Calibration and Reliability in Groundwater Modelling: From Uncertainty to Decision Making
(Proceedings of ModelCARE'2005, The Hague, The Netherlands, June 2005). IAHS Publ. 304, 2006.

Comparison of stochastic and regression based methods for quantification of predictive uncertainty of model-simulated wellhead protection zones in heterogeneous aquifers

STEEN CHRISTENSEN[1], CATHERINE MOORE[2] & JOHN DOHERTY[2]

[1] *Department of Earth Sciences, University of Aarhus, Ny Munkegade b. 520, DK-8000 Aarhus C, Denmark*
sc@geo.au.dk

[2] *School of Engineering, University of Queensland, Brisbane, Queensland, Australia*

Abstract For a synthetic case we computed three types of individual prediction intervals for the location of the aquifer entry point of a particle that moves through a heterogeneous aquifer and ends up in a pumping well. (a) The nonlinear regression-based interval (Cooley, 2004) was found to be nearly accurate and required a few hundred model calls to be computed. (b) The linearized regression-based interval (Cooley, 2004) required just over a hundred model calls and also appeared to be nearly correct. (c) The calibration-constrained Monte Carlo interval (Doherty, 2003) was found to be narrower than the regression-based intervals but required about half a million model calls. It is unclear whether or not this type of prediction interval is accurate.

Keywords accuracy; computational requirements; Monte Carlo method; prediction interval; predictive uncertainty; regression based method; wellhead protection zone

INTRODUCTION

We study some stochastic and regression based methods that quantify the predictive uncertainty related to incomplete representation of spatial and temporal variations in hydrological and hydrogeological variables, and parameter uncertainty. A synthetic model case is used to compare the particle path uncertainty quantified by the various methods and to compare the computational requirements of the methods. In our example the hydraulic conductivity of the aquifer is a correlated Gaussian field which is described by m-vector $\boldsymbol{\beta}$, where m is the number of grid elements of the numerical model, and the hydraulic conductivity is assumed to be constant within each element. In the groundwater model used for prediction and for quantification of predictive uncertainty $\boldsymbol{\beta}$ is substituted by a parameter vector, $\boldsymbol{\theta}$, of dimension $p << m$, i.e. $\boldsymbol{\beta}$ is substituted by $\boldsymbol{\gamma\theta}$ where $\boldsymbol{\gamma}$ is an $m \times p$ interpolation or averaging matrix. This model simplification increases the uncertainty of model prediction because the simplified model neglects part of the actual hydrogeological heterogeneity. In our example we define a large number of base parameters as hydraulic conductivities at pilot points, which is reduced to a small number of super parameters, $\boldsymbol{\theta}$, that are to be estimated. Using this parameterization method was found, in this case, to produce much less prediction uncertainty than if using zonation, and with the regression-based methods of Cooley (2004) it forms an exact approach for quantifying uncertainty of model predictions.

METHODS USED TO QUANTIFY PREDICTIVE UNCERTAINTY

Regression based methods

We us the regression based methods described in the recent works of Christensen & Cooley (2003, 2004), Cooley (2004). Very briefly described, the methodology is that the *p* vector of model parameters, $\boldsymbol{\theta}$ is estimated by nonlinear regression, i.e. by minimizing the objective function:

$$S(\boldsymbol{\theta}) = [\mathbf{Y} - \mathbf{f}(\gamma\boldsymbol{\theta})]^T \boldsymbol{\omega} [\mathbf{Y} - \mathbf{f}(\gamma\boldsymbol{\theta})] \tag{1}$$

where $\mathbf{Y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\varepsilon}$ is the *n*-vector of observations, $\boldsymbol{\varepsilon}$ is the random independent observation error with variance σ_ε^2, $\mathbf{f}(\gamma\boldsymbol{\theta})$ is the *n*-vector of values simulated using $\boldsymbol{\gamma\theta}$ instead of $\boldsymbol{\beta}$, $\boldsymbol{\omega}$ is the $n \times n$ weight matrix, and ^{T} indicates transpose. The weight matrix $\boldsymbol{\omega}$ used in the following is an estimate of $\boldsymbol{\Omega}^{-1}$, where $\boldsymbol{\Omega}$ is the second moment matrix of the true model errors (Gauss-Markov estimation).

Cooley (2004) has shown that the limits of an individual prediction interval for a variable $Y_p = g(\boldsymbol{\beta}) + \boldsymbol{\varepsilon}_g$ can be computed as the extreme values:

$$\left(\min_{\boldsymbol{\theta}}\left(g(\gamma\boldsymbol{\theta}) + \upsilon\right), \max_{\boldsymbol{\theta}}\left(g(\gamma\boldsymbol{\theta}) + \upsilon\right)\right) \tag{2a}$$

subject to:

$$S(\boldsymbol{\theta}) \leq S(\hat{\boldsymbol{\theta}})\left(1 + c_p \frac{t_{\alpha/2}^2(n-p)}{n-p}\right) - \omega_p \upsilon^2 \tag{2b}$$

where $\hat{\boldsymbol{\theta}}$ is the regression estimate obtained by minimizing (1), $t_{\alpha/2}(n-p)$ is the $(1 - \alpha/2) \times 100$ percentile of the *t* distribution with $n - p$ degrees of freedom, c_p is a correction factor given by Christensen & Cooley (2004, eqn 55), $\upsilon = Y_p - g(\gamma\boldsymbol{\theta})$ is an estimate of the prediction error, and ω_p^{-1} is the variance of the true model error of the prediction. Cooley (2004) has shown that for nonlinear problems $\boldsymbol{\theta}$ and υ can be computed by solving equation (2) using the procedure given by Vecchia & Cooley (1987). It is noticed that computing each of the extreme values in (2a) corresponds to solving a nonlinear regression problem.

If $\mathbf{f}(\gamma\boldsymbol{\theta})$ and $g(\gamma\boldsymbol{\theta})$ are linearized then solving (2) reduces to Cooley (2004):

$$Y_p = g(\gamma\hat{\boldsymbol{\theta}}) \pm t_{\alpha/2}(n-p)\left(c_p S(\hat{\boldsymbol{\theta}})/(n-p)\right)^{\frac{1}{2}} \left(\mathbf{Z}(\mathbf{X}^T\boldsymbol{\omega}\mathbf{X})^{-1}\mathbf{Z}^T + \omega_p^{-1}\right)^{\frac{1}{2}} \tag{3}$$

which has the form of a standard linear prediction interval. In equation (3) $\mathbf{X} = [\partial f_i/\partial\theta_j]$ and $\mathbf{Z} = [\partial g/\partial\theta_j]$.

Monte Carlo methods

The stochastic method we use is the calibration-constrained Monte Carlo method described by Doherty (2003). Each of the Monte Carlo runs includes the following steps:

1. Generate a stochastic hydraulic conductivity field.
2. Estimate a field of factors that when multiplied (“warped”) with the field generated in step 1 produces a modified hydraulic conductivity field that ensures that $\Phi_m(\boldsymbol{\theta}) = [\mathbf{Y} - \mathbf{f}(\boldsymbol{\gamma}\boldsymbol{\theta})]^T [\mathbf{Y} - \mathbf{f}(\boldsymbol{\gamma}\boldsymbol{\theta})] = n\sigma_\varepsilon^2$, where $\boldsymbol{\theta}$ is a *p*-vector of pilot point parameters from which the factor field is spatially interpolated, and $\boldsymbol{\gamma}$ is the interpolation matrix.
3. Use the field estimated in step 2 to simulate the prediction Y_p.

Steps 1 to 3 are repeated a large number of times to estimate the cumulative probability distribution for the prediction, Y_p. The lower limit of the 95% prediction interval corresponds to the prediction at the 2.5% level of cumulative probability, and the upper limit corresponds to the prediction at the 97.5% level of cumulative probability.

To ensure that $\Phi_m(\boldsymbol{\theta}) = n\sigma_\varepsilon^2$ the dimension *p* of $\boldsymbol{\theta}$ is larger than the dimension *n* of $\mathbf{Y}$. To ensure that the estimation procedure used in step 2 converges in as few model runs as possible, and to ensure minimal deviation between the hydraulic conductivity fields computed in step 1 and step 2 of the procedure, we use the mathematical Tichonov regularization technique described by Doherty (2003, 2004) which imposes “maximum homogeneity” on the parameters, $\boldsymbol{\theta}$.

THE EXAMPLE

The dimensions of the two-dimensional flow domain (Fig. 1) are 13.5 by 12, divided into uniform structural elements of size 0.2 × 0.2, except for the left-most column of elements which have width 0.1, and the top row and bottom row of elements which have height 0.1. The transmissivity is constant within each structural element. There is a pumping well in the centre of the domain where groundwater is abstracted at a rate of 1. Boundary conditions include no flow across the top and bottom boundaries, a constant head along the right boundary, and a constant flux across the left boundary (simulated as recharge equal to 6.2152 over the left-most column of cells). The observations, $\mathbf{Y}$, consist of simulated hydraulic head at 30 locations (Fig. 1) perturbed by independent random error with a small variance, $\sigma_\varepsilon^2 = 0.01$.

The vector $\boldsymbol{\beta} \sim \mathrm{N}(\mathbf{0}, \sigma_\beta^2 \mathbf{V}_\beta)$ consists of spatially varying $\log_{10}$-transmissivity with exponential covariance, $\sigma_\beta^2 \mathbf{V}_\beta$, having a correlation scale of 3.0 in the *x*-direction and 0.3 in the *y*-direction, and $\sigma_\beta^2 = 0.7544$. The $\log_{10}$-transmissivity within the element with the pumping well is known to be 0.8686. We generated 2000 independent realizations of $\boldsymbol{\beta}$. These realizations were used to compute $\boldsymbol{\Omega}$ and ω_p (Christensen & Cooley, 2003) to be used with the regression-based methods, and to generate the cumulative probability distribution of the prediction, $Y_p = g(\boldsymbol{\beta})$, where $g(\boldsymbol{\beta})$ in this example is the entry point of a particle that ends up in the pumping well (Fig. 1).

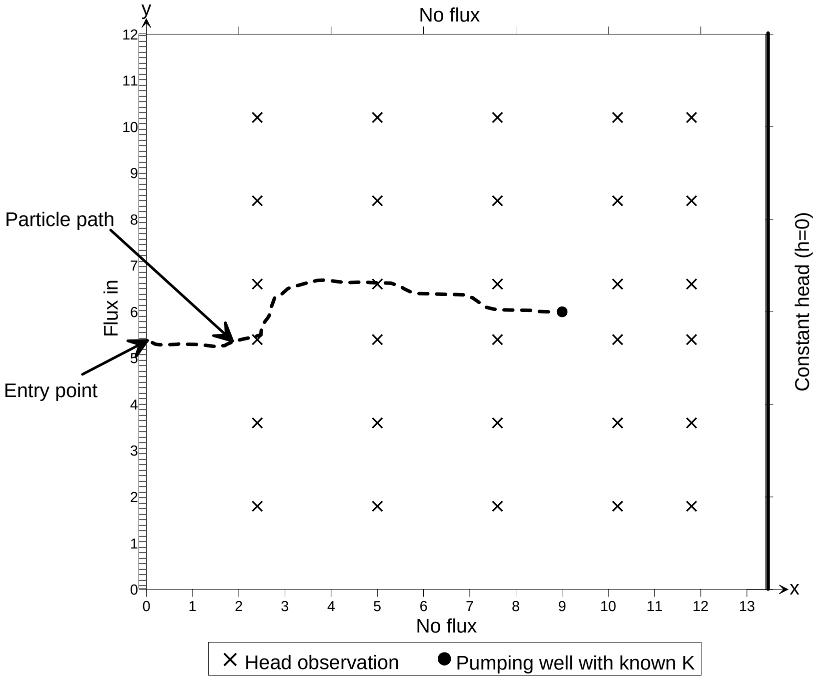

Fig. 1 Model domain, boundary conditions, location of pumping well and head observations, and example of particle path and entry point.

Table 1 Y coordinate (*Ytrue*) and level of cumulative probability of the particle entry point location for the five selected hydraulic conductivity field realizations.

Realization #	1556	1430	1558	479	502
Y_{true} of entry point	4.061	5.215	6.010	6.829	7.906
Level of cum. prob.	5%	25%	50%	75%	95%

The entry point is calculated by backward particle tracking using MODFLOW-2000 (Harbaugh *et al.*, 2000) with the ADV2 package (Anderman & Hill, 2001). On the basis of the cumulative probability distribution of Y_p we chose five different realizations of β (Table 1) for which we computed and compared 95% prediction intervals for the location of the entry point using the various methods.

For the regression-based methods we parameterized the model by defining 182 pilot points uniformly distributed over the model domain. The $\log_{10}$-transmissivity within the elements of the pilot points we call the base parameters. Because the number of base parameters is greater than the number of observations, the number of parameters that are actually estimated by nonlinear regression is reduced by using a singular value decomposition (SVD) technique similar to that of the SVD-Assist technique implemented in PEST and described by Doherty (2004). We call this reduced set of parameters the super parameters. For the regression based calculations of prediction intervals we used only nine super parameters. The correction factor, c_p = 0.812, to be used in (2b) was computed using the Corfac-2k program of Christensen & Cooley (2004). The parameter estimation and the calculation of prediction interval limits were carried out using PEST (Doherty, 2004) with MODFLOW-2000 (Harbaugh *et al.*, 2000) as the forward-problem simulator.

It was not mentioned in the conference proceedings version of this manuscript (Christensen *et al.*, 2005) that when using super parameters to parameterize a groundwater model then $\mathbf{\Omega}$ is a function of the weight matrix $\boldsymbol{\omega}$. We therefore used the following procedure to estimate $\boldsymbol{\omega} \approx \mathbf{\Omega}^{-1}$: (i) Begin with using $\boldsymbol{\omega} = \mathbf{I}$ in equation (1), carry out SVD of (1) to define super parameters, and compute the corresponding second moment matrix $\mathbf{\Omega}_0$ by the Monte Carlo method. (ii) Then set $\boldsymbol{\omega} = \mathbf{\Omega}_0^{-1}$, repeat the SVD of (1) to define the final super parameters, and repeat the Monte Carlo analysis to compute the corresponding second moment matrix $\mathbf{\Omega}$. In the example the result is that $\boldsymbol{\omega} = \mathbf{\Omega}_0^{-1} \approx \mathbf{\Omega}^{-1}$. Step (ii) was not carried through in Christensen *et al.* (2005), and the results presented here therefore differ slightly from those of Christensen *et al.* (2005).

For the calibration-constrained Monte Carlo method we parameterized the model by defining 195 pilot points uniformly distributed over the model domain that locate the base parameters from which the hydraulic conductivity multiplier field was interpolated. In estimating the base parameters we used Tichonov regularization to achieve maximum homogeneity of the parameter values. In estimation, the 195 base parameters were reduced to 50 super parameters by using PEST with its SVD-Assist functionality (Doherty, 2004).

When computing the regression based nonlinear prediction intervals we thus had to estimate nine super parameters three times, first to find the set of parameters that minimize equation (1), and then to estimate the set of parameters that gives each of the limits defined by equation (2). When computing the regression based linearized prediction intervals only the minimization of (1) had to be carried out. For the calibration-constrained Monte Carlo method we had to estimate 50 super parameters in step 2 of each of the 1000 Monte Carlo runs that were found necessary to obtain stabilized results.

RESULTS

Table 2 shows the calculated nonlinear 95% prediction intervals for the five chosen hydraulic conductivity realizations. The width of the prediction interval varies between 4.059 and 7.365, and in all five cases the true entry point location (Y_{true} in Table 1) falls inside the 95% prediction interval. The total number of model calls necessary to minimize equation (1) and compute the two limits defined by equation (2) is seen to vary between 222 and 469.

Table 2 The 95% prediction interval limits and total number of model calls for nonlinear intervals computed by the regression based method.

Realization #	1556	1430	1558	479	502
Upper limit	8.014	7.823	8.111	8.415	8.355
Lower limit	3.932	0.458	3.966	4.327	4.296
Width of interval	4.082	7.365	4.145	4.088	4.059
Number of model calls	222	469	267	224	244

Nonlinear intervals were actually calculated for 1998 realizations of $\boldsymbol{\beta}$. In 1863 cases (a frequency of 93.2%) Y_{true} is inside the 95% prediction interval, in 79 cases (4.0%) Y_{true} is above the interval, and in 56 cases (2.8%) it is below. The nonlinear prediction intervals thus appear to be nearly accurate.

Table 3 shows the calculated linearized 95% prediction intervals. It is noticed that the linearized intervals are sometimes wider, and sometimes narrower than the corresponding nonlinear intervals in Table 2. It is also noticed that the total number of model calls necessary to compute the linearized intervals as expected is about a third of the model calls necessary to compute the nonlinear intervals, and that in all five cases the true value of the entry point location falls inside the linearized interval. The linearized intervals were computed for 1998 realizations of $\boldsymbol{\beta}$, and these results showed that in 1868 cases (a frequency of 93.4%) Y_{true} is inside the 95% prediction interval. This indicates that the linearized intervals are also nearly accurate and tend to be slightly wider than the corresponding nonlinear intervals.

Table 3 The 95% prediction interval limits and total number of model calls for linearized intervals computed by the regression based method.

Realization #	1556	1430	1558	479	502
Upper limit	8.012	7.857	8.113	8.420	9.126
Lower limit	3.939	3.646	3.977	4.322	3.524
Width of interval	4.073	4.211	4.135	4.098	5.602
Number of model calls	83	104	104	83	104

Table 4 The 95% prediction interval limits and total number of model calls for intervals computed by the calibration constrained Monte Carlo method (warping 1000 fields using Tichonov regularization).

Realization #	1556	1430	1558	479	502
Upper limit	7.153	8.835	7.364	9.218	8.156
Lower limit	4.135	4.618	4.976	5.708	5.265
Width of interval	3.018	4.217	2.388	3.510	2.891
Number of model calls	418 187	485 411	(not avail.)	585 161	445 226

Table 4 shows the 95% prediction intervals computed by the calibration-constrained Monte Carlo method. For all five realizations the width of the Monte Carlo computed interval is narrower than or similar to both types of regression based intervals. The width of the intervals in Table 4 is, for example, between 57% and 86% of the width of the corresponding intervals in Table 2. For realization 1556 the true value of the entry point location falls below the interval. That the true entry point location falls outside the 95% prediction interval in one out of five cases could indicate that the intervals are inaccurate (too narrow). However, with only five cases there is no statistical significance for drawing such a conclusion, and it remains unclear whether or not the Monte Carlo based prediction intervals are accurate in this example. The total number of model calls necessary to compute prediction intervals by the Monte Carlo method is of the order of 1000 times greater than the number of model calls used for the regression-based intervals. (For realization 1556 we also computed the prediction interval without using PEST's SVD-Assist functionality. In this case the required number of model calls tripled.)

CONCLUSIONS

For a synthetic case we computed three types of individual prediction intervals for the location of the aquifer entry point of a particle that moves through a heterogeneous aquifer and ends up in a pumping well.

(a) The nonlinear regression-based interval (Cooley, 2004) was found to be nearly accurate and required a few hundred model calls to be computed.

(b) The linearized regression-based interval (Cooley, 2004) required just over a hundred model calls and also appeared to be nearly correct.

(c) The calibration-constrained Monte Carlo interval (Doherty, 2003) was found to be narrower than the regression-based intervals but required about half a million model calls. It is unclear whether or not the Monte Carlo based prediction intervals are accurate.

REFERENCES

Anderman, E. R. & Hill, M. C. (2001) MODFLOW-2000, the US Geological Survey modular ground-water model -- Documentation of the Advective-Transport Observation (ADV2) Package, version 2. *US Geol. Survey Open-File Report 01-54.*

Christensen, S. & Cooley, R. L. (2003) Experiences gained in testing a theory for modelling groundwater flow in heterogeneous media. In: *Calibration and Reliability in Groundwater Modelling: A Few Steps Closer to Reality* (ed. by K. Kovar & Z. Hrkal) (Proc. Int. Conf. ModelCARE'2002, Prague, June 2002), 22–27. IAHS Publ. 277. IAHS Press, Wallingford, UK.

Christensen, S. & Cooley, R. L. (2004) User guide to the UNC1NLI1 package and three utility programs for computation of nonlinear confidence and prediction intervals using MODFLOW-2000. *US Geol. Survey, Techniques and Methods Report 2004-1349.*

Christensen, S., Moore, C. & Doherty, J. (2005) Comparison of stochastic and regression based methods for quantification of predictive uncertainty of model-simulated wellhead protection zones in heterogeneous aquifers. In: *Calibration and Reliability in Groundwater Modelling: From Uncertainty to Decision Making*, 440–447. Pre-published Proc. Int. Conf. ModelCARE' 2005, The Hague, The Netherlands.

Cooley, R. L. (2004) A theory for modeling ground-water flow in heterogeneous media. *US Geol. Survey Professional Paper 1679.*

Doherty, J. (2003) Ground water model calibration using pilot points and regularization. *Ground Water* **42**(2), 170–177.

Doherty, J. (2004) *Manual for PEST*, 5th edition. Brisbane, Australia: Watermark Numerical Computing. Downloadable from www.sspa.com/pest.

Harbaugh, A. W., Banta, E. R., Hill, M. C. & McDonald, M. G. (2000) MODFLOW-2000, the US Geological Survey modular ground-water model—User guide to modularization concepts and the Ground-Water Flow Process. *US Geol. Survey Open-File Report 00-92.*

Vecchia, A. V. & Cooley, R. L. (1987) Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model. *Water Resour. Res.* **23**(7), 1237–1250.