CALCULATING MODFLOW ANALYTICAL SENSITIVITIES USING ADIFOR FOR EFFECTIVE AND EFFICIENT ESTIMATION OF UNCERTAINTIES AND USE OF AQUIFER SENSITIVITIES FOR OPTIMAL OPERATION

Amir Gamliel, Ph.D., P.E., P.G., U.S. Army Corps of Engineers, Interagency Modeling Center, 3301 Gun Club Road, West Palm Beach, Florida 33406, amir.gamliel@usace.army.mil
Maged Hussein, Ph.D., P.E.,U.S. Army Corps of Engineers, Interagency Modeling Center, 3301 Gun Club Road, West Palm Beach, Florida 33406, maged.m.hussein@usace.army.mil

Abstract

A modified MODFLOW code was developed to calculate the analytical sensitivities of both the field parameters and the operational parameters using the automatic differentiation system ADIFOR. The analytical sensitivities were then used as input parameters for PEST to calculate estimates of the uncertainties. In addition, the manuscript provides comparison between the analytical sensitivities and numerical estimates of the sensitivities as calculated by using PSET. Analytical calculations of the sensitivities are known to be more accurate than numerical estimates, (usually) takes less time to compute, and the results of analytical computations are independent of numerical assumptions such as the magnitude of the perturbation. Furthermore, using analytical sensitivities yields more accurate uncertainties than using the numerical estimates of the sensitivities. The paper also demonstrates the use of sensitivity analysis with respect to operation decision parameters such as magnitude of pumping rates from wells for optimal management of water resources.

INTRODUCTION

Both sensitivity and uncertainty analysis require accurate estimates of the derivatives of the model output with respect to the uncertain parameters. This manuscript demonstrates the use of the automatic differentiation utility ADIFOR to accurately and efficiently calculate analytical derivatives of model output with respect to input parameters. These derivatives can be used for automatic model calibration, sensitivity analysis, and uncertainty analysis. The Modular Three-Dimensional Finite-Difference Ground-Water Flow Model (MODFLOW) was used herein to demonstrate the utilization of the methodology in groundwater applications. Among the available codes for numerical estimation of model sensitivities, the authors arbitrarily chose to use the Model-Independent Parameter Estimation code (PEST) to perform sensitivity analysis using the numerical approach.

SENSITIVITY ANALYSIS

Scientists and engineers usually perform sensitivity analysis with respect to four types of parameters: field and state parameters, parameters related to boundary conditions, decision parameters, and parameters related to the numerical algorithm. Examples of field parameters for groundwater modeling include viscosity, transmissivity, conductivity, dispersivity, density, heat capacity, chemical absorption, and reaction rates. State variables include pressure, velocity, and concentration. Boundary conditions parameters are state variables that are specified along the

boundary of the area of interest and include pressure, velocity, and concentration. Decision parameters include allocation of resources to maximize or minimize system output, subject to given constrains. Examples of decision parameters in groundwater modeling include pumping and injection rates, drain elevation, drain conductance, etc. Optimization of the decision parameters results in maximum utilization and preservation of the aquifer. Grid spacing is an example of a numerical model parameter. This work presents, as an example, sensitivity of an independent variable (pressure head in an aquifer system) with respect to two field parameters (transmissivity and vertical conductivity), one boundary condition parameter (specified head along a river), a decision parameter (pumping rate of a well system), and grid spacing as an example of a parameter related to the numerical algorithm.

Automatic Differentiation

Code development for analytical derivatives is very difficult and tedious for complex numerical models which involve thousands of lines of computer code. Examples of manually developed analytical derivatives in computer codes include Hill (1992 and 1998), Harbaugh, et al. (2000) and Manganelli, et al. (2002). A more efficient alternative is using an Automatic Differentiation (AD) technique. AD is a technique for augmenting modeling codes with partial derivative computations. It exploits the fact that every code executes a sequence of elementary arithmetic operations. Martins, et al., (2000) and Griewank (2000) discuss the mathematical and computational techniques used for general purpose AD computer programs. The Automatic Differentiation of Fortran (ADIFOR 3.0) (Fagan and Carle, 2001) was selected for the demonstration in this paper.

In this effort we utilized ADIFOR on the popular groundwater flow modeling code MODFLOW. ADIFOR allowed us to calculate a wide range of partial derivatives of state variables including head, discharge, and budget terms among others with respect to any input parameter. Unlike the approach used by Harbaugh, et al. (2000) in MODFLOW-PES, where analytical derivatives with respect some of the field parameters could not be calculated, ADIFOR generated derivatives of the dependent variables with respect to boundary conditions, decision parameters, and grid spacing. To demonstrate the utility of the approach a series of simple hypothetical tests were conducted and compared to the numerical perturbation method.

The authors performed a number of tests viz. nonlinear problems, complex combination of boundary conditions, and ground water/surface water interaction problems. In some of these tests, numerical derivatives deviate considerably when compared to the analytical derivatives obtained through automatic differentiation. The issue of reliability of numerical derivatives has been a recognized limitation in inverse modeling, sensitivity analysis, and uncertainty analysis (Doherty 2002, Hill 1998 and Poeter and Hill 1998). For illustration purposes, the next section details a groundwater simulation example having a simple homogeneous domain, where the transmissivity or vertical conductance could be lumped into one parameter. Overall, calculations of analytical derivatives were noticeably faster than calculation of numerical derivatives by PEST. The cost saving in run time was considerably greater for simulations of non-homogenous conditions.

Example

Figure 1 illustrates the first test problem, which consists of two isotropic and homogenous confined aquifers that are separated by a leaky confining unit. The horizontal hydraulic conductivity of the two aquifers is set at 10 ft/day while the vertical conductivity is set at1 ft/day. The upper aquifer receives recharge of 0.004 ft/day. Both the upper and lower aquifers drain into a river whose bed resistance is 1000 ft^2/day and having a water elevation of 75 ft. The bottom aquifer contains two wells that pump at a rate of 35000 ft^3/day. The problem was discretized into 2 layers and 10 by 15 square cells of 500 ft side length. The quasi 3-dimensional approach of MODFLOW (Mcdonald and Harbaugh 1988) was used to represent leakage between the two aquifers through the semi-confining unit. Figure 2 shows the resulting head distribution in the upper and lower aquifers.

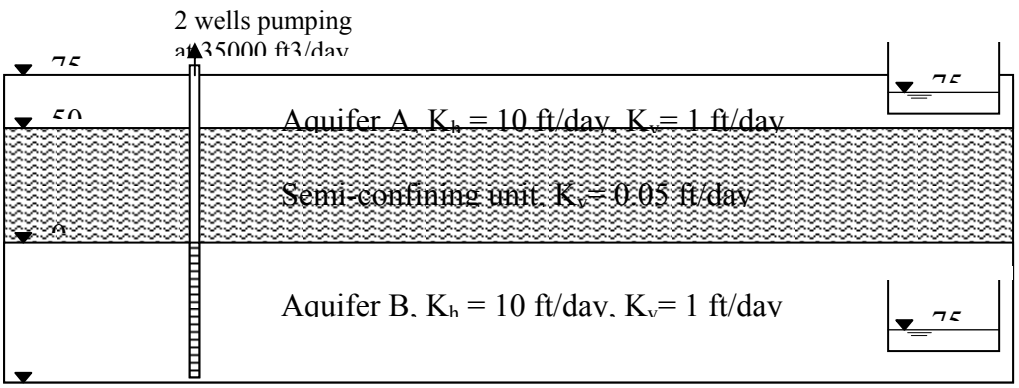
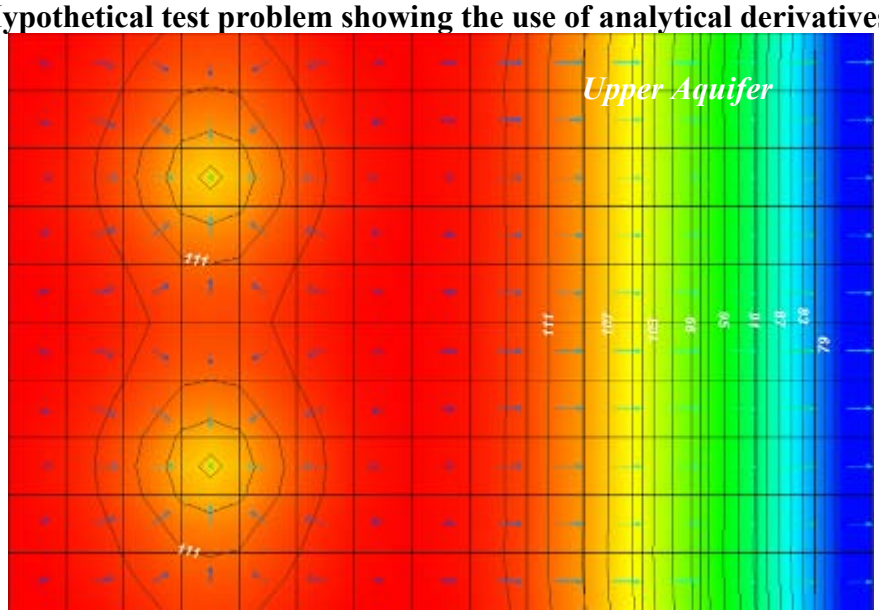


Figure 1. Hypothetical test problem showing the use of analytical derivatives approach.



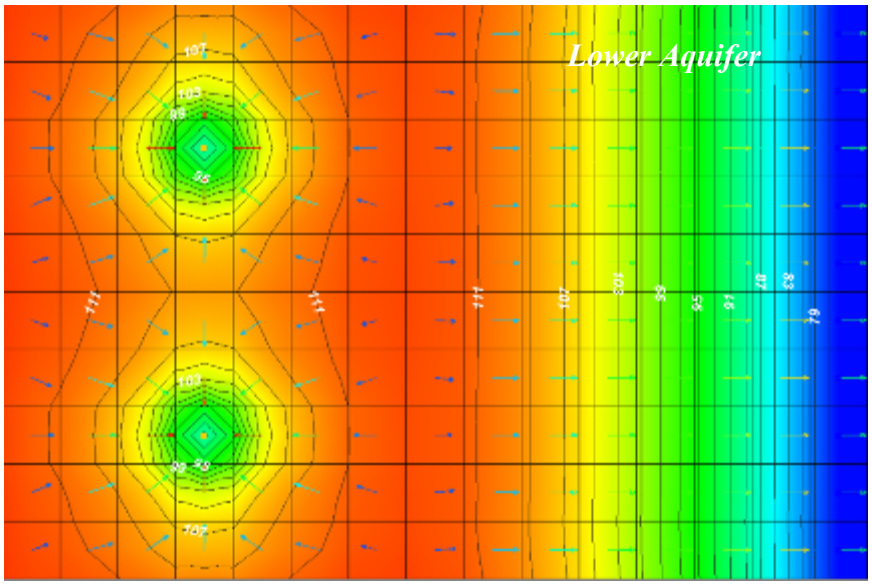


Figure 2. Head distribution and flow field in the upper and lower aquifers.

Under the assumption that parameter uncertainty is the dominant source of uncertainty in the above problem (hence ignoring any uncertainty in the conceptual model), the following parameters can be identified as uncertain; transmissivities of the two aquifers, vertical conductance, and conductance of the river bed of the confining unit as field parameters; water level in the river and recharge as boundary conditions; pumping rate as a decision parameter; and the analyst decision for grid spacing. For demonstration purposes, this work elected to perform sensitivity and uncertainty analyses with respect to transmissivity, vertical conductivity, water level at the river, and pumping rate. The aquifer pressure head was used as the dependent variable. The authors also employed ADIFOR to calculate the analytical sensitivity of model prediction of the pressure head with respect to grid spacing along the 'x axis', which could not be done with a perturbation method without altering the computational grid. Except for the dependency of the pressure head with respect to the pumping rate, the problem is linear and the values of the derivatives calculated accurately by numerical differentiation should be relatively close to the values calculated analytically.

Model Sensitivities

Despite the fact that the test problem is simple and predominantly linear, numerical differentiation produced considerably different derivatives for different perturbation intervals. Figure 3 depicts the result of both analytical and numerical calculations of the derivatives of head with respect to transmissivity, vertical conductance, and pumping rate and the duration of the computation time. The figure also illustrates the changes in the numerical derivatives with the change in the perturbations (increments) used for the numerical calculation. The authors changed the example problem for the calculation of the derivatives of aquifer head with respect to river stage by fixing the boundary head on the left side of Figure 1 to 100 ft. Otherwise, the derivatives of aquifer heads with respect to river stage are unity over the entire domain. Figure 4 depicts the changes for the derivatives of aquifer heads with respect to river head for both the analytical and numerical calculations. Both Figures 3 and 4 exhibit that the numerical derivatives

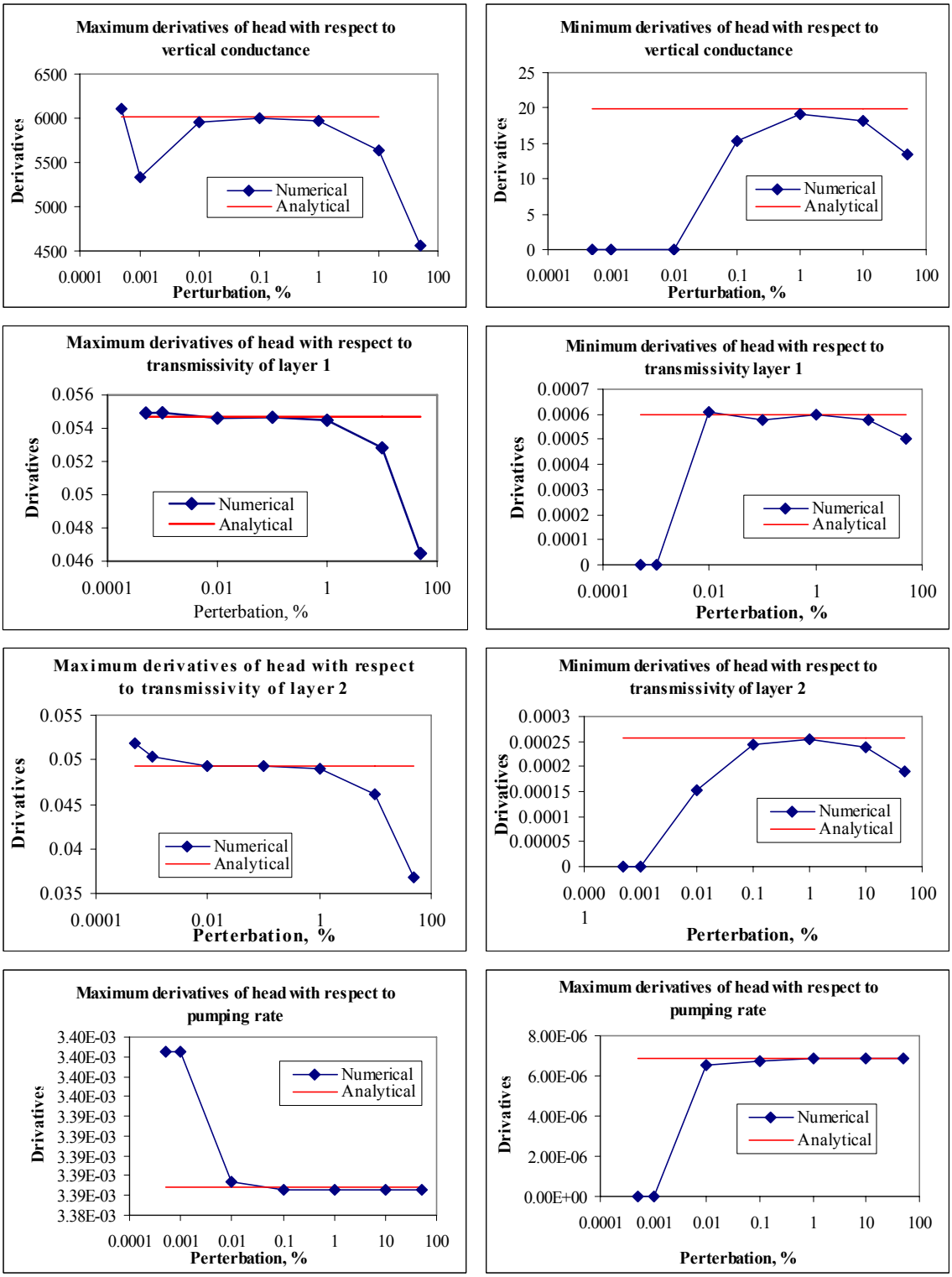


Figure 3. Comparison between analytical numerical derivatives using different perturbations (increments) showing the reliability of the numerical derivative calculations in a simple problem.

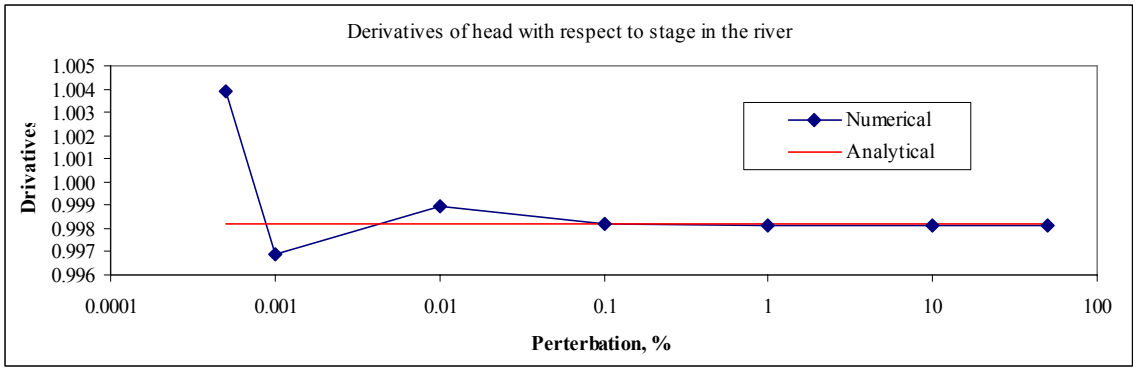


Figure 4. Comparison between analytical and numerical derivatives using different perturbations (increments) for boundary conditions.

vary considerably over the range of chosen perturbations. However, the values of the numerical derivatives are close to the analytical derivatives for a given range of perturbation.

Figures 3 and 4 show that if one must use a numerical method to calculate estimates of the derivatives, a perturbation of 1% of the independent variable should be used unless problem specific data recommend otherwise.

Analytical calculations of pressure head with respect to grid spacing resulted in about 0.3 ft/ft for the two model cells that contain the wells. The result indicates that if the grid spacing were increased by 1 ft, the resulted model predictions for the pressure heads for these grid cells would increase by 0.3 ft, or reduction in the influence of the pumping activity by 0.3 ft.

Table 1 illustrates that calculation of the analytical derivatives was noticeably faster than the calculations of the numerical derivatives by PEST using numerical estimates based on perturbations. For execution time comparison, all model runs were made on a DELL Precision M60 laptop, equipped with a 1.5 GH Pentium processor, 2 GB PC2100 RAM, and a main board with 266MH external bus speed. The cost saving in run time was considerably higher when the above problem was slightly modified by considering a non-homogenous hydraulic conductivity field and requesting PEST and MODFLOW (with the analytical derivatives) to estimate the derivatives of each head with respect to the transmissivity of each grid cell. The analytical derivatives approach provided a considerable cost saving over PEST with respect to runtime for the small example discussed herein.

UNCERTAINTY ANALYSIS

Users of computer models employ uncertainty analysis mainly for the following purposes: examination of how much confidence one should have in the model prediction; study of the effect of errors in field parameters, used as input data, on model calibration and verification; investigation of the propagation of errors in the input data on the predicted results; improvement of model calibration with availability of additional field data; and decreasing uncertainty of model results and obtain further calibration of the input parameters by minimization of uncertainty-based error functions (inverse problem).

Table 1. Comparison of computer run time between analytical and numerical estimates of the derivatives with respect to pressure heads

Independent Variable	Analytical Calculations (seconds)	Numerical Estimations (seconds)
Vertical conductance (homogeneous-scalar)	0.13	0.90
Vertical conductance (heterogeneous-matrix)	4.77	22.54
Transmissivity (homogeneous-scalar)	0.20	1.31
Pumping rate (homogeneous-scalar)	0.18	0.50
River stage (homogeneous-scalar)	0.20	1.20
Grid spacing (homogeneous-scalar)	0.18	

The uncertainty analysis methods can be grouped in four categories: respond surface or statistical methods, fussy logic methods, cross validation methods, and minimization of a generalized Baysian loss function methods. For the respond surface method, the analyst first chooses a small subset of the system parameters using their "engineering judgment." In the second step the analyst selects a perturbation pattern for the selected parameters from experimental design theories. In the third step the analyst uses the perturbation pattern to perform multiple code runs that provide new response (output) values. In the fourth step the analyst employs simple functional forms to fit the response data to a multidimensional response surface based on a pre-determined probabilistic distribution that simplifies the original complexity of the model in question. In the fifth step the analyst utilizes sophisticated statistical algorithms (e.g., moment matching, Monte Carlo) to estimate statistical properties and uncertainty distributions of the responses.

For the minimization of a generalized Bayesian loss function the analyst solves an optimization problem (minimization of a constrained Lagrangian function) of the residual vector multiplies with some form of the covariance matrix. The optimization problem consistently combines field measurements with model outputs to simultaneously obtain best estimates for model parameters and reduce uncertainties in model outputs. This method considers all model parameters; thus, it guarantees that no important effects are overlooked by generating a full set of sensitivities. A full set means that the sensitivities with respect to all parameters are computed, without making an *a-priori* judgment as to which one is important.

To demonstrate the use of analytical derivatives in uncertainty analysis, the model described in Figure 1 is used to predict total discharge to the river from both aquifers. The discharge is estimated at 80,000 ft3/day by MODFLOW. The modified MODFLOW is used to calculate both the derivatives of head at observation points and the total discharge to the river. The derivatives calculated using the modified MODFLOW are then fed to PEST to calculate the uncertainty in predicted discharge to the river. PEST calculates upper and lower bounds of the predicted discharge. The upper and lower bounds of the discharge are ~85,500 ft3/day and ~76,300 ft3/day respectively. These bounds are estimated as the upper and lower critical points at a predefined value of the objective function (Doherty 2002). The same uncertainty calculations were performed using PEST numerical derivatives. Both methods produced similar results. The advantage of using analytical derivatives was limited to the cost saving in runtime and accuracy.

For example, in the case of numerical derivatives, MODFLOW needed to be modified to write the head and flux values with greater precision in order for PEST to converge to the predefined value of the objective function. This issue was irrelevant in the case of analytical derivatives supplied by the modified version of MODFLOW. The analytical derivatives are calculated accurately by MODFLOW and provided to PEST directly and the solution converged faster to the predefined value of the objective function.

CONCLUSION

The presented work illustrates the great promise offered by automatic differentiation for analytically computing derivatives. The results of this work demonstrated that analytical calculations are more accurate, take less time to compute, and their values are not functions of the size of perturbations the analyst has chosen to use, or the method of differentiation. As a result, automatic differentiation provides great benefits for both sensitivity and uncertainty analysis; especially in the case of problems with complex parameter structure, spatial and temporal variations in parameters, and highly nonlinear problems. In addition, using ADIFOR facilitates analytical calculations of the dependent variable with respect to numerical grid spacing without modifying the grid. This capability enables modelers to decide upon grid spacing using systematic and logical procedures.

References

Cooley, R. L., and R.L. Naff, 1990, Regression modeling of ground-water flow: U.S. Geological Survey Techniques in Water Resources Investigations, book 3, chap, B4, 232#p.

Doherty, J. 2002. PEST; Model-Independent Parameter Estimation; User Manual: 5th edition. Watermark Numerical Computing, Brisbane, Australia.

Griewank, A., 2000. ”\Evaluating derivatives: principles and techniques of algorithmic differentiation, SIAM, Philadelphia, 369 P.

Fagan, M., and A. Carle, 2001. Adifor 3.0: Design and Implementation., Rice University Technical Report, 2001.

Harbaugh A.W., Banta E.R., Hill M.C., and McDonald, M.G., 2000. MODFLOW-2000, 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.

Hill, M.C., 1992, A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional, ground-water flow model using nonlinear regression, U.S. Geological Survey Open-File Report 91-484, 358 p.

Hill, M.C., 1998. Methods and guidelines for effective model calibration: U.S. Geological Survey, Water-Resources Investigations Report 98-4005.

Mcdonald, M. G. & A. W. Harbaugh. 1988. A Modular Three-Dimensional Finite-Difference Ground-Water Flow Model, U.S. Geological Survey Techniques of Water Resources Investigations, Book 6, Chapter A!, 586 p.

Manganalli, S., Valadino Ceci, and Walter Vecchiato, 2002. Sensitivity Analysis of Volatility: A New Tool for Risk Management, European Central Bank, Working Papar No. 194.

Martins J.R.R.A., Ilan Kroo, and Juan J. Alonso, 2000. An Automated Method for Sensitivity Analysis Using Complex Variables, AIAA-2000-0689.

Park, C.K., K.J. Park, K.S. Lee, D.S. Bae, and H.J. Cho, . Development of an Efficient Algorithm for Dynamic and Design Sensitivity Analysis of Railway Vehicle Systems.