Application of Automatic Differentiation in Groundwater Sensitivity Analysis

Baalousha, H.

Hawke's Bay Regional Council, Environmental Management Group.
Email: Baalousha@web.de

Keywords: *Sensitivity Analysis, Groundwater Modelling, Automatic Differentiation, uncertainty analysis.*

EXTENDED ABSTRACT

Derivatives of functions are used very often in groundwater modelling. Computation of sensitivity, optimization problems and inverse modelling require evaluating of derivatives. The most commonly used method for derivative estimation is the finite difference method. This method, however, suffers from poor accuracy and its result is highly dependent on the step size of the finite difference. In addition, the method is slow and requires solving the function many times to obtain the derivative.

Automatic differentiation is a powerful technique to compute the derivatives of a function given by a piece of code. The derivative of a two-dimensional finite element groundwater flow and contaminant transport model (MCB2D) was obtained using Automatic Differentiation of Fortran (ADIFOR) and was used for groundwater sensitivity analysis.

Two input parameters were considered as uncertain (random): hydraulic conductivity and groundwater recharge. Sensitivity analysis was done to see the effect of the hydraulic conductivity and groundwater recharge on the model output. Sensitivity analysis has shown that the model results are more sensitive to changes in hydraulic conductivity than groundwater recharge.

Uncertainty analysis was also done to investigate the effect of uncertainty in model input parameters on the model output. The uncertainty in input parameters was changed by considering different values of coefficient of variation (COV).

Results of uncertainty analysis show that the uncertainty in hydraulic conductivity has more impact on the model results than the uncertainty in groundwater recharge. As the uncertainty in model parameters increase, the model underestimates the output.

The results achieved in this study have revealed that automatic differentiation is an efficient tool for sensitivity analysis. Automatic differentiation is easy to apply on most groundwater models and does not require knowledge of the model code. The application of automatic differentiation is very promising in groundwater modelling as it reduces the uncertainty of classical differentiation methods.

1. INTRODUCTION

Groundwater models are very efficient tools for simulating groundwater movement and contaminant transport. Despite the power of the numerical models, the accuracy in their output is uncertain because of uncertainty in hydrogeological groundwater models (Baalousha 2003, Baalousha 2006). The sources of uncertainty in groundwater modelling can be classified into three categories: natural uncertainty, model uncertainty and parameter uncertainty. Natural uncertainty is the inherent variation in the physical system; it is stochastic, irreducible uncertainty. This randomness can exist in time or in space. Model uncertainty is a result of estimations in the mathematical equations used in the model and is referred to as conceptual error (Hua Lei and Schilling, 1996). Parameter uncertainty is associated with input data, which is used in any model. It is a result of errors in measurements and data collection.

Given the above-mentioned sources of uncertainty, the modelling process turns into a complicated task. The tendency has been to use groundwater models in a deterministic way, assuming the input parameters are accurate and representative of the reality. Thus, the accuracy of the output of deterministic models is questionable.

As the number of input parameters in groundwater modelling is usually large, it is important to concentrate on those parameters that have a greater influence on the model results. Sensitivity analysis has widely been used to find out the importance of each input parameter in groundwater modelling. That is, to estimate the rate of change in the output of a model with respect to changes in model input. It has also been used to allocate and design sampling sites. This analysis is very helpful in designing site investigation wells since the sampling points should be located at points of high sensitivity. Moreover, sensitivity analysis can help in modelling process to pay more attention on the important input variables, and thus, improve the calibration process. Sensitivity analysis is also important to understand the general behaviour of a model. As a result, uncertainty of a model output can be reduced if sensitivity analysis is being carried out.

The most widely used method for computing derivatives in groundwater models is the finite difference method. The accuracy of finite difference method depends on the step size of the difference. In one hand, if the step size is too large, the truncation error will be large. On the other hand, a small step size will result in large cancellation error. As a result, it is difficult to determine the appropriate step size for to achieve a higher accuracy. In groundwater problems, finite difference method is the common way to obtain the derivative of functions for sensitivity analysis and optimization. However, the finite difference method is not accurate, its accuracy is dependent on the increment size which affects the convergence rate (Baalousha 2006).

Therefore, the uncertainty resulting from the finite difference estimation with the other sources of uncertainty adds more errors into the model output.

Automatic differentiation is a good alternative for evaluating the gradient vector instead of using the crude finite difference method or manual method. The advantages of automatic differentiation are that it is easy to implement and does not require any knowledge of the original code contents.

In this study, automatic differentiation was used for sensitivity analysis in a groundwater flow and contaminant transport model.

2. AUTOMATIC DIFFERENTIATION

Automatic differentiation is used to evaluate the derivative of the models codes. The advantages of automatic differentiation are that it is easy to implement, does not require any knowledge of the original code contents and can be applied to any model code. The accuracy of automatic differentiation is up to machine precision.

ADIFOR "Automatic Differentiation of Fortran" (Bishof et. al. 1996, Bischof et.al. 2002) is a Fortran pre-processor to generate a code that computes the partial derivatives of dependent variables with respect to pre-defined independent variables. There are two approaches for computing derivatives of functions using automatic differentiation: the forward mode and the reverse mode (Bischof et.al. 2002). As most groundwater models codes are written in Fortran, ADIFOR is the most appropriate automatic differentiation tool.

The idea of automatic differentiation is dependent on the fact that any computer code, regardless to its length, is composed of set of mathematical operations such as summation, multiplication, etc. Using the chain rule of calculus on all mathematical operation in the code, ADIFOR obtains the derivative of a dependent variable with respect to the independent one.

2.1. Forward Mode

Given a simple composition $f(x)=g(y(x))$ the chain rule gives:

$$\frac{df}{dx}=\frac{dg}{dy}\frac{dy}{dx} \quad \text{Equation 1}$$

Forward mode traverses the chain rule from right to left, that is dy/dx is computed first and then dg/dy is computed. Forward mode is superior for functions f: $\Re \rightarrow \Re^m$ with m>>1. So the forward mode is appropriate when the number of input variable is low.

2.2. Reverse Mode

The reverse mode traverses the chain rule from left to right. Reverse mode is superior to forward mode for functions f: $\Re^n \rightarrow \Re$ with n>>1. Therefore, reverse mode is good for problems with many input parameters.

In this study, ADIFOR was used to generate the derivative code of the two-dimensional finite element groundwater flow and contaminant transport model (MCB2D) (Sun 96). The resulted code by ADIFOR is a Fortran subroutine including the derivative code for the pre-defined input parameters and the original model code. ADIFOR 2.0 uses the forward mode to obtain the derivative. The input uncertain parameters, for which the derivative of the model output should be obtained, are identified for the model as the hydraulic conductivity and the groundwater recharge.

Hydraulic conductivity and groundwater recharge have the highest uncertainty among other input parameters. This is the reason why they were considered as random variables.
The resulting derivative code of the original model was obtained and the required gradient vector was evaluated with a very good accuracy and fewer computations and time in comparison to the finite difference method.

2.3. Contaminant Transport model

A groundwater flow and contaminant transport model was used in this study and coupled with automatic differentiation. MCB2D is a two-dimensional finite element groundwater flow and transport model written in Fortran computer language. The model couples groundwater flow with contaminant transport using a Multiple Cell Balance Method (Sun 96) to solve the two-dimensional advection dispersion equation. After identification of uncertain input parameters (groundwater recharge and hydraulic conductivity in this case study), the derivative code was obtained using ADIFOR.

A case study from the northern area of the Gaza Strip (Figure 1), Palestine was used to demonstrate the use of automatic differentiation in contaminant transport model.

The finite element mesh consists of 532 nodes and 977 elements (Figure 2). A wastewater treatment plant in the area of study was considered as a point source of pollution. The aquifer in the area is unconfined with groundwater depth varying between 20 and 35 meters (above mean sea level).

The geology of the area is composed of calcareous sandstone, and gravel with high hydraulic conductivity. Values of aquifer parameters were obtained from pumping test data and from the literature (Melloul and Bachmat 1975, Yakirevich et. al. 1998). Values of groundwater recharge were obtained from the literature (Baalousha 2005, IWACO and WRAP 1995, Melloul A., Bachmat 1975). Pumping data and groundwater levels were obtained from the Palestinian Water Authority (PWA). Statistical analysis of the collected data was carried out to find the statistical distributions of the input parameters (mean, variance and probability distribution).

Steady state conditions were assumed at the beginning of simulation (1995) and the output was used for transient simulation (between 1995-2005).

2.4. Sensitivity analysis

Sensitivity of the model output (C) with respect to each input variable can be computed as follows:

$$S_{x_i} = \frac{\partial C}{\partial x_i} \qquad \text{Equation 2}$$

where S_{xi} is the sensitivity of the parameter x_i . In this case, the sensitivity parameters are the groundwater recharge (R), and the hydraulic conductivity (K). The mode was run to obtain the dimensionless concentration of pollutant. That is, the dimensionless concentration is (C/C_0), where C is the pollution concentration at any point and C_0 is the pollution concentration at the source.

Figure 3 and Figure 4 show the sensitivity of model output (concentration of pollutant) with respect to hydraulic conductivity and groundwater recharge respectively at the end of transient simulation period.

From sensitivity figures, the following conclusion can be drawn:

- In general, the model output is more sensitive to changes in hydraulic conductivity than changes in groundwater recharge.
- It is also clear that the sensitivity of hydraulic conductivity is high at the sources of pollution and in the contamination path.
- The negative values of sensitivity indicate that a decrease in the value of parameter leads to higher probability of exceedance.
- The sensitivity of groundwater recharge is small in general compared to the sensitivity of hydraulic conductivity.

2.5. Uncertainty Analysis

Uncertainty analysis of model input was carried out to investigate the effect of uncertainty in model parameters on the model output. The uncertainty measure is the coefficient of variation for each random variable. Coefficient of variation (CV) is a statistical measure of the deviation of a variable from its mean, and it is used to determine the degree of relative dispersion of the population. That is, CV is the standard deviation divided by the mean value of a population.

Different formulations of coefficient of variation (COV) for each input parameter were set up and the model output was obtained at each formulation. Different values of COV were used for this purpose. For hydraulic conductivity, COV was given different values as shown in Figure 5 and the model was run keeping the other parameters constant. The same procedure was followed for groundwater recharge. Finally, both hydraulic conductivity and groundwater recharge were given different values of COV simultaneously and the results of probability of failure were obtained for each case. Figure 5 shows model output at different values of COV.

The first line in Figure 5 (lower line) shows the effect of uncertainty in hydraulic conductivity alone; the upper line (top) shows effect of uncertainty in the groundwater recharge and the middle line effect of uncertainty in both groundwater recharge and hydraulic conductivity. For all cases, it was found that the concentration of contamination decreases as the uncertainty of parameter increases and vice versa. So the increase in uncertainty of model parameters results in underestimation of model output.

It was found that the model output is more sensitive to uncertainty in hydraulic conductivity than the uncertainty in groundwater recharge. In case of low uncertainty (low COV), both hydraulic conductivity and groundwater recharge have the same degree of influence on the model output.

2.6. Conclusions and Recommendations

Automatic differentiation is a powerful technique for computing derivative codes of groundwater models up to many orders and with precision of the machine code. Automatic differentiation of Fortran (ADIFOR) is good and suites groundwater models as it works with Fortran-written programs.

Results of sensitivity analysis reveal that the hydraulic conductivity has greater effect on the model results than the groundwater recharge. Uncertainty of model input parameters plays a big role in contaminant transport modelling.
Based on the results of uncertainty analysis, it was found that the model output decreases as uncertainty in either parameter (hydraulic conductivity and groundwater recharge) increases. The model is more sensitive to likely changes in uncertainty of hydraulic conductivity than groundwater recharge.

2.7. References

Baalousha, H. (2003), Risk Assessment and Uncertainty Analysis in Groundwater Modelling. Ph.D. Dissertation. Fak. 03 Fak. für Bauingenieurwesen. Signatur:Sz8578. Germany.

Baalousha, H. (2005), Using CRD method for quantification of groundwater recharge in the Gaza Strip, Palestine, *Environmental Geology,* (48), 889-900.

Baalousha, H. (2006), Groundwater Pollution Risk Using a Modified Latin Hypercube Sampling, *Hydroinformatics Journal,* 8(3), 223–234

Bischof C., P. Khademi, A. Mauer and A. Carle (1996), ADIFOR2.0: Automatic differentiating of FORTRAN77 programs, Technical report, IEEE Computational Science and Engineering 1996.

Bischof, C., M. Bücker and B. Lang, (2002), Automatic Differentiation for Computational Finance Computational Methods in Decision-Making, Economics and Finance, Kluwer Academic Publishers

Hua Lei, J. and W. Schilling, (1996), Preliminary Uncertainty Analysis for Assessing the Predictive Uncertainty of Hydrologic Models, *Water Science and Technology,* 33 (2), 79–90.

IWACO and WRAP. (1995), Groundwater Resources Assessment of The Gaza Strip, Technical Report, IWACO Euroconsult and Water Resources Action Program (WRAP)

Kuczera G. and E. Prent (1998), Monte Carlo Assessment of Parameter Uncertainty in Conceptual Catchment Models: the Metropolis Algorithm, *Journal of Hydrology,* (211), 69-85.

Melloul A. and Y. Bachmat (1975), Evaluation of Hydrological Situation as a Basis for Operational Management of the Coastal Plain Aquifer of the Gaza Strip, Report 50 pp, Hydrological Service of Israel.

Sun Ne-Zheng. (1996), Mathematical Modeling of Groundwater Pollution, New- York: Springer.

Yakirevich A., A. Melloul, S. Sorek, S. Shaath and V. Borisov. (1998), Simulation of Seawater Intrusion into the Khan Yunis Area of the Gaza Strip Coastal Aquifer, *Hydrogeology Journal*, 6(4), 549-559.

Figure 1 The study area

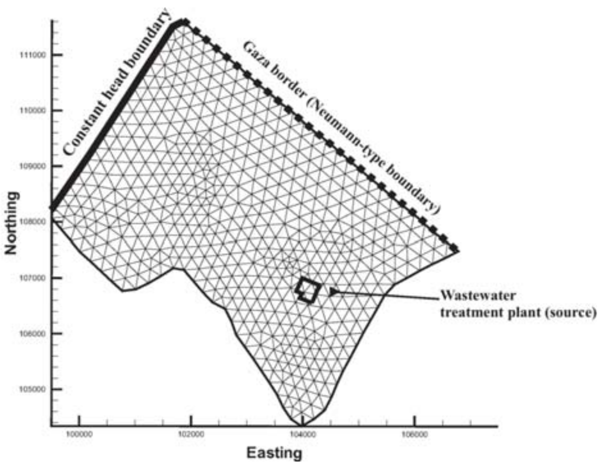

Figure 2 2D finite element mesh and boundary conditions

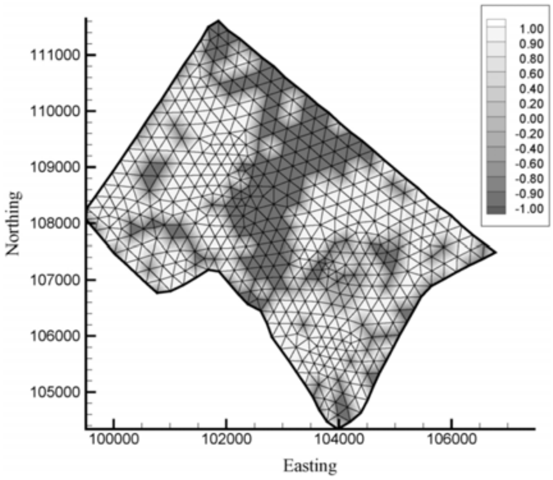

Figure 3 Sensitivity of contaminant transport model with respect to hydraulic conductivity

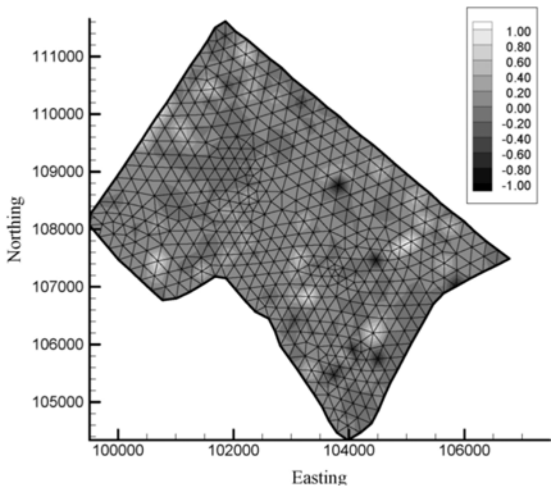

Figure 4 Sensitivity of contaminant transport model with respect to groundwater recharge.

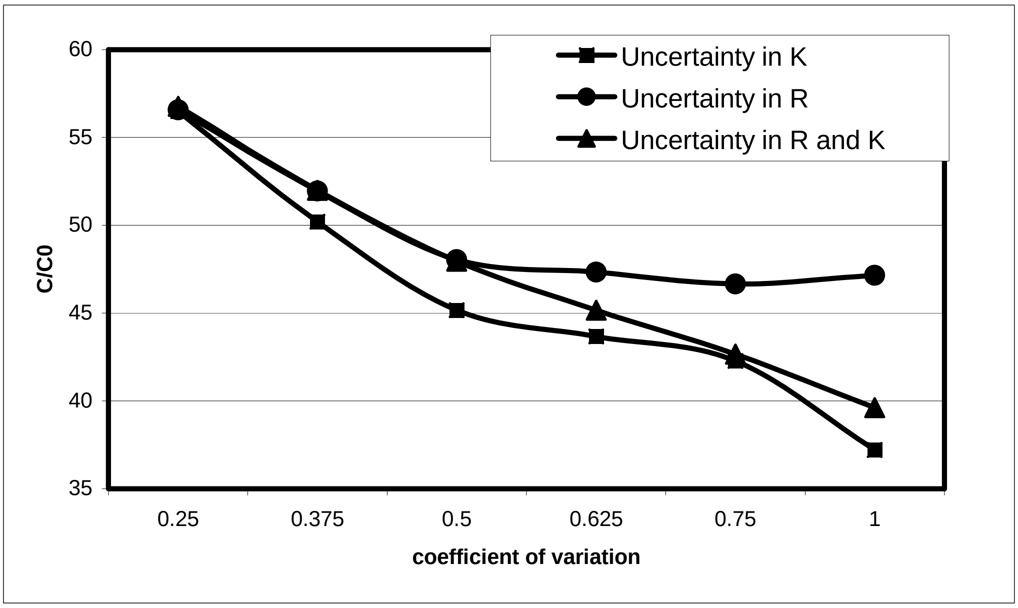

Figure 5 Effect of uncertainty in model input parameters on the results