Chapter 1. INTRODUCTION

Ground-water models almost never perfectly represent real systems or fit the data used for their calibration. This results in uncertainties in both estimates of parameters and predictions of model-related variables. The uncertainty of estimated parameters and predicted variables can be quantified by using either confidence or prediction intervals. However, computation of such intervals is not straightforward because model-computed hydraulic heads and quantities such as flows that are functions of hydraulic heads generally are nonlinear functions of the model parameters.

Vecchia and Cooley (1987) and Clark (1987) independently derived similar methodologies that can be used for the computation of confidence and prediction intervals for both the estimated parameters and the output from a nonlinear regression model with normally distributed observation errors. Vecchia and Cooley (1987) showed that the same computer program can be used both to estimate the parameters by nonlinear regression and to compute confidence limits from which the desired confidence intervals are obtained. Computation of each confidence limit is a separate optimization problem in which the limit is obtained as an extreme value on the edge of the parameter likelihood region. Beven and Binley (1992) and Brooks and others (1994) presented methods that are similar in intent to the method of Vecchia and Cooley (1987). For brevity we will term confidence intervals computed using the nonlinear regression model as nonlinear confidence intervals.

The nonlinear regression model can be linearized using a first-order Taylor series expansion. Confidence intervals computed using the linearized model are the standard ones found in most regression texts (for example, Seber and Wild, 1989, chapter 5) and are termed here linear confidence intervals. Hill (1994) presented two computer codes, BCINT and YCINT, that can be used to compute linear confidence intervals for parameters and linear confidence and prediction intervals for hydraulic heads and flows from MODFLOWP. Hill and others (2000) developed YCINT-2000 as a modification of YCINT to be used with MODFLOW-2000 to compute confidence and prediction intervals for hydraulic heads and flows; linear confidence intervals for the parameters are computed by MODFLOW-2000 at the end of the Parameter-Estimation Process.

Synthetic case studies (Cooley and Vecchia, 1987; Vecchia and Cooley, 1987; Hill, 1989; and Cooley, 1993a, b) show the following. Corresponding linear and nonlinear confidence intervals often are offset, and the nonlinear intervals generally are larger (Hill, 1989; Cooley, 1993b); the variability in sizes of nonlinear confidence intervals generally is larger than the variability in sizes of corresponding linear confidence intervals (Hill, 1989; Cooley, 1993b); the differences between the sizes of corresponding nonlinear and linear confidence intervals increase as the sizes of the intervals increase (Cooley, 1993a); and use of prior information can have a significant effect on the sizes of confidence intervals (Cooley, 1993b). Field case studies (Christensen and Cooley, 1996; Christensen and others 1998; and Christensen and Cooley, 1999a) support the conclusions from the synthetic studies. However, the studies of Christensen and Cooley (1996; 1999a) show that linear confidence intervals sometimes are larger than the corresponding nonlinear intervals.

Donaldson and Schnabel (1987) and Cooley (1997) present synthetic case studies that investigate the coverage probability of nonlinear confidence intervals as well as of linear confidence intervals. In the studied cases the coverage probability of nonlinear confidence intervals is correct, whereas the coverage probability of corresponding linear confidence intervals often is incorrect.

Christensen and Cooley (1999b) tested the accuracy of 95 percent individual prediction intervals for hydraulic heads, streamflow gains, and effective transmissivities computed by ground-water models of two Danish aquifers. An individual prediction interval is for a random variable $Y_p = g(\boldsymbol{\beta}) + \varepsilon_p$, where $g(\boldsymbol{\beta})$is a model-related variable (for example, hydraulic head, flux, or a model parameter), and ε_p is a normally distributed error that is a composite of small-scale model error and other errors including measurement errors. Small-scale model errors result from smaller-scale variability in geology and hydrology than those represented in the model. Such errors often are significant in ground-water modeling because the model commonly is constructed using model parameters that represent only important, or large-scale, geologic and hydrologic features. Christensen and Cooley (1999b) tested the accuracy of the prediction intervals by a cross-validation method and by comparing new field measurements, which were not used to develop or calibrate the models, to corresponding prediction intervals. They concluded that for the ground-water models of the two real aquifers the linear as well as the nonlinear individual prediction intervals appear to be accurate. However, their results indicate that for the predicted values of streamflow gains some of the nonlinear prediction intervals may be conservative without using a correction factor. In a somewhat similar study Hamilton and Wiens (1987) found that for a confidence interval the correction factor can be either larger or smaller than unity.

Inspired by the experiences from the previously mentioned studies, Cooley (2004) developed and tested a regression theory for modeling ground-water flow in heterogeneous systems. The theory addresses the small-scale model error that usually is present in ground-water models of heterogeneous systems where the model only uses average system characteristics having the same form as the drift. Important aspects of the theory are summarized and further tested by Christensen and Cooley (2003) and Cooley and Christensen (written communication, 2003). Cooley (2004) showed, among other things, that small-scale model error can result in significant correlations among the true errors of the regression model; that the model error often contributes bias in estimated model parameters, whereas predictions of hydraulic head often are unbiased; that the statistic used to compute a confidence or prediction interval often should be multiplied by a correction factor that depends on covariances and interrelations among all of the types of system characteristics that give rise to model parameters. The synthetic case studies by Cooley (2004), Christensen and Cooley (2003), and Cooley and Christensen (written communication, 2003) show that correction factors for confidence intervals can be very large, whereas the correction factors for prediction intervals appear to be relatively close to one.

The results of Cooley (1997), Christensen and Cooley (2003), Cooley and Christensen (written communication, 2003), and Cooley (2004) strongly indicate that linear confidence intervals for estimated parameters and predictions of ground-water flow models may be inaccurate in many cases, whereas corresponding nonlinear confidence intervals seem to be correct or nearly correct in at least the synthetic cases without small-scale model error. Nonlinear prediction intervals seem to be correct for the field cases. This motivated development of the Uncertainty (UNC) Process (or simply UNC for simplicity) that can be incorporated into MODFLOW-2000 (Harbaugh and others, 2000) and can be used to compute nonlinear confidence and prediction intervals for any model parameter and for most types of data that can be used as regression data by MODFLOW-2000 (hydraulic head, differences between hydraulic heads, head dependent flow, and differences between head-dependent flows). UNC implements the computation method of Vecchia and Cooley (1987) into MODFLOW-2000.

Use of UNC is intended to be straightforward and flexible. This is accomplished by minimizing the amount of extra input needed for the computation of intervals compared to the input needed by MODFLOW-2000 for parameter estimation. Options have been added to facilitate the search for a global extreme value within the likelihood region (corresponding to the upper or lower limit of a confidence or prediction interval), and so convergence problems resulting from parameter correlations can sometimes be avoided. Any number and any type of confidence or prediction intervals can be computed in the same run. The output from UNC is similar to the output from the Parameter-Estimation Process of MODFLOW-2000. However, the results from the computation of interval limits also can be summarized in a separate easy-to-overview output file.

The procedures used by the UNC Process to compute nonlinear confidence and prediction intervals assume the model errors to be random and normally distributed. Whether or not this is the case is indicated by residuals analysis. The work of Cooley (2004) and Cooley and Christensen (written communication, 2003) indicate that some deviation from these ideals can be tolerated. The graphical procedures by Cooley and Naff (1990, p. 167-171) together with the results produced by the RESAN2-2k program are valuable for the analysis. Intrinsic nonlinearity also is assumed to be small for the UNC Process. Whether or not this may be the case can be investigated by using the RESAN2-2k and BEALE2-2k programs and by analyzing weighted residuals computed by UNC.

Purpose and Scope

This report introduces and documents the UNC Process, which is a new Process in MODFLOW-2000 that calculates uncertainty measures for model parameters and for predictions produced by the model. Uncertainty measures can be computed by various methods, but when regression is applied to calibrate a model (for example, when using the Parameter-Estimation

Process of MODFLOW-2000), it is advantageous to also use regression-based methods to quantify uncertainty. For this reason the UNC Process computes (1) confidence intervals for parameters of the Parameter-Estimation Process and (2) confidence and prediction intervals for most types of functions that can be computed by a MODFLOW-2000 model calibrated by the Parameter-Estimation Process. The report also documents two programs, RESAN2-2k and BEALE2-2k, that are valuable for the evaluation of results from the Parameter-Estimation Process and a third program, CORFAC-2k, that computes correction factors, including approximations that require minimal geostatistical information, for UNC. The theory behind UNC and the RESAN2-2k, BEALE2-2k, and CORFAC-2k programs is summarized in chapter 2 and appendixes A and B. (For a thorough description the reader is referred to Cooley, 2004.) Guidance in application of, and input instructions for, the programs are given in chapters 3 to 6. Appendix C includes input and output files for an example problem. Source files for MODFLOW-2000 (including the programs documented in this report) are available at the Internet address listed in the preface of the report.

Users of this report need to be familiar with the Ground-Water Flow Process as well as the Observation, Sensitivity, and Parameter-Estimation Processes of MODFLOW-2000. The Ground-Water Flow Process is described by Harbaugh and others (2000), and the three other Processes are described by Hill and others (2000). The user also should be familiar with basic statistics and the application of nonlinear regression. The book by Cooley and Naff (1990) discusses these subjects with the focus on application in ground-water modeling.

Acknowledgments

The Danish Natural Science Research Council provided support for Steen Christensen to travel to the United States in 1996 and 2001 to work with Richard Cooley to develop the methods presented in this report. Support for Richard Cooley after June, 2001, was provided by a post retirement contract with the U.S. Geological Survey. Subroutines RANUM and RNDV in RESAN2-2k were modified from programs given by Press and others (1986).