Groundwater Resources Management under Uncertainty

Harald Kunstmann*, Wolfgang Kinzelbach* & Gerrit van Tonder**

*Institute of Hydromechanics and Water Resources Management, Swiss Federal Institute of Technology Zürich, ETH Hönggerberg, CH 8093 Zürich, Switzerland

Email: kunstmann@ihw.baum.ethz.ch

**Institute for Groundwater Studies, University of the Free State, Box 339, Bloemfontein 9300, Republic of South Africa

Abstract

Groundwater resources and aquifer parameters are usually poorly known and model results suffer from uncertainty. This uncertainty is acceptable only if it can be quantified. The quantification allows to take conservative decisions and to determine when a decision is only possible after further data collection. Mote Carlo techniques are suited for this analysis but usually require a huge computational effort. An alternative and computationally efficient approach is the First Order Second Moment Technique (FOSM) which directly propagates the uncertainty originating from parameter uncertainty into the result. We formulate the FOSM method for both the groundwater flow and transport equations in a way suitable for application to field studies. The FOSM method is applied to the computation of stochastic catchment zones and results are compared to straightforward Monte Carlo simulation. Finally uncertainties arising from the often encountered ill-posedness of the inverse problem and the incorporation of conditioning head information into the method are discussed.

1 Introduction

Groundwater models are widely used although their predictive power is limited due to the uncertainty of the model parameters. The reliability of predictions depends on the accuracy of available knowledge of aquifer

structure, boundary conditions, aquifer parameters and the future hydrological input parameters. Lacking aquifer parameters can be estimated by fitting modeled piezometric heads (and/or concentrations) to measurements made in the past. However, even a perfectly calibrated model cannot entirely remove uncertainty because the inverse problem does not necessarily have a unique solution.

Imperfect knowledge of aquifer parameters can be expressed by giving sets or distributions of input parameters instead of single values, e.g. a set of transmissivities and/or recharge rates. If the resulting range of the possible drawdowns (and/or concentrations) at an observation point is known, the probability of failure of a proposed management (and/or remediation) scheme can be estimated. Conservative scenarios with maximum allowed failure rates can then be designed. In this process stochastic modeling is required which contrary to deterministic modeling yields probabilities and percentiles rather than exact numbers.

The *Monte Carlo Method* is widely used in stochastic modeling. It is a versatile method that in principle can always be applied. It consists of a large number of deterministic calculations for single realizations of the problem and a statistical analysis of all results. However, the computational effort may become huge before results converge and the number of realizations necessary is not known in advance.

An alternative and computationally efficient approach is the *First Order Second Moment Technique* (FOSM) which directly propagates the uncertainty originating from parameter uncertainty. It can be applied in all cases where the parameter variances are moderate.

Assuming that imperfect knowledge of transmissivities and recharge rates is of major concern, we applied the FOSM-Method to the propagation of parameter uncertainty into both piezometric heads and contaminant concentrations.

2 Propagation of Uncertainties by the *First Order Second Moment* Technique

2.1 Flow Equation

The numerical solution of the two-dimensional groundwater flow equation for confined aquifers

$$\vec{\nabla} \cdot (T \nabla h) + q = S \frac{\partial h}{\partial t} \tag{1}$$

is usually obtained by *Finite Element-* or *Finite Difference* schemes for the unknown piezometric heads $\vec{h}(t + \Delta t)$ at given nodal parameters $\vec{p}$ (i.e.: $\vec{T}$: transmissivities, $\vec{q}$: recharge rates, $\vec{S}$: storage coefficients) and initial heads $\vec{h}(t)$. The aquifer parameters $\vec{p}$ are assumed to be uncertain. They are

described as having a mean value $\hat{\vec{p}}$ and a covariance $Cov(\hat{\vec{p}})$ (e.g. Gaussian).

In a first order approximation the mean heads $\hat{\vec{h}}$ (the first moments) are obtained by the solution of the PDE using the mean values of the aquifer parameters $\hat{\vec{p}}$. The propagation of parameter uncertainties into the head uncertainties (given by the covariance matrix $Cov(\vec{h})$) can be approximated when expanding $\vec{h}$ into a Taylor series to first order

$$\vec{h}(\vec{p}) \overset{First\ Order}{=} \vec{h}(\hat{\vec{p}}) + (D_{\vec{p}^T}\vec{h})\Big|_{\vec{p}=\hat{\vec{p}}} \cdot (\vec{p} - \hat{\vec{p}}) \tag{2}$$

and inserting it into

$$Cov_{hh} \equiv \quad Cov(\vec{h}) \quad = E\left[(\vec{h} - \hat{\vec{h}})(\vec{h} - \hat{\vec{h}})\right] \tag{3}$$

This finally yields

$$Cov_{hh} \overset{First\ Order}{=} (D_{\vec{p}^T}\vec{h})\Big|_{\vec{p}=\hat{\vec{p}}} \cdot Cov(\vec{p}) \cdot (D_{\vec{p}^T}\vec{h})^T\Big|_{\vec{p}=\hat{\vec{p}}} \tag{4}$$

(*Dettinger & Wilson, 1981*), (*Townley & Wilson, 1985*).

$(D_{\vec{p}^T}\vec{h})\Big|_{\vec{p}=\hat{\vec{p}}}$ is the 1st derivative of $\vec{h}$ with respect to the transpose of $\vec{p}$, evaluated at $\vec{p} = \hat{\vec{p}}$. It is a matrix that expresses the sensitivity of h_i with respect to p_j. If the input parameters, T and q, are assumed to be lognormally distributed (i.e. $Y = ln(T)$ and $Z=ln(q)$ are normally distributed) this finally leads in steady state simulation to

$$Cov_{hh}\Big|_{Steady\ State} \overset{First\ Order}{=} D_{hY} \cdot Cov_{YY} \cdot D_{hY}^{\ T} + D_{hZ} \cdot Cov_{ZZ} \cdot D_{hZ}^{\ T} \tag{5}$$

$$Cov_{hY}\Big|_{Steady\ State} \overset{First\ Order}{=} D_{hY} \cdot Cov_{YY} \tag{6}$$

$$Cov_{hZ}\Big|_{Steady\ State} \overset{First\ Order}{=} D_{hZ} \cdot Cov_{ZZ}\,. \tag{7}$$

The diagonal of Cov_{hh} contains the standard devation of the piezometric heads σ_h at each grid point: $\sigma_{h_k} = \sqrt{Cov_{h_k h_k}}$.

2.2 Transport equation

In analogy to (2.1) the FOSM method can also be applied to the groundwater transport equation

$$n_m \frac{\partial c}{\partial t} + n_m\left(\vec{\nabla}\cdot(\vec{v}c) - \vec{\nabla}\cdot\left(\left(D_{mol} + D_{disp}\right)\cdot\vec{\nabla}c\right)\right) - qc_m = 0, \text{ with } \vec{v} = -\frac{1}{n_m}\frac{e^Y}{m}\vec{\nabla}h \tag{8),(9}$$

In steady state, the uncertainty of the concentration is given by its covariance matrix

$$\begin{aligned} Cov_{cc}\Big|_{Steady\ State} \overset{First\ Order}{=}\ & D_{ch} \cdot Cov_{hh} \cdot D_{ch}^{\ T} + + D_{cY} \cdot Cov_{YY} \cdot D_{cY}^{\ T} + D_{cZ} \cdot Cov_{ZZ} \cdot D_{cZ}^{\ T} + \\ & + D_{ch} \cdot Cov_{hY} \cdot D_{cY}^{\ T} + D_{cY} \cdot Cov_{Yh} \cdot D_{ch}^{\ T} + \\ & + D_{ch} \cdot Cov_{hZ} \cdot D_{cZ}^{\ T} + D_{cZ} \cdot Cov_{Zh} \cdot D_{ch}^{\ T} \end{aligned} \tag{10}$$

Computation of Cov_{cc} requires the evaluation of Cov_{hh}, Cov_{hY}, and Cov_{hZ}, which are determined by the covariance propagation of the groundwater flow equation as described in (2.1). The uncertainty of the concentrations, σ_c, is, again, given by the diagonal elements of Cov_{cc}.

2.3 Calculation of Sensitivity Matrices

None of the sensitivity matrices D requires the repeated evaluation of the flow or transport equation, which is the reason for the computational efficiency of the FOSM method. This can be seen from the following (illustrated for the transport equation):

Discretization of the transport equation in a *Finite Difference* scheme leads to

$$\begin{aligned} &A_{j,i}c_{j-1,i-1}(k\Delta t) + B_{j,i}c_{j,i-1}(k\Delta t) + C_{j,i}c_{j-1,i}(k\Delta t) + \\ &D_{j,i}c_{j,i}(k\Delta t) + E_{j,i}c_{j,i+1}(k\Delta t) + F_{j,i}c_{j+1,i}(k\Delta t) + \\ &G_{j,i}c_{j+1,i+1}(k\Delta t) + H_{j,i}c_{j-1,i+1}(k\Delta t) + I_{j,i}c_{j+1,i-1}(k\Delta t) = K_{j,i}c_{j,i}((k-1)\Delta t) + L_{j,i} \end{aligned} \tag{11}$$

which can be summarized in tensor notation as

$$\mathrm{A}\cdot\vec{c}(k\Delta t) = \mathrm{B}\cdot\vec{c}((k-1)\Delta t) + \vec{d} \tag{12}$$

and solved according to

$$\vec{c}(k\Delta t) = \mathrm{A}^{-1}\cdot\mathrm{B}\cdot\vec{c}((k-1)\Delta t) + \mathrm{A}^{-1}\cdot\vec{d} \tag{13}$$

The sensitivity matrix of the concentrations with respect to the parameters (D_{cp}) can be derived with Vetter's calculus (*Vetter, 1973*):

$$D_{cp} = D_{\vec{p}}\left[\vec{c}(k\Delta t)\right] = D_{\vec{p}}\left[\mathrm{A}^{-1}\cdot\mathrm{B}\cdot\vec{c}((k-1)\Delta t) + \mathrm{A}^{-1}\cdot\vec{d}\right] = \ldots = -\mathrm{A}^{-1}\cdot D_{\vec{p}}\left[\mathrm{A}\vec{c}(k\Delta t)\right] \tag{14}$$

Defining $\vec{M} = \mathrm{A}\vec{c}(k\Delta t)$, which is in fact

$$\vec{M}\Big|_{l=(j,i)} = A_{j,i}c_{j-1,i-1}(k\Delta t) + \mathrm{B}_{j,i}c_{j,i-1}(k\Delta t) + \ldots + I_{j,i}c_{j+1,i-1}(k\Delta t) \tag{15}$$

yields

$$D_{cp} = -\mathrm{A}^{-1}\cdot D_{\vec{p}}\Big|_{\vec{p}=\vec{p}}\left[\vec{M}\right] \tag{16}$$

with

$$D_{\vec{p}}\Big|_{\vec{p}=\vec{p}}\left[\vec{M}\right]\Big|_{l,s} = \frac{\partial A_{(j,i)=l}}{\partial p_s}c_{j-1,i-1}(k\Delta t) + \frac{\partial B_{(j,i)=l}}{\partial p_s}c_{j,i-1}(k\Delta t) + \ldots + \frac{\partial I_{(j,i)=l}}{\partial p_s}c_{j+1,i-1}(k\Delta t) \tag{17}$$

The matrices $A,B,\ldots I$, are simple coefficient matrices whose derivatives $\partial A_{(j,i)=l}/\partial p_s, \partial B_{(j,i)=l}/\partial p_s, \ldots, \partial I_{(j,i)=l}/\partial p_s$ can directly be calculated, either analytically or using a secant approximation. D_{hZ} and D_{hY} are also used in Inverse Modeling algorithms. The gain in CPU time savings by the direct evaluation of these sensitivities is enormous.

2.4 Application: Calculation of Stochastic Catchment Zones

Catchment zones can be constructed by particle tracking in a given flow field. Alternatively, backward tracking from a well can be used. This procedure can also be replaced by reversing the flow field and injecting fluid with concentration 1 into the well. If transport is calculated without

dispersion, the plume with concentration 1 is congruent to the catchment. If dispersion is taken into account the backward calculation (*Kolmogorov Backward Equation*, see (*Uffink, 1989*)) has only a probabilistic meaning. The concentration now indicates the probability to reach the well from a certain point in the flow field under the given dispersion. We will define the isoline $c=0.5$ as the boundary of the catchment zone. Then stochastic catchment zones of a well under uncertain flow parameters T and q are obtained by calculating first and second moments $\hat{c}$ and σ_c. Assuming that the distribution of the computed concentration is normal, one can assign to each grid point a probability of belonging to the catchment zone of the well: The expected catchment zone is defined by the concentration isoline $\hat{c}=0.5$. The real catchment zone however can differ from it and with 95.5% probability its border lies between the isolines $\hat{c}-2\sigma_c=0.5$ and $\hat{c}+2\sigma_c=0.5$.

Figures 1 and 2 show the mean concentration and its standard deviation in a synthetic test case. The domain is discretizd by 1024 nodes and consists of three zones for both recharge and transmissivity. The input uncertainties of the logarithmic quantities were assumed to be $\sigma_Y=0.6$, $\sigma_Z=0.6$:

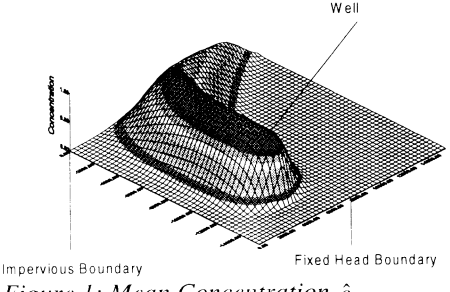

Figure 1: Mean Concentration $\hat{c}$

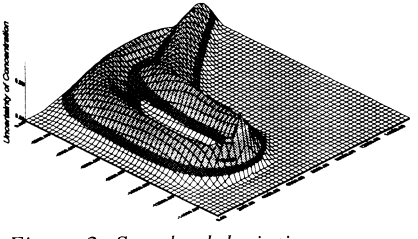

Figure 2: Standard deviation σ_c

The resulting confidence intervals for the borderline of the catchment zones can be seen in figures 3 and 4 both for the FOSM and the Monte Carlo method:

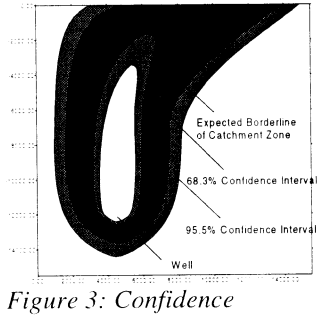

Figure 3: Confidence intervals of catchment zone (FOSM)

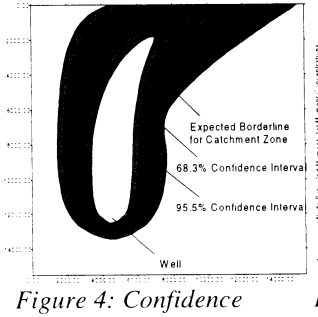

Figure 4: Confidence intervals of catchment zone (Monte Carlo)

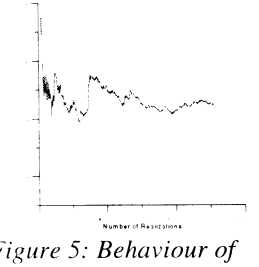

Figure 5: Behaviour of moments' convergence (Monte Carlo)

Transactions on Modelling and Simulation vol 17, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

The results of both methods differ only insignificantly. Convergence of first and second moments in the Monte Carlo calculation is illustrated by the behaviour of the ratio of both quantities shown in figure 5. Comparison of calculation times reveals CPU time savings by a factor of 10 in the FOSM method over the Monte Carlo method (if 1024 realizations are assumed to be sufficient for convergence).

2.5 Strategy For Field Case Applications

The computational efficiency of the FOSM method is limited by the size of the matrices involved. To guarantee the feasibility in real case studies the following features were implemented into the numerical algorithms of our code "UFLOW" (*Uncertainty & Flow, H. Kunstmann, ETH-Zürich, 1998*):

- Reduction of the number of uncertain parameters by zonation and/or interpolation between (uncertain) pivot points
- Calculation of covariance propagation on a coarser grid than the one used for the calculation of first moments
- Calculation only of elements of the covariance matrices that are needed (e.g. diagonal elements or elements where the first moment is greater than a threshold value (e.g. zero for concentrations))
- Utilization of symmetry and sparseness properties of matrices involved

The dimension of Cov_{hh} is crucial to the advantage of FOSM over the Monte Carlo Method. Therefore the number of nodes is presently limited to several thousands active cells. The FOSM method for flow and transport equations works very well, if the size of perturbations of the uncertain input parameters can be restricted by $\sigma(\ln(parameter)) \leq .5$; FOSM still works satisfactorily if $\sigma(\ln(parameter)) \leq 1$.

3. Quantification of Uncertainties Arising from the Non-Uniquenes of the Solution of the Inverse Problem

3.1 Principle

The questions remains, how the uncertainties σ_T and σ_q of the input parameters transmissivity T and recharge q can be determined. One can make qualified guesses on both their mean values and ranges based on pumping tests and water balance studies. In addition to that one usually estimates T and q by Inverse Modeling techniques. It is known however that the calibration of a model to observed heads (and/or concentrations) does not necessarily have a unique solution for T and q: one can possibly find a set of combinations of T and q that lead to a practically identical model output at the measurement locations. From our experience this type of uncertainty is the most important one in groundwater modelling. The

question therefore is: which set of recharges $q \in \{q \pm \Delta q\}$ will for a given set of possible transmissivities $T \in \{T \pm \Delta T\}$ lead to an identical model output at the measurement locations. Assuming an uncertainty ΔT_k of T at node k, the real value T' of the transmissivity can be $T' = T + \Delta T$. If the transmissivity is changed from T to T' at node k, the modeled head h at an observation node m will change. This change however can be compensated, if the recharge at a node l changes from value q to a value $q' = q + \Delta q$. The relationship between ΔT_k and the corresponding Δq at node l can be deduced from the total differential of h_m, which is given by

$$\Delta h_m = \frac{\partial h_m}{\partial T_k} \Delta T_k + \frac{\partial h_m}{\partial q_l} \Delta q_l \tag{18}$$

The condition of compensation $\Delta h_m = 0$ leads to

$$\Delta q_l = -\left(\frac{\partial h_m}{\partial T_k} / \frac{\partial h_m}{\partial q_l}\right) \Delta T_k \tag{19}$$

If there are several measurement locations, a shift of T to $T' = T + \Delta T$ will affect the calculated head values at *all* observation points. A generalization of the introductory example shows how to deal with this situation

3.2 Generalization of the Methodology

Assuming N measured heads, M_Y values of uncertain log-transmissivities, and M_Z values of uncertain log-recharge values (e.g. defined by zonation or by pivot point values) one can in analogy to chapter 3.1 establish the following linear system of equations:

$$\begin{aligned}
\Delta h_1 &= \sum_{k=1}^{M_Y} \frac{\partial h_1}{\partial Y_k} \Delta Y_k + \sum_{l=1}^{M_Z} \frac{\partial h_1}{\partial Z_l} \Delta Z_l \stackrel{!}{=} 0 \\
\Delta h_2 &= \sum_{k=1}^{M_Y} \frac{\partial h_2}{\partial Y_k} \Delta Y_k + \sum_{l=1}^{M_Z} \frac{\partial h_2}{\partial Z_l} \Delta Z_l \stackrel{!}{=} 0 \\
\Delta h_N &= \sum_{k=1}^{M_Y} \frac{\partial h_N}{\partial Y_k} \Delta Y_k + \sum_{l=1}^{M_Z} \frac{\partial h_N}{\partial Z_l} \Delta Z_l \stackrel{!}{=} 0
\end{aligned} \tag{20}$$

The sensitivities $\partial h_n / \partial Y_k$ and $\partial h_n / \partial Z_l$ are the entries of the sensitivity matrices D_{hY} and D_{hZ} as introduced in chapter 2.1. If the question of interest is, for example, the uncertainty range of log-recharge that corresponds to the uncertainty of log transmissivities, this linear system of equations must be solved for the unknowns $(\Delta Z_1, \ldots, \Delta Z_{M_Z})$. Three cases can be distinguished:

- $N = M_Z \Rightarrow$ there is one unique solution to the equation system.
- $N < M_Z \Rightarrow$ $(\Delta Z_1, \ldots, \Delta Z_{M_Z})$ are linearly dependent of each other; M_Z -N log-recharge uncertainties must be fixed by using additional *a priori* uncertainty information
- $N > M_Z \Rightarrow$ a regressive solution for $(\Delta Z_1, \ldots, \Delta Z_{M_Z})$ can be found.

If there are concentration measurements available one can additionally demand $\Delta\vec{c} = 0$ and obtain another set of equations restricting $(\Delta Z_1, ..., \Delta Z_{M_Z})$.

3.3 Outlook for the Methodology

The proposed methodology can be used in the conditioning of stochastic simulations and is easily incorporated into the FOSM method. The incorporation of head (or concentration) information into the Monte Carlo approach is much more tedious and costly. Whereas in the test example presented above, ΔY and ΔZ were assumed independent (in an unconditioned approach), they can be linked to each other if measurements of heads (or concentrations) are available. Equations (5) and (10) have therefore to be generalized to contain the additional terms $D_{hY} \cdot Cov_{YZ} \cdot D_{hZ}^T + D_{hZ} \cdot Cov_{ZY} \cdot D_{hY}^T$, $D_{cY} \cdot Cov_{YZ} \cdot D_{cZ}^T + D_{cZ} \cdot Cov_{ZY} \cdot D_{cY}^T$, respectively. Cov_{ZY} is given by

$$Cov_{ZY} = -(D_{hZ}^T \cdot D_{hZ})^{-1} \cdot D_{hZ}^T D_{hY} \cdot Cov_{YY}. \quad (21)$$

These terms reduce the variance of the result. It is in this fashion that head (or concentration) measurements serve the improvement of uncertain models. The value of a measurement can be quantified by determining the amount of uncertainty reduction it brings into the final results.

In general all model computations should be subjected to a stochastic analysis of uncertainty. Far too often deterministic results are uncritically used for design purposes. Only quantified uncertainty gives the kind of safety required in rational management decisions.

References

[1] Dettinger M. D. & Wilson J. L., *First Order Analysis of Uncertainty in Numerical Models of Groundwater Flow, Part 1, Mathematical Development*, Water Resources Research, Vol. 17, No, 1, 1981

[2] Townley L. R. & Wilson, J. L., *Computationally Efficient Algorithms for Parameter Estimation and Uncertainty Propagation in Numerical Models of Groundwater Flow*, Water Resources Research, Vol. 21, No, 12, 1985

[3] Uffink, G., *Application of Kolmogorov's backward equation in random walk simulations of groundwater contaminant transport*, in Contaminant Transport in Groundwater, edited by Kobus & Kinzelbach, Balkema, Rotterdam, 1989

[4] Vetter, W. J., *Matrix calculus operations and Taylor expansions*, SIAM Rev., 15(2), 352-369, 1973