty may be associated with the parameter uncertainty?
- Is the flow model over-parameterized?

The general conclusions were:

- The aquifer data (estimates of SAS hydraulic conductivity, ICU leakance, UFAS transmissivity, UFAS anisotropy, and SAS and UFAS specific yield and storativity) is probably representative of the part of the aquifer tested. However, the actual UFAS thickness represented in the model may be much greater requiring higher than observed transmissivities to calibrate the model. There are significant uncertainties associated with all the characterization data.

- The model has reasonably incorporated the transmissivity and storativity field data. Similarly, the scarce anisotropy data is reasonably honored. Leakance is, however, more problematic, and because of the presence of sinkholes probably has a very large uncertainty associated with it.
- The flow model uncertainty associated with the parameter uncertainty ranges from 0.5 to 1.2 ft, at a minimum.
- Because the model average water level and discharge predictions are reasonable during the validation period, the model does not appear to be over-parameterized.

SENSITIVITY ANALYSIS

Sensitivity analysis involves varying model input parameters and evaluating how model results change with these variations. Sensitivity analysis provides valuable understanding of both model implementation and the underlying physical processes thus providing insight into both system and model behavior. Furthermore, results from sensitivity analyses can guide both model calibration and the prioritization of future data collection plans.

Surface Water Model Parameters

The analysis for the principal surface-water model parameters focused on determining the sensitivity of model results to the HSPF parameters. To perform this analysis, the CNTB-121 model was run several times, varying one parameter at a time, and the model results were examined. Two model runs were performed after all parameters were adjusted simultaneously. This analysis provided some insight into whether each parameter has the effect to change model predictions significantly. It also provided some insight into which parameters should be more carefully considered to re-calibrate the integrated model.

The results were analyzed at the 25 flow gage sites that were used for calibration, covering the period 1976-1996. Most of the parameter changes that were explored caused significant changes in simulated discharge throughout the CNTB area, especially for low flow conditions. The correlation between observed and simulated stream

discharge at each calibration location was virtually the same for all sensitivity runs. Furthermore, the model predictions for the groundwater system during dry conditions were also affected for several of the surface water model parameters. Since all tested changes are recommended in order to make the parameter values more realistic, a recalibration effort of the HSPF component of ISGW will be required. These changes as follows:

1. For physically-based parameters (NSUR, SLSUR, LZ RAT) use DEM and land-use coverages to calculate them, rather than allowing them to be adjusted during calibration
2. Limit other parameters (USZN, ..) to more plausible values.

Making the above changes will increase confidence and defendability of the ISGW model by basing parameters, where possible, on accepted physiographic data.

Groundwater Model Parameters

Model calibration and validation statistics were used to compare and contrast the impact of selected groundwater parameter changes as a means toward investigating model conceptualization. This also provided a quantitative assessment of the model parameters most affecting model calibration and prediction. The emphasis was on providing insight into model behavior to guide future re-calibration efforts. For all the parameters investigated in this sensitivity analysis, tables of the calibration metrics are provided for select changes in model parameters. In addition, for some parameters, a spatial map of sensitivity was generated.

Within the limits of the perturbations evaluated in this sensitivity analysis, the most sensitive parameters were determined to be the SAS specific yield, UFAS transmissivity, and ICU leakance with respect to SAS head calibration and prediction. Since there is little SAS specific yield data available within the study area, it will be difficult to reduce the model's uncertainty unless additional data for the SAS specific yield as well as UFAS transmissivity and ICU leakance is obtained. While it is true that there are physically plausible limits on specific yield, conceptually it can be used to adjust for water balance errors. Compensating errors in leakance and transmissivity must be introduced to

balance the impact of any error in the water balance due to uncertain specific yield estimates.

There is inherent uncertainty in the parameter values estimated from pumping tests. Because water budget data was used in model calibration these uncertainties do not necessarily mean that the model head error will be of the same order of magnitude. For instance, leakance is probably subject to an order of magnitude uncertainty. The model is obviously calibrated to within an order of magnitude, at least in terms of the average measures used.

Sensitivity Analysis Summary

The two codes linked to create ISGW, HSPF and MODFLOW, require specification of a large number of parameters. A qualitative summary of the sensitivity analysis results is presented in the table below. The parameters with the highest sensitivity for model predictions are UZSN (upper zone nominal storage), SAS and UFAS specific yield, and ICU leakance. Three of the four most sensitive parameters are those that govern the storage of water in the hydrologic system. This is intuitively correct since the CNTB-121 model is transient. The ramification of this sensitivity is that water balance errors can be compensated for by adjusting storage parameters, which are not exhaustively, and in the case of UFAS specific yield not at all, known. Thus, some element of uncertainty will always exist with respect to these parameters' impact on predictions.

Sensitivity Analysis Summary

Model Parameter	Model	Comment
INFILT	HSPF	Low overall sensitivity. Mostly affects local surface runoff values.
UZSN	HSPF	High overall sensitivity
NSUR	HSPF	Low overall sensitivity. Mostly affects local surface runoff values.
LZETP	HSPF	Low overall sensitivity
SLSUR	HSPF	Medium overall sensitivity
LSUR	HSPF	Medium overall sensitivity
SAS Hydraulic Conductivity	MODFLOW	Low overall sensitivity
UFAS Transmissivity	MODFLOW	Medium overall sensitivity
UFAS Anisotropy	MODFLOW	Low overall sensitivity. Local sensitivity at Crossbar Ranch wellfield.
SAS and UFAS Specific Yield	MODFLOW	High overall sensitivity
ICU Leakance	MODFLOW	High overall sensitivity
UFAS Storativity	MODFLOW	Low overall sensitivity

Suggested Subsequent Work

Recommendations of work subsequent to this study have been developed by considering the preliminary suggestions of Jacobs and Gelhar (2000), the findings of this study, and the short term (operational) and long term (planning) needs of Tampa Bay Water. Our recommended scope of work (SOW) requires two distinct sets of tasks. The general objective of our first set of tasks focuses on recalibrating the ISGW model to provide Tampa Bay Water with a wellfield management tool capable of making adequate forecasts of long-term water resource assessments as well as forecasts of short-term operational impacts to the water resource system. If, at the end of the re-calibration, the model is still unable to make short-term operational forecasts with sufficient accuracy,

the second set of tasks is suggested to overcome the model's shortcomings for operational use.

The suggested first set of tasks includes:

- Task 1. Define the surface water and groundwater model performance measures (calibration metrics) for the full range of the hydrologic conditions represented by the model
- Task 2. Update the precipitation implementation method
- Task 3. Develop improved spatial and temporal estimates of ET
- Task 4. Update the HSPF surface water model parameters
- Task 5. Update the MODFLOW groundwater model parameters
- Task 6. Recalibrate the ISGW model using automated calibration techniques that provide a direct method of quantifying uncertainty. Evaluate model the uncertainty in predictions using either direct or indirect methods, depending on the final model's computational needs.

Once these recommendations are implemented in a re-calibration of the model, the predictive uncertainty of the model may then be re-assessed. These changes should enable the model to provide improved predictions. If the changes do not improve the short-term operational predictions of the model then it is our opinion that the model should not be used directly in Tampa Bay Water's operational program. Instead, another approach should be considered.

Our second set of recommendations provides some recommendations for an alternative approach. These recommendations assume that Tampa Bay Water would desire to continue to use the ISGW model for long-term planning only and would apply alternative models for operational purposes. Our recommendations are to follow one of the following courses of action:

- 1) Develop either local numerical or analytical groundwater models with detailed spatial discretization over each well field that would integrate the boundary fluxes provided by the regional ISGW model onto the local wellfield model domain. This process could be automated/streamlined to facilitate the nested calculation that

would be necessary for operational purposes. An advantage to this approach is that continuity with ISGW long-term forecasts would be maintained. However, depending on the local accuracy, the local models may require extensive modification. Furthermore, the effects of pumping at one wellfield on water levels at other wellfields must be evaluated to define local areas associated with each group of pumping wells.

- 2) Develop multiple linear regression or artificial neural network (ANN) models for each of the wellfields. ANNs have been shown to provide robust predictions of highly non-linear responses such as surface water impacts due to deeper aquifer pumping. This approach has the advantage that once explanatory variables and network have been defined no further evaluation is required, and the ANN can be periodically retrained to gain accuracy over time. The ANN can be trained using the field data, with no intermediate physics-based model necessary.

The first approach is the one classically used. In principle, if the local models were continually improved they would gain predictive accuracy in much the same way as an ANN would. However, the labor involved would be much greater. On the other hand, once ANNs are developed and their accuracy was reasonable, collected data over time could be used to make the ANNs more accurate with minimal labor.

CONCLUSIONS

The principal findings from these analyses are summarized as follows:

- The CNTB model provides reasonable predictions of mean or average hydrologic conditions over a time scale of years.
- Since the model has reasonable mean predictions of hydrologic states over the validation time period (1997 through 1999), the model is not considered overparameterized.
- The CNTB model's predictions of the hydrologic system during dry or drought conditions (low water levels, low stream discharge, etc) are poor.
- The accuracy of the CNTB model under dry, average or wet hydrologic conditions during the verification or predictive data set (1997 through 1999) is consistent with

the accuracy of the CNTB model calibration data set (1971 through 1996) under dry, average or wet hydrologic conditions.

- The prediction accuracy of water-level changes within the CNTB model is very poor on weekly and monthly time scales and acceptable on a yearly time scale.
- Potential ET within the model is very poorly represented. The Holdridge method used to determine PET input to the model underestimates PET during the summer months by as much as 50% and overestimates PET in the winter months by as much as 30%. In addition, there appears to be very little correlation between ET and land use. This results in poor spatial and temporal variability of ET in the model.
- The disaggregating method for distributing daily rainfall at rain gauges to hourly rainfall has not been substantiated. The Thiessen polygon method for distributing the hourly rainfall across the subbasins leads to significant estimation errors. Combined, these effects lead to significant errors and uncertainty in the rainfall used in the model.
- NEXRAD radar could help reduce the errors associated with the Thiessen polygon method by defining the rain/no rain boundary. This would restrict the distribution of rainfall data to only those areas or subbasins of the model that lay within the rain boundary.
- Several HSPF model parameters have unreasonable values and/or lack strong correlation to land topography. These parameters include overland slope, hydraulic length and Manning's roughness. In addition, the upper zone nominal storage, infiltration and lower zone potential ET parameters are outside of the practical ranges for these parameters.
- Substantial uncertainty exists in the observed transmissivity values subsequently used to define transmissivity in the model. The source of uncertainty stems from difficult to interpret pumping tests, scale issues related to fractured rock conditions, and screened intervals that only represent part of the UFAS.
- Anisotropy in the fractured and dissolved UFAS is probably ubiquitous, but little data exists on its value, direction, and trends. Consequently large uncertainty exists in anisotropy within the UFAS transmissivity field.
- Substantial uncertainty exists in leakance between the SAS and UFAS.
- The UFAS water level uncertainty associated with the parameter uncertainty (heterogeneous transmissivity values) ranges from 0.5 to 1.2 ft, at a minimum.

- The model requires recalibration prior to any formal uncertainty analysis of model output. Therefore, the subsequent SOW involves model recalibration followed by uncertainty quantification.