RECLAMATION
Managing Water in the West

Framework to Effectively Quantify and Communicate Groundwater Model Uncertainty to Management and Clients

Mission Statements

The mission of the Department of the Interior is to protect and provide access to our Nation's natural and cultural heritage and honor our trust responsibilities to Indian Tribes and our commitments to island communities.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Framework to Effectively Quantify and Communicate Groundwater Model Uncertainty to Management and Clients

prepared by

River and Reservoir Operations
Jennifer Johnson, P.E., Hydraulic Engineer

Acknowledgments

Contents

This page is intentionally left blank.

1 Introduction

A groundwater model is a simplified representation of a complex natural system. While recognizing that not everything can be known about a natural system, the goal, when developing a model, is to minimize the unknowns and to represent the real system as closely as possible. However, in all cases, a certain amount of uncertainty is inherent in the modeling process.

An acknowledgement of uncertainty in a model is sometimes interpreted by nontechnical clients as admission that a model is not adequate to address the problems it was created to solve. Uncertainty should not be thought of as a limitation of the model. Rather, it should be thought of as a tool to describe what is not known about a system, so that appropriate decisions are made with the model results. In addition, an investigation of model uncertainly often improves the knowledge of the overall system and can guide where additional information should be collected.

1.1 Groundwater model uncertainty – What is it?

Uncertainty in a model can be defined as the difference between the model and the complex physical system and processes that the model represents. It can be determined in terms of the parameters used to describe the system or the accuracy in its predictions. Since a mathematical model is a simplification of the complex system and processes, there will always be some difference between the model and reality. By quantifying and recognizing that difference, the usefulness of the model can be determined.

There are four main areas in which uncertainty can arise and be evaluated in a groundwater model. These areas are:

Conceptual Framework Uncertainty: The conceptual framework describes the current understanding of the physical properties of the site that is being analyzed. This includes, but is not limited to, the stratigraphy and geologic structure of the hydrogeologic units, the location of any surface water features (streams, lakes, drains), and the location of wells. More importantly, the conceptual model describes how all of the features are related and how they interact. Since it is not possible to completely know the physical attributes of the site, uncertainty can arise (Brendecke, 2009).

Model Parameter Uncertainty: Model parameter uncertainty relates to how the conceptual model can be represented within the constructs of the mathematical

model. It also relates to uncertainty in the input parameters and the observations used in the model.

Calibration Uncertainty: Calibration uncertainty relates to how well the model solution matches observed values.

Predictive Uncertainty: Predictive uncertainty refers to how well the model can evaluate the question at hand and can be impacted by the uniqueness of the calibrated model.

1.2 Groundwater model uncertainty – Why is it important?

Groundwater models are used to make decisions, to analyze risk, and to manage water systems. While no model can be 100 percent correct, when properly constructed and evaluated, a model can be a useful and informative tool. Evaluating the uncertainty that exists within a model reinforces the output from the model and makes it more useable to the end user.

There have been many articles written about the use of groundwater models in litigation and as decision making tools (Randazzo, 2005; Cosgrove et al., 2008; Hall et al., 2005; Tandon and Kilburg, 2005). Many of the authors agree that groundwater models, as imperfect representations of complex natural systems, can be discredited if the modeler does not acknowledge the uncertainty in a model. As stated by one author: "Adversarial proceedings are a lousy place to verify models, but a modeler will minimize the effect of the opposition's claims of uncertainty by actually quantifying the uncertainty" (Myers, 2007).

Decisions being made about a natural system require an understanding of that system. A manager or client may ask for a single number as an output from a model. However, that number is more powerful if the manager understands how the output was determined and the effect of model uncertainty.

1.3 Groundwater model uncertainty: how should we talk about it?

Uncertainty, as a word, tends to have a negative connotation when it comes to decision making. Modelers commonly use the word uncertainty to describe what is not known about a model, and to the modeler, this is not negative at all. Rather, the modeler believes that understanding the unknowns in a model gives them a better understanding of the system. It is the responsibility of the modeler to inform the decision maker of the usefulness of model uncertainty and how it ultimately will lead to better decisions being made with the model output.

1.4 Purpose of this document

This document is designed to illustrate ways that uncertainty can be recognized and quantified in groundwater modeling applications. It is not all inclusive since there are numerous modeling methods and aspects of uncertainty. Rather, it describes the major types of uncertainty, some methods for handling uncertainty, and ways to communicate what is found in the final model report (as reports are the most common way that modeling results are documented and shared with decision makers). At the very minimum, the purpose of this document is to illustrate how understanding model uncertainty can be used as a tool to better understand the system that is being modeled and to ensure that decision makers have all of the information to help them make a more informed decision.

This page is intentionally left blank.

2 Conceptual Framework Uncertainty

"Conceptual modeling starts the moment a hydrologist visits a site, and evolves to form the basis of analytical and numerical models" (Poeter, 2006). The first step in developing any model is to develop the conceptual framework, an analyzed compilation of all available data and knowledge about the modeled area. It is used to determine how a system is modeled, what type of model is used, and what additional data is collected to understand the system. This uncertainty can be evaluated at all stages of the pre-modeling, modeling, and post-modeling process and should be included in all parts of the model report.

2.1 Conceptual framework

Developing the conceptual framework involves piecing together all of the available data for the area of interest and looking critically at how the system behaves. Developing a strong conceptual framework is an important aspect of reducing uncertainty in a groundwater model. By developing a good understanding of system behaviors before applying a mathematical model to the system, the modeler avoids accepting model output that is not representative of the real system.

Limitations in the data collection phase can lead to uncertainty in the conceptual framework. Since groundwater systems are continuous and underground, it can be difficult to define all of the features. Some data that may not be well defined include the configuration of stratigraphic units, the extent of individual stratigraphic layers, or the location and behavior of structural features such as faults. In addition, some features may be known at individual points (from drillers well logs), but may not be known between those points.

Understanding the limitations of the conceptual framework can indicate where the modeler should use the mathematical model to quantify uncertainty.

2.2 Methods for evaluating conceptual framework uncertainty

Determining where uncertainty exists in the conceptual framework is largely a qualitative analysis. All of the available data and system knowledge is compiled to determine if there is enough information to make an informed decision. In the unlikely case that there is enough information, (which is almost never the case

due to system complexity, lack of time, lack of funds, etc.) the modeler will apply a mathematical algorithm and generate a solution to the question at hand.

In the more likely scenario that there exist unknowns in the conceptual framework, the modeler will apply a mathematical algorithm, evaluate the unknowns and determine if a range of suitable solutions can be produced (discussed more in the calibration and prediction sections) and if more information should be collected. Since it is common that the conceptual model will change as the numerical model is applied and as the modeler becomes more familiar with the system, the conceptual framework should be reevaluated as more is learned about the system.

In some cases, more than one conceptual model may be appropriate to explain unknowns in the system. Model averaging tools like MMA (Poeter and Hill, 2007) can help to evaluate predictions using multiple conceptual models. Developing more than one conceptual model can be costly as it typically requires duplicating the modeling process for each conceptual model, but it has been argued that multiple conceptual models can give the most accurate modeling results (Poeter, 2007).

2.3 Communicating conceptual model uncertainty

The modeling report typically consists of, at the very minimum, a statement of the problem, a description of the hydrogeology, a description of how the model is developed, and the conclusions derived from the model. The conceptual model can be discussed in each of these sections, but is most commonly described in the hydrogeology section. It is also beneficial to describe the relationship of the conceptual framework to the representation in the numerical model as there are typically differences between the conceptual framework and the numerical model representation. Some differences that may be included in the discussion are the model representation of layering versus the complex layering system that likely exists in reality or location of specific features versus their placement in the model.

3 Model Parameter Uncertainty

Model parameter uncertainty describes uncertainty in both the quality of the data being used in the model and impacts from applying that data to a mathematical model.

3.1 Parameter Uncertainty

The quality of data that goes into a model is directly proportional to the quality of the prediction that can be made with the model (affectionately known as "garbage in, garbage out")[1]. When collecting data, whether by hand or with an automated instrument, there can be a degree of error. A person collecting data by hand with a measurement device or improper use of measurement devices can cause error in the readings. Most measurement devices report the tolerances of their readings in the user's manual. If the model intends to calculate results that are within the tolerance of an instrument's reading, the tolerances should be included in the uncertainty evaluation.

Recharge is an example of a common parameter used in groundwater modeling that can have a varying degree of uncertainty. By definition, recharge is the amount of water that infiltrates the surface, moves past the root zone and reaches the aquifer. This water can be from many sources including precipitation, on-farm irrigation, leakage from septic systems, and leakage from canals. Many factors impact recharge including soil types that affect the seepage rate, vegetation types that consumptively use the water, or the slope of the land surface that affects the surface runoff rate. In addition, precipitation is typically measured at a single point location and must be extrapolated to cover the entire model area. The variability in each of these aspects of recharge may increase or decrease the total amount of recharge water and the uncertainties in each aspect can increase or decrease the total uncertainty in the recharge value itself.

No matter what mathematical model is used, the data being applied will need to be altered in some way to fit the model. Grid based models only calculate a solution for the center of a grid cell. Since most data is not available at the exact resolution of the defined grid, it will need to be modified to fit the grid. Non-grid based models typically require modification to the data to fit the processes that are calculated.

Groundwater models are typically calibrated to measured values of aquifer water levels or discharges to streams. These observed values are typically measured at an exact location. However, the model solves for water levels and flows at the

---

[1] Appendix A includes a list of model input data and where such data can be obtained.

center of each grid cell, which is rarely at the measured location. In addition, observed data is collected at a specific time which generally does not coincide with the model time steps. These location and time discrepancies can combine to create uncertainty in the model results. The effects of location and time "error" can be minimized using interpolation, bilinear interpolation within the grid and linear interpolation for time. To further assist in calibration, weights can be assigned that indicate the level of confidence in each observation. A higher weight would indicate more confidence in the accuracy of the representation of an observation.

3.2 Methods for evaluating uncertainty in parameters

The most useful method for evaluating uncertainty in input parameters and observations is to determine the relative sensitivity of the parameters. If a set of parameters is considered sensitive, it will impact the calibration and therefore errors in that data should be minimized.

Sensitivity describes how the model (specifically a model observation) reacts to changes in a given parameter. A parameter is considered sensitive if a small change in the parameter value causes a large change in the model results at observed water level and stream flow locations (or other observation type). If a parameter is sensitive, it is more likely to be approximated correctly during the calibration process than a parameter that is insensitive. Since the sensitive parameter will have a large degree of influence on the calibration, errors in the parameter should be well known to ensure a correct calibration. Insensitive parameters, by definition, will not impact the calibration. However, changes in insensitive parameters may impact predictions made with the model and should be explored (ways to explore impact and uncertainty in insensitive parameters is described in section 5.2).

It is helpful to perform a sensitivity analysis both before and after the model is considered calibrated. If using an automated calibration tool, parameters that are considered insensitive before calibration should be fixed during the calibration process (to a reasonable value). Changing insensitive parameters during calibration will not contribute to the final calibrated product and may prevent the parameter estimation process from finding a solution. In some cases, changing some parameters will change the sensitivity of others, so performing a sensitivity analysis after calibration will verify that those parameters that were considered insensitive before calibration stayed insensitive.

There are a couple of methods that are used to determine parameter sensitivity. The first is the manual method in which each parameter in the model is changed separately by increasing and decreasing the parameter value by a fixed amount or percentage (bracketed) and the change in the observation is recorded. For

example, the hydraulic conductivity of a particular layer or zone may be increased by 10% and decreased by 10%. The changes in the observations are recorded.[2]

Another effective method for calculating sensitivities is to utilize an automated method, such as the OPR method that accompanies MODFLOW2000 (Harbaugh, et al, 2000) or the calculated sensitivities in PEST (Doherty, 2004).[3] The automated methods produce comparable output statistics such as dimensionless scaled sensitivities (DSS), composite scaled sensitivities (CSS), and parameter correlation coefficients (PCC). These parameters not only let the modeler determine how sensitive a parameter is with respect to another parameter, but also how sensitive a parameter is with respect to an observation. DSS indicate the amount of change in one observation given a one percent change in a parameter value. CSS indicate the amount of change in all of the observations given a one percent change in a parameter value. PCC indicate the level of correlation for all possible pairs of model parameters. Parameters with large DSS and CSS values are considered sensitive with respect to the observations and are therefore considered important to the model calculation. PCC values that are near -1.00 and 1.00 indicate that the parameters cannot be estimated uniquely (Hill and Tiedeman, 2007).

Sensitivity should be evaluated for all parameters represented in the model, including those parameters that will be represented with measured data. If a parameter will be represented with measured data and is determined to be sensitive, any suspected errors in that data should be closely investigated and resolved, if possible. If the results of the sensitivity analysis show that a parameter is sensitive with respect to the observed values and the parameter is not correlated to another parameter, an automated calibration process should estimate the parameter correctly and uniquely. If the parameter is determined to be insensitive and will be represented with measured data, the errors in the data are not as important to the calibration. However, a quick test should be run after the model is calibrated to determine if changing those parameters impacts the prediction being made with the model. Parameters estimated by the calibration process that are determined to be insensitive should also be explored to determine if changes in the parameter impacts the prediction being made with the model (a process to evaluate these parameters is described in section 5.2). Parameters that are correlated should also be explored since they will not be able to be estimated uniquely.

---

[2] The book "Effective Groundwater Model Calibration: With analysis of Data, Sensitivities, Predictions, and Uncertainty" includes a good description of how sensitivities are calculated in MODFLOW2000 and PEST.

3.3 Communicating uncertainty related to parameters

Model reports typically include a section that describes the data that will be used in the model. This section should describe the methods that were used to collect or develop the data (or the source of the data if it was acquired or developed by another entity) including all uncertainties that are associated with the collection or development. This includes tools, equations, and assumptions that went into generating the data.

The sensitivity analysis should be discussed in a separate section. Most readers have a general understanding of sensitivity, but it is still a good idea to describe what goes into a sensitivity analysis and the significance it has to the reader. Also, the specific technique used to conduct the sensitivity analysis for the specific project should be included. In the sensitivity section, the plots showing the DSS and CSS should be presented along with a detailed description of their significance. Lists of the sensitive and insensitive parameters are useful to help the reader further understand the results of the sensitivity analysis. It is also helpful to discuss the confidence in the data relative to the sensitive parameters since the sensitive parameters will impact the results.

The sensitivity analysis is an important tool when describing the uncertainty of the model to decision makers, which is why it is so important that a sensitivity analysis be conducted and properly described in the model report. When decision makers understand the parameters that are vital to developing the most reliable model result, they can better understand why and where new data could be collected to make a better decision. Also, it helps them to understand why inaccurate or unreliable data may be contributing to lower confidence in model predictions.

4 Model Calibration Uncertainty

Calibration uncertainty is the most explored area of groundwater model uncertainty. Calibration uncertainty refers to how well the model calculations match observations (most often water level and stream flow observations) and many tools are available to explore and quantify the uncertainty in this portion of the modeling process.

4.1 Calibration Uncertainty

Calibration is the process by which unknown model parameters are adjusted to match observed water levels and stream flows (or other observed values). This can be done manually or using automated processes such as UCODE (Poeter and Hill, 1998) or PEST (Doherty, 2004). Calibration, whether by manual or automated methods, has the goal of minimizing the difference between model calculated values and observed values. Determining calibration uncertainty can be as simple as determining the goodness of fit (how well the model solution fits the observed water levels and stream flows).

4.2 Methods for evaluating calibration uncertainty

Quantifying calibration uncertainty incorporates uncertainty from the conceptual model and uncertainty from the input parameters. A way to address uncertainty in the model calibration is to assess the goodness of fit of the model calculation to the observations. Since the difference between the observation and what the model calculates at that point is easily obtainable from the model, many statistics and presentations can be used to illustrate how well calibrated the model is. While there are many other statistics available, it is generally accepted that the following should be in the calibration section of every groundwater modeling report:

> **Root- mean squared error**- The root mean squared error (RMSE) should be calculated for each type of observation used in the model (water level elevation, flow, etc.). If the error is defined as the difference between an observation and the calculated value at that location, the RMSE quantifies the absolute mean of all of the error values in the model. The RMSE uses the same units of the observation (for example, length units for water level elevation).
>
> The equation for RMSE is

$$RMSE = \sqrt{(X_i - \overline{X})^2}$$

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n}(X_i)$, n = number of observations, and X_i = the ith observation.

Percent error- The percent error is the RMSE of the model divided by the total change in the observation type being calculated with the RMSE; for example, the RMSE of water levels divided by the total change in observed water levels. For a well calibrated model, this value should be less than 10 percent.

The equation for percent error is:

$$\%Error = \frac{RMSE}{\Delta observed}$$

Plot of observed versus calculated – A plot of the observed versus calculated values can quickly illustrate how well the model matches the observations. If the modeled values perfectly match the observations, the scatter plot will line up along a 1:1 straight line. Since no model will match the observations perfectly (due to variation that exists in natural systems), there will be some deviation from the straight line. This deviation can be quantified with a linear regression (R^2) calculation. The closer the R^2 is to 1, the better calibrated the model is.

An example plot of observed versus calculated values is shown in Figure 4-1. Note that most of the values line up along the 1:1 straight line. There are two separate groups of numbers along the line as a result of plotting the results from two different model layers.
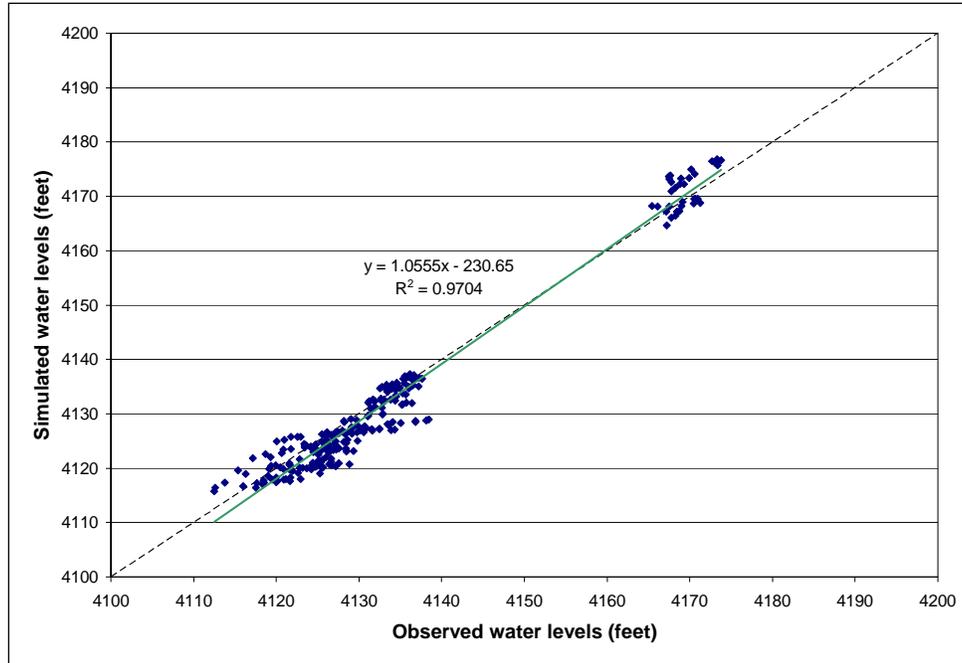
Figure 4-1: Example plot of observed versus simulated water levels.

There are many other statistics and measures of goodness of fit that can be calculated and for a more complete list the reader is referred to Hill and Tiedeman (2007).

If there is a difference in the confidence level of the observations that are being used, weights can be assigned to the observations that indicate the level of confidence. Weights allow the modeler to place more importance on observations that have a high degree of confidence and are calculated using the tolerance in the observation relative to its location in the groundwater model. The level of confidence in the observations may be an indication of confidence in the measurement itself or how the measurement is represented in the model. The automated process (PEST, UCODE or MODFLOW 2000) will adjust parameters to meet the observations, and will place more importance on those observations with higher weights.

4.3 Communicating calibration uncertainty

Communicating calibration uncertainty requires a technical discussion of how well the model matches observed flow and water level values. As described in section 4.2, there are minimum requirements when it comes to reporting how well the model is calibrated. Statistics and graphs help to illustrate the level of calibrations, but each should be well explained and should include a description of what constitutes a well calibrated model and how the current model compares.

This page is intentionally left blank.

5 Model Predictive Uncertainty

Predictive uncertainty is considered by some to be the most important type of uncertainty that should be evaluated; however, it cannot be the only uncertainty that is evaluated. Determining predictive uncertainty requires information from the other uncertainty evaluations to make a proper determination of the uncertainty in the prediction. Predictive uncertainty ultimately determines how well the model will be able to evaluate the effect of changes in the system.

5.1 Predictive Uncertainty

Models are typically designed for two purposes, explaining how a complex system behaves or predicting how a complex system might react to change. Predictive uncertainty refers to the uncertainty that arises when the model is used to predict how the modeled complex system might react to change. If the conceptual model is accurately defined, the data is well understood and the model is well calibrated, predictive uncertainty has the potential to be minimal.

5.2 Methods for evaluating predictive uncertainty

Once the model is calibrated and a sensitivity analysis has been conducted, it is typically used to evaluate scenarios (changed conditions) and make predictions.

The results of the sensitivity analysis can be used to evaluate predictive uncertainty. Parameters that were considered not sensitive during the calibration process may become sensitive when running the predictive analysis. Since the parameters were not sensitive to calibration, estimating the most accurate parameter value is not possible. When running a predictive scenario, the model should be tested to determine if the results of the prediction are sensitive to the changes in the previously determined insensitive input parameters. This can be done by simply changing the value of the insensitive parameter and looking at the relative change in the model results at the prediction location. If the prediction changes, confidence intervals should be developed to describe the range of possible solutions to the predictive scenario.

Monte Carlo analysis can be used to generate confidence intervals for estimated parameters or for model predictions. Monte Carlo analysis involves running the model with a range of values for a given estimated parameter to determine the values that will satisfy the observations. It can also be used to run a range of values for a parameter to develop confidence intervals that represent the changes in the observations or predictions. Detailed information about the Monte Carlo

method can be found in Skinner (1999), Vose (2000), Bedford and Cook (2001), and Hill and Tiedeman (2007).

Non-unique solutions occur when there is more than one option for an unknown parameter that is being solved during the calibration process. To determine if a solution is unique, Hill and Tiedeman (2007) pose two questions: "(1) given the constructed model, are the observations and prior information sufficient to have estimated the one and only set of parameter values that provide the best fit? and (2) does a set of parameter values exist that produces a better fit than that achieved?" If the solution is non-unique, the possibility exists that a better prediction can be made with one version of the model versus another.

Solution uniqueness can be addressed by evaluating the parameter correlation coefficients and by repeating the calibration process with different starting values. If parameters are highly correlated (a PCC above 0.95), there exists the possibility that the calibrated solution is non-unique. If repeating the calibration process with different starting values results in significantly different parameter estimates, then the calibrated solution is likely non-unique.

One method for increasing solution uniqueness that has recently become popular is to use PEST calibration with pilot points and Tikhonov regularization. Pilot points are a way to characterize the spatial distribution of parameters (such as hydraulic conductivity) within the grid, eliminating the need for lumping the parameter into piecewise homogeneous zones. Parameters are estimated at the pilot points and then interpolated to the remaining cells (in this case, the pilot points are interpolated using kriging). Since the pilot points are at discreet locations, PEST has the ability to make large changes at each point to best match an observation, which can lead to large variations in a parameter over short distances. Tikhonov regularization constrains the PEST calibration process so that PEST does not calculate unrealistic parameters simply to meet the observations. It has been argued that using pilot points with Tikhonov regularization calculates the most unique parameter distribution possible and reduces uncertainty in the model results (Fienen and others, 2009).

It has been argued that finding a unique model solution can give the wrong impression that there is more known about a system than really is and that developing confidence intervals on the prediction acknowledges the unknowns in the system. By either using confidence intervals on the predictive solution or by trying to find a unique model to make the prediction, the modeler is exploring the parameter space to find the best possible solution to the predictive question. Both options are valid and documented, and the validity of either approach can be discredited if the modeler does not correctly document the uncertainty in their model.

5.3 Communicating predictive uncertainty

Since predictive uncertainty is the most important uncertainty to decision makers, it should be presented and communicated in a clear concise way. Predictive uncertainty should be considered low only if, in addition to a unique solution and well calibrated model, parameter error is considered low and the conceptual model is well defined. Any document should include the reason for determining that predictive uncertainty is low, since this will mean that any assessments made with that particular model will be correct and accurate.

If the model does not have a unique solution, there are a couple of options for communicating the bounds of the solution. Depending on the question being asked, the solution can be presented using graphics or maps. For example, if the question involves explaining when water will reach a certain elevation at different locations, the result may be presented as a number with plus or minus confidence intervals or with the use of a chart that visually represents the number with confidence intervals. Some questions are better answered using a map. If the question is where will the water surface be at a particular elevation and a particular point in time, a colored map can illustrate the location of the water surface at difference points in time. Due to the simple visual nature of maps and graphs, they can often communicate information to decision makers more quickly than can numbers or text.

This page is intentionally left blank.

6 Conclusion

Groundwater model uncertainty is a large topic that has been investigated and will continue to be investigated by many. There are four main topics in groundwater model uncertainty that should be addressed in all groundwater reports: conceptual model uncertainty, parameter uncertainty, calibration uncertainty, and predictive uncertainty.

It is the responsibility of the groundwater modeler to understand and evaluate groundwater model uncertainty. It is also the responsibility of the modeler to inform the decision maker of the uncertainty associated with the model and its output. The concept of uncertainty in a groundwater model should not be considered negative, since when properly analyzed, it can be used to inform better decisions.

Groundwater model uncertainty can be used as a tool to inform the modeler and the decision maker about the system. It can be used to determine where new data should be collected, which can ultimately save money. It can also be used to place confidence intervals on model output so that appropriate decisions are made with the model.

The consequences of not evaluating and understanding uncertainty far outweigh the difficulty that may be associated with determining and explaining uncertainty. As more modelers bring the topic to management, it will become better understood and easier to explain.

It is important to note that a well calibrated model and a unique solution do not necessarily mean that uncertainty in a groundwater model has been eliminated completely. If the conceptual model is poorly defined without investigation of unknowns through a rigorous method, such as Monte Carlo analysis, there will be inherent uncertainty within a model that will not be accounted for. If the data was poorly collected and the errors in the data are not properly explored, uncertainty will not be accounted for. The groundwater world is a complex place and there are many opportunities to not account for uncertainty. However, if the modeler is diligent and accounts for the four types of uncertainty and uses the tools to evaluate the unknowns, uncertainty can be reduced to an acceptable level and, ultimately, the model will be a better decision making tool.

This page is intentionally left blank.

7 References

Bedford, T. and Cook, RM (2001). *Probabilistic Rick Analysis, Foundations and Methods.* Cambridge, UK, Cambridge University Press.

Brendecke, Charles (2009). Comments on Model Uncertainty Memo. For consideration by the Eastern Snake Hydrologic Modeling Committee, Jan 13-14, 2009.

Cosgrove, D.M., Johnson, G.S., Tuthill, D.R.(2008) The Role of Uncertainty in the Use of Ground Water Models for Administration of Water Rights, Journal of Contemporary Water Research and Education, Issue 140, P 30-36.

Doherty, John (2004). PEST: Model Independent Parameter Estimation, User Manual: 5th Edition. Watermark Numerical Computing. 336 pgs.

Fienen, M., Muffels, C., and Hunt, R., 2009, On Constraining Pilot Point Calibration with Regularization in PEST, *Groundwater* 47.6: 835-844.

Hall G H; Bennett V; Church J N, (2005). Uncertainty And Defensibility In Modeling; Proceedings Of The 2005 NGWA Ground Water And Environmental Law Conference, July 21-22, 2005, Baltimore; P259-265

Harbaugh, Arlen, Edward Banta, Mary Hill, and Michael McDonald (2000). MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model – User Guide to Modularization Concepts and the Ground-Water Flow Process. U.S. Geological Survey Open-File Report 00-92. 130 pgs.

Hill, Mary and Clair Tiedeman (2007). Effective Groundwater Model Calibration: with analysis of Data, Sensitivities, Predictions, and Uncertainty, Wiley, New Jersey, 455 pages.

Myers T, 2007, Minimizing The Insufficient Data Argument: Uncertainty Analysis In Adversarial Decision Making; 2007 NGWAGround Water And Environmental Law Conference July 24-26, 2007, Dublin, OH; P127-140

Poeter, Eileen (2007). All Models are Wrong, How Do We Know Which Are Useful, *Groundwater* 45.4: 390-391.

Poeter, Eileen and Mary Hill (2007). MMA, A Computer Code for Multi-Model Analysis, USGS Techniques and Methods 6-E3, Boulder, CO.

Poeter, Eileen and Mary Hill (1998). Documentation of UCODE, A Computer Code for Universal Inverse Modeling. U.S. Geological Survey Water-Resources Investigations Report 98-4080. 122 pgs.

Randazzo P J, (2005). A Groundwater Modeler's Guide To Survival In The Courtroom, Proceedings Of The 2005 NGWAGround Water And Environmental Law Conference July 21-22, 2005, Baltimore; P246-247

Skinner, DC (1999). *Introduction to Decision Analysis, 2ⁿᵈ ed*. Gainesville, FL, Probabilistic Publishing.

Tandon V;Kilburg J A (2005). The Role Of Model Prediction Uncertainty In Resolving Water Resource Conflicts Associated With Mine Development, Proceedings Of The 2005 NGWA Ground Water And Environmental Law Conference, July 21-22, 2005, Baltimore; P266-281

Tikhonov, A.N., 1963a, Solution of incorrectly formulated problems and the regularization method, *Soviet Mathematics Doklady* 4, 1035-1038.

Tikhonov, A.N., 1963b, Regularization of incorrectly posed problems, *Soviet Mathematics Doklady* 4, 1624-1637.

Van Heeswijk, Marijke and Daniel Smith (2002). Simulation of Ground-Water Flow System at Naval Submarine Base Bangor and Vicinity, Kitsap County, Washington, USGS Report No. 02-4261, Tacoma Washington, 142 pages.

Vose, D (2000). *Risk Analysis, A Quantitative Guide, 2ⁿᵈ Ed.* West Essex, UK, Wiley.

Appendix A: Uncertainty Flow Chart and Description
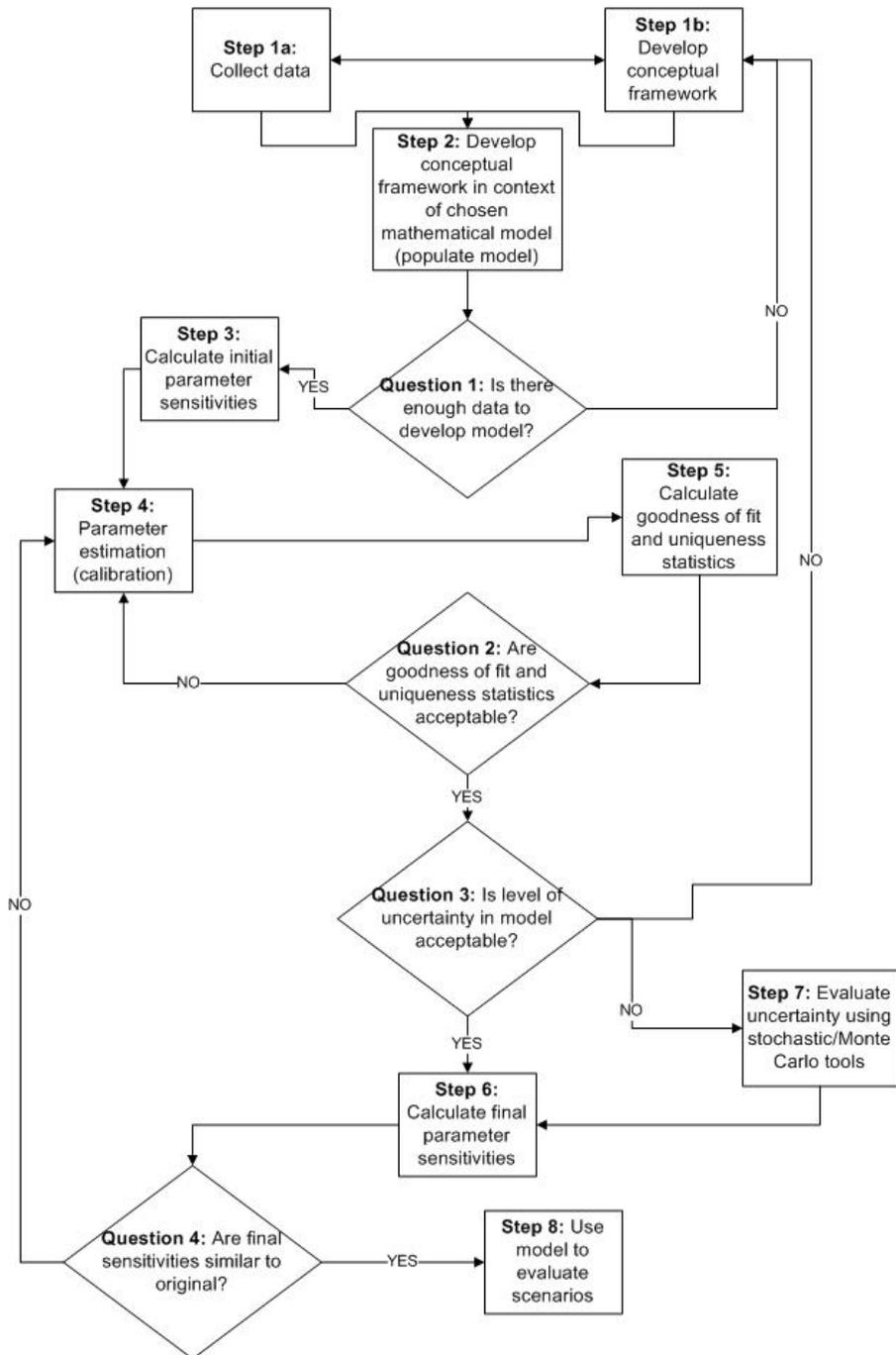
Step 1: Step 1 has two parts, collect all available data and develop conceptual framework. These two parts of step 1 should be performed iteratively since one informs the other. For example, while collecting new data you may redefine a portion of the conceptual model, or while defining a part of a conceptual model you may find a data gap. Note that not all data gaps can be filled by existing data, but may be estimated during the calibration process. This process helps contribute to the knowledge of the system and to find out what is not known about the system (uncertainty).

Step 1a: Collect all available data. Begin your groundwater modeling study by collecting all of the available data about the site you are attempting to analyze. If possible, try to obtain most of the data in GIS format or convert the data to GIS format; this will allow for more efficient model development. The table below describes the type of data to look for and places to find it.

Data type	Description	Examples of where to find data
Well logs	A well log can provide information about the geologic stratigraphy (layering) as well as depth to water.	Well logs are generally site specific data and may be obtained from knowledgeable staff (Reclamation geology staff usually has knowledge of well log locations). State water management agencies sometimes have online databases of well logs.
Well locations and measurements	Well locations should be in some coordinate format if possible (sometimes it is only possible to know the township, range and section where a well is located). Depth to water and screened elevations of water are also useful information.	Well locations and measurements can sometimes be obtained from the same location as well logs. Some well locations and measurements are project specific, so study personnel may have the best information.
Geologic Map	Geologic maps can provide information that will inform the conceptual model design as well as the calibration process. For example, the map may show geologic faults, which can be a location of high or low permeability.	General geologic maps can be obtained from the USGS or state geologic agencies. Reclamation geology staff sometimes has geologic maps for specific locations which tend to be more detailed than USGS or state maps.

Surface elevation data	DEMs, LiDAR, and topographic maps can provide information about the ground surface.	Surface elevation data can usually be obtained from online resources or from Reclamation GIS staff.
Hydrography	Hydrography provides the geographic location and elevation of streams.	Hydrography can be obtained from the USGS website.
Recharge information	Recharge information should describe the net areal recharge to the aquifer (this should not include recharge from streams or lakes unless they are not otherwise defined in the model). This may include recharge from precipitation, on-farm infiltration, or septic systems.	Recharge information can be collected from many sources since recharge is such a diverse dataset. The goal when defining an aerial recharge dataset is to account for only the water that actually recharges the aquifer; water that is used during the ET process or is surface runoff should not be counted as recharge. Precipitation information can be obtained from Hydromet or PRISM and should be corrected for the amount that actually infiltrates to the aquifer. On-farm infiltration should be calculated based on amount of water applied to the land minus crop requirement (this can be a very complex calculation and may be estimated if data is not available). Recharge due to septic system infiltration can be estimated from M&I water records.

Step 1b: Begin defining conceptual framework. The conceptual framework describes the current understanding of the physical properties of the site you are attempting to analyze. This includes, but is not limited to, the geographic structure of the aquifer(s), the location of any surface water features (streams, lakes, drains), and the location of wells. When defining the geographic structure, it is helpful to build a visual conceptual model of the geologic features in a software package such as GMS or Rockworks. Such software packages allow you to include stratigraphic information that can be derived from boreholes, develop cross-sections, and interpolate stratigraphic layering based on known information. While developing this conceptual model, you may find that you do

not have all of the data necessary, requiring an assessment of the feasibility of collecting more data. Not being able to collect more data gives a location for you to begin defining uncertainty in your model.

Step 2: Develop conceptual framework in context of chosen mathematical model. Whether using an analytic or numeric model, the conceptual framework will need to be adjusted to fit into the context of the mathematical model. The adjustments may include changing data resolution, converting the data to consistent units, or making assumptions for unknown data.

Step 3: Calculate initial parameter sensitivities. Calculation of parameter sensitivity is one of the most important aspects of evaluating the level of uncertainty in a groundwater model. A parameter is considered sensitive if a small change in the parameter value causes a large change in the model solution with respect to the observations. Conversely, an insensitive parameter is one that causes a small or no change in the model solution with a large change in the parameter value. If a parameter is sensitive to changes, it is likely that the calibration process will estimate the parameter with a good degree of certainty. However, insensitive parameters may be estimated anywhere in a range, since the changes made to the parameter do not show up in the calibrated solution. If the prediction being made by the model is impacted by the insensitive parameter, a poor estimation of the insensitive parameter during the calibration process will affect the results, even though the model calibration matches observed water levels and stream flows. Sensitivity analysis can be conducted manually or by using an automated parameter estimation tool. It is recommended that the modeler use an automated procedure since many automated parameter estimation tools include a sensitivity analysis procedure and output the relevant statistics. Since automated calibration procedures may not converge if insensitive parameters are left to adjust, it is common to fix the parameters (this is the reason for checking sensitivity before calibration). Calculating uncertainty with respect to insensitive parameters will be addressed in step 7.

Step 4: Parameter estimation (calibration). Calibration can be done using by manual "trial and error" method or by using automated parameter estimation software, such as PEST, UCODE, or MODFLOW2000. Although the manual method can allow the modeler to get a feel for how changes to parameters impact the calibration, it is generally recommended that modelers use automated procedures. Automated procedures, although computationally expensive, inherently provide information that allows the modeler to evaluate uncertainty.

Step 5: Calculate goodness of fit and uniqueness statistics. There are many types of statistics that will quantify goodness of fit and uniqueness. At the minimum, root mean squared error of water levels or flow, percent error over the total change in water level or flow, and a linear regression plot of the observed versus calculated values should be evaluated and presented. Additional statistics are well explained in Effective Groundwater model Calibration by Hill and

Tiedeman (2007). For a well calibrated model, it is generally accepted that percent error should be less than 10 percent and the linear regression (R squared value) should be greater than 0.9.

Step 6: Calculate final parameter sensitivities. This step is required since sensitivities may change as parameters are adjusted during the calibration process. If parameters that were insensitive before the calibration become sensitive after calibration, the calibration procedure should be rerun without the parameters fixed.

Step 7: Evaluate model uncertainty using stochastic/Monte Carlo methods. By the time the modeler gets to this step, he/she should have evaluated uncertainty in the conceptual model and input data. The sensitivity analysis should have indicated which parameters could be successfully estimated and which may not be. The stochastic/Monte Carlo analysis should be used to evaluate correlated and insensitive parameters with respect to the prediction being made by the model (this is explained more in the main document).

Step 8: Use model to evaluate scenarios. After the model has been calibrated and analyzed for uncertainty, it can be used to make predictions. If it has been determined that uncertainty is small (and this has been documented in the model report), the model output can be used as is. If uncertainty is a concern, the stochastic/Monte Carlo analysis should be used to develop error bands for the predictions.

This page is intentionally left blank.

Appendix B: Memo Minidoka Test Case

Before a case was tested for uncertainty, the authors of the report "Framework to effectively quantify and communicate groundwater model uncertainty to management and clients" attempted to survey Reclamation groundwater modelers about their use of uncertainty in their models. A meeting was held at the 2009 Reclamation Construction and Geology conference to introduce the topic and a new SharePoint site was developed that would facilitate the survey. Only one response to the survey was received. The comments from the one survey were incorporated into the model document and test case.

The test case model was a modified version of the Eastern Snake Plain MODFLOW model. The modified model was to be used to evaluate the impacts of raising Lake Walcott 5 feet on the regional aquifer and nearby streams. In addition, the model was used to evaluate the impacts due to modified operations after replacing the spillway at the dam (without raising the lake elevation). Since only a portion of the model was modified for the study, uncertainty was only evaluated with respect to the changes that were made to the model. Those changes include refining the model grid to smaller cells near Minidoka dam, adding a model layer to represent a conductive sandy layer on the north side of the reservoir, and modifying the location of the river features to match the new grid. The model was recalibrated to determine hydraulic conductivity values for the new model layer and to refine the conductances of the river features.

Since the new portion of the model was well defined geologically and hydrogeologically, uncertainty in the conceptual model was considered small. In addition, any new data that was used in the model was considered to be within bounds that would not impact the output of the model. A sensitivity analysis was conducted on the parameters that were used in the calibration process. Insensitive parameters were identified and tested to determine if changes in their values impacted the prediction. Since they did not, uncertainty in the model was considered small, and the model solution was presented without error bands. This was discussed in a section of the model report called Model Uncertainty.

After the client and reviewers read the report, they were sent a survey to evaluate the effectiveness of uncertainty communication within the document. One reviewer, a modeler, asked for additional statistics to prove the uncertainty was low. Another reviewer, a manager, thought the description of uncertainty was clear and acceptable.

Test Case using Flow Chart

This section describes the development of the Minidoka model in the context of the steps listed in appendix A. A complete description of model development can be found in the report called "Groundwater Modeling Analysis of Potential Minidoka Pool Raise" (USBR, 2009).

Problem: Evaluate the impacts on the Eastern Snake Plain Aquifer from replacing Minidoka spillway and from keeping Lake Walcott at a higher elevation for a longer period each year. (Figure B-1 shows the extent of the regional aquifer and the location of Minidoka Dam and Lake Walcott.)
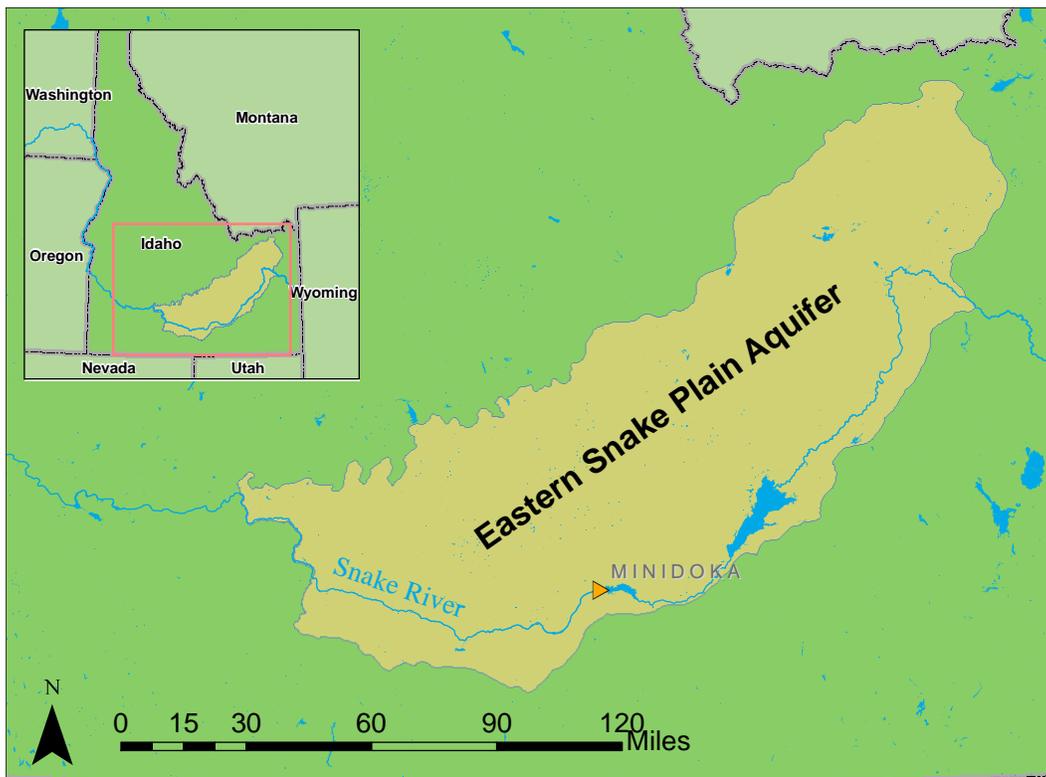

Figure B-1: Eastern Snake Plain Aquifer and Minidoka Dam (orange triangle).

Step 1: Collect all available data and develop conceptual model. This step began by evaluating models and data that existed for the Eastern Snake Plain Aquifer (ESPA) and the region surrounding Lake Walcott. Since the groundwater evaluation was only a small part of a larger investigation, a large amount of data had already been collected and compiled. In addition, a model of the ESPA already existed, so it could be used, with some small improvements, to answer the study questions.

The conceptual understanding of the aquifer near Lake Walcott used in the ESPA model development is that the Snake River flows into Lake Walcott, which may or may not be in contact with the regional aquifer. The river downstream of

Minidoka dam is considered perched, not in contact with the regional aquifer, so it is not included in the model. Although it was conceptually understood that a perched aquifer layer on the north side of the reservoir allowed seepage from the reservoir, this was not included in the regional model.

For the Minidoka study, the conceptual model was updated to include the perched layer on the north side of the river. The conceptual understanding is that water flows through this perched layer, under the North Side canal to seepage channels that are measured along the Snake River. Another addition to the conceptual model was the North Side canal, since seepage measured in the seepage channels would be a combination of seepage from both Lake Walcott and the North Side canal. The last addition to the conceptual model was that a small section of the Snake River downstream of Minidoka dam was in contact with the regional aquifer and therefore was included in the model.

Step 2: Develop conceptual framework in context of chosen mathematical model. Since the mathematical model was previously developed, many of the decisions that would normally be made during this step were already made. Some of the representations were updated to accommodate the new understanding of the conceptual model.

The existing ESPA model simulates the aquifer as a single, confined layer with cells that were one mile square. A single recharge parameter represents both recharge from all sources and discharge from pumping wells. Since the model was a regional model and only represented a single aquifer layer, refinement was necessary near the area of interest to better address the study questions.

The model was updated to have smaller cell sizes near Lake Walcott and an additional model layer that represented the sandy layer on the north side of the Lake. MODFLOW river cells were added to represent the North Side canal in the new model layer.

Step 3: Calculate initial parameter sensitivities. Parameter sensitivities were calculated to determine which parameters should be estimated during the calibration process. Only the final sensitivities were reported for this model.

Step 4: Parameter estimation (calibration). The Minidoka model was an improvement of an existing model that had previously been calibrated. So, only new features added to the model were included in the recalibration process, as it was assumed that the original ESPA calibration was sufficient to answer the study questions. The parameters that were adjusted during the calibration process include: the conductance of Lake Walcott, the conductance of the Snake River below Minidoka dam, the conductance of North Side canal, hydraulic conductivity of the new model layer, the storativity of the new model layer, and the general head boundary along Lake Walcott in the new model layer. Some of the calibrated parameters were divided into zones.

The parameters were estimated using the zones and the parameter estimation software PEST. Pilot points were not used to estimate hydraulic conductivity and storativity in this model since the area being estimated was small (however, it is recommended in most cases that pilot points and regularization be used because the combination produces the most unique solution).

Step 5: Calculate goodness of fit and uniqueness statistics. The quality of the calibration was determined by examining the RMSE and percent change of both the water level elevation measurements and the seepage measurements. Also, a regression for both data sets was presented. The RMSE for water level elevation was 2.8 feet which was 4.5 percent of the total change in head over the modeled time period. The RMSE for seepage was 4,148.51 ft/day, which is 5.4 percent of the total change in seepage. The R-squared value for the water level regression and seepage was 0.97 and 0.93, respectively. Since the percent change of both data sets was less than ten percent and the R-squared values were greater than 0.90, the model was considered well calibrated.

Some of the parameters that were estimated were correlated with a correlation factor of greater than 0.95, which can indicate non-uniqueness in the solution. These parameters were evaluated by starting the calibration process with different starting values, which produced similar results, indicating that the solution was unique.

Step 6: Calculate final parameter sensitivities. This section is taken directly from the model report and describes the final sensitivities in this model.

> Determining the sensitivity of model parameters using Fit-Independent Statistics can help quantify model uncertainty (Hill and Tiedeman, 2007). Sensitive parameters are those that cause a large change in the model solution when their values are varied; while insensitive parameters can be varied by large amounts and not affect the solution. The sensitivities are presented in two ways, composite scaled sensitivities (CSS) and dimensionless scaled sensitivities (DSS). DSS values indicate the importance of an observation with respect to a parameter and are scaled to a dimensionless value so that it can be compared with other parameters that may have a different dimension. CSS values are DSS values that are combined with respect to each parameter to give an indication of the overall importance of that parameter. CSS values were plotted in a bar chart to show the relative sensitivity of each parameter (Figure B-2).
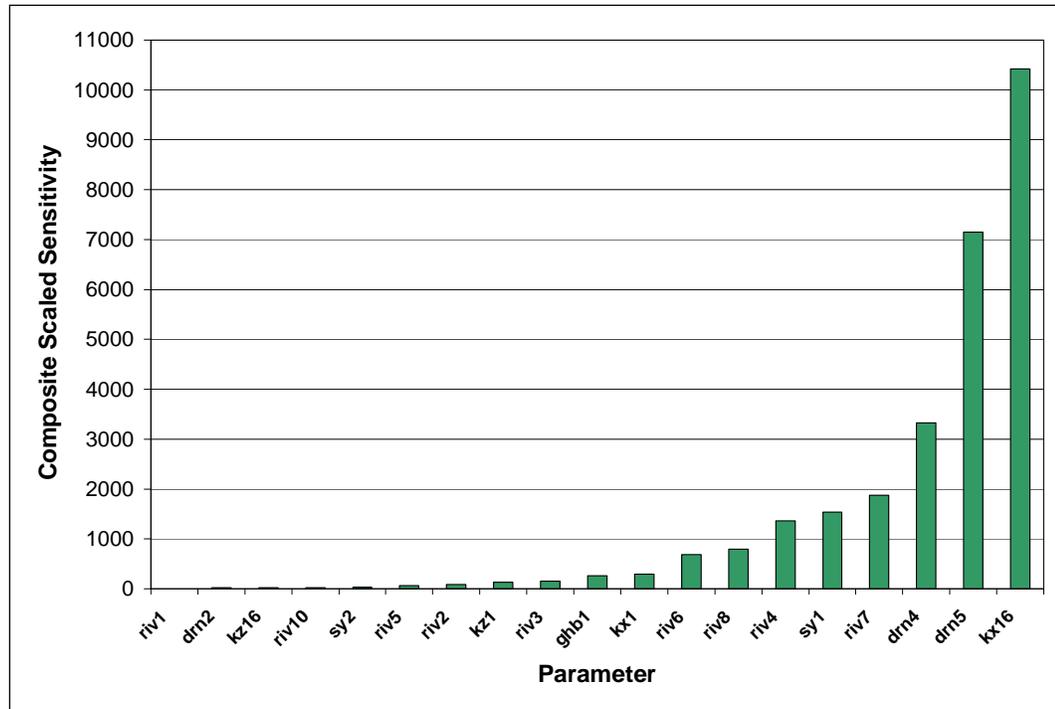
Figure B-2: Plot of CSS values for Minidoka groundwater model.

Parameters that are less than 1 percent of the largest CSS value are considered not sensitive. Since the maximum CSS for this model is about 10,400 for parameter horizontal hydraulic conductivity in zone 16 (kx16), the sensitive parameters are those with CSS greater than 104. The sensitive parameters in this model are:

- vertical hydraulic conductivity zone 1 (kz1),
- river conductance zone 3 (riv3),
- General Head Boundary Zone 1 (ghb1),
- horizontal hydraulic conductivity zone 1 (kx1),
- river conductance zone 6 (riv6),
- river conductance zone 8 (riv8),
- river conductance zone 4 (riv4),
- storativity zone 1 (sy1),
- river conductance zone 7 (riv7),
- drain conductance zone 4 (drn4),
- drain conductance zone 5 (drn5), and
- horizontal hydraulic conductivity zone 16 (kx16).

DSS values were plotted in a line graph to show the relative importance of each parameter to an observation based on its sensitivity (Figure B-3). Observations with DSS values that are large with respect to the other observations are considered sensitive with respect to the corresponding parameter. For example, the parameter River Zone 2 is important to the estimation of the piezometers in layer 2 but not the piezometers in layer 1 nor

the seepage estimates because the relative DSS values are large for the piezometers in layer 2 and small for the piezometers in layer 1 and the seepage estimates.
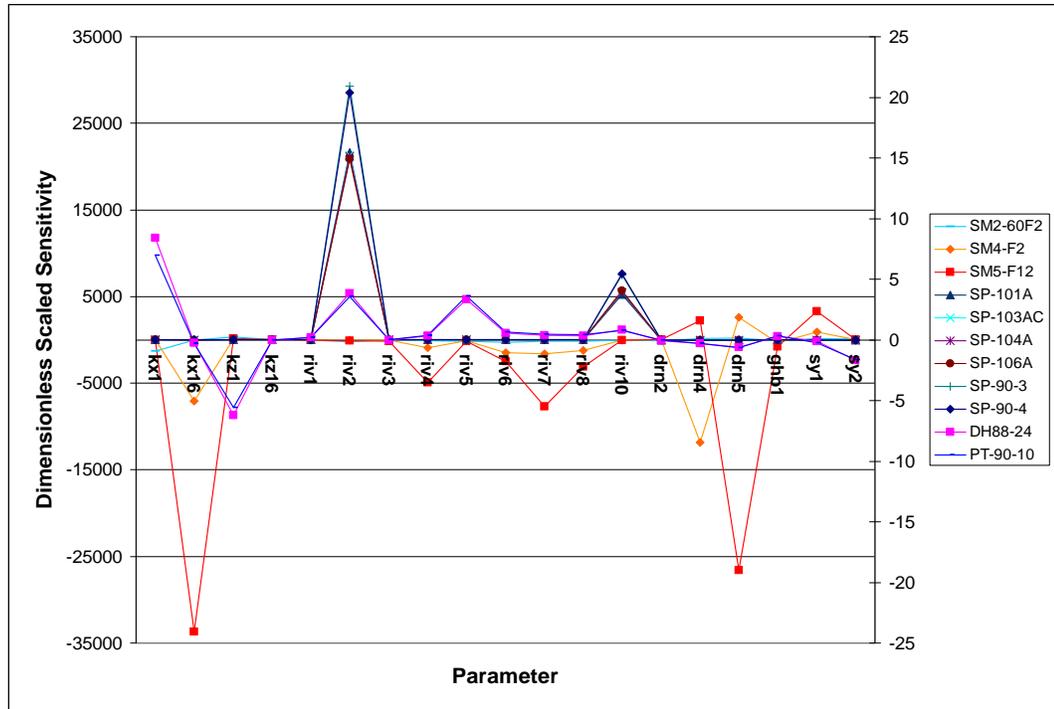
Figure B-3: Plot of dimensionless scaled sensitivities for each estimated parameter in the model. The left y-axis represents the DSS for observations SM2, SM4, and SM5. The remaining observations are reflected on the right y-axis.

The CSS and DSS values can help to determine which parameters are important to developing a well calibrated model. Estimates of parameters with low DSS and CSS values are can be discounted for a well calibrated model because a low DSS means that none of the observations depend on the parameter and a low CSS means that the model is not sensitive to changes in the parameter. The parameters in this model that do not have a high DSS (or high importance) are the vertical hydraulic conductivity for zone 16, river conductance for zones 1 and 3, drain conductance for zone 2 and GHB conductance for zone 1. Of these parameters, GHB for zone 1 and river conductance for zone 3 are considered slightly sensitive, so changes may result in a different solution, but not at the observation points. It is often recommended that such parameters be set with a fixed value during the calibration process since changes will have little effect on the solution (Hill and Tiedeman, 2007).

Parameters that have low CSS values and high DSS values can contribute to the uncertainty of a model. In the Minidoka model, this condition is true for river conductance for zones 2, 5, and 10 and specific yield for zone 1. To account for those parameters that contribute to the uncertainty of the model, it

is common to use a range of values for each parameter during the prediction phase to give a range of possible solutions to the problem. Since in this model the parameters with large DSS values are all on the low end of the CSS chart, changes to the parameters did not affect the solution to any large degree, therefore, it was not necessary to use a range of values. Therefore uncertainty within the modified portion of the model is considered small.

Step 7: Evaluate model uncertainty using stochastic/Monte Carlo methods.
Since uncertainty in this model was considered small, this step was not necessary.

Step 8: Use model to evaluate scenarios. The model was used to evaluate the study questions.