Statistical and Stochastic Approaches to Assess Reasonable Calibrated Parameters in a Complex Multi-Aquifer System

PHATCHARASAK ARLAI¹, MANFRED KOCH¹ AND SUCHARIT KUNTANAKULWONG²

¹Department of Geohydraulics and Engineering Hydrology, University of Kassel, Kurt-Wolters-Str.3, D34109 Kassel, Germany
²Department of Water Resources Engineering, Chulalongkorn University, Bangkok, Thailand

ABSTRACT

Effective assessment of reasonable calibrated aquifer parameters can be mainly divided into three categories: firstly, the conventional trial-and-error forward analysis, secondly, the calculation of sensitivity- and correlation coefficients through nonlinear inverse regression and, thirdly, the pure stochastical approach applying well-known stochastical formulae. We apply these three approaches to the Bangkok multi-aquifer system that consists of eight complex water bearing layers underneath the Chao Phraya river basin. First, using the MODFLOW code in the conventional forward trial-and-error manner, the hydraulic parameters, namely, the transmissivity for this aquifer system are estimated by visually comparing observed and calculated heads. Next, the estimated parameters are assessed by combining MODFLOW with the automated nonlinear regression program, UCODE. Sensitivity and correlation coefficients for each of the parameters are calculated zone-wise. The results indicate that the estimated parameters are well-determined and unique in each of the sub-zones. In a third step, the full stochastic approach is used. Applying a random field generator, realizations of a logarithmic transmissivity field $Y=\ln T$ with various sets of variances $\sigma^2_Y$ and correlation lengths $(\lambda_x, \lambda_y)$ for each layer that characterize the possible stochastic range of the $\ln T$–field in the multi-aquifer system are simulated. Using these Monte Carlo MODFLOW simulations we investigate how $\sigma^2_Y$ contributes to the $\sigma^2_H$ of the observed head and/or residuals. Stochastic theory predicts that $\sigma^2_H$ and $\sigma^2_Y$ are related to each other as $\sigma^2_H = \sigma^2_Y \times \frac{\lambda^2}{\lambda_x^2 + \lambda_y^2}$. Finally, we investigate which factors affect the residual error of the model estimation. Obviously, both transmissivity variations and errors in the head measurements are mostly responsible for a non-zero estimated residual head. Hence, the variances of head that are obtained from stochastically generated transmissivities and the intrinsic errors of the head measurements were determined. The results show that the stochastically predicted variances of the head are still somewhat lower than the variances of the residual head, indicating additional uncertainties in the fitted model. Indeed, the pumping rates turn out to be very evasive. To investigate the effects of the latter on the residual head variance, Monte Carlo simulations with randomly disturbed pumping rates of varying magnitudes are performed. The results show that pumping plays a smaller but still significant role for the estimation of the residual error, as the residual head variances obtained from stochastic pumping are lower than those of the stochastic transmissivity field.

1. INTRODUCTION

Prediction reliability is a fundamental problem in groundwater modeling studies (Anderson and Woessner, 1992; Poeter and Hill, 1997; Yobbi, 2000). Usually a trial-
and-error approach is employed to estimate the relevant geohydraulic parameters of the aquifer system under question during the calibration process, but this may turn out to deliver non-unique solutions, i.e. different combinations of parameters produce identical head distributions, forsaking a successful prediction of the behaviour of the groundwater system under future possible stresses. Another problem of this common trial-and-error method is that the solution may only deliver a local instead of a global minimum of the, generally, multidimensional head distribution. As neither the observed heads nor the geohydraulic parameters of the aquifer system are known precisely, this might not be too much of a drawback in practical studies. Nevertheless, better results and a more thorough investigation of the response surface can be achieved through the additional use of statistical (Poeter and Hill, 1997; Yobbi, 2000) and, namely, methods based on stochastical theory (Gelhar, 1993), with the latter also being able to quantify the effects of errors in both the objective and subjective data.

Here we present an application of these approaches to the complex Bangkok multi-aquifer system, Thailand, which is the major groundwater resource for about a third of the population of that country. In addition to the conventional trial-and-error modeling approach with the well-known MODFLOW model, as embedded in the PMWIN-WINDOWS-surface (Chiang and Kinzelbach, 2001), the calibration for the relevant geohydraulic parameters, namely, transmissivity and leakance fields, but also uncertain pumping rates, was done by using the automated non-linear regression program, UCODE (Poeter and Hill, 1998) in conjunction with MODFLOW. Advantages of non-linear regression inverse modeling are that, firstly, the data shortcomings and further needs can be evaluated, secondly, better confidence in estimates and predictions are obtained and, thirdly, parameter correlations and low sensitivities which are indicators of non-unique solutions can be computed, providing a better understanding of the overall system behaviour.

However, neither the trial-and error, nor the inverse method, both of which are still deterministic techniques, work satisfactorily when the “objective” head and/or pumping data and/or the “subjective” aquifer calibration parameters (hydraulic conductivity, transmissivity) are prone to a high degree of uncertainty, as is the case in the present application. For this reason stochastic modeling using Monte-Carlo simulations of realizations of the $Y=\ln T$ transmissivity field and, subsequently, of the pumping rates are performed. A formulae based on Gelhar’s stochastic theory (Gelhar, 1993) that predicts how the uncertainty (variance) $\sigma_Y^2$ projects into the variance $\sigma_H^2$ of the head and/or the residual head is validated through the MC simulations. Using this stochastic approach we are able to provide the best calibration of the aquifer system under the present constraints of erroneous head data, local transmissivities and pumping rates.

2. STUDY AREA AND MODEL IMPLEMENTATION

The Bangkok multi-aquifer system is located underneath the lower Chao Praya river basin which is bordered in the east, north and west by ridges of hills and mountains and in the south by the Gulf of Thailand (Figure1). Hydrogeologically, the aquifer system can be divided into a topmost soft and stiff clay layer and eight lower principle confined aquifers (Kokusai Kogyo, 1995). The groundwater flow model for the Bangkok multilayered aquifers is implemented by the quasi 3D finite-difference model MODFLOW (McDonald and Harbaugh, 1988), with 9 aquifer layers whereby the topmost clay layer is treated as an unconfined aquifer and the 8 lower ones as confined. The model is divided into 55 rows and 52 columns with grid sizes varying from 2*2 km$^2$ to 16*16 km$^2$, following the approach of Kokusai Kogyo (1995) (Figure1).
FIGURE 1: 3D hydrogeological map and FD grid in the fifth layer of the model.

The top boundary of the model is specified as constant head, representing the water table. The main recharges into the aquifer system are at the outcropping basin flanks and are simulated also as constant head that is set equal to the terrain altitude. Because the topmost clay layer has a thickness that varies from 15 to 30 meters, recharge inside the basin is basically zero. The bottom of the 9th layer is assigned as a no-flow boundary. Cells in the southern 55th row of the model that are connected to the Gulf of Thailand are treated as constant head at sea level. The initial transmissivities and vertical leakances have been taken from a former study of Kokusai Kogyo (1995) and new geological profiles (Kasetsart University, 2003). Head and pumping data have been provided by the Water Resources System Research Unit, Chulalongkorn University. Using the three different approaches mentioned, the piezometric data from about 179 monitoring wells in 1999 are used to calibrate mainly, among others, zonal transmissivities and leakances under steady state flow conditions and form a basis of an ongoing study of saltwater intrusion and management.

3. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE BANGKOK AQUIFER MODEL

3.1 Conventional (forward) trial-and-error approach

A qualitative evaluation of the calibration success is provided by visual inspection of similar patterns between computed and observed heads in the 3rd, 4th and 5th model layers which are the prime productive groundwater layers and where most of the observation wells are located (Figure 2). A quantitative assessment is carried out by the analysis of the scatter plot of measured against computed heads (Figure 3a) and by a measure of the residual error quantified by (1) the mean error (ME), (2) the mean absolute error (MAE) and, (3) the root mean squared error (RMS) (Figure 3b). The scatter plot reveals a well-posed calibration, since all points are close to the diagonal line, with the coefficient of determination $R^2$ being close to one. Moreover, the diagrams of the sensitivities for ME, MAE and RMS (Figure 3b) disclose that this set of calibrated hydraulic parameters is optimal in the sense that it provides the best result among the manifold of slightly perturbed parameter values.

One point of particular importance to the modeler should be to examine how large a model error, i.e. residual is acceptable for a calibrated model to be considered satisfactory. Usually, the maximally acceptable error depends on the magnitude of head...
FIGURE 2: Observed versus steady-state computed heads for 1999 in (a) layer 3, (b) layer 4 and (c) layer 5.

FIGURE 3: (a) Observed versus computed heads for 1999, with upper and lower 95% confidence limits; (b) ME, MAE, RMS obtained when transmissivities T and vertical leakances Vk are varied percentally from their optimal calibrated values.

change over the model domain and is formulated as the ratio of the RMS to the former (Anderson and Woessner, 1992) In our case these ratios are 4%, 3% and 4% in layers 3, 4 and 5, respectively, which is quite satisfactory given the uncertainties in the data as will be discussed further down.

3.2 Statistical regression (inverse) approach

Combining the forward MODFLOW model with the nonlinear regression (inverse) model UCODE (Poeter and Hill, 1998) offers a more versatile exploration of the model calibration space than is possible by the trial-and-error method above. In addition to allow for area zoning of hydraulic parameters, UCODE computes sensitivity and correlation coefficients for each of these, resulting in an estimate of the uncertainties in the calibration and providing a scrutiny of the overall piezometric response surface.
In the present application, the model domain is divided into 5, 6 and 5 sub-zones for the transmissivity and 5, 5 and 4 sub-zones for the vertical leakance in layers 3, 4 and 5, respectively. Exemplarily, Figure 4 shows the zoning used for layer 3 which is the major production aquifer. The *composite scaled sensitivity* ($cs$) computed by UCODE expresses how well the parameters are designated by the observations and reflects how well the parameters can be calibrated. In order to interpret this sensitivity easier, the *relative composite scaled sensitivity* ($rcs$) obtained by normalizing all $cs$-values by the largest one will be represented. A parameter with larger $rcs$ is likely to have a smaller uncertainty, a broader confidence interval and, thus, be more informative for the calibration. Although the lower cut-off value for $rcs$ to characterize a parameter value as totally uncertain is rather arbitrary, Yobi (2000) sets the former to 0.02.

**FIGURE 4:** Set of sub-zones of transmissivity in layer 3, Phra Padeang Aquifer

**FIGURE 6:** Relative composite scale sensitivity analysis for (a) transmissivity and (b) vertical leakance in the various sub-zones.

**FIGURE 7:** Correlation coefficient matrix of (a) transmissivity and (b) vertical leakance in layer 3, 4 and 5.
The \( \text{rcs} \) for the 16 sub-zones of the transmissivity and the 14 sub-zones of the vertical leakance are analyzed and displayed in Figure 6. Figure 7 shows, additionally, the matrix of the correlation coefficients (\( cc \)) between the various subzonal parameters providing information on which of the subzones can be determined independently of the other. It can be seen that the \( cc \) are unity along the diagonal lines and that they have only small values in the off-diagonal elements of the matrix, particularly for the transmissivity, whereas for the leakance the zoning is slightly less optimal. According to the statistical criteria assumed previously we would consider our zoning choice for the vertical leakance and, more so, for the transmissivity as satisfactory.

### 3.3 Stochastic modeling using MC-simulations and validation of stochastic theory

Although the two deterministic modeling approaches used previously, namely, the trial-and error and the nonlinear inverse methods have resulted in satisfactory calibrations, none of them has been able to perfectly fit the observed “objective” piezometric head data, leaving a nonzero residual as quantified by the RMS (see Figure 3b). There are two reasons for this: (1) the “objective” head and/or the pumping data are not exactly measured and/or (2) the particular, “deterministic” calibration parameters obtained represent only a local instead of a global minimum of the piezometric response surface. On the other hand, one has to assume that the Bangkok aquifer system is more heterogeneous than is pictured by the zonal calibrated transmissivity \( T \) and leakance fields \( V_k \) obtained so far. Moreover, given that local estimates of \( T \) from pump tests and of \( V_k \) from geological borehole profiles are available that point to a rather heterogeneous subsurface structure, one would like to condition the model calibration on these “subjective” model observations.

For this purpose we are presenting in the following a stochastic modeling approach using Monte-Carlo simulations of realizations of the \( Y = \ln T \) transmissivity field and partly conditioned on the observed “subjective” \( T \)-data. We will look into the question as to how the uncertainty (variance) \( \sigma_Y^2 \) projects into the variance \( \sigma_H^2 \) of the heads and/or the residual heads. Indeed, using stochastic theory, Gelhar (1993) developed the following relationship between these two variances (assuming steady state saturated groundwater flow):

\[
\sigma_H^2 = C \sigma_Y^2 \lambda^2 J^2
\]

where \( C \) = a coefficient that depends on the dimensionality of the flow (\( C = 0.46 \) for 2D); \( \sigma_Y^2 \) = the variance of \( \ln K \) or \( \ln T \); \( \lambda \) = an integral or correlation scale; and \( J \) = average hydraulic gradient. Eq. (1) shows that the residual head variance \( \sigma_H^2 \) is directly proportional to the variance \( \sigma_Y^2 \) and the square of the correlation scale \( \lambda^2 \) of the transmissivity field. Thus, the head variance \( \sigma_H^2 \) provides a lower bound for a calibration target which means that the model should be calibrated such that the simulated \( \sigma_H^2 \) (sim) attains approximately the theoretical value \( \sigma_H^2 \) (Gelh.) of Eq. (1).

Firstly, we validate Eq.(1) by comparing it with MC simulations of random realizations of a logarithmic transmissivity field \( Y = \ln T \). Using a classical random field generator (Chiang and Kinzelbach, 2001), 180 realizations of \( Y = \ln T \) with a set of variances \( \sigma_Y^2 \) for each layer, namely, \( \sigma_Y^2 = 0.55, 0.77, 0.59 \), in the layers 3, 4 and 5, respectively, and two sets of correlation lengths (\( \lambda_x, \lambda_y \)) (with \( x \) and \( y \) corresponding to the EW and NS-direction, respectively) that represent 63% and 95% of the sill of the observed variograms, namely, \( \lambda_x = 9000, 6000, 5500 \) and \( \lambda_y = 12500, 23000, 7500 \) for the 63%-sill and \( \lambda_x = 26000, 21500, 17500 \) and \( \lambda_y = 33000, 56000, 22500 \) for the
FIGURE 8: Simulated variograms of head in layer 3 (a), 4 (b) and 5 (c) with variances $\sigma^2_Y$ and correlation lengths $\lambda_Y$ as specified in Table 1.

TABLE 1: Comparison of $\sigma^2_H$ obtained from Gelhar’s Eq. (1) and MC-simulations.

<table>
<thead>
<tr>
<th>Layer</th>
<th>C</th>
<th>$\sigma^2_Y$</th>
<th>$\lambda_Y$ (m)</th>
<th>$J$</th>
<th>$\sigma^2_H$ (Gel.) (m$^2$)</th>
<th>$\sigma^2_H$ (sim)(m$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.46</td>
<td>0.55</td>
<td>33000</td>
<td>3.24E-04</td>
<td>28.9</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>0.46</td>
<td>0.77</td>
<td>56000</td>
<td>2.37E-04</td>
<td>62.3</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>0.46</td>
<td>0.59</td>
<td>22500</td>
<td>5.00E-04</td>
<td>34.1</td>
<td>35</td>
</tr>
</tbody>
</table>

95%-sill—which appears to characterize the possible stochastic range of the ln T-field in the multi-aquifer system—are performed. Using these MC-Carlo simulations we investigate how the input $\sigma^2_Y$ projects through the groundwater system onto the head variance $\sigma^2_H$. It turns out that the $\sigma^2_H$ obtained for the 95%-sill correlation length conforms well with that of the stochastic formula, while that for the 65%-sill length results in biased (too low) values. The variograms of the MC-simulated heads are shown in Figure 8. The results obtained theoretically and from the MC-simulations are summarized in Table 1 and they show rather unanimously that the $\sigma^2_H$ of Gelhar conform well with the $\sigma^2_H$ simulated.

For further clarification and in order to alleviate the problem of the poorly known flow factor $J^2$ in Eq. (1), results of MC-calculations with $\sigma^2_Y$ twice as big as those in Table 1 are presented in Figure 9 and Table 2. As expected, the simulated $\sigma^2_H$ are about two times higher than in the first case. This clearly provides further evidence of the applicability of the analytical stochastic theory of Gelhar (1993) in the present case.

As mentioned earlier, the lower bound of an acceptable error of calibration should be equal to $\sigma^2_H$ calculated with Gelhar’s formula. The accepted error targets for $\sigma^2_H$ computed stochastically are 5.36, 7.89 and 5.84 m in layers 3, 4, and 5, respectively.

FIGURE 9: Similar to Figure 8, but with $\sigma^2_Y$ in layer 3 (a), 4 (b) and 5 (c) twice as big and as specified in Table 2.
TABLE 2: Similar to Table 1, but with $\sigma_Y^2$ twice as big.

<table>
<thead>
<tr>
<th>layer</th>
<th>C</th>
<th>$\sigma_Y^2$</th>
<th>$\lambda_y$ (m)</th>
<th>$J$</th>
<th>$\sigma_H^2$ (Gel)(m$^2$)</th>
<th>$\sigma_H^2$ (sim)(m$^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.46</td>
<td>1.10</td>
<td>33000</td>
<td>3.24E-04</td>
<td>57.9</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>0.46</td>
<td>1.54</td>
<td>56000</td>
<td>2.37E-04</td>
<td>124.6</td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>0.46</td>
<td>1.17</td>
<td>22500</td>
<td>5.00E-04</td>
<td>68.1</td>
<td>70</td>
</tr>
</tbody>
</table>

The set of calibrated parameters obtained using the trial and error approach, resulted in MAE values 1.97, 2.14 and 2.11m, respectively, i.e. are lower than the ones above obtained from the stochastic representation of the aquifer’s transmissivity. Basically this means that the deterministic trial and error approach resulted in a calibration which is not conditioned enough on the transmissivity-information available.

Finally, we investigate which factors affect the model residual errors the most. In addition to the stochastic transmissivity variations $\sigma_Y^2$ which result in head variations $\sigma_H^2$ as given by Eq. (1), there will be intrinsic errors in the head measurements (noise) $\sigma_n^2$ that should add up to the first ones. Hence, the $\sigma_H^2$ obtained from stochastically generated transmissivities and the intrinsic errors of the head measurements were determined. The average $\sigma_H$ of all monitoring stations are 3.73, 11.23, 7.17 m in the layers 3, 4 and 5, respectively which is higher than predicted by Eq. (1). A major reason for this might be that erroneous pumping rates are used. Indeed, pumping rates in the study area are not always correctly reported by well owners. To investigate the effects of varying pumping rates on the residual head variance, 180 Monte Carlo simulations with randomly disturbed pumping rates of varying magnitudes (30 - 80 % of the reference value) are performed. The average $\sigma_H$ of all monitoring stations result in 3.40, 7.53, 6.23 m in layers 3, 4 and 5, respectively. It is evident that although the residual head variances due to stochastic pumping are lower than those due the stochastic transmissivity field, pumping is still a significant factor contributing to the model error.

REFERENCES


