Experience gained in testing a theory for modelling groundwater flow in heterogeneous media

STEEN CHRISTENSEN
Department of Earth Sciences, University of Aarhus, Ny Munkegade b. 320, DK-8000 Aarhus, Denmark
stein@geo.au.dk

RICHARD L. COOLEY
US Geological Survey, Box 25046, Mail Stop 413, Denver Federal Center, Denver, Colorado 80225, USA

Abstract Usually, small-scale model error is present in groundwater modelling because the model only represents average system characteristics having the same form as the drift, and small-scale variability is neglected. These errors cause the true errors of a regression model to be correlated. Theory and an example show that the errors also contribute to bias in the estimates of model parameters. This bias originates from model nonlinearity. In spite of this bias, predictions of hydraulic head are nearly unbiased if the model intrinsic nonlinearity is small. Individual confidence and prediction intervals are accurate if the t-statistic is multiplied by a correction factor. The correction factor can be computed from the true error second moment matrix, which can be determined when the stochastic properties of the system characteristics are known.

Key words bias; confidence interval; correction; model error; prediction interval; regression; small-scale variability; true error

INTRODUCTION

A groundwater flow system can be characterized by hydrogeological variables such as hydraulic head, hydraulic conductivity, recharge, discharge, hydraulic heads and fluxes along internal and external boundaries, and point sources and sinks. In a model these system characteristics can be conceptualized as being discretely variable (in space and/or time) because discrete variation can be at as small a scale as desired (depending only on model discretization), allowing it to be virtually the same as continuous variation. All of these system characteristics can be assembled into an m-vector \( p \). Because \( p \) includes all scales of variation, any model function of \( p \) is almost free of model error, assuming that the model accurately represents the physical processes.

Seen from a practical modelling perspective, vector \( p \) represents small-scale variability that cannot be explicitly represented in a model and larger-scale variability (the drift) that can. In addition, vector \( p \) is unknown. Because \( p \) has much too large a dimension to be estimated from a limited number of uncertain observations, the vector of model parameters to be estimated is reduced from one of large dimension, \( p \), to one of much smaller dimension, \( \theta^* \). The \( p \)-vector \( \theta^* \) represents the spatial and/or temporal average of \( p \) and has the same form as the drift; it does not represent the small-scale variation. A model function \( f(\beta) \), which contains small-scale variation from \( p \), would be represented in terms of \( \theta^* \) as \( f(\gamma\theta^*) \), where \( \gamma \) is an \( m \times p \) interpolation or averaging
matrix, so that \( f(\mathbf{\beta}) = f(\mathbf{\gamma}\mathbf{\theta}_*) + e \). The error \( e \) introduced into the model’s representation of the hydrogeological variables by using \( \mathbf{\theta}_* \) is termed small-scale model error. Small-scale model error leads to a vector of model functions \( f(\mathbf{\gamma}\mathbf{\theta}_*) \) that are in error compared to the true function values \( f(\mathbf{\beta}) = f(\mathbf{\gamma}\mathbf{\theta}_*) + e \). Furthermore, because the errors, \( e \), are spatially correlated, it causes a violation of ordinary regression assumptions if \( \mathbf{\theta}_* \) is estimated using nonlinear regression.

Cooley (2002) developed a theory for modelling groundwater flow in heterogeneous media. The essence of the theory is that a model that only represents averages of the hydrogeological variables produces useful predictions, that the parameter values, \( \mathbf{\theta}_* \), can be estimated by nonlinear regression, and that the uncertainty caused by small-scale model error can be quantified by using modifications of regression based methods. The theoretical results of Cooley (2002) have been extensively tested and here we briefly summarize some of them and present results obtained for a test case.

**SUMMARY OF THEORY**

Following Kitanidis (1995), the system characteristics, \( \mathbf{\beta} \), are assumed to vary randomly as:

\[
\mathbf{\beta} \sim N(\mathbf{\gamma}\overline{\mathbf{\theta}}, \sigma_\beta^2 \mathbf{\Sigma}_\beta)
\]

where \( \overline{\mathbf{\theta}} \) is a \( p \)-vector of drift parameters, \( \sigma_\beta \) is a scalar, and \( \sigma_\beta^2 \mathbf{\Sigma}_\beta \) is the covariance of \( \mathbf{\beta} \). In groundwater modelling we are interested in estimating \( \mathbf{\theta}_* \) that characterizes the actual realization of \( \mathbf{\beta} \), and not the expected value, \( \overline{\mathbf{\theta}} \), of \( \mathbf{\beta} \). When \( \mathbf{\beta} \) consists only of spatially varying characteristics, \( \mathbf{\theta}_* \) is defined to be the spatial average of \( \mathbf{\beta} \) that has the same form as the drift. An estimate of \( \mathbf{\theta}_* \), \( \hat{\mathbf{\theta}} \), can be found by minimizing:

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

where \( Y \) is an \( n \)-vector of observations \( (m>>n>p) \), \( f(\mathbf{\gamma}\mathbf{\theta}) \) is the corresponding \( n \)-vector of values simulated using \( \mathbf{\gamma}\mathbf{\theta} \) instead of \( \mathbf{\beta} \), \( \omega \) is an arbitrary \( n \times n \) weight matrix, and \( I \) indicates transpose. For the observations we have \( Y = f(\mathbf{\beta}) + e \), where \( e \sim N(0, \mathbf{V}_e\sigma_e^2) \) are observation errors with the covariance \( \mathbf{V}_e\sigma_e^2 \). Cooley (2002, eq. 4-35) showed that \( \hat{\mathbf{\mathbf{\theta}}} \) obtained this way is a biased estimate of \( \mathbf{\theta}_* \) when \( f(\mathbf{\gamma}\mathbf{\theta}) \) is a nonlinear function of \( \mathbf{\theta} \), as is usually the case for groundwater models. Model errors contribute to the bias.

Predictions can be made with the groundwater model. In practice variables of interest \( g(\mathbf{\beta}) \) and \( g(\mathbf{\gamma}\mathbf{\theta}_*) \) are predicted by \( g(\hat{\mathbf{\theta}}) \). Cooley (2002, eq. 4-46) showed that \( g(\mathbf{\gamma}\hat{\mathbf{\theta}}) \) is an unbiased prediction of \( g(\mathbf{\gamma}\mathbf{\theta}_*) \) if the combined intrinsic nonlinearity is small. This is the case if there exist unique transformations of \( \mathbf{\theta} \) and \( \mathbf{\beta} \) that nearly linearize \( f(\mathbf{\gamma}\mathbf{\theta}) \) and \( g(\mathbf{\gamma}\mathbf{\theta}) \), and \( f(\mathbf{\beta}) \) and \( g(\mathbf{\beta}) \), simultaneously. Cooley (2002) describes different methods that can be used to determine the magnitude of the combined intrinsic nonlinearity.

Cooley (2002) showed that the limits of an individual confidence interval for \( g(\mathbf{\gamma}\mathbf{\theta}_*) \) can be calculated as the extreme values:
\[
\left( \min_0 (g(\gamma \theta)), \max_0 (g(\gamma \theta)) \right)
\]
subject to:
\[
S(\theta) \leq S(\hat{\theta}) \left( 1 + c_c \frac{t_{\alpha/2}(n - p)}{n - p} \right)
\]
where \( t_{\alpha/2}(n - p) \) is the \((1 - \alpha/2) \times 100\) percentile of the \(t\) distribution with \(n - p\) degrees of freedom, and \(c_c\) is a correction factor given by Cooley (2002, eq. 5-50). A procedure for solving (3) for nonlinear problems is provided by Vecchia & Cooley (1987). If the intrinsic nonlinearity is small, then \(c_c\) is a function of \(\sigma^2_{\theta} = \frac{\sqrt{\text{E}(\lambda(Y - f(y^*_0)) (Y - f(y^*_0)))^T \sigma^2_v}}{\sigma^2_v}\) is proportional to the second moment matrix of the true error vector \(Y - f(\gamma \theta_0) = \epsilon + f(\beta) - f(\gamma \theta_0)\). The true error is a sum of observation error and true model error. If \(\epsilon = \Omega^{-1}\), then \(c_c = 1.0\); otherwise, testing has shown that \(c_c\) can be large if small-scale model error is significant.

Cooley (2002) shows that the calculation of individual prediction interval limits for \(g(\beta)\), or for \(g(\beta) + \epsilon_p\) where \(\epsilon_p\) is an observation error, is similar to the computation of confidence interval limits, equation (3). However, for prediction intervals the correction factor, \(c_p\), is less dependent upon \(\epsilon^2_{\theta}\) than \(c_c\) (Cooley, 2002, eq. 5-96). Testing has shown that \(c_p\) is often close to unity.

If the stochastic properties of \(\beta\) are known, then \(\Omega\) can be either computed accurately from Cooley (2002, eq. 3-32), which would not be straightforward, or approximated by Cooley (2002, eq. 3-33), or computed from Monte Carlo simulations. Knowing, or having an estimate of, \(\Omega\), then \(c_c\), or \(c_p\), can be computed. If the stochastic properties are not known, and therefore \(\Omega\) is unknown, then the correction factors can be bounded. They can also perhaps be refined using the cross-validation method developed by Christensen & Cooley (1999).

**A TEST EXAMPLE**

One of our test examples is somewhat similar to that of Guadagnini & Neuman (1999). The dimensions of the two-dimensional flow domain are 18 by 8, divided into 90 by 40 uniform structural elements (Fig. 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 3.1076 over the left-most column of cells).

The vector \(\beta\) of unknown system characteristics consists of spatially varying log-transmissivity, which is a stationary random field with \(\beta \sim N(\theta \beta, \sigma^2_{\beta} \mathbf{V}_\beta)\), where \(\theta = 0\) is the expected value of each log-transmissivity, \(\mathbf{I}\) is a vector with all elements equal to one, and the covariance, \(\sigma^2_{\beta} \mathbf{V}_\beta\), is exponential with a correlation scale of 1.0 and \(\sigma^2_{\beta} = 4.0\). In this study 1000 independent realizations of \(\beta\), which has dimension \(m = 3600\), were generated using a simulator described by Kitanidis (1997, p. 237). For each realization the \(n\)-vector of observations, \(Y\), consists of \(n = 36\) observations of
Experience gained in testing a theory for modelling groundwater flow in heterogeneous media

Figure 1: Model domain, boundary conditions, and locations of head observations and head predictions.

No flux + Head observation ○ Predicted head • Pumping well
Constant flux

The hydraulic head observed at almost uniformly distributed grid points (Fig. 1). The observation errors are normally distributed, \( \varepsilon \sim N(0, \sigma^2 I) \), with a variance of \( \sigma^2 = 1.0 \).

The true error second moment matrix, \( \Omega \), was computed from the 1000 realizations of \( Y - f(\theta, 1) \), where \( \theta^* \) is the average of all elements in \( \beta \) for the realization.

For each realization the estimate of \( \theta^* \), \( \hat{\theta} \), was computed by minimizing equation (2). This was done using two different weight matrices, \( \omega = \Omega^{-1} \) (Gauss-Markov estimation) and \( \omega = \hat{\omega} \), where \( \hat{\omega} \) is a diagonal matrix with the diagonal elements being equal to the diagonal elements of \( \Omega^{-1} \). For the realizations we found estimates \( \hat{\theta} \) of the expected values to be \( \hat{E}[\theta] = -0.013 \), \( \hat{E}[^{\theta} ] = -0.075 \) for \( \omega = \Omega^{-1} \), and \( \hat{E}[\hat{\theta}] = -0.054 \) for \( \omega = \hat{\omega} \). Thus \( \hat{\theta} \) found using either of the weight matrices is a biased estimate of \( \theta^* \).

For each realization, and for both weight matrices, we computed a prediction of the hydraulic head at the six locations in Fig. 1, \( g_i(\hat{\theta}1) \) for \( i = 1, 6 \). Table 1 summarizes the estimated expected values of the predictions, \( g_i(\hat{\theta}1) \), the true values, \( g_i(\beta) \), the values computed with the spatial average log-transmissivity, \( g_i(\theta, 1) \), and the values computed with the expected value of log-transmissivity, \( g_i(\hat{\theta}1) \). For both weight matrices note that \( g_i(\hat{\theta}1) \) is a biased prediction of the true head \( g_i(\beta) \) near the pumping well, whereas the predictions away from the well, \( g_i(\hat{\theta}1) \) for \( i = 2 \ldots 6 \), are nearly unbiased. For most predictions \( g_i(\hat{\theta}1) \) is closer to \( g_i(\beta) \) than to \( g_i(\theta, 1) \), and is much closer to \( g_i(\beta) \) than to \( g_i(\hat{\theta}1) \). The model produced an unbiased fit to \( Y \).

For each realization we computed the 95% confidence interval for \( \theta^* \). The interval was computed using \( \omega = \Omega^{-1} \) and \( c_c = 1.0 \), \( \omega = \hat{\omega} \) and \( c_c = 1.0 \), as well as \( \omega = \hat{\omega} \) and \( c_c = 31.32 \), which is the theoretically computed correction factor (Cooley, 2002, eq. 5-50) for this specific case. Table 2 summarizes the results. Note that for \( \omega = \Omega^{-1} \) the confidence interval is accurate in the sense that \( \theta^* \) is inside the confidence interval for 949 realizations, a frequency of almost 95%. For \( \omega = \hat{\omega} \) the uncorrected confidence interval \( (c_c = 1.0) \) is very inaccurate, whereas the corrected interval \( (c_c = 31.32) \)
Table 1: Estimated expected values of hydraulic head at six locations computed from 1000 realizations of $\beta$ and $\gamma$.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$g_i(\hat{\theta}_1)$ $\omega = \Omega^{-1}$</th>
<th>$g_i(\hat{\theta})$ $\omega = \hat{\omega}$</th>
<th>$g_i(\hat{\theta}.1)$</th>
<th>$g_i(\hat{\theta_1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.9639483</td>
<td>4.9069397</td>
<td>4.5117315</td>
<td>4.5341798</td>
</tr>
<tr>
<td>3</td>
<td>7.2055296</td>
<td>7.1227775</td>
<td>7.0681220</td>
<td>6.5816896</td>
</tr>
<tr>
<td>4</td>
<td>3.7505155</td>
<td>3.7074426</td>
<td>3.6554755</td>
<td>3.4258035</td>
</tr>
<tr>
<td>5</td>
<td>2.2393110</td>
<td>2.2135935</td>
<td>2.1090300</td>
<td>2.0454360</td>
</tr>
<tr>
<td>6</td>
<td>5.4417089</td>
<td>5.3792135</td>
<td>5.3888899</td>
<td>4.9705769</td>
</tr>
</tbody>
</table>

Table 2: Number of realizations where $\theta_i$ was below, within, and above the confidence intervals computed using different weight matrices, $\omega$, and correction factors, $c_c$.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$c_c$</th>
<th>Number of realizations:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>below</td>
<td>within</td>
<td>above</td>
<td></td>
</tr>
<tr>
<td>$\Omega^{-1}$</td>
<td>1.0</td>
<td>6</td>
<td>949</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>$\hat{\omega}$</td>
<td>1.0</td>
<td>284</td>
<td>352</td>
<td>364</td>
<td></td>
</tr>
<tr>
<td>$\hat{\omega}$</td>
<td>31.32</td>
<td>9</td>
<td>979</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Number of realizations where $g_i(\hat{\beta}) + \varepsilon_{pi}$ at each of six locations was below, within, and above the uncorrected ($c_p = 1.0$) and the corrected prediction intervals computed using $\omega = \hat{\omega}$.

<table>
<thead>
<tr>
<th>$i$</th>
<th>Uncorrected interval ($c_p = 1.0$):</th>
<th>Corrected interval:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>below</td>
<td>within</td>
<td>above</td>
<td>$c_p$</td>
<td>below</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>981</td>
<td>1</td>
<td>1.229</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>988</td>
<td>8</td>
<td>0.5545</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>984</td>
<td>7</td>
<td>0.5601</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>969</td>
<td>16</td>
<td>1.047</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>932</td>
<td>29</td>
<td>1.392</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>970</td>
<td>14</td>
<td>0.7683</td>
<td>32</td>
</tr>
</tbody>
</table>

appears to be a little conservative (meaning that the interval is a little too wide and therefore has an actual probability of containing $\theta_i$ being larger than the nominal probability of 95%).

Similar testing was done for predictions of $g_i(\hat{\beta}) + \varepsilon_{pi}$, where $\varepsilon_{pi}$ is a random error with a variance of 1.0, at the six locations in Fig. 1. Uncorrected and corrected prediction intervals were computed with $\omega = \hat{\omega}$. The results summarized in Table 3 show that the correction factors, $c_p$, are fairly close to 1.0; some are smaller, and others are larger than 1.0. The corrected prediction intervals appear to be a little more accurate than the uncorrected prediction intervals except at location 1 near the well. At this location both intervals are conservative.

Similar results have been obtained using other boundary conditions, other observation sets $\gamma$, and other values of $\sigma_p^2$ and $\sigma_c^2$. 
CONCLUSIONS

The theory and the test have shown that small-scale model error can result in significant correlations among the true errors of the regression model. The model error also contributes to bias in the estimate, $\hat{\theta}$, of the model parameter vector $\theta^*$. The bias originates from model nonlinearity.

The test case predictions of hydraulic head, $g_*(\hat{\theta})$, were essentially unbiased except near a strong sink. The theory shows that a prediction will only be biased if the combined intrinsic nonlinearity is significant. In the test case this seems to be the case only for the hydraulic head near the sink.

Individual nonlinear confidence and prediction intervals were found to be accurate when the weight matrix used in the regression equals the inverse of the true error second moment matrix.

If a weight matrix other than the inverse of the true error second moment matrix is used, the $t$-statistic used to compute the individual confidence or prediction interval should be multiplied by a correction factor. We used correction factors given by the theory and found them to produce accurate or slightly conservative (slightly too wide) intervals. The correction factor found for the confidence interval was very large, whereas the correction factors found for prediction intervals were relatively close to one.

The correction factors given by the theory depend on the true error second moment matrix, which can only be determined if the first two statistical moments of $\beta$ are known or can be estimated. Approximate bounds for the correction factors must be used when these moments are unavailable.

The synthetic experimental results thus are in agreement with the theoretically developed results of Cooley (2002).

REFERENCES


