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

by

Phatcharasak Arlai\textsuperscript{1}, Manfred Koch\textsuperscript{1}, Sucharit Koontanakulvong\textsuperscript{2}

\textsuperscript{1}Department of Geohydraulics and Engineering Hydrology, Faculty of Civil Engineering, University of Kassel, Germany and

\textsuperscript{2}Department of Water Resources Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
PRESENTATION CONTENTS

1. Introduction
   1.1 Aspects of reliability in groundwater modeling
   1.2 Techniques to check for reliability

2. Study Area and Model Implementation

3. Theory

4. Effective Approaches to assess reliable Parameters in the Groundwater Model
   4.1 Conventional trial-and-error forward analysis,
   4.2 Calculation of sensitivity- and correlation coefficients through inverse regression
   4.3 Pure stochastical approach applying well-known stochastical formulae.

5. Summary
1. INTRODUCTION

1. Even more questions for the prediction reliability in groundwater modeling project??

2. What is the reliability in groundwater modeling??
   - Qualitative
   - Quantitative
   - Sensitivity
   - Uniqueness


3. The calibrated parameters from steady state are tested to whether they can give reliable calibrated parameters or not
1. INTRODUCTION

Non-uniqueness problem:

Different combinations of parameter values match the observations equally well.

Poeter and Hill, 1997
How to check for the reliability of calibrated parameters:

**Conventional Misfit**: Qualitative and Quantitative
- Contour plots between observed and measured piezometric head
- Mean Error
- Mean Absolute Error
- Root Mean Square Error

**Statistical Analysis**: Sensitivity and Uniqueness; Math Model + Inverse Model
- Relative Composite Scaled Sensitivity
- Correlation Coefficient

**Stochastic Analysis**:
- Monte Carlo simulation to validate stochastic formula
- Monte Carlo simulation to determine the uncertainty factors that contribute to model error
2. STUDY AREA AND MODEL IMPLEMENTATION

Study Area

3D – Geological Map
2. STUDY AREA AND MODEL IMPLEMENTATION

Profile of Bangkok Aquifers System

Conceptual Model
2. STUDY AREA AND MODEL IMPLEMENTATION

**Model Area:**

- $208 \times 256$ km$^2$
- $53248$ km$^2$

**Model Grid:**

- One layer = $52 \times 55$
  - $2860$
- 9 layers = $25740$

**FD Grid in layer 5**

**3D FD Grid**
3.THEORY

3.1 Sensitivity Analysis

Scaled Sensitivity

\[ ss_{ij} = \sum_{k=1}^{ND} \left( \frac{\partial y_k'}{\partial b_j} \right) b_j \omega_{ik}^{1/2} \]

Composite Scaled Sensitivity

\[ css_j = \left[ \sum_{j=1}^{ND} (ss_{ij})^2 / b_j \right]^{1/2} / ND \]

Relative Composite Scaled Sensitivity

\[ rcss_j = (css_j) / \text{Max}(css) \]

Criterion: Relative Composite Scaled Sensitivity > 0.02
3.2 Correlation Coefficient Matrices

Correlation Coefficient =

\[ \text{Correlation Coefficient} = \frac{\text{cov}(i,j)}{\sqrt{\text{var}(i)\text{var}(j)}} \] (Poeter and Hill, 1997; Hill, 1998)

Can the existing parameters result in a unique solution?

Criterion: Correlation coefficient matrices < 0.95
3.3. 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

\( \sigma_H^2 \) = estimated head variance  
\( C \) = a coefficient that depends on the dimensionality of the flow  
\( (C=0.46 \text{ for } 2D) \)  
\( \sigma_Y^2 \) = the variance of \( \ln K \) or \( \ln T \)  
\( \lambda \) = an integral or correlation scale  
\( J \) = average hydraulic gradient
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUNDWATER MODEL

4.0 Initial calibration and check of sensitivity of model parameters

Data from 179 monitoring wells are applied to calibrate Transmissivity and Vertical Leakance for steady state in 1999 by cooperating between trial-error-method and non-linear regression inverse model, as water levels in 1999 are more or less constant.

Sensitivity of homogeneous T and Vk in each layer
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUNDWATER MODEL

4.1 Conventional (Trial and Error) forward analysis of the misfit error
(a) Qualitative assessment

--- observed heads

--- modeled heads

Layer 3

Layer 4

Layer 5
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUND WATER MODEL

4.1 Conventional analysis of the misfit error

(b) Quantitative assessment

Kogkosai Kogyo, 1995

AIT, 1997

current model
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUNDWATER MODEL

4.1 Conventional analysis of the misfit error
(b) Quantitative assessment

The acceptable error target (AET) for ground-water modeling

\[
AET = \frac{\text{RMS}}{\text{Total Head Loss}} \quad (\text{Anderson and Woessner, 1992})
\]

AET at layer 3: 4 %
AET at layer 4: 3%
AET at layer 5: 4%
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUND WATER MODEL

4.2 Statistical (inverse) analysis:

Using the nonlinear regression model UCODE (Poeter and Hill, 1998) to compute sensitivity and correlation coefficients for each of the parameter zones, resulting in an estimate of the uncertainties/uniqueness in the calibration and providing a scrutiny of the overall piezometric response surface.

An example of parameter-subzone calibration

Transmissivity:
16 Sub-zones

Vertical Leakance:
14 Sub-zones
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUNDWATER MODEL

4.2 Statistical (inverse) analysis:

(a) Sensitivity Analysis, 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, Yobbi (2000) sets the former to 0.02.

Relative Composite Scaled Sensitivity of Transmissivity

Relative Composite Scaled Sensitivity of Vertical Leakance
4.2 Statistical (inverse) analysis:

(b) 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. This computation discloses the uniqueness. Hill (1998) specifies \(cc= 0.95\) to discriminate between parameters that are uniquely determined \((cc<0.95)\) or not \((cc>0.95)\).
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THEGROUND-WATER MODEL

4.3 Stochastic analysis: (a) validate pure stochastic formula (Gelhar, 1992); 1 time of $\sigma_Y^2$

$$\sigma_H^2 = C * \sigma_Y^2 * \lambda^2 * \mathcal{P}$$

<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>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>

A realization of heterogeneous transmissivity in each layer

Layer 3

Layer 4

Layer 5
4. EFFECTIVE APPROACHES TO ASSESS THE RELIABLE PARAMETERS IN THE GROUNDWATER MODEL

4.3 Stochastic analysis: (a) validate the pure stochastic formula; 1 time of $\sigma_Y^2$

$$\sigma_H^2 = C \times \sigma_Y^2 \times \lambda^2 \times \rho$$

Correlation length at 95% of sill fits better than at 63% of one to Gelhar’s formulation.

<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.) (m²)</th>
<th>$\sigma_H^2$ (sim)(m²)</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>
4.3 Stochastic analysis: (b) validate the pure stochastic formula; 2 times of \( \sigma_Y^2 \)

\[
2\sigma_H^2 = C \times 2\sigma_Y^2 \times \lambda^2 \cdot \varphi^2
\]

<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) (m²)</th>
<th>( \sigma_H^2 ) (sim) (m²)</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>

Validation shows that stochastic formula conforms well with the computed data from the Monte Carlo simulations using random realizations of the transmissivity field in the study area***
4.3 Stochastic Analysis: investigate which calibration parameters (transmissivity, pumping rates) are responsible for the systematic misfit error

- 90 realizations of lnT are generated and run with Monte Carlo (MC) simulation.s
- lower bound of an acceptable error of calibration should be equal to $\sigma_H^2$ calculated with Gelhar’s formula. The accepted error targets for $\sigma_H$ computed stochastically are 5.36, 7.89 and 5.84 m in layers 3, 4, and 5, respectively.
- set of calibrated parameters 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, indicating an “overfit-calibration” of the deterministic model.
- 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.
- 90 realizations of MC simulations with randomly disturbed pumping rates of varying magnitudes (30 - 80 % of the reference value) are performed.
- 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.
5. SUMMARY

5.1 Effective assessment of reasonable calibrated aquifer parameters can be mainly divided into three categories:
(a) conventional trial-and-error forward analysis,
(b) calculation of sensitivity- and correlation coefficients through inverse regression
(c) pure stochastical approach applying well-known stochastical formulae.

5.2 Estimated parameters are assessed by combining MODFLOW with the automated non-linear regression program, UCODE. Sensitivity and correlation coefficients for each of the parameters are calculated zone-wise. Results indicate that estimated parameters are well-determined and unique in each of the sub-zones.

5.3 Using Monte Carlo MODFLOW simulations investigation of how $\sigma_Y^2$ contributes to the $\sigma_H^2$ of the observed head and/or residuals: Stochastic theory predicts that $\sigma_H^2$ and $\sigma_Y^2$ are related to each other as $\sigma_H^2 \sim \sigma_Y^2 \times \lambda^2$ and $\lambda$ should be obtained from 95% of sill. MC results agree with stochastic formula and show that residual modelled heads are due to the stochastic transmissivity field.

5.4 Investigation of which hydraulic factors affect the residual error of the model estimation: MC simulations with randomly disturbed pumping rates are performed. 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.
5. SUMMARY (cont.)

5.5 Outlook and future work:

Investigation of **sustainable yield management** for preventing saltwater intrusion from the Gulf of Thailand and/or the surficial clay layers in the basin (*Arlai et al.*, 2006, *IAHR-meeting, Toulouse, France, June 12-14, 2006*)

Simulated salt plumes in (a) layer 3, (b) layer 4 and (c) layer 5 in 2032.