Appendix 3-B

Model Calibration Plan, Updated 2013 version and May 2016 Recalibration Plan

Appendix 3-B

COHYST Model Calibration Plan

Prepared for the COHYST Sponsors

Final Draft, November 2010

Lee Wilson
Larry Land
Duane Woodward
Doug Hallam
John Engel
Marc Groff

Introduction

The COHYST Model consists of four components -- a watershed response model, a groundwater model, a surface water model, and an integrated model. Refer to the 2010 Work Plan for brief descriptions of each model, and their interrelationships. A primary application of the COHYST Model will be to quantify time and space impacts to stream flow and the water table that will likely result from changes in the use and management of water. As such the primary basis for calibration of the model will be its ability to reasonably reproduce historic changes in stream flow and water table throughout the study area and over an extended timeframe.

The calibration will performed initially for conditions in 1985-1997, and the models so calibrated will be verified against conditions in 1998-2005 and recalibrated if necessary. The determination of whether the models "reasonably reproduce historic changes" will be reached as follows.

- Results from each model will be compared to observed data using statistical techniques such as calculating the sum of the squared differences between observation and simulated values to determine whether a calibration adjustment improved model performance. There will be no predetermined error bar for an acceptable result.
- Errors and residuals for each model will be plotted as time-series and as maps, to identify locations and times when the model results systematically depart from observations.
- Assessment of model calibrations will ultimately depend on the professional judgment of the modeling team, the COHYST Technical Committee, and the COHYST Sponsors. Judgment will require that the model results make sense -- that the simulated and observed hydrology are similar in matters such as timing of change, magnitude of change, and rate of change. For example, if the average model baseflow and calculated baseflow values are similar, but the changes in baseflow are timed very differently in the model than in the data (as was the case for the prior COHYST models), the groundwater model will not be considered calibrated.
- Fundamental to this approach is that calibration is a combination of rigorous quantitative comparisons and the flexibility to make changes when such are indicated by the calibration results and the modelers understanding of system hydrology. The modeling team recognizes that there are uncertainties in all aspects of the models, and that calibration can simply mean that different uncertainties cancel each other out. This is not inappropriate, as the result should be to bracket and constrain the net uncertainty. Moreover, by calibrating the models individually, but then collectively, the chances for a major modeling error should be reduced.

The technical team will propose to end the calibration process when there is agreement that the models are fully functional and hydrologically reasonable, and substantive improvements in calibration are not likely. This judgment will be subject to peer review and approval by the Sponsors.

The following sections provide more detail on the independent calibrations of each model and also identify quality assurance checks that will be used to verify that the interactions (handshakes) between models are correctly transferring data.

Watershed response model

A water use model (CROPSIM) will calculate the field-scale water budget based on climate, land use, crop types and irrigation practices. Specifically, the fate of water reaching the land surface (from precipitation, canal deliveries or well pumping) will be partitioned into consumptive use, infiltration/recharge and field-edge runoff. Calibration of the watershed functions will be based primarily on rational adjustments to the annual water budget calculations and monthly distributions to improve the calibration of the surface and ground water models.

To the extent data are reasonably available, calibration also will occur through matching model results and data on partitioning. For example, evapotranspiration (consumptive use) will be compared to previous research estimates published in such documents as University of Nebraska-Lincoln NebGuides. Groundwater pumping will be compared to recorded measurements and estimates from Natural Resource Districts.

Field-edge runoff will be further partitioned into additional consumptive use, infiltration/recharge, and volumes which ultimately become streamflow. Calibration of these additional partitioned volumes will be based on the overall water budget comparison and on consistency with observed runoff data. Gages to be used in calibration may include any of those listed in the Work Plan, but it is expected that tributary gages will prove most useful.

Groundwater model

The groundwater component of the Project will consist of a MODFLOW simulation of the Central Platte River encompassing the area previously included in the COHYST central and eastern model units. Parameters calibrated in those models will be recalibrated only if the model does not reasonably represent baseflows, well-specific hydrographs, and regional water level maps.

The model will "warm-up" based on watershed response model inputs for 1979-1985, with the UNL-CSD-1979 water table map of the region being used to establish starting heads. A map based on the actual 1985 water levels will be compared to the results of warm-up model to ensure that the 1985 starting heads are reasonable. This comparison will not lead to any recalibration per se, but it may cause the 1985 starting heads to be adjusted. The water levels simulated for 1985 will be used as the starting heads for the 1985-1997 calibration run.

Calibration of the groundwater model will consider baseflows, long-term well hydrographs, water level maps, and variable well hydrographs.

- Baseflows. The Work Plan specifies that DNR will develop estimates of monthly baseflows throughout the model area. Based on HDR's QA/QC review of DNR's baseflow extractions, the initial evaluation will use data generated by the WHAT method, which will be subject to check by other methods as needed and potentially adjusted if judged unreasonable. It is likely that the overall match between modeled and adjusted baseflows will be determined through logical and defensible changes to water budget terms, e.g. estimates of net pumping from the watershed response model and canal recharge from the surface water model. However, the process will recognize that the baseflows are estimates; adjusting the baseflow values will be considered if it would improve model calibration while retaining consistency with hydrologic conditions. Adjustments to aquifer parameters may be made if there are time and spatial patterns in model results that cannot be corrected by adjusting the water budget.

- Water level trends. DNR will develop criteria to filter water-level records from more 1,922 wells to identify data that can be used to identify long-term trends in water levels 1985-1997. The location and magnitude of such changes (water level declines or mounding) will depend primarily on water budget inputs, and thus will be used to evaluate those inputs. Further, there may be times and locations where adjustments to aquifer parameters (most likely storage coefficient) will both make sense and improve the fit between model results and data. The long-term well data will also be used to evaluate and possibly adjust the modeled boundary inputs.
- GIS evaluations. Modeled water levels will be mapped and compared to data-based water level maps for 1997, and potentially, years between 1985-1997. The GIS will be used to superpose one map on another, and to subtract one set of contours from another to produce a differences map. The water budget (and possibly aquifer parameters) may be adjusted for times and locations where differences between observed and simulated data are large in comparison to the natural variability of the water table, and especially where (and if) the simulated slope and aspect of the water table is illogical.
- Transient responses. A small number of well hydrographs with monthly data will be used to evaluate how well the model simulates transient responses to pumping. In this case, adjustments of aquifer parameters (most likely hydraulic conductivity) may be considered if the result is to substantively and logically improve the fit between modeled and observed drawdowns without damaging other aspects of the calibration.

Surface water model

The surface water model will use STELLA software to represent key features of the surface water system and to simulate the effects on that system from natural inflows, reservoir and irrigation district operations, and surface-groundwater interactions. Calibration will be based primarily on how well the model outputs match observed reservoir levels, gaged flows, and canal diversions. Each individual canal will be calibrated independently, with possible additional calibration based on the river segment and system-level water balances.

Initial model conditions will represent actual data from January 1, 1985 as to reservoir levels and gaged stream flow. Initial calibrations will use the baseflow and runoff estimates from the baseflow separation work for the mainstem. Final calibration will consider dynamic input from the watershed and groundwater models.

While both the actual data and the model outputs are daily, the calibration will consider monthly values. Qualitative (statistical) as well as quantitative comparisons will be made to identify systematic errors in the water budget. Direct comparisons of historic and predicted values will be made, as well as cumulative difference plots to evaluate potential systemic differences. Adjustments of canal seepage losses and returns, along with reservoir seepage, are expected to be the primary means to improve calibration.

In evaluating model performance, it is expected that some aspects of the forecasting model (e.g. operational rules) may need to be replaced with observed data. This is because actual historic operations reflected real-time decisions that were not always perfectly aligned with a fixed protocol. Thus, if historic simulations of surface flows do not match observed flows, the team must first look to records of actual operations before attempting to improve the calibration by adjustment of model parameters.

Integrated model

The MODFLOW and STELLA models will be integrated so that appropriate water-budget quantities can be fed from each model to the other. Initially, this requires verification that the transfer of information preserves the water balance, e.g. that the water exchanges between surface model nodes equals the exchanges in the equivalent groundwater model nodes. The two should be sufficiently similar to conclude that no water is lost or double counted in the data transfer.

Calibration of the integrated model will be achieved primarily through adjustments to the watershed response model that improve the separate and collective ability of the groundwater and surface water models to produce a reasonable water budget for the model area as a whole, and for selected sub-areas. Parameters in the separate models (e.g. canal seepage coefficients, aquifer parameters) may also be adjusted in order to achieve better calibration of the integrated model, provided that the result is not an unacceptable change in the calibration of the standalone models. Outputs of the watershed response models must produce quantities for water demand and losses required for the surface water model to match observed canal diversion and stream gage data. Similarly, the outputs must produce quantities of demand and recharge that cause the groundwater model to match changes in water levels and baseflows.

The calibration procedure will be iterated until the model provides a fully functioning simulation that is judged to be sufficient for answering the variety of regulatory and management questions posed by the Project sponsors, or until it is determined that this objective cannot be reached for certain questions.

The Calibration Process

The calibration process will be implemented according to the principles outlined above. The individual components/models will be developed and preliminarily calibrated independently and the output compared and shared with the other models with input parameters for the subsequent iteration further constrained by an assessment of the previous output. The following task list is illustrative.

Process Tasks - 1st iteration:

- Partition current best estimate of historic regional ET, recharge, and runoff using the watershed response model.

- Develop pumping estimates using watershed response model and metered data.
- Partition the runoff into a range of recharge and discharge into major streams for guidance.
- Using runoff from the watershed response model, conduct a preliminary calibration of STELLA to historic streamflow and surface water system operations data.
- Using pumping from metered data and the watershed response model, recharge from the watershed response model, and return flow (recharge) from STELLA conduct a preliminary calibration of MODFLOW model to historic head and streamflow measurements.
- Evaluate model outputs to identify water budget components in STELLA and MODFLOW that are likely to need adjustment.
- Refine constraints for seepage return flow (loss) from canals (STELLA), recharge (MODFLOW), and pumping (MODFLOW) to add constraints the watershed response model.

Process Tasks – 2nd iteration:

- Revise key parameters in the partitioning code, which may require revisions to parameters in the watershed response model. Run the partitioning code to develop recharge and pumping estimates for groundwater model and/or runoff for the surface water model.
- Recalibrate STELLA to historic streamflow and surface water system operations data with revised information from the post processing program and baseflow from MODFLOW.
- Using pumping and recharge derived from the portioning code, along with constrained STELLA seepage (recharge), calibrate MODFLOW model to historic head and streamflow measurements.
- Evaluate model outputs with technical team.
- Reevaluate the water budget components, with emphasis on the dependent parameters in both STELLA and MODFLOW.
- As needed, refine constraints for seepage from canals (STELLA), recharge (MODFLOW), and pumping (MODFLOW) to add constraints to the watershed response model.
- Repeat *Process Tasks – 2nd iteration* as determined by technical team in evaluation noted above.

Verification

The model will be verified by running the calibrated 1985-1997 model for the 1998-2007 time period, and evaluating model outputs against observed water level and total streamflow observations. If there are significant discrepancies, the team will make calibration adjustments to the model such that 1998-2005 results are improved without unacceptable effects on the validity of the model to simulate 1985-1997 simulation.

PLAN FOR RECALIBRATION OF THE COHYST MODEL
May 2016

The COHYST Model was built to quantify time and space impacts to stream flow and aquifer water levels that are reasonably expected to result from changes in the use and management of water in the Platte River in central Nebraska. As such the primary basis for calibration of the model is its ability to reproduce historic changes in stream flow and water table throughout the study area and over an extended time frame.

The COHYST Model consists of a watershed response model, a groundwater model and a surface water model that together complete an integrated model. The initial COHYST model was completed in 2013; that included calibration and documentation. Since then there have been numerous significant changes to the models, including:

- the watershed model was restructured to use actual monthly data;
- the surface water model now uses a demand shift to better estimate surface diversions;
- the groundwater model has drains and a modified representation of evapotranspiration.

Calibration of the prior models resulted in a good match for normal to wet conditions but there were substantial and consistent errors in simulating low flow conditions in June–November. The models also contained systematic regional errors in the simulated water level drawdown. However, based on statistical measures, overall calibration was acceptable.

The new calibration will be performed for conditions in the period from 1990-2005, with 1985-1989 run as a model warm-up. This does not preclude the possibility of eventual model verification based on conditions in the period from 2006-2010. The determination of whether the models reasonably reproduce historic changes will be reached using the same criteria applied in 2013, as follows.

- Results from each model will be compared to observed data using statistical techniques such as calculating the sum of the squared differences between observation and simulated values to determine whether a calibration adjustment improved model performance. There is no predetermined error bar for an acceptable result.
- Errors and residuals for each model will be plotted as time-series and as maps, to identify locations and times when the model results systematically depart from observations.
- Assessment of model calibrations will ultimately depend on the professional judgment of the modeling team, the COHYST Technical Committee, and the COHYST Sponsors.
- Fundamental to this approach is that calibration is a combination of rigorous quantitative comparisons and the flexibility to make changes when such are indicated by the calibration results and the modelers' understanding of system hydrology.

The integrated model will be calibrated through a four-step process. The modelers will:

1) use current results from the watershed model and available measured data to construct standalone versions of the groundwater model and surface water model. The surface water model will employ historic reach gains and losses. The groundwater model will employ historic inflows, diversions and returns.

2) iteratively adjust the surface water model and the groundwater model to match historical flows and water levels.

3) modify the watershed model to address problems identified in the surface water model and groundwater model.

4) complete an integrated simulation and identify features that need to be adjusted to improve the integrated model's match to historic conditions.

Steps two through four will be repeated as necessary. The modeling team will propose to end the calibration process when there is agreement that the models are fully functional and hydrologically reasonable and that further substantive improvements in calibration are unlikely. The judgment of the modeling team will be subject to peer review by the Technical Committee and approval by the Sponsors.

The following sections provide more detail on the independent calibration of each model. Note that calibration begins with outputs from the watershed model as it is currently configured.

The surface water model simulates stream flows, reservoir operations, and irrigation diversions and returns. Calibration of the surface water model is primarily based on adjusting the model rules for operation of the irrigation and power facilities in the basin to match observed values for each individual reservoir, canal and gaging station. Initial simulations with the surface water model will use historic reach gains and losses and input from the watershed model and groundwater model will be added in subsequent steps.

Statistical evaluations as well as professional judgment will be used to compare the observed conditions to simulated values and to identify and correct any systematic errors in the water budget. Cumulative difference graphs provide an illustrative tool for that comparison.

An effort will be made to improve estimates of canal seepage losses and returns, which are considered the least constrained component of the current model. Care will be taken to recognize times in which historical surface water operations reflected real-time decisions that were exceptions to the protocols assumed in the model, and to place little or no weight on failed calibrations at such times.

The groundwater model simulates aquifer water levels and exchanges with surface water. Aquifer hydraulic conductivity will be calibrated by the groundwater model through an optimization process using the USGS UCODE parameter estimation program. Optimization may also be used to account for a long recharge lag if the initial results contain a systematic regional error associated with the depth to water. Concurrent with optimization, the aquifer specific yield and riparian evapotranspiration rates will be adjusted to achieve an adequate simulation of low-flow conditions on the Platte River. Additional runs will be used to test alternative methods for defining the stream bed thickness.

Water levels and tributary baseflows will provide the optimization targets. Water level targets will be changed from the targets used in the 2013 calibration by using a higher density of targets in areas near the Platte River and a lower density in areas distant from the Platte River. Tributary baseflows on Platte River tributaries will be added to the targets used in the 2013 calibration and the groundwater model will be modified as necessary to calculate baseflow data for the calibration.

The watershed response model partitions the fate of precipitation and irrigation into consumptive use, deep percolation and runoff. The model will be adjusted within reasonable limits to improve the calibration of the surface and ground water models.

Insights also can be obtained by continuing and expanding the current practice of comparing watershed model results to known data. For example, evapotranspiration (consumptive use) can be compared to research estimates published in such documents as University of Nebraska-Lincoln NebGuides; groundwater pumping can be compared to recorded measurements and estimates from Natural Resource Districts.

To address dry river, efforts will be made to improve model performance during periods of very low stream flow, even if overall calibration statistics do not improve. Calibration targets during low flow conditions may be given extra weight. However, failure to reasonably predict flows during dry periods will not be considered a reason to reject the model.

If resources allow, verification of the model may be conducted by running the calibrated 1985-2005 model for the 2006-2010 time period, recognizing that the land use data set available as of 2006 is markedly different from that available for 1985-2005. Issues raised through verification will first be addressed by changes to the model which improve the 2006-2010 match without significant degradation of the 1985-2005 calibration. Remaining differences between modeled and simulated conditions for 2006-2010 will be evaluated and explained to the extent feasible.