Experimental Designs for Model Discrimination and Uncertainty Reduction in Groundwater Modeling

Hai Van Pham
Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Part of the Civil and Environmental Engineering Commons

Recommended Citation
A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Civil and Environmental Engineering

by
Hai Van Pham
B.S., Water Resources University, 2006
M.S., Dongguk University, 2011
August 2015
Acknowledgments

I would like to express my sincere gratitude to my advisor, Dr. Frank Tsai, for his excellent guidance, support, and supervision throughout all the stages of this dissertation. I am thankful to my committee members Dr. Qin Chen, Dr. Carol Wicks, Dr. Karsten Thompson, and Dr. Krishna Paudel for their time and support. Special thanks go to my friends and colleagues in Dr. Tsai’s research group students Dr. Ahmed Elshall, Dr. Nima Chitsazan, Dr. Ehsan Beigi, Mr. Amir Mani, Mr. An Li, and Mr. Ramazan Karakullukcu for their support and help during the last four years.

I acknowledge Dr. Frank Tsai, Dr. Jeffrey Hanor, Dr. Le Yan, Dr. Ahmed Elshall, Dr. Ming Ye, and Dr. Nima Chitsazan for their contributions to and for being coauthors on the five peer-reviewed manuscripts and the 16 conference proceedings, which were carried out during the study period. I would like to thank to Dr. Lorelei Patrick for giving useful suggestions and for careful edits of my manuscripts; Mrs. Janet Labatut and Mrs. Julie Mueller for their administrative support.

I acknowledge the USGS, the Louisiana Department of Natural Resources, and the Louisiana Geological Survey for providing well log data, the LSU High Performance Computing for providing computing resources and technical assistance. I acknowledgments the financial supports from the National Science Foundation under Grant No. 1045064 and the LSU Economic Development Assistantship (EDA).

Finally, I would like to dedicate this work to my parents for everything they have done for me; to my wonderful wife Hanh, and my sons, Minh Huy and Khoi Nguyen for their constant patience, love and support.
# Table of Contents

Acknowledgments ........................................................................................................ ii

List of Tables ................................................................................................................ vi

List of Figures ................................................................................................................. vii

List of Abbreviations ..................................................................................................... xi

List of Symbols ............................................................................................................... xii

Abstract ......................................................................................................................... xiv

1. Introduction .................................................................................................................. 1

2. Literature Review ....................................................................................................... 7
   2.1. Uncertainty in Groundwater Head Prediction ................................................... 7
   2.2. Experimental Designs ......................................................................................... 9
   2.3. Model Discrimination Criteria ......................................................................... 10
   2.4. MODFLOW Grid Generation ........................................................................... 11

3. Methodology ............................................................................................................... 14
   3.1. MODFLOW Grid Generation ........................................................................... 14
       3.1.1. Well Log Interpretation .......................................................................... 14
       3.1.2. Indicator Kriging for Construction of Hydrostratigraphic Architecture .... 16
       3.1.3. Upscale Hydrostratigraphic Architecture to Structured Grid ................. 17
       3.1.4. Upscale Hydrostratigraphic Architecture to Unstructured Grid ............. 21
   3.2. High-performance Computing for Groundwater Model Calibration ............ 24
       3.2.1. Groundwater Model Calibration Using CMA-ES .................................. 24
       3.2.2. Parallel CMA-ES for High Performance Computing ............................... 25
   3.3. Model Discrimination Criteria ......................................................................... 27
   3.4. Experimental Design using BMA Prediction ...................................................... 28
       3.4.1. Time-Sequential Sampling Approach ....................................................... 29
       3.4.2. Update Posterior Model Probabilities Using BMA Mean Predictions ....... 31
       3.4.3. Proposition Discrimination and Uncertainty Analysis ......................... 33
   3.5. Optimal Observation Network Design Using Information Theory and BMA .... 36
       3.5.1. Shannon Entropy and Discrimination Function ....................................... 36
       3.5.2. \( \text{Pr}(M_i | \Delta_{\text{obs}}) \) Calculation .......................................................... 38
       3.5.3. \( \text{Pr}(M_i | \Delta_{\text{obs}}, \Delta_{\text{new}}) \) Calculation ............................................... 39
       3.5.4. \( \Delta', \Delta_i, \sigma_i^2, \text{and } \sigma_i^2 \) Calculations ........................................ 39
       3.5.5. Future Data Predictions by BMA for \( \Delta_{\text{new}} \) ....................................... 40
       3.5.6. \( \sigma^2 \) Estimation of by BMA ................................................................. 42
References .................................................................................................................. 107
Appendix A: Derivation of a Multi-Data Discrimination Function........................................... 121
Appendix B: Open Access Permission ....................................................................................... 124
Vita........................................................................................................................................... 131
List of Tables

Table 5-1: RMSE and the estimated parameter values of 18 candidate models.......................... 62

Table 6-1: RMSE, weighted sum of squared errors, BIC values, and posterior model probabilities for the 18 groundwater models using variance window with 5% significance levels and $4\sigma_D$... 76
List of Figures

Figure 1.1: Map of the study area. .......................................................... 2

Figure 3.1: Interpretation of fluvial facies for a well log: A. amalgamated braided channel-fill with brackish water, B. channel-fill point bar sand with brackish water, C. stacked/amalgamated channel-fill with very salty water, D. floodplain, and E. natural levee. ....................... 15

Figure 3.2: Translation of well log positions from (a) dipped domain to (b) non-dipped domain. .................................................................................................................. 16

Figure 3.3: A schematic for thin sand and thin clay elimination in vertical columns. The black dots are bed boundaries. Segments between two consecutive bed boundaries are sand or clay. ....... 18

Figure 3.4: Illustration of bed boundary projections: (a) six more bed boundaries are added to the vertical column (i, j) from its four adjacent columns; and (b) only three bed boundaries are added to the vertical column (i, j) after applying the thickness threshold........................................ 19

Figure 3.5: Illustration of the ruler algorithm: (a) a vertical column with a distinct number of layers up to each bed boundary; and (b) a vertical column where two bed boundaries have the same number of layers. ........................................................................................................ 20

Figure 3.6: Illustration for conceptual model development and grid generation for: (a) a non-uniform structured grid with the grid cell size of 200 m by 200 m, and (b) a quadtree unstructured grid with the minimum cell size of 200 m by 200 m. ................................................................................. 23

Figure 3.7: Embarrassingly parallel scheme of CMA-ES to run on high-performance computing systems. .............................................................................................................. 26

Figure 4.1: Map of the well logs in the UTM (m) coordinate system. The box is the study area. 43

Figure 4.2: Construction of hydrostratigraphic architecture from well logs: (a) Distribution of boreholes and results of well log interpretations, and (b) hydrostratigraphic architecture......... 44

Figure 4.3: Distributions of the number of sand and clay segments in vertical columns .......... 46

Figure 4.4: Three-dimensional MODFLOW structured grid for the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand: (a) Boreholes and the MODFLOW structured grid, and (b) structured grid at the selected cross sections in Figure 4.1. Clay is blanked. The vertical exaggeration is 20 times. ................................................................. 48

Figure 4.5: Two-dimensional quadtree unstructured grid for the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand. .............................................................. 49

Figure 4.6: MODFLOW-USG unstructured grid: (a) boreholes and the MODFLOW-USG quadtree unstructured grid, and (b) unstructured grid at the selected cross sections in Figure 4.1. Clay is blanked. The vertical exaggeration is 20 times. ................................................................. 50
Figure 5.1: A three-dimensional MODFLOW grid for the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand. The generalized parameterization method was used to generate the hydrostratigraphic architecture. Clay cells are blanked. The vertical exaggeration is 20 times. .................................................................................................................. 54

Figure 5.2: Distribution of the cumulative pumpage in the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand from 1975 to 2010. .......................................................... 55

Figure 5.3: Illustration of model components for MODFLOW model. ........................................... 56

Figure 5.4: Sand distributions at layer 36 using (a) IZ, (b) GP and (c) IK methods, showing uncertainty in geological architecture construction. ......................................................... 57

Figure 5.5: The number of iterations for different population sizes required to reach several target fitting errors. ........................................................................................................ 60

Figure 5.6: A comparison of RMSE of 18 candidate groundwater models. ........................................ 62

Figure 5.7: Observed heads, simulated heads and one-standard deviation bounds at selected USGS observation wells in the “1,200-foot” sand. ............................................................. 64

Figure 5.8: Observed heads, simulated heads and one-standard deviation bounds at selected USGS observation wells in the “1,500-1,700-foot” sands. ..................................................... 65

Figure 5.9: Distribution of simulated groundwater heads for the Industrial District given by the best model GP-B-1.5-HC4 on (a) 1/1/1975; and (b) 12/31/2010. ..................................................... 66

Figure 5.10: Distribution of simulated groundwater heads for the Lula wells and Government St. wells pumping areas given by the best model GP-B-1.5-HC4 on (a) 1/1/1975; and (b) 12/31/2010. ......................................................................................... 67

Figure 5.11: Top view of simulated heads given by the best model on 12/31/2010 for (a) the “1,200-foot” sand; and (b) the “1,500-1,700-foot” sands. .......................................................... 68

Figure 5.12: Groundwater budget for the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand in between the two faults in December 2010: (a) inflow; and (b) outflow. ................. 69

Figure 5.13: Groundwater budget for the “1,500-1,700-foot” sands in between the two faults in December 2010: (a) inflow; and (b) outflow. .................................................................................. 70

Figure 5.14: Head differences across the two faults: (a) a cross-section showing vertical distribution of groundwater heads at the end of model calibration. The locations with black dots were used to calculate the head differences; and (b) simulated head differences across the BR fault and the DSS fault. ........................................................................................................ 71

Figure 5.15: Predicted heads and one-standard deviation bounds of the best model, the second best model and other 16 groundwater models at 19 USGS observation wells. .......... 73
Figure 6.1: Map of the study area in the UTM (m) coordinate system, which includes location of the pumping wells, the USGS observation wells, the BR fault, and the DSS fault.  

Figure 6.2: The BMA tree of posterior model probabilities and conditional posterior model probabilities before the observation network design. Posterior model probabilities and conditional posterior model probabilities less than 1% are not shown in the figure.  

Figure 6.3: Change of posterior model probabilities of 18 groundwater models after 60 months of continuous groundwater head observation: (a) using the eight USGS observation wells used in model calibration; (b) using one USGS observation well at different locations; and (c) using a different number of best USGS observation wells.  

Figure 6.4: The least months to reach the design criterion of 80% posterior model probability given a number of best selected USGS observation wells.  

Figure 6.5: The frequency of the USGS observation wells to be selected for the best observation network designs given the criterion of 80% posterior model probability.  

Figure 6.6: Change of posterior model probabilities of 18 groundwater models over time based on the best selected 6 USGS observation wells, EB-146, EB-413, EB-652, EB-789B, EB-946, and EB-996.  

Figure 6.7: The BMA tree of posterior model probabilities and conditional posterior model probabilities in December 2015 after the observation network design with 6 best selected USGS observation wells. Posterior model probabilities and conditional posterior model probabilities less than 1% are not shown in the figure.  

Figure 6.8: Comparisons of the total variances of predicted groundwater heads at 19 USGS observation wells on 12/31/2015 without and with the observation network design of 6 best selected USGS observation wells after 60 months of observation.  

Figure 7.1: Posterior probabilities of 18 models given the existing observation data in 2010.  

Figure 7.2: Distributions of discrimination function values averaged over layers for the first sampling round considering homoscedastic and heteroscedastic future observations. F values smaller than 0.6 are blanked. The black dots are the locations of the USGS observation wells.  

Figure 7.3: Distributions of discrimination function values averaged over layers after the first sampling round under different sources of model structure uncertainty.  

Figure 7.4: Designs using one USGS well after the first sampling round: (a) discrimination function value at the USGS wells; and (b) expected posterior model probabilities and one-standard deviation bounds using data from EB-291.  

Figure 7.5: Design using one best USGS well at different sampling rounds: (a) evolution of discrimination function values using data from EB-291 or from EB-652; and (b) expected posterior model probabilities and one-standard deviation bounds of three best models found at the end of the design.
Figure 7.6: Best observation locations given a desired number of USGS wells for the first four sampling rounds. ................................................................. 98

Figure 7.7: Discrimination function values, 99%-identifiable groundwater models, and sampling rounds given a desired number of USGS observation wells.......................................................... 99

Figure 7.8: Expected posterior model probabilities and one-standard deviation bounds of the four identified groundwater models in Figure 7.7................................................................. 100

Figure 7.9: Total variances of predicted groundwater heads on December 31, 2015 at 19 USGS observation wells before and after the optimal observation network design.............................. 101
List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIC</td>
<td>Bayesian information criterion</td>
</tr>
<tr>
<td>BMA</td>
<td>Bayesian model averaging</td>
</tr>
<tr>
<td>BR</td>
<td>Baton Rouge</td>
</tr>
<tr>
<td>CHD</td>
<td>The time-variant specified head</td>
</tr>
<tr>
<td>CMA-ES</td>
<td>The covariance matrix adaptation-evolution strategy</td>
</tr>
<tr>
<td>DSS</td>
<td>Denham Springs-Scotlandville</td>
</tr>
<tr>
<td>GMS</td>
<td>Groundwater modeling system</td>
</tr>
<tr>
<td>HBMA</td>
<td>Hierarchical Bayesian model averaging</td>
</tr>
<tr>
<td>RMSE</td>
<td>The root mean square error</td>
</tr>
<tr>
<td>SP</td>
<td>Spontaneous potential</td>
</tr>
<tr>
<td>TSS</td>
<td>Time-sequential sampling (TSS)</td>
</tr>
<tr>
<td>USGS</td>
<td>The United States geological survey</td>
</tr>
</tbody>
</table>
List of Symbols

\( \Delta_{\text{obs}} \) the existing observation data
\( \Delta_{\text{new}} \) the future observation data at potential observation locations and time using a design \( D \)
\( \Delta_{\text{new}}^D \) the BMA mean prediction
\( \Delta_1, \Delta_2, \ldots, \Delta_k \) the total data collected after the first round, the second round, \ldots, and the \( k \)th round of data collection
\( \Delta_i = \{ \Delta_{n,i} \} \) the future observation data predicted by models \( M_i \) at potential observation locations \( n = 1, 2, \ldots, N \)
\( \Delta_i' \) simulated quantities by model \( M_i \) at the existing data domain
\( \Sigma_i^2 = \sigma_i^2 + \sigma_i'^2 \) the diagonal covariance matrix of future observation data
\( \Sigma_i'^2 = \sigma_i'^2 + \sigma_i''^2 \) the diagonal covariance matrix of existing observation data
\( \alpha \) the scaling factor for the variance window
\( \beta_i \) model parameters of model \( M_i \)
\( \hat{\beta}_i \) the maximum likelihood estimate of model parameters \( \beta_i \) of model \( M_p \)
\( \gamma' \) a probability threshold
\( \delta \) the sand and clay cutoff value
\( \eta \) the expected values of future observation data
\( \lambda_i \) the indicator kriging weights
\( \nu(\mathbf{x}_0) \) the expected value of the indicator at an unobserved location \( \mathbf{x}_0 \)
\( \sigma_j^2 \) the variance of prediction errors at the \( j \)th observation in the historical data space
\( \sigma_{m,n}^2 \) the variance of prediction errors at the \( n \)th observation of the \( m \)th sampling round in the prediction space
\( \sigma_D \) coefficient of variance window
\( \sigma_i^2 = \{ \sigma_{n,i}^2 \} \) the variances of future observation data at potential observation locations \( n = 1, 2, \ldots, N \)
\( \sigma_i'^2 \) the diagonal covariance matrix of the existing observation data
\( \sigma_i^2 = \{ \sigma_{n,i}^2 \} \) the variances of predicted future observation data by model \( M_i \) at potential observation locations \( n = 1, 2, \ldots, N \)
\( \sigma_i''^2 \) the variances of simulated quantities by model \( M_i \) at the existing data domain
\( (\text{BIC}_{i,k})_{\text{min}} \) the minimum BIC value of the models.
a potential design

the optimal design to determine potential locations and time of observations that produces the highest posterior model probability

the discrimination function

indicator function

\[ M = \{ M_i ; i = 1, 2, \ldots, m \} \]
a set of m candidate models

\[ M_B \]
the base models at level B

\[ M_{B-1} \]
the BMA models at level B-1

\[ M_B^{(i)} \]
the models at level B under the same propositions of level B-1

the number of existing observation data

the number of future observation data

the number of well logs for a horizontal surface

the number of potential observation locations

number of models

the posterior model probabilities calculated using \( \Delta^{\text{obs}} \)

the marginal likelihood

the prior model probability

the sum of weighted squared errors.

the number of realizations of parameters

the number of unknown model parameters to be estimated

the expected entropy change before and after the design

entropy of a system

an observed head at the observation \( i^{\text{th}} \)

a calculated head \( i \) given model parameters \( \beta \) for model \( i \).

the number of sampling rounds

a HBMA level

the expected marginal likelihood

the time for the first round, the second round, … , and the \( k^{\text{th}} \) round of data collection
Abstract

Groundwater systems are complex and subject to multiple interpretations due to a lack of sufficient information. Different propositions (or alternatives) are often proposed to represent uncertain model components resulted in many conceptual models using the same data. Yet considering too many models may lead to high prediction uncertainty and may lose the purpose of model development. To decrease the prediction uncertainty due to conceptual model uncertainty, two experimental designs are proposed.

The first experimental design intends to identify model propositions in each uncertain model component. A discrimination criterion is developed based on posterior model probability. Bayesian model averaging (BMA) is used to predict future observation data. The experimental design aims to find the optimal number and location of future observations and the number of sampling rounds such that the desired discrimination criterion is met. Hierarchical Bayesian model averaging (HBMA) is adopted to assess if highly probable propositions can be identified.

The second experimental design is to discriminate conceptual models and in turn, reduce the number of models. The Box-Hill discrimination function derived for one additional observation was modified to account for multiple independent spatiotemporal observations. The BMA method is used to predict future observation data and quantify conceptual and parametric prediction uncertainty. The design goal is to find optimal locations and the number of sampling rounds such that the Box-Hill discrimination function value is maximized, and the highest posterior probability of a model satisfies a desired probability threshold.

The experimental designs are implemented to plan new head observation networks based on existing USGS wells in the Baton Rouge area, Louisiana. The sources of uncertainty that create
multiple groundwater models are geological architecture, boundary condition, and fault permeability architecture. All possible design solutions are enumerated using a multi-core supercomputer. The result shows that each highly probable proposition can be identified for each uncertain model component once the discrimination criterion is achieved. Heteroscedasticity (unequal variances) for future groundwater heads should be considered in the design procedure to account for various sources of future observation uncertainty. The variances of head predictions are significantly decreased by reducing posterior model probabilities of unimportant models.
1. Introduction

Groundwater resources are vital to the Baton Rouge area, southeastern Louisiana. In 2010, total water use was 182 (million gallon per day) (MGD) (690,000 m³/d) of which 106 (MGD) (401,300 m³/d) is from groundwater accounting for 87.8%. More importantly, 100% public water supply for the Baton Rouge area comes from groundwater [Sargent, 2012]. Due to excessive groundwater withdrawal, the fresh water aquifers underneath the Baton Rouge area are being contaminated by saltwater intrusion.

The study area is shown in Figure 1.1 including parts of East Baton Rouge Parish and West Baton Rouge Parish. The aquifer system of the study area belongs to the Southern Hills regional aquifer system and is extended to a depth of 3000 ft. (914 m). The aquifer system consists of a succession of south-dipping siliciclastic sandy units and mudstones of Upper Miocene through Pleistocene age. As a result of fluvial deposition, the aquifer system is highly complex. Sand deposition is non-uniform due to spatial and temporal variations in fluvial processes as well as large amount of missing sand possibly due to erosional unconformity [Chamberlain et al., 2013]. This study focuses on the “1,200-foot” sand, the “1,500-foot” sand and the “1,700-foot” sand. The sand units were classified and named by their approximate depth below ground level in the Baton Rouge Industrial District [Meyer and Turcan, 1955].

The Baton Rouge fault system consists of the east-west trending Baton Rouge (BR) fault and Denham Springs-Scotlandville (DSS) fault. These two faults crosscut the aquifer/aquitard sequence in the study area [McCulloh and Heinrich, 2013]. The BR fault separates the freshwater to the north from the saline water to the south [Rollo, 1966]. The presence of saltwater in the north of the BR fault has been documented in several studies [Morgan and Winner, 1964; Whiteman, 1979; Tomaszewski, 1996; Lovelace, 2007]. The source of saline water is likely from the expulsion
of over-pressured brine fluids, extending vertically upward above the top of Gabriel salt dome all the way to the water table [Anderson et al., 2013]. Bense and Person [2006] suggested the BR fault is low-permeability faults and acts as horizontal flow barriers.

Figure 1.1: Map of the study area.

Excessive groundwater withdrawals in the area between two faults have caused declination of groundwater levels north of the BR fault. The USGS reported as much as 53.34 m (175 feet) decline in groundwater level in the “1,500-foot” sand and as much as 83.82 m (275 feet) decline in groundwater level in the “2,000-foot” sand in the Baton Rouge area since 1945. The declination of groundwater level north of the BR fault has caused changing of naturally flow direction that is
from north to south. The flow direction is now reversed and saltwater at south of the BR fault is moving north toward pumping stations.

To better understand the impact of groundwater withdrawals that caused water level decline and saltwater water intrusion, this study developed three-dimensional groundwater flow models, which included the “1,200-foot”, the “1,500-foot”, and the “1,700-foot” sand and the two faults. The ultimate goal of this study is to identify a reliable groundwater model for further saltwater intrusion study, according to the four following research steps:

The first research step is to develop a grid generation technique to create structured and unstructured MODFLOW grids from well logs. The technique maintains high vertical resolution of hydrofacies geometries with a reasonable number of non-uniform boundary-fitted layers. This is done by vertically upscaling a very-fine geological architecture into a computational grid for flow and transport simulations. In addition, the grid generation technique avoids correlation error by using manual correlation and avoids gridded error by using a pre-defined grid for the flow and transport simulation purposes. Another advantage of the grid generation technique is its ability to facilitate simulation model development and model structural uncertainty analysis by rapidly reconstructing grids for different hydrofacies interpolation methods or when new well log data become available. While the technique is mainly for generating both structured and unstructured grids for USGS MODFLOW, it can also be adopted to generate finite-element grids, which is outside the scope of this study. The grid generation technique is applied to generation of both MODFLOW structured grids and MODFLOW-USG unstructured grids for a real-world fluvial aquifer system of the Baton Rouge area, southeastern Louisiana, including two geological faults.

The second research step is to calibrate complex groundwater models and quantify their related uncertainty by adopting the Covariance Matrix Adaptation-Evolution Strategy (CMA-ES)
[Hansen and Ostermeier, 2001; Hansen et al., 2003] and the “embarrassingly parallel” in a high-performance computing system. The CMA-ES can obtain a near global solution of rugged, non-separable and noisy function that is typically encountered in solving inverse problems in groundwater modeling. Owing to the population-based method, the global search efficiency can be enhanced by increasing the population size. However, this may result in greater computational burden as it increases the number of function evaluations in each iteration. This challenge is overcome by adopting the parallel CMA-ES [Elshall et al., 2015] to run in a high-performance computing system. Along with the parameter estimation, CMA-ES adapts the full covariance matrix of the estimated parameters that can be used for Monte Carlo sampling to quantify the parameters-related uncertainty.

The lack of hydrogeological data and knowledge often results in different propositions (or alternatives) to represent uncertain model components and creates many candidate groundwater models using the same data. Uncertainty of groundwater head prediction may become unnecessarily high. The third research step is to introduce an experimental design to identify propositions in each uncertain model component and decrease the prediction uncertainty by reducing conceptual model uncertainty. A discrimination criterion is developed based on posterior model probability that directly uses data to evaluate model importance. Bayesian model averaging (BMA) is used to predict future observation data. The experimental design aims to find the optimal number and location of future observations and the number of sampling rounds such that the desired discrimination criterion is met. Hierarchical Bayesian model averaging (HBMA) is adopted to assess if highly probable propositions can be identified, and the conceptual model uncertainty can be reduced by the experimental design. The experimental design is implemented
to design a new head observation network based on existing USGS observation wells in the Baton Rouge area, Louisiana.

The fourth research step is to expand the Box-Hill discrimination function to design an optimal observation network to discriminate conceptual models and in turn, reduce the number of models. The Box-Hill discrimination function was derived for one additional observation. This study modifies the discrimination function to account for multiple independent spatiotemporal observations. The BMA method is used to incorporate existing observation data and quantify future observation uncertainty arising from conceptual and parametric uncertainties in the discrimination function. In addition, the BMA method is adopted to predict future observation data. The design goal is to find optimal locations and the number of sampling rounds such that the Box-Hill discrimination function value is maximized, and the highest posterior probability of a model satisfies a desired probability threshold. The optimal observation network design is implemented in a groundwater study in the Baton Rouge, Louisiana area to collect additional groundwater heads from USGS observation wells.

This dissertation is organized in eight chapters. Chapter 1 presents a general introduction of the study area and four research steps to obtain the final goal of this study. Chapter 2 presents a literature review about the uncertainty in groundwater head prediction, experimental designs, and techniques to generate MODFLOW grids. Chapter 3 presents techniques and mathematical formulations that will be used in this study. Chapter 4 presents the first research step that is to generate MODFLOW grid from well log data. Chapter 5 presents the second research step that is to calibrate complex groundwater models and quantify uncertainty using the parallel CMA-ES. Chapter 6 presents an experimental design to identify model propositions under conceptual model uncertainty. Chapter 7 presents an optimal observation network design for model discrimination
and uncertainty reduction using the BMA method and information theory. The final chapter, chapter 8 presents conclusion remarks.

Most of this work is published or submitted for publication in Pham and Tsai [2015], Pham and Tsai (submitted), and Pham and Tsai (submitted). The grid generation technique was used to generate MODFLOW grids in Elshall and Tsai [2014], Chitsazan et al. [2014] and Elshall et al. [2015]. The groundwater flow models developed in this study were used to develop saltwater intrusion prediction models in Chitsazan [2014] and to design a hydraulic barrier to protect pumping wells from saltwater intrusion in Chitsazan et al. [2014].
2. Literature Review

2.1. Uncertainty in Groundwater Head Prediction

Conceptualization of groundwater systems is a very difficult task due to its complexity, multiple interpretations, and insufficient data. Uncertainty always exists in the conceptualization process, and many alternatives are developed to represent uncertain model components of groundwater systems. For example, geological architecture can be one uncertain model component, and many hydrostratigraphy modeling techniques may be proposed to construct different geological architectures. As a result, multiple conceptual models are often developed to represent the groundwater systems. Identification of a reliable groundwater model for future applications is the ultimate goal of model development. This goal, however, remains challenging due to aforementioned difficulties.

Over the past several decades, many studies have focused on understanding, quantifying, and reducing head prediction uncertainty arising from model parameter uncertainty given a conceptual model [Sun, 1994; Carrera et al., 2005]. Considering only one model may lead to statistical bias and underestimation of uncertainty because groundwater systems are often complex and have multiple interpretations [Neuman, 2003; Poeter and Anderson, 2005; Refsgaard et al., 2006a].

In recent years, conceptual model uncertainty has received much attention in groundwater applications [Hsu and Yeh, 1989; Beven and Freer, 2001; Neuman and Wierenga, 2003; Troldborg et al., 2007; Rojas et al., 2008, 2010; Singh et al., 2010; Ye et al., 2010; Gupta et al., 2012]. Many studies have shown that contribution of conceptual model uncertainty to predictive uncertainty is significantly larger than that of model parameter uncertainty [Carrera and Neuman, 1986; Bredehoeft, 2003, 2005; Neuman, 2003; Neuman and Wierenga, 2003; Poeter and Anderson,
Common strategies for assessing conceptual model uncertainty are model selection and identification to find the best model, model elimination to eliminate unimportant models, model averaging to obtain an ensemble prediction of alternative conceptual models, and model discrimination [Burnham and Anderson, 2002; Neuman, 2003; Neuman and Wierenga, 2003; Poeter and Anderson, 2005; Refsgaard et al., 2006; Rojas et al., 2008, 2009, 2010a, b, c; Tsai and Li, 2008a, b; Li and Tsai, 2009; Tsai, 2010; Ye et al., 2008, 2004, 2005; Singh et al., 2010; Trolldborg et al., 2010; Seifert et al., 2012; Foglia et al., 2007, 2013; Tsai and Elshall, 2013; Elshall and Tsai, 2014; Chitsazan and Tsai, 2014a, b; Usunoff et al., 1992].

As Poeter and Anderson [2005] pointed out, if one model is clearly superior to other candidate models represented by its best fit to observations, one should use that model for prediction and reject other models. If none of the models are dominant, their mean predictions are preferentially utilized to avoid biased predictions from using a single conceptual model.

Bayesian model averaging [Draper, 1995; Kass and Raftery, 1995; Raftery et al., 1997; Hoeting et al., 1999] is often employed to conduct multi-model prediction studies because BMA employs probabilistic techniques to derive consensus predictions from a set of candidate models based on their corresponding posterior model probabilities. Averaged predictions from BMA are less biased than predictions obtained from individual models [Raftery and Zheng, 2003; Ye et al., 2004; Ajami et al., 2006]. Moreover, BMA is able to study uncertainty propagation from model parameter uncertainty and model structure uncertainty to model prediction uncertainty, thereby distinguishing prediction uncertainty arising from individual models, between models, and between methods [Hsu and Yeh, 1989; Tsai and Li, 2008; Li and Tsai, 2009].
Yet considering too many conceptual models requires a great deal of computationally intensive when conducting prediction and uncertainty analysis. Prediction results using BMA can become very high and may lose the purpose of model development [Bredehoeft, 2005; Højberg and Refsgaard, 2005]. This concern highlights the importance of conducting an experimental design to discriminate groundwater models, identify highly probable model and model propositions, and in turn, reduce conceptual model uncertainty (model choice uncertainty) in prediction.

2.2. Experimental Designs

In the field of groundwater hydrology, experimental designs have been studied extensively. Readers are referred to several in-depth review articles [Loaiciga et al., 1992; Hassan, 2003; Minsker, 2003; Kollat et al., 2011]. Predominant studies are focused on improving parameter identification [Hsu and Yeh, 1989; Cleveland and Yeh, 1990; Sun and Yeh, 1990, 2007; Wagner, 1995; Altmann-Dieses et al., 2002; Sciortino et al., 2002; Chang et al., 2005; McPhee and Yeh, 2006], reducing model prediction uncertainty [McKinney and Loucks, 1992; Wagner, 1995; Chadalavada and Datta, 2008; Janssen et al., 2008; Nowak et al., 2010], minimizing cost [Criminisi et al., 1997; Herrera and Pinder, 2005; Zhang et al., 2005; Sun and Yeh, 2007], minimizing decision errors [Nowak et al., 2012], detecting of plumes [Meyer and Brill, 1988; Storck et al., 1997; Dhar and Datta, 2007; Kim and Lee, 2007; Dokou and Pinder, 2009], and combinations of these in multi-objective formulations [Knopman and Voss, 1989; Dhar and Datta, 2007; Kollat et al., 2011; Dhar and Patil, 2012; Alzraiee et al., 2013].

Few studies have used experimental designs to discriminate candidate models and identify the “true” model [Knopman and Voss, 1988, 1989; Knopman et al., 1991; Usunoff et al., 1992; Yakirevich et al., 2013]. Knopman and Voss [1988] investigated the theoretical discrimination
power of designs and suggested sample locations where predictions of candidate models were the most different. They proved the efficiency of the method in a sampling design for a solute transport study at Cape Cod, Massachusetts in a subsequent paper [Knopman et al., 1991]. Usunoff et al. [1992] hypothesized that the “true” model is close to one of the candidate models and no admissible parameter sets of remaining candidate models could produce similar predictions. They used an experimental design to discriminate models based on the distance of model predictions. Yakirevich et al. Yakirevich et al. [2013] recently applied the Kulback-Leibler information to discriminate flow and transport models.

A variety of methodologies have been introduced in experimental designs for groundwater modeling [Loaiciga et al., 1992; Mogheir et al., 2006]. Among these methods, physically based simulation approach [Knopman and Voss, 1989; Cleveland and Yeh, 1990; Hudak and Loaiciga, 1992; McKinney and Loucks, 1992; Meyer et al., 1994; Cieniawski et al., 1995; Wagner, 1995; Storck et al., 1997; Reed et al., 2000; Dhar and Datta, 2007; Sun and Yeh, 2007] and information theory (entropy-based method) [Mogheir and Singh, 2002; Poeter and Anderson, 2005; Mogheir et al., 2006; Alfonso et al., 2010] are commonly employed owing to their flexibility in examining design scenarios and design constrains.

2.3. Model Discrimination Criteria

Significant efforts have been invested in the development of different discrimination criteria for experimental designs in various fields [Hunter and Reiner, 1965; Box and Hill, 1967; Atkinson and Fedorov, 1975; Hill, 1978; Atkinson, 1981; Guido Buzzi-Ferraris, 1990; Leon and Atkinson, 1991; Pukelsheim and Rosenberger, 1993; Schwaab et al., 2008; Michalik et al., 2010; Alberton et al., 2012; Donckels et al., 2012; Yakirevich et al., 2013], but only a few criteria have been applied to groundwater modeling, e.g., the differences between model outputs [Knopman and
Model discrimination criteria should serve not only for model discrimination, but also for model identification. As discussed in Box and Hill [1967], the ultimate goal of maximizing information from the system should aim to make the posterior probability of one model to be 1 and others to be zero.

2.4. MODFLOW Grid Generation

Numerical models have been widely used to simulate and predict subsurface flow and solute transport, especially for modeling highly complex aquifer systems. One of the challenges, besides estimation of spatially variable hydraulic parameters, is to construct better model grids that are consistent with the geometries of modelled hydrofacies. Errors from inaccurate model grids that fail to capture hydrofacies geometries often result in incorrect estimated hydraulic parameters during model calibration.

In the past decades, many methods have been developed to model hydrofacies for different heterogeneity scales using various input data sets (e.g., well logs, pumping test, and seismic data). Among them, widely used methods in hydrogeology are the two-point variogram statistics, such as indicator geostatistics [Journel, 1983; Johnson and Dreiss, 1989; Johnson, 1995; Proce et al., 2004]; transition probability-based indicator geostatistics [Carle and Fogg, 1996; Lee et al., 2007; Koch et al., 2014]; and multiple-point simulation (MPS) [Strebelle, 2002; Journel, 2005; dell’Arciprete et al., 2012]. Reviews of these methods can be found in Koltermann and Gorelick [1996], Marsily et al. [2005] and Hu and Chugunova [2008]. While the applications of these methods are site specific and are subject to user preference and expertise, it has been well understood that different hydrofacies methods generate significantly different spatial distributions of hydraulic
properties [Alabert and Modot, 1993; Gómez-Hernández and Wen, 1998; Western et al., 2001; Zinn and Harvey, 2003; Zhang et al., 2006; Lee et al., 2007; Bianchi et al., 2011; Berg and Illman, 2015] and consequent flow and solute transport responses. Conducting uncertainty analysis with respect to different hydrostratigraphic characterizations is crucial [Neuman and Wierenga, 2003; Bredhoeft, 2005; Troldborg et al., 2007; Rojas et al., 2008; Seifert et al., 2008; Gupta et al., 2012; Neuman et al., 2012; Refsgaard et al., 2012], but is very challenging for flow and transport modeling since computational grids need to be modified to fit different hydrofacies models [Refsgaard et al., 2012; Chitsazan et al., 2014; Elshall and Tsai, 2014; Koch et al., 2014].

The literature shows a lack of a general approach to accurately convert a hydrostratigraphic architecture into a grid. Two approaches are commonly used: one is the solid approach and the other is the pre-defined grid approach. Before implementing either approach, one needs to obtain geological information (e.g., lithology, bed boundary elevation, formation dip, etc.) from well logs. Readers are referred to some classic books for well log interpretation techniques [Schlumberger, 1972; Hilchie, 1982; Bassouni, 1994], which were used to interpret well logs for this study.

Using the solid approach, one needs to manually correlate well logs and label distinct hydrofacies for each well log. Once the well log correlation is established, interpolation methods are applied to generate surfaces using the same types of hydrofacies. These surfaces represent the hydrofacies boundaries and result in a solid model. Jones et al. [2002] and Lemon and Jones [2003] developed a grid generator to generate computational grids for MODFLOW [Harbaugh, 2005] from the solid model. The beauty of this approach is the creation of non-uniform computational layers that match well with the generated surfaces, including pinch-outs. Due to its simplicity, the algorithm was adopted in several commercial software packages, e.g., Groundwater Modeling System (GMS) [Aquaveo, 2014] and RockWorks [RockWare, 2014], among others.
The biggest challenge in using this approach is performing manual correlation between well logs, which is subjective and can become laborious and impractical when dealing with a huge number of well logs in areas known to be highly complex. (e.g., fluvial depositional environments). Manually correlating well logs often results in inconsistencies with geological deposition, forces correlation of unrelated hydrofacies, and produces erroneous hydraulic connections of discontinuous hydrofacies.

The pre-defined grid approach is to estimate hydrofacies at a pre-defined grid. Mostly, this approach generates uniform, relatively coarse layers. Pre-defined grids are directly used for flow and transport modeling. Examples of this approach include using T-PROGS [Carle, 1999] and geostatistical tools (e.g., GSLIB [Deutsch and Journel, 1997]). This approach does not force generating surfaces of hydrofacies, and therefore, avoids the issues caused by manual correlation. Since geostatistical approaches have been well received in the literature, they are commonly adopted by commercial software, for example, T-PROGS with GMS [Jones et al., 2005; Faulkner et al., 2012; Aquaveo, 2014], among others. However, the greatest concern of using pre-defined grids is of losing the vertical resolution of hydrofacies geometries if layers are not fine enough. Using fine layers can improve vertical resolution to better capture vertical hydrofacies geometry, but will significantly increase flow and transport computation time since pre-defined grids are directed used for simulation.
3. Methodology

3.1. MODFLOW Grid Generation

3.1.1. Well Log Interpretation

The primary sources of information used to establish hydrofacies geometries are wire-line spontaneous potential (SP) and electrical resistivity logs for boreholes. Spontaneous potential and resistivity log responses are controlled largely by the ratio of sand to clay minerals. They have long been used to interpret sedimentary depositional environments. Galloway [1977] used SP and resistivity curve morphologies to identify fluvial facies for channel fill, levee, crevasse splay and floodplain, and established a meandering stream facies. Kerr and Jirik [1990] adapted Galloway’s [1977] facies model and provided examples of SP and resistivity responses that match known fluvial facies for the middle Frio formation, South Texas. Sands deposited by braided streams produce jagged, wedge-shaped curve morphologies [Miall, 2010]. Based on these established relationships between log responses and fluvial facies, Chamberlain et al. [2013] used SP and resistivity data to study depositional environments of siliciclastic sediments in the Baton Rouge area.

Following Chamberlain et al. [2013], this study uses SP, resistivity, and gamma ray (when available) to identify the location of sand facies at depth. Figure 3.1 shows a typical SP-resistivity log in the Baton Rouge area. Based on deviations from a visually estimated shale baseline, boundaries of sands can be drawn on inflection points of SP curves. A cutoff value generally fell between 10 and 35 ohm-m for resistivity curves is assigned to determine boundaries of sands. Low long-normal resistivity generally indicates the occurrence of salty water. When gamma ray curve is available, low gamma ray response generally indicates a sand facies. Sand boundaries can be well identified by correlating SP, resistivity and gamma ray curves [Schlumberger, 1972; Hilchie,
For example, seven sand facies are picked and many thin sands are ignored as shown in Figure 3.1. Non-sand intervals are assumed to be clay (shale or mudstone) facies. This study analyzed 583 well logs in the Baton Rouge area based on this approach.

Figure 3.1: Interpretation of fluvial facies for a well log: A. amalgamated braided channel-fill with brackish water, B. channel-fill point bar sand with brackish water, C. stacked/amalgamated channel-fill with very salty water, D. floodplain, and E. natural levee.

Well log interpretation is inevitably subjected to an individual’s experience and the purpose of the work. This study does not intend to discuss the uncertainty of computational grids due to different log interpretations. Moreover, it is possible to use the established relationships from Galloway [1977], Kerr and Jirik [1990], and Miall [2010] to infer different fluvial facies
underneath Baton Rouge. For example, Figure 3.1 shows some identified fluvial facies based on the established relationships. However, identifying specific fluvial facies is not the scope of this work. Instead, this study focuses on sand and clay facies identification in well logs.

3.1.2. Indicator Kriging for Construction of Hydrostratigraphic Architecture

This study constructs hydrostratigraphic architecture of sand and clay facies using information from the well logs. The well log data are first transformed into binary indicator values. The indicator value for sand facies is 1 and for clay facies is 0. To honor a regional geological dip as shown in Figure 3.2(a), indicator kriging needs to be performed on inclined surfaces where indicator data are obtained at the intersections with well logs. To make it easier for operating indicator kriging, all well logs are translated vertically to a non-dipped domain as illustrated in Figure 3.2(b).

![Figure 3.2: Translation of well log positions from (a) dipped domain to (b) non-dipped domain.](image)

The vertical translation distance depends on the dip angle and the distance from well log location to a strike that serves as a pivot. Then, indicator kriging is performed on horizontal surfaces. Tomography-type hydrofacies geometries can be achieved by assembling a large number of horizontal surfaces. This study conducts indicator kriging at horizontal surfaces of one-foot
intervals. It is noted that the grid generation technique in this study is not limited to indicator kriging. Any geostatistical methods can be used to estimate hydrofacies for a surface. The resulting indicator data from horizontal surfaces are used to compute experimental variograms. Then, a variogram model can be derived by fitting to the experimental variograms.

The expected value of the indicator at an unobserved location is obtained by

\[ v(x_0) = \sum_{i=1}^{N_L} \lambda_i I(x_i), \]

(1)

where \( v(x_0) \) is the expected value at an unobserved location \( x_0 \), \( N_L \) is the number of well logs for a horizontal surface, and \( \lambda_i \) are the indicator kriging weights. Indicator kriging has been well documented in the literature. Readers are referred to Olea [1999] for more information.

The expected value of indicator function represents the probability that facies at a location \( x \) fall into sand facies or clay facies. By giving a cutoff \( \delta \) as follows, distributed sand and clay facies on a horizontal surface can be achieved.

\[ I(x) = \begin{cases} 1 : \text{sand if } v(x) \geq \delta, \\ 0 : \text{clay if } v(x) < \delta. \end{cases} \]

(2)

Determination of a defensible cutoff value is challenging. A value of 0.5 is commonly used for a neutral section. However, a better cutoff can be determined in a calibration process where facies estimates are subject to additional information, e.g., driller’s logs, total volume of sand or clay facies from electrical logs, etc. [Elshall et al., 2013].

3.1.3. Upscale Hydrostratigraphic Architecture to Structured Grid

Once a hydrostratigraphic architecture with very fine vertical discretization is generated by indicator kriging, the following three steps are developed to upscale the hydrostratigraphic architecture to a MODFLOW structured grid by merging the same hydrofacies in the vertical
direction to reduce the number of layers. The structured grid for MODFLOW uses rectangular grid
cells in the two-dimensional grid. The three-dimensional model domain is discretized into rows,
columns, and layers, ordered in a Cartesian coordinate system. Each grid cell (except for those at
model boundaries) is connected to six surrounding cells. The key feature is that all upscaled layers
in the structured grid must be continuous throughout the model domain.

Step 1: Eliminate thin sand and thin clay

To avoid an overwhelming number of computation layers in MODFLOW, it is
recommended to eliminate thin sand and thin clay in each vertical column of the hydrostratigraphic
architecture before generating a MODFLOW grid. For the purpose of illustration, a vertical
column of the architecture is represented by a vertical line that passes through its center. Figure
3.3 shows two typical situations that involve thin sand or thin clay. Given a criterion of the
minimum thickness to define thin sand and thin clay, thin sand in thick clay or thin clay in thick
sand are eliminated as shown in Figure 3.3(a). For a sequence of thin sand and thin clay shown in
Figure 3.3(b), eliminate sand or clay, whichever has total thickness less than 50% in the sequence.
After eliminating thin sand and thin clay, bed boundaries of different facies in each vertical column
are assigned indices as basic information to form MODFLOW layer boundaries.

![Figure 3.3: A schematic for thin sand and thin clay elimination in vertical columns. The black
dots are bed boundaries. Segments between two consecutive bed boundaries are sand or clay.](image)
Step 2: Project neighboring bed boundaries

To account for the continuity of MODFLOW layers through neighboring columns, additional bed boundary information is added to a vertical column by projecting the bed boundaries of its four adjacent vertical columns to the column. Figure 3.4 illustrates this procedure. For example, from Step 1 column (i, j) has 4 bed boundaries shown in Figure 3.4(a). After projection, column (i, j) gains 6 additional bed boundaries from its four adjacent columns. The bed boundaries of column (i, j) increases to 10. This is an important step in order to preserve the continuity of flow pathways, especially through geological faults, pinch-out areas, or narrow connections between thick sands.

A new bed boundary may be deleted if the thickness between two consecutive bed boundaries is less than a thickness threshold. The smaller the thickness threshold is, the more the bed boundaries are added to vertical columns, which increases MODFLOW layers. By applying a thickness threshold, the final number of bed boundaries of column (i, j) reduces to 7 as shown in

![Figure 3.4: Illustration of bed boundary projections: (a) six more bed boundaries are added to the vertical column (i, j) from its four adjacent columns; and (b) only three bed boundaries are added to the vertical column (i, j) after applying the thickness threshold.](image)
Figure 3.4(b). The bed boundary indices are then reassigned. Up to this step, the minimum number of MODFLOW layers can be determined.

Step 3: Determine MODFLOW grid

Given a desired number of MODFLOW layers, this study introduces a “ruler” algorithm to assign MODFLOW layer indices to each vertical column. Again, the layer boundaries are required to match the bed boundaries. As shown in Figure 3.5(a), the start and end of the ruler match the top and bottom boundaries of a vertical column, respectively. The number of major ticks in the ruler represents the number of MODFLOW layers. The number of layers up to a bed boundary for a vertical column is obtained by comparing its bed boundary location to the ruler. For example, a bed boundary located between 0 and 1.5 in the ruler indicates one layer up to the bed boundary, between 1.5 and 2.5 indicates two layers up to the bed boundary, between 2.5 and 3.5 indicates three layers up to the bed boundary, and so forth.

Figure 3.5: Illustration of the ruler algorithm: (a) a vertical column with a distinct number of layers up to each bed boundary; and (b) a vertical column where two bed boundaries have the same number of layers.
When the thickness between consecutive bed boundaries is small, the ruler algorithm is likely to assign two or more bed boundaries with the same number of layers up to their bed boundaries as shown in Figure 3.5(b). In this case, the ruler algorithm will adjust the numbers to make sure that each bed boundary has a distinct layer index. In the last step, equal thickness of layers is given to segments that need to be divided into two or more layers based on the final assignment of the layer indices to the bed boundaries.

A MATLAB code is developed to generate MODFLOW grids using the above three steps. The input data to the code are: (1) a three-dimensional hydrostratigraphic architecture, (2) a formation dip, (3) a criterion for eliminating thin sand and thin clay, (4) a thickness threshold for deleting new bed boundaries, and (5) a desired number of layers. The code is able to efficiently generate MODFLOW grids with different vertical resolutions. The code output includes MODFLOW grids, indicators of facies for computational cells, and active/inactive status for computational cells.

3.1.4. Upscale Hydrostratigraphic Architecture to Unstructured Grid

As presented above, the most difficult task in developing a MODFLOW structured grid is to identify indices for continuous layers (assign distinct bed boundary indices for sand and clay segments in each vertical column). In a structured grid, cell indices are identified by continuous row, column, and layer indices. There is no need to define the connectivity between cells because each cell is physically connected to the six surrounding cells.

In an unstructured grid, cell indices are specific and depend on the user’s specification. One needs to define both the number of connections and a list of the connected nodes for each cell. The number of connections in an unstructured grid may vary for each cell. The layers are not necessarily continuous or defined because MODGLOW-USG does not require information about
the specific grid type or reference location [Panday et al., 2013]. The grid information is characterized by the connectivity, connection lengths, connection flow areas, cell areas, and cell volumes. Due to grid flexibility, MODFLOW-USG is able to simulate groundwater flow using a wide range of grid types (e.g., nesting model grids, quadtree or octree grids) to locally refine areas of interest, such as well sites, rivers, etc., and to better represent discontinuous (pinching) hydrofacies and faults compared to MODFLOW. Moreover, by owing a smaller number of cells, the simulation time and computer memory used by MODFLOW-USG are significantly reduced.

An example shown in Figure 3.6 includes a fault and four pumping wells. Figure 3.6(a) shows a non-uniform structured grid of 504 cells, where the minimum cell size is a 200 m by 200 m. Figure 3.6(b) shows a quadtree unstructured grid of 168 cells, which has the same minimum cell size 200 m by 200 m. Both grids have the areas around the wells and the fault refined using the same cell size. Results showed that the unstructured grid used significantly fewer cells than the non-uniform structure grid.

A MODFLOW-USG grid requires that (1) top and bottom cell faces are horizontal, (2) side faces are vertical, and (3) the vertices that define the top and bottom cell faces must have the same x and y locations, such that cells are prismatic in the vertical direction (e.g., triangular, rectangular, cube, pentagonal, or hexagonal prism). This information is provided to MODFLOW-USG.

To construct a MODFLOW-USG unstructured grid, the study first generates a two-dimensional unstructured grid (e.g., a rectangular nested grid, a rectangular quadtree grid, etc.) discussed in [Panday et al., 2013]. The same horizontal discretization is applied to all surfaces. Once a hydrostratigraphic architecture with very fine vertical discretization is generated by indicator kriging, this study upscales the hydrostratigraphic architecture to a MODFLOW-USG grid by the following steps: (1) thin sand and thin clay in each vertical column of the
hydrostratigraphic architecture are eliminated; (2) merge the same hydrofacies in each vertical column to reduce the number of prisms. Prisms of sand and clay cells are determined by the sand and clay segments and the cross-sectional area of the vertical column; and (3) A MODFLOW-USG grid is constructed by assembling all prisms.

Figure 3.6: Illustration for conceptual model development and grid generation for: (a) a non-uniform structured grid with the grid cell size of 200 m by 200 m, and (b) a quadtree unstructured grid with the minimum cell size of 200 m by 200 m.

Constructing an unstructured grid is much simpler than constructing a structured grid from a hydrostratigraphic architecture. An unstructured grid has also an advantage in vertical discretization. An unstructured grid does not need layer continuity throughout the domain and uses coarse discretization to capture better facies geometries in the vertical direction. A MATLAB code is written to prepare the needed information (connectivity, connection lengths, connection flow areas, and cell areas) for unstructured grid generation for MODFLOW-USG.
3.2. High-performance Computing for Groundwater Model Calibration

3.2.1. Groundwater Model Calibration Using CMA-ES

Groundwater model calibration is a process of finding model parameters that give the best match between model outputs and related measured values (e.g., heads, concentrations). The simplest method for model calibration is the trial and error in which one manually modifies input parameters and checks the fitting error after running the simulation model. This method is easy to apply but is time-consuming, especially for models with numerous parameters and long simulation time. Moreover, the trial-and-error method may not guarantee of finding the best solutions because different user’s manipulations may produce dissimilar solutions. Optimization methods such as gradient-based methods and global-search methods are more commonly used due to their ability to handle with a high number of model parameters and the accuracy of solutions. Gradient-based methods fast converge to the solution but may fail to reach a near global solution due to their stagnation at a local minimum solution. Global-search methods are able to identify the near global solution but are computationally intensive. Reviews of methods for model calibration can be found in many books [e.g., Sun, 1994 and Hill and Tiedeman, 2007].

This study adopts the CMA-ES [Hansen and Ostermeier, 2001; Hansen et al., 2003] to calibrate groundwater model and estimate model parameters. CMA-ES is a global-search optimization method and has a capability of avoiding entrapment in a local optimum. In groundwater modeling, CMA-ES has successfully applied to estimate model parameters [Hansen and Ostermeier, 2001; Keating et al., 2010; Tsai and Elshall, 2013a] and optimize groundwater management problems [Bayer et al., 2010]. Moreover, CMA-ES provides a full covariance matrix of estimated parameters, which can later be used to assess head prediction uncertainty. Another
advantage of using CMA-ES is its easiness to be implemented in a core-based high-performance computing system with little overhead.

3.2.2. Parallel CMA-ES for High Performance Computing

For a complex groundwater model that generally takes hours to run, using the CMA-ES [Hansen and Ostermeier, 2001; Hansen et al., 2003] to calibrate the model is impractical due to the prohibitive computational cost. This study resolves the issue by implementing the parallel CMA-ES [Elshall et al., 2015] in a high-performance computing cluster using an embarrassingly parallel master/slave technique.

The parallel implementation of CMA-ES using “embarrassingly parallel” is shown in Figure 3.7. Embarrassingly parallel considers individual solutions as independent tasks. In the first step, the CMA-ES generates solutions at the master node. These solutions are distributed to the slave nodes to run the simulation models and calculate fitting errors accordingly. In the second step, the slave nodes pass the fitting errors to the CMA-ES at the master node to generate new solutions for the next iteration. Step 1 and step 2 are repeated until reaching the stopping criterion. The scheme can be implemented with a simple Bash script without the need of any shared or distributed memory programming languages such as Open MP or MPI. Consequently, the embarrassingly parallel problems are easy to parallel and have the minimal overhead since the individual tasks require no communication of results between tasks.

The parallel CMA-ES has been successfully implemented to both high-performance computing systems of Louisiana State University and Louisiana optical network initiative. To obtain the most efficiency of parallel implementation, the number of processors was always chosen the same as the population size in this study.
Figure 3.7: Embarrassingly parallel scheme of CMA-ES to run on high-performance computing systems.
3.3. Model Discrimination Criteria

One intuitive discrimination criterion is the use of posterior model probability. Posterior model probability describes the importance of a model based on data and its prior probability. A model with the highest posterior probability is generally called the best model among other candidate models. If current data do not support the best model with a superior posterior model probability, the data are not sufficient to discriminate the best model from other models. If a superior posterior model probability of the best model can be found by adding new data, the best model is not only distinguishable from other models, but also identifiable for a system.

This study introduces a discrimination criterion using posterior model probability to identify a model for a system. Let \( M = \{ M_i; i = 1, 2, \ldots, P \} \) be a set of \( m \) candidate models, which are constructed to represent a system. The multiple models are the result of different sources of model structure uncertainty of the system. Posterior model probabilities are denoted as \( \Pr(M_i | \Delta^{\text{obs}}), i = 1, 2, \ldots, P \), where \( \Delta^{\text{obs}} \) are the existing observation data. The sum of the posterior model probabilities is one. The best model is said to be \( \gamma \)-identifiable for the system if its posterior model probability is superior to others and is over a probability threshold \( \gamma \):

\[
\max \left\{ \Pr(M_i | \Delta^{\text{obs}}), i = 1, 2, \ldots, P \right\} \geq \gamma ,
\]

(3)

Given a situation in which the existing observation data cannot achieve this probability threshold, experimental designs are introduced as follows to collect additional data until the \( \gamma \) probability threshold is met.
3.4. Experimental Design using BMA Prediction

If none of the candidate models meets the probability threshold, the existing observation data $\Delta^{\text{obs}}$ is insufficient to identify a highly probable model. An experimental design is conducted to seek potential locations and time to collect future observation data through an observation network such that the $\gamma$-identifiable model can be identified. The experimental design is to maximize the maximum posterior model probability across all the candidate models and is defined as follows:

$$D^* = \arg\max_D \left[ \max \left\{ \Pr( M, i | \Delta^{\text{obs}}, \Delta^{\text{new}}_D), i = 1, 2, \ldots, P \right\} \geq \gamma \right],$$

where $D^*$ is the optimal design to determine potential locations and time of observations that produces the highest posterior model probability; $\Delta^{\text{new}}_D$ is the future observation data at potential observation locations and time using a design $D$; $\Delta = \left\{ \Delta^{\text{obs}}, \Delta^{\text{new}}_D \right\}$ are the total data. The optimal design should have the maximum posterior model probability $\Pr( M, i | \Delta^{\text{obs}}, \Delta^{\text{new}}_D \geq \gamma$.

Since the future observation data $\Delta^{\text{new}}_D$ are unknown, these data may be predicted deterministically or stochastically. It is understood that uncertainties in the future observation data may affect posterior model probability calculations, thereby affecting the experimental design results; however this is not investigated in the current study. In the following sampling approach, the deterministic approach is used and BMA mean predictions are proposed as the future observation data in the experimental design because the BMA mean prediction presents unbiased estimation.

---

¹This section is reproduced with modifications from Pham and Tsai [2015]. It is reprinted by permission from the publisher.
3.4.1. Time-Sequential Sampling Approach

Future observation data at different locations and times provide different information about model predictions and its ability to discriminate models. A time-sequential sampling (TSS) approach is used to sequentially collect spatiotemporal future observation data and update posterior model probabilities over time. To best predict future observation data, $\Delta^{\text{new}}$, using all candidate models, BMA [Draper, 1995; Kass and Raftery, 1995; Raftery et al., 1997; Hoeting et al., 1999] is adopted to obtain mean predictions. Given a number of observation locations, the experimental design with the TSS approach is as follows:

Step 1: For the first round of data collection at time $t_1$, the future observation data presented by the BMA mean predictions are

$$E[\Delta^m_D(t_1) | \Delta^{\text{obs}}] = \sum_{i=1}^{m} E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i, M_i] \Pr(M_i | \Delta^{\text{obs}}_i)$$

where $E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i]$ are the BMA means at time $t_1$, $E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i, M_i]$ are the means of the future observation data predicted by model $i$ at time $t_1$ given the existing observation data $\Delta^{\text{obs}}_i$ and a design $D$, and $\Pr(M_i | \Delta^{\text{obs}}_i)$ is the posterior model probability of $M_i$ calculated using the existing observation data. This study adopts the large sample assumption of Draper [1995], that $E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i, M_i]$ is approximated by $E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i, \hat{\beta}_i]$, where $\hat{\beta}_i$ is the maximum likelihood estimate of model parameters $\beta_i$ of model $M_i$. The total data $\Delta_i = \{\Delta^{\text{obs}}, E[\Delta^m_D(t_1) | \Delta^{\text{obs}}_i]\}$ are used to update the posterior model probabilities $\Pr(M_i | \Delta_i)$, which will be explained in the next section. Either an optimization or an enumeration approach can be used to find the best number and location of observations. If the design $D$ satisfies equation (4), data collection is completed. Otherwise, a second round is needed to collect additional data.
Step 2: For the second round of data collection at time $t_2$, the future observation data presented by the BMA mean predictions are

$$E\left[\Delta_{D}^{\text{new}} (t_2) | \Delta_{\text{obs}}\right] = \sum_{i=1}^{m} E\left[\Delta_{D}^{\text{new}} (t_2) | \Delta_{\text{obs}}, M_i\right] \Pr\left(M_i | \Delta_{\text{obs}}\right),$$

where $E\left[\Delta_{D}^{\text{new}} (t_2) | \Delta_{\text{obs}}, M_i\right]$ are the means of the future observation data predicted by model $i$ at time $t_2$ given existing observation data $\Delta_{\text{obs}}$ and the design $D$, $E\left[\Delta_{D}^{\text{new}} (t_2) | \Delta_{\text{obs}}\right]$ are the BMA mean predictions at time $t_2$. The total data $\Delta_2 = \{\Delta_1, E\left[\Delta_{D}^{\text{new}} (t_2) | \Delta_{\text{obs}}\right]\}$ are used to update the posterior model probabilities $\Pr\left(M_i | \Delta_2\right)$. If equation (4) is met, the design is completed. Otherwise, the next round of data collection is needed.

Step 3: Repeat Step 2 until equation (4) is satisfied. For the $k^{th}$ round of data collection at time $t_k$, the future observation data presented by the BMA mean predictions are

$$E\left[\Delta_{D}^{\text{new}} (t_k) | \Delta_{\text{obs}}\right] = \sum_{i=1}^{m} E\left[\Delta_{D}^{\text{new}} (t_k) | \Delta_{\text{obs}}, M_i\right] \Pr\left(M_i | \Delta_{\text{obs}}\right),$$

where $E\left[\Delta_{D}^{\text{new}} (t_k) | \Delta_{\text{obs}}, M_i\right]$ are the mean predictions of the future observation data predicted by model $i$ at time $t_k$ given the existing observation data $\Delta_{\text{obs}}$ and the design $D$, $E\left[\Delta_{D}^{\text{new}} (t_k) | \Delta_{\text{obs}}\right]$ are the BMA mean predictions at time $t_k$. The total data $\Delta_k = \{\Delta_{k-1}, E\left[\Delta_{D}^{\text{new}} (t_k) | \Delta_{\text{obs}}\right]\}$ are used to update the posterior model probabilities $\Pr\left(M_i | \Delta_k\right)$.

The decision variables of the experimental design are the number and location of future observations and the number of sampling rounds. Different types of future observation data $\Delta_{D}^{\text{new}}$ can be collected from existing and new observation locations. From an economic perspective, existing observation locations should be considered first to design a new observation network.
before a new observation location is introduced. In case the discrimination criterion cannot be achieved, one may need to increase observation period, increase sampling frequency, seek new observation locations, or apply all of these strategies. Moreover, if the future observation data can be collected from the field between sampling rounds, all candidate models should be re-calibrated using both the existing data and the new data. \( \Lambda^{\text{obs}} \) and \( \Pr(M_i | \Lambda^{\text{obs}}) \) should be updated by including the new data. This study does not conduct model re-calibration because the new observation data are not available. The BMA mean predictions cannot be used to re-calibrate the models.

### 3.4.2. Update Posterior Model Probabilities Using BMA Mean Predictions

From Bayes’ theorem, the posterior model probability of a candidate model after the \( k^{\text{th}} \) round of data collection is

\[
\Pr(M_i | \Lambda_k) = \frac{\Pr(\Lambda_k | M_i)\Pr(M_i)}{\sum_{j=1}^{m} \Pr(\Lambda_k | M_j)\Pr(M_j)}, \tag{8}
\]

where \( \Pr(\Lambda_k | M_j) \) is the marginal likelihood and \( \Pr(M_j) \) is the prior model probability. The marginal likelihood is

\[
\Pr(\Lambda_k | M_j) = \int_{\beta} \Pr(\Lambda_k | \beta_j, M_j)\Pr(\beta_j | M_j) d\beta_j, \tag{9}
\]

where \( \beta_j \) is the model parameters of model \( M_j \). Laplace approximation can be used to obtain an estimate of the marginal likelihood [Raftery, 1995] as

\[
\Pr(\Lambda_k | M_j) \approx \exp\left[ -\frac{1}{2} \text{BIC}_{t_k} \right], \tag{10}
\]

where the Bayesian information criterion (BIC) at the \( k^{\text{th}} \) round of data collection at time \( t_k \) is
\[
\text{BIC}_{i,k} = -2 \ln \Pr \left( \Delta_k \mid \hat{\beta}_i, M_i \right) + S_i \ln \left( N_0 + kN_w \right),
\]

where \( N_0 \) is the number of existing observation data; \( k \) is the number of sampling rounds; \( N_w \) is the number of potential observation locations; and \( S_i \) is the number of unknown model parameters to be estimated.

Using equation (10), equal prior model probabilities and the variance window [Tsai and Li, 2008] to account for the effect of data size, the posterior model probability in equation (8) is estimated by

\[
\Pr \left( M_i \mid \Delta_k \right) = \frac{\exp \left\{ -\frac{1}{2} \alpha \left[ \text{BIC}_{i,k} - \left( \text{BIC}_{i,k} \right)_{\text{min}} \right] \right\}}{\sum_{j=1}^{m} \exp \left\{ -\frac{1}{2} \alpha \left[ \text{BIC}_{j,k} - \left( \text{BIC}_{j,k} \right)_{\text{min}} \right] \right\}},
\]

where \( \alpha \) is the scaling factor for the variance window and \( \left( \text{BIC}_{i,k} \right)_{\text{min}} \) is the minimum BIC value of the models. The scaling factor depends on (1) the choice of a BIC difference corresponding to a significance level in Occam’s window and (2) the choice of a variance window size in terms of the standard deviation of a chi-square distribution of the BIC, denoted as \( \sigma_p \). Tsai and Li [2008] provide more details on the statistical meaning and selection of the scaling factor.

Considering \( \Pr \left( \Delta_k \mid \hat{\beta}_i, M_i \right) \) to be a multivariate Gaussian distribution and assuming independent observation data for \( \Delta_k \), which includes the existing data and BMA mean prediction data, equation (11) becomes

\[
\text{BIC}_{i,k} = Q_{i,k} + \left( N_0 + kN_w \right) \ln(2\pi) + S_i \ln \left( N_0 + kN_w \right),
\]

where \( Q_{i,k} \) is the sum of weighted squared errors. \( Q_{i,k} \) is
where $\Delta_j^{\text{obs}}$ and $\Delta_j^{\text{cal}}(\hat{\beta}_i)$ are the $j^{th}$ observed value and calculated value using model $i$, respectively. $\Delta_n^{\text{cal}}(t_m; \hat{\beta}_i)$ is the prediction at location $n$ by model $i$ at the $m^{th}$ round of data collection; and $E[\Delta_{D,n}^{\text{new}}(t_m)|\Delta^{\text{obs}}]$ is the BMA mean prediction at location $n$ and the $m^{th}$ round of data collection given the existing observation data $\Delta^{\text{obs}}$. The first term at the right side of equation (14) is the sum of weighted squared errors in the historical data space. The second term is the sum of weighted squared errors for $k$ rounds of data collection in the prediction space. The variances $\sigma_j^2$ and $\sigma_{m,n}^2$ are calculated by Monte Carlo simulation, which samples a sufficient number of realizations of the model parameters using the following equations:

$$
\sigma_j^2 = \frac{1}{P \times R} \sum_{i=1}^{P} \sum_{r=1}^{R} (\Delta_j^{\text{cal}}(\beta_i^{(r)}) - \Delta_j^{\text{obs}})^2 ,
$$

$$
\sigma_{m,n}^2 = \frac{1}{P \times R} \sum_{i=1}^{P} \sum_{r=1}^{R} (\Delta_n^{\text{cal}}(t_m; \beta_i^{(r)}) - E[\Delta_{D,n}^{\text{new}}(t_m)|\Delta^{\text{obs}}])^2 ,
$$

where $\beta_i^{(r)}$ is the $r^{th}$ realization of model parameters for model $i$, and $R$ is the total number of realizations.

### 3.4.3. Proposition Discrimination and Uncertainty Analysis

This study adopts the HBMA method of Tsai and Elshall [2013] to analyze proposition discrimination and prediction uncertainty reduction given the experimental design. The HBMA is different from the BMA in terms of providing more insights on model selection, model averaging and uncertainty propagation through a BMA tree. The HBMA is able to evaluate competing propositions of each uncertainty model component, to prioritize different sources of uncertainty in
a hierarchical order, and to understand the uncertainty propagation. Each source of uncertainty
develops a level of the hierarchy and suggests a number of propositions to represent an uncertain
model component. Combinations of propositions of all levels create candidate models
\( \mathbf{M} = \{ \mathbf{M}_i; i = 1, 2, ..., P \} \), which are the base models at the base level of the hierarchy. A number of
BMA models are developed bottom up from the base level to the top-most level by averaging
models under the same propositions of their upper level. The top-most level is called the hierarch
level. The BMA model at the hierarchy level is called the hierarch model. For operational purposes,
this study defines the base level of the hierarchy as level B. The level number decreases as the
level goes from the base level B to the hierarch level. \( \mathbf{M}_B \) are the base models at level B. \( \mathbf{M}_{B-1} \)
are the BMA models at level B-1, and so forth.

After the \( k^{th} \) round of data collection, the posterior model probabilities of the base models
are \( \text{Pr}(\mathbf{M}_B | \Delta_k) \), which are \( \text{Pr}(\mathbf{M}_i | \Delta_k), i = 1, 2, ..., P \) as discussed previously. The posterior
model probabilities at level B-1 are the sums of the posterior model probabilities at level B under
the same propositions of level B-1. That is
\[
\text{Pr}(\mathbf{M}_{B-1} | \Delta_k) = \sum_i \text{Pr}(\mathbf{M}_{B}^{(i)} | \Delta_k),
\]
(17)
where \( \mathbf{M}_{B}^{(i)} \) are the models at level B under the same propositions of level B-1. Therefore, the
posterior model probabilities at any level \( l \) is
\[
\text{Pr}(\mathbf{M}_l | \Delta_k) = \sum_i \text{Pr}(\mathbf{M}_l^{(i)} | \Delta_k).
\]
(18)
The conditional posterior model probabilities at level \( l+1 \) under a model at level \( l \) is
\[
\text{Pr}(\mathbf{M}_{l+1}^{(i)} | \Delta_k, \mathbf{M}_l) = \frac{\text{Pr}(\mathbf{M}_{l+1}^{(i)} | \Delta_k)}{\text{Pr}(\mathbf{M}_l | \Delta_k)}.
\]
(19)
Equations (18) and (19) develop a BMA tree of posterior model probabilities and conditional posterior model probabilities. The $\gamma$-identifiable propositions can be assessed for all uncertain model components by using equations (18) and (19) during the experimental design. Total prediction variances are evaluated by BMA at any level $l$ by

$$\text{Var}(\Delta | \Delta_k, M_i) = \sum_i \text{Var}(\Delta | \Delta_k, M_{i+1}) \Pr(M_{i+1} | \Delta_k, M_i)$$

$$+ \sum_i \left[ \text{E}(\Delta | \Delta_k, M_{i+1}) - \text{E}(\Delta | \Delta_k, M_i) \right]^2 \Pr(M_{i+1} | \Delta_k, M_i),$$

where $\text{Var}(\Delta | \Delta_k, M_i)$ are the total BMA prediction variances of $\Delta$ when models $M_i$ are used. The first term at the right side of equation (20) is the within-model variances of $\Delta$ using models at level $l + 1$ and the second term is the between-model variances of $\Delta$ using models at level $l + 1$. Calculations of both variances are referred to Li and Tsai [2009] and Tsai and Elshall [2013]. Variance reduction can be assessed by comparing the total variances before and after the experimental design.

Prioritization of the sources of model structure uncertainty that form the order of uncertain model components in the BMA tree is subject to the analyst’s preference. However, prioritization will not affect the optimal experimental design since this study maximizes the maximum posterior model probability across all the base models, not the BMA models. Equation (18) shows the addition of the posterior model probabilities from the base level to the hierarch level. Once the optimal experimental design achieves the desired discrimination criterion, $\gamma$-identifiable propositions for all uncertain model components are identified at the same time, which compose the best model.
3.5. Optimal Observation Network Design Using Information Theory and BMA

3.5.1. Shannon Entropy and Discrimination Function

This study adopts the discrimination function of Box and Hill [1967] based on the expected decrease in Shannon’s entropy to develop an optimal observation network design for model discrimination purposes. Information theory, as defined by Shannon [1948], provides measures of how much information is produced by a set of candidate models, \( M \). According to Shannon [1948], the entropy of a system, \( H(M) \), is defined as

\[
H(M) = -\sum_{i=1}^{p} \Pr(M_i | \Delta^{obs}) \ln \left[ \Pr(M_i | \Delta^{obs}) \right]
\]

where \( \ln \left[ \Pr(M_i | \Delta^{obs}) \right] \) is the information of model \( M_i \). Negative entropy (-H) is the average amount of information provided by all candidate models. The least information corresponds to the maximum entropy when all models have equal posterior model probability. The maximum information gain from the system corresponds to the minimum entropy when one model has 100% posterior model probability and other models have zero posterior model probability. Therefore, to obtain the maximum information from an optimal observation network design, it needs to maximize the expected entropy change before and after the design, which is the entropy before the design minus the expected entropy after the design as follows:

\[
X = -\sum_{i=1}^{p} \Pr(M_i | \Delta^{obs}) \ln \left[ \Pr(M_i | \Delta^{obs}) \right] - (-1) \int \sum_{i=1}^{p} \Pr(M_i | \Delta^{new}_D) \ln \left[ \Pr(M_i | \Delta^{new}_D) \right] q\left(\Delta^{new}_D\right) d\Delta^{new}_D
\]

where \( \Delta^{new}_D \) are the new observation data, \( \Pr(M_i | \Delta^{new}_D) \) is the posterior model probability given the new data, and \( q\left(\Delta^{new}_D\right) \) is the expected marginal likelihood, which is
\[
q(\Delta_D^{new}) = \sum_{i=1}^{p} \Pr(\Delta_D^{new} | M_i) \Pr(M_i | \Delta^{obs})
\]

(23)

where \( \Pr(\Delta_D^{new} | M_i) \) is the marginal likelihood for model \( M_i \). Therefore, the expected entropy change before and after the observation network design is

\[
X = \sum_{i=1}^{p} \Pr(M_i | \Delta^{obs}) \int \Pr(\Delta_D^{new} | M_i) \ln \frac{\Pr(\Delta_D^{new} | M_i)}{q(\Delta_D^{new})} d\Delta_D^{new}.
\]

(24)

Instead of maximizing the expected entropy change \( X \), Box and Hill [1967] found that \( X \) is a lower bound of a discrimination function. This study follows Box and Hill’s approach to maximize the discrimination function value for the optimal observation network design. Appendix A develops a multi-data discrimination function to account for multiple spatial-temporal independent data from an observation network design. The discrimination function is

\[
F = \frac{1}{2} \sum_{i=1}^{p} \sum_{j=i+1}^{p} \Pr(M_i | \Delta^{obs}) \Pr(M_j | \Delta^{obs})
\]

\[
\sum_{n=1}^{N} \left[ \left( \frac{\sigma_{n,j}^2 - \sigma_{n,i}^2}{\sigma_n^2 + \sigma_{n,i}^2} \right)^2 + \left( \Delta_{n,i} - \Delta_{n,j} \right)^2 \left( \frac{1}{\sigma_n^2 + \sigma_{n,i}^2} + \frac{1}{\sigma_n^2 + \sigma_{n,j}^2} \right) \right]
\]

(25)

where \( N \) is the total number of future observation data collected by an observation network; \( \Delta_i = \{\Delta_{n,j}\} \) and \( \Delta_j = \{\Delta_{n,j}\} \) are the future observation data predicted by models \( M_i \) and \( M_j \), respectively; \( \sigma_{i}^2 = \{\sigma_{n,j}^2\} \) and \( \sigma_{j}^2 = \{\sigma_{n,j}^2\} \) are the variances of predicted future observation data by models \( M_i \) and \( M_j \), respectively; and \( \sigma^2 = \{\sigma_n^2\} \) are the variances of future observation data. The optimal observation network design is to collect new observation data in order to maximize the discrimination function value as follows

\[
\max F
\]

(26)

until the following \( \gamma \) probability threshold

37
max \{ \Pr \left( M_i \mid \Delta_{\text{obs}}^i, \Delta_{D}^{\text{new}} \right), i = 1, 2, \ldots, P \} \geq \gamma \quad (27)

is met. The design involves the posterior model probabilities given existing observation data as well as new data. In the following sections, the methods to calculate \( \Pr \left( M_i \mid \Delta_{\text{obs}}^i \right) \) and \( \Pr \left( M_i \mid \Delta_{\text{obs}}^i, \Delta_D^{\text{new}} \right) \) are discussed.

### 3.5.2. \( \Pr \left( M_i \mid \Delta_{\text{obs}}^i \right) \) Calculation

From Bayes’ theorem, the posterior model probability of model \( M_i \) given existing observation data is

\[
\Pr \left( M_i \mid \Delta_{\text{obs}}^i \right) = \frac{\Pr \left( \Delta_{\text{obs}}^i \mid M_i \right) \Pr \left( M_i \right)}{\sum_{j=1}^{m} \Pr \left( \Delta_{\text{obs}}^i \mid M_j \right) \Pr \left( M_j \right)} \quad (28)
\]

where \( \Pr \left( \Delta_{\text{obs}}^i \mid M_i \right) \) is the marginal likelihood function and \( \Pr \left( M_i \right) \) is the prior model probability for model \( M_i \). Following the derivations in the Appendix (A4)-(A7) for existing observation data \( \Delta_{\text{obs}}^i \), the marginal likelihood function is

\[
\Pr \left( \Delta_{\text{obs}}^i \mid M_i \right) = (2\pi)^{N_0 / 2} \left| \Sigma_i^{-1} \right|^{-1} \exp \left[ -\frac{1}{2} \left( \Delta_{\text{obs}}^i - \Delta_i' \right)^T \Sigma_i^{-1} \left( \Delta_{\text{obs}}^i - \Delta_i' \right) \right] \quad (29)
\]

where \( N_0 \) is the number of existing observation data; \( \Sigma_i'^{-1} = \sigma_i'^2 + \sigma_i'^2 \) is the diagonal covariance matrix; \( \sigma_i'^2 \) is the diagonal covariance matrix of the existing observation data; and \( \Delta_i' \) and \( \sigma_i'^2 \) are the simulated quantities and their variances by model \( M_i \) at the existing data domain. Given \( \Delta_{\text{obs}}^i \), \( \Delta_i' \), \( \sigma_i'^2 \) and \( \sigma_i'^2 \), the marginal likelihood value \( \Pr \left( \Delta_{\text{obs}}^i \mid M_i \right) \) can be calculated.
3.5.3. Pr\left(M_i \mid \Delta_{D}^{\text{obs}}, \Delta_{D}^{\text{new}}\right) Calculation

Similarly, from Bayes’ theorem, the posterior model probability of model \( M_i \) can be updated when new observation data are available

\[
\Pr\left(M_i \mid \Delta_{D}^{\text{obs}}, \Delta_{D}^{\text{new}}\right) = \frac{\Pr\left(\Delta_{D}^{\text{new}} \mid M_i, \Delta_{D}^{\text{obs}}\right) \Pr\left(M_i \mid \Delta_{D}^{\text{obs}}\right)}{\sum_{j=1}^{m} \Pr\left(\Delta_{D}^{\text{new}} \mid M_j, \Delta_{D}^{\text{obs}}\right) \Pr\left(M_j \mid \Delta_{D}^{\text{obs}}\right)}
\]  

(30)

where \( \Pr\left(M_i \mid \Delta_{D}^{\text{obs}}\right) \) becomes the prior model probabilities. The marginal likelihood function \( \Pr\left(\Delta_{D}^{\text{new}} \mid M_i, \Delta_{D}^{\text{obs}}\right) \) is the same as \( \Pr\left(\Delta_{D}^{\text{new}} \mid M_i\right) \) in (A7):

\[
\Pr\left(\Delta_{D}^{\text{new}} \mid M_i, \Delta_{D}^{\text{obs}}\right) = \left(2\pi\right)^{-\frac{N}{2}} \left|\Sigma_i\right|^{-1} \exp\left[-\frac{1}{2} \left(\Delta_{D}^{\text{new}} - \Delta_i\right)^T \Sigma_i^{-2} \left(\Delta_{D}^{\text{new}} - \Delta_i\right)\right]
\]

(31)

where \( \Sigma_i^2 = \sigma_i^2 + \sigma_i^2 \). Given \( \Delta_{D}^{\text{new}}, \Delta_i, \sigma_i^2 \), and \( \sigma_i^2 \), the marginal likelihood value \( \Pr\left(\Delta_{D}^{\text{new}} \mid M_i, \Delta_{D}^{\text{obs}}\right) \) can be calculated.

3.5.4. \( \Delta_i', \Delta_i, \sigma_i^2 \), and \( \sigma_i^2 \) Calculations

This section discusses the techniques needed to conduct predictions and estimate variances of predicted quantities at the existing data domain and the future data domain. The maximum likelihood approach in Draper [1995] was adopted. In this approach, the simulated quantities \( \Delta_i' \) by model \( M_i \) at the existing data domain are approximated by \( \Delta_i'(\hat{\beta}_i) \) where \( \hat{\beta}_i \) is the maximum likelihood estimate of model parameters \( \beta_i \) of model \( M_i \). The predictions \( \Delta_i \) by model \( M_i \) at the future data domain are approximated by \( \Delta_i(\hat{\beta}_i) \).

Methods to quantify uncertainties due to input data and parameter uncertainties using a single conceptual model have been well developed in the literature [Dettinger and Wilson, 1981;
The commonly used methods are Monte Carlo approaches [Hill, 1989; Kuczera and Parent, 1998; Bates and Campbell, 2001; Marshall et al., 2004] and the first and second moment methods [Dettinger and Wilson, 1981; Helsel and Hirsch, 2002; Kunstmann et al., 2002; Glasgow et al., 2003; Stauffer et al., 2004]. This study adopts the Monte Carlo simulation approach as follows to estimate variances of simulated quantities at the existing data domain due to parameter estimation errors:

\[
\sigma^2_i = \frac{1}{R-1} \sum_{r=1}^{R} \left[ \Delta_i(\beta_{i}^{(r)}) - \Delta_i(\hat{\beta}_i) \right]^2
\]

where R is the total number of realizations of parameters \( \beta_{i}^{(r)} \) of model \( M_i \). Through the post-analysis, the statistics of the model parameters \( \beta_i \) can be obtained after model calibration [Sun, 1994]. This study will adopt the CMA-ES [Hansen and Ostermeier, 2001; Hansen et al., 2003] to calibrate models and uses the statistics of estimated model parameters from the CMA-ES to conduct Monte Carlo simulation [Elshall et al. 2014]. The Monte Carlo simulation approach is also used to estimate prediction variances \( \sigma^2 \) by model \( M_i \) for the future data domain.

3.5.5. Future Data Predictions by BMA for \( \Delta^\text{new}_D \)

Since future observation data \( \Delta^\text{new}_D \) are not known, future observation data predictions and their uncertainty in terms of \( \sigma^2 \) would be best estimated using predictions of all candidate models. Two multimodel approaches are commonly used to predict future observation data and their uncertainty: generalized likelihood uncertainty estimation [Beven and Binley, 1992; Beven and Freer, 2001] and Bayesian model averaging [Neuman, 2003; Ye et al., 2004, 2008; Poeter and Anderson, 2005; Tsai and Li, 2008; Li and Tsai, 2009; Tsai and Elshall, 2013b; Chitsazan et al.,
This study proposes the BMA mean predictions to represent future observation data as follows [Draper, 1995; Hoeting et al., 1999]

\[
\bar{\Delta}_{D}^{\text{new}} = \sum_{i=1}^{P} \Delta_{i}(\hat{\beta}_{i}) \Pr(M_{i} | \Delta^{\text{obs}}).
\] (33)

BMA employs probabilistic techniques to derive consensus predictions from a set of candidate models based on the existing observation data and their corresponding posterior model probabilities. Averaged predictions from BMA are less biased than predictions obtained from individual models [Raftery and Zheng, 2003; Ye et al., 2004; Ajami et al., 2006], and therefore, will be the best representative for future observation data. By substituting \( \Delta_{D}^{\text{new}} \) with \( \bar{\Delta}_{D}^{\text{new}} \) in equation (30), the posterior model probability can be evaluated as \( \Pr(M_{i} | \Delta^{\text{obs}}, \bar{\Delta}_{D}^{\text{new}}) \).

The uncertainty of using the BMA mean \( \bar{\Delta}_{D}^{\text{new}} \) in equation (33) to represent future observation data in calculating \( \Pr(M_{i} | \Delta^{\text{obs}}, \bar{\Delta}_{D}^{\text{new}}) \) could have an impact on the results of an optimal observation network design. Theoretically, it is possible to consider an exhaustive set of future observation data predictions by Monte Carlo simulation on model parameters \( \beta_{i} \) for all candidate models. However, this is not possible in practice since the optimal observation network design in equations (26)-(27) has already been computationally challenging. Instead of including the posterior model probability uncertainty in the design problem, this study conducts a post analysis using Monte Carlo simulation on \( \Delta_{i}(\beta_{i}) \) to quantify the uncertainty in \( \Pr(M_{i} | \Delta^{\text{obs}}, \bar{\Delta}_{D}^{\text{new}}(\beta_{i})) \) with respect to \( \Pr(M_{i} | \Delta^{\text{obs}}, \bar{\Delta}_{D}^{\text{new}}(\hat{\beta}_{i})) \) after an optimal observation network design is achieved.
3.5.6. $\sigma^2$ Estimation of by BMA

In the literature, future observation data are often considered homoscedastic, where variances of future observations $\sigma^2$ are the same in the discrimination function [Box and Hill, 1967; Yarirevich et al., 2013]. However, in reality, future observation data are heteroscedastic. At different locations and times, future observation data have different variances. To calculate the variances of future observation data that include model parameter uncertainty and model structure uncertainty in the optimal observation network design, the BMA method was adopted as follows [Hoeting et al. 1999; Draper 1995]

$$\sigma^2 = \sum_{i=1}^{p} \sigma_i^2 \Pr(M_i | \Delta_{obs}) + \sum_{i=1}^{p} \left[ \Delta_i(\hat{\beta}_i) - \bar{\Delta}_D^{new} \right]^2 \Pr(M_i | \Delta_{obs})$$  \hspace{1cm} (34)

The first term in the right hand side of equation (34) is the within model variance that is due to the uncertainty in model parameters. The second term in the right hand side of equation (34) is the between-model variance that comes from model structure errors (e.g., the uncertainty in geological architectures, boundary conditions, etc.).
4. Construction of Structured and Unstructured MODFLOW Grids from Well Logs

4.1. Hydrostratigraphic Architecture Construction Using Well Logs

The study analyzed wireline well logs from 583 boreholes shown in Figure 4.1 to construct a hydrostratigraphic architecture for the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand and the “2,000-foot” sand of the Baton Rouge aquifer system. Well log interpretation is given in Figure 4.2(a). The number of sand and clay segments in each well log ranges from 3 to 59. It is impractical to manually build correlations between boreholes using the solid approach.

Figure 4.1: Map of the well logs in the UTM (m) coordinate system. The box is the study area.
Figure 4.2: Construction of hydrostratigraphic architecture from well logs: (a) Distribution of boreholes and results of well log interpretations, and (b) hydrostratigraphic architecture.
The indicator kriging was adopted to construct a hydrostratigraphic architecture with the dip angle of the Baton Rouge aquifers is 0.29 degrees and a cutoff value is 0.40 [Elshall et al. 2013]. The resulted hydrostratigraphic architecture is illustrated in Figure 4.2(b). The model domain in the planar direction is discretized into 93 rows and 137 columns with a cell size 200 by 200 m, resulting in 12,741 cells. In the vertical direction, the aquifer system is discretized by every one foot (0.304 m) from elevation -1210 ft (-368.8 m) to -2500 ft (-762.0 m) below the National Geodetic Vertical Datum of 1929 (NGVD29). The elevation range covers the four sand aquifers. As a result, the hydrostratigraphic architecture has 12,741 vertical columns and 1,290 layers.

4.2. Results and Discussion

4.2.1. Eliminate Thin Sand and Thin Clay

Before eliminating thin sand and thin clay, the number of sand and clay segments in each vertical column is shown in Figure 4.3(a). The number of segments in vertical columns ranges from 4 to 53, indicating at least 53 layers needed for constructing a MODFLOW grid. The minimum segment thickness is 2 ft. (0.61 m) and the maximum segment thickness is 936 ft. (285 m). The total number of sand and clay segments in the model domain is 265,287, of which 83,756 segments (31.6%) are less than 10 ft. thick.

By eliminating sand and clay segments less than 10 ft. (3.05 m) thick, the total number of sand and clay segments is reduced to 158,707, a 59.8% reduction. Using 10-foot thickness as a criterion to eliminate thin sand and thin clay is subjective. Guidance was given by a USGS report [Rollo, 1969] that marks areas where aquifers are less than 20 feet thick. Figure 4.3(b) shows the distribution of the number of segments after eliminating thin sand and thin clay. The number of sand and clay segments in each vertical column ranges from 3 to 26. The minimum segment thickness is 10 ft. (3.05 m) and the maximum segment thickness is 984 ft. (300 m).
4.2.2. Project Neighboring Bed Boundaries

The second step is to preserve continuity of layers by projecting the bed boundaries of four adjacent vertical columns to their respective column. After the bed boundary projection, the number of segments in vertical columns increases, ranging from 5 to 76. Figure 4.3(c) shows the distribution of sand and clay segments of vertical columns after the projection. The minimum segment thickness is 1 ft. (0.305 m) and the maximum segment thickness is 921 ft. (281 m). The total number of sand and clay segments increases to 299,724.

Given a minimum thickness threshold of 15 ft. to eliminate new bed boundaries, Figure 4.3(d) shows the distribution of sand and clay segments of vertical columns. The number of sand and clay segments in vertical columns ranges from 5 to 42. The minimum segment thickness is 10
ft. (3.05 m) and the maximum segment thickness is 928 ft. (283 m). The total number of sand and clay segments decreases to 199,039. As a result of Figure 4.3(c) and Figure 4.3(d), the number of the MODFLOW layers should be between 42 and 76. This study found that projection step increases the number of sand and clay segments in vertical columns, especially in the pinch-out areas and the areas adjacent to the faults. Increasing sand and clay segments in these areas is important to preserve hydraulic connections in these areas.

4.2.3. Structured Grid

This study used the constructed hydrostratigraphic architecture from the previous section to generate a MODFLOW structured grid. Using the developed method, a grid of 968,316 cells given by 93 rows, 137 columns and 76 layers shown in Figure 4.4(a) accurately matches the complex hydrostratigraphic architecture and preserves layer continuity. Each cell is 200 m by 200 m with cell thickness ranging from 3.05 m to 13.4 m. The average thickness of the layers is 5.2 meters. The “1,200-foot”, the “1,500-foot”, and the “1,700-foot” sands are from layer 6 to layer 46. The “2,000-foot” sand is from layer 47 to layer 76.

Figure 4.4(b) shows a closer look at the generated grid for two cross sections AA’ shown in Figure 4.1. The complexity of the Baton Rouge aquifer system shows unconformed sand and clay sequences, isolated sands, discontinuity, especially along the east-west direction, varying thicknesses, complex interconnections, pinch-outs and geological faults. Results show that the method not only accurately converts the hydrostratigraphic architecture into a MODFLOW grid, but also preserves all narrow connections including those around the faults.

Different complex three-dimensional MODFLOW grids similar to Figure 4.4 can be regenerated using: (1) different criteria to eliminate thin sand and thin clay, (2) different thickness thresholds to delete new bed boundaries, and (3) different number of layers. Since
hydrostratigraphic architectures usually carry facies indices (e.g., “1” for the sand facies and “0” for the clay facies in this study), the facies properties are also automatically assigned to new grids.

Figure 4.4: Three-dimensional MODFLOW structured grid for the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand: (a) Boreholes and the MODFLOW structured grid, and (b) structured grid at the selected cross sections in Figure 4.1. Clay is blanked. The vertical exaggeration is 20 times.
4.2.4. Unstructured Grid

This study used the developed method to generate an MODFLOW-USG unstructured grid from 583 well logs shown in Figure 4.1. First, a two-dimensional unstructured grid shown in Figure 4.5 is generated. The base grid has the cell size of 800 m by 800 m. The quadtree is adopted to refine the areas around 87 pumping wells and two faults by decreasing cell size to 400 m and to 200 m. The total number of cells in Figure 4.5 is 2,201. The next step is to apply indicator kriging to estimate sand and clay facies in each vertical column at 1-foot interval. Afterwards, thin sand and thin clay in each vertical column are eliminated.

![Figure 4.5: Two-dimensional quadtree unstructured grid for the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand.](image)

The MODFLOW-USG unstructured grid shown in Figure 4.6 uses 27,264 cells and accurately matches the complex hydrostratigraphic architecture. In the horizontal direction, the unstructured grid is coarser than the structure grid at some areas but along the area of interest such as pumping wells and faults, the cell sizes are similar. The number of cells in the unstructured grid
is significantly smaller than 968,316 needed for the structured grid. The cell thickness ranges from 3.04 m (10 feet) to 282 m (926 feet). The minimum number of connections to a cell is two that is the same as in structured grid. The maximum number of connections to a cell is 28 that is much larger than six for the structured grid.

Figure 4.6: MODFLOW-USG unstructured grid: (a) boreholes and the MODFLOW-USG quadtree unstructured grid, and (b) unstructured grid at the selected cross sections in Figure 4.1. Clay is blanked. The vertical exaggeration is 20 times.
In the vertical direction, the unstructured grid is much coarser than the structured grid. As shown in Figure 3.5(b) and Figure 3.6(b). In this cross section, the number of cells in each vertical column of the unstructured grid varies from 7 to 14 that is significant smaller than 76 cells of the structured grid. In both horizontal and vertical direction, the unstructured grid is coarser than the structured grid but the unstructured grid is still able to accurately capture facies geometries of the complex hydrostratigraphic architecture.

4.3. Conclusions

This study develops a technique to generate structured and unstructured grids for MODFLOW and MODFLOW-USG, respectively. Using regional geological dip to correlate a large number of well logs results in better facies architecture and reflects the depositional environment. The technique begins with a simple, but important step to construct hydrostratigraphic architecture with fine vertical discretization such that complex facies geometries (e.g., pinch-outs, coalescence, and displacement at faults) can be delineated. Then, the upscaling approach efficiently upscales the complex hydrostratigraphic architecture to a simulation grid for flow and transport simulations without loss of accuracy in facies geometries.

The grid generation technique successfully generates MODFLOW grids for the sequence of the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand in the Baton Rouge area. The grid generation technique is able to reveal the narrow hydraulic connections through the BR fault and the DSS fault. To maintain the hydrofacies geometries in the structured grid, 76 model layers are needed to fit bed boundaries of complex hydrostratigraphic architectures. The unstructured grid with a quadtree approach significantly reduces the number of computational cells in comparison with the structured grid.
The MODFLOW grids can be re-generated automatically using different criteria for eliminating thin sand and thin clay, thickness thresholds for eliminating new bed boundaries, and number of layers. Moreover, the MODFLOW grid can be easily updated when new well log data become available. Inadequate interpretation of hydrostratigraphic architecture often leads to an erroneous groundwater model structure, which results in unreliable predictions of flow and transport. Model structure errors can be significantly reduced by incorporating as many well log data as possible to preserve available geological information in the groundwater model.

The methods in Section 3.1 and Section 3.2 were applied to generate computational grids and calibrate groundwater flow models for the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand in the Baton Rouge area, southeastern Louisiana. As the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand between the two faults are interconnected, these sands should be modeled together. The “2,000-foot” sand is separated from the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand by a thick clay layer and will not be considered in the model. Three model propositions were proposed for uncertain geological architecture, boundary conditions and the fault architecture, respectively. As a result, 18 conceptual models were developed and will be calibrated in this section. It is noted that three geological architectures from Elshall et al. [2013] were used to generate MODFLOW grids. These architectures were constructed using three different indicator geostatistical approaches: indicator zonation, generalized parameterization, and indicator kriging.

5.1. Model Development

This study uses the USGS MODFLOW [Harbaugh, 2005] to develop groundwater models for the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand of the Baton Rouge aquifer system. These sands are extended from −877 ft. (−267.3 m) to −1967 ft. (−599.5 m) below the National Geodetic Vertical Datum of 1929. Given a hydrostratigraphic architecture, the study area was discretized into 93 rows, 137 columns and 45 layers as shown in Figure 5.1. The cell size is 200 m by 200 m. The layer thicknesses varies from 3.0 m to 12.8 m. The simulation period is from
1/1/1975 to 12/31/2015 and is divided into calibration and prediction periods. The calibration period is from 1/1/1975 to 1/1/2010, and the prediction period is from 1/1/2010 to 12/31/2015.

Figure 5.1: A three-dimensional MODFLOW grid for the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand. The generalized parameterization method was used to generate the hydrostratigraphic architecture. Clay cells are blanked. The vertical exaggeration is 20 times.

The study area includes 87 pumping wells and the connector well (EB-1293) (Figure 5.2). Monthly pumping data are available from the Louisiana capital area groundwater conservation commission from 1975 to 2010. The accumulative pumpage of these pumping wells from 1/1/1975 to 12/31/2010 is shown in Figure 5.2. The average pumping rate of these wells was 112,556 m$^3$/day (29.73 MGD) of which the average pumping rate from the Government St. pumping wells in the same period was 7,570 m$^3$/day (1.98 MGD). The connector well (EB-1293) is injecting about 2,600 m$^3$/day (0.69 MGD) of groundwater from the “800-foot” sand to the “1,500-foot” sand.
Figure 5.2: Distribution of the cumulative pumpage in the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand from 1975 to 2010.

The MODFLOW well package (WEL) was used to simulate these pumping wells and injection wells. The time-dependent boundary conditions were assigned to all active cells at the boundaries of the model area using the time-variant specified head (CHD) package. The BR fault and DSS fault were discretized and simulated using the horizontal flow barrier (HFB) package. The discretization of these model components are illustrated in Figure 5.3. It is noted that the permeability of these faults is characterized by the hydraulic characteristic, which is hydraulic conductivity per unit width of the fault [Hsieh and Freckleton, 1993].

The initial head on 1/1/1975 and the time-varied head boundary condition were estimated by the measured heads in the USGS observation wells. Detailed approach of determining the initial and boundary conditions for the study area can be found in Tsai and Li [2008]. In this study, 2756 groundwater head data measured from 1975 - 2010 from 20 USGS wells (Figure 5.2) were
collected and will be used for model calibration. The estimating model parameters are hydraulic conductivity [m/d], specific storage [1/m], and fault hydraulic characteristic [1/d].

Figure 5.3: Illustration of model components for MODFLOW model.

5.2. Sources of Model Structure Uncertainty

Model structure uncertainty has been recognized by many authors to be the main source of prediction uncertainty [Usunoff et al., 1992; Beven and Freer, 2001; Neuman and Wierenga, 2003; Refsgaard et al., 2006b]. This study considers three sources of uncertainty in developing conceptual groundwater models. The first uncertainty source is from the geological architecture which was the most important source of uncertainty in many studies [e.g., Harrar et al., 2003; Højberg and Refsgaard, 2005; Troldborg et al., 2007; Refsgaard et al., 2012; Chitsazan et al., 2014]. Different conceptualizations and characterization approaches may result in hydrostratigraphic architectures that are different in terms of sand and clay distribution, their interconnections, and sand unit displacement on the faults. As a result, the groundwater models
developed based on these architectures may lead to the different flow pathways in within each sand unit and across the faults.

To take into account the geological structure uncertainty in groundwater modeling, this study employed three hydrostratigraphic architectures from Elshall et al. [2013]. These geological architecture models were constructed from 583 electrical well logs using: (1) indicator kriging (IK) that provides smooth sand/clay interfaces; (2) indicator zonation (IZ) that divides the model domain into non-overlapping zonation structures; and (3) generalized parameterization (GP) that combines IZ method and IK method. The complexity of the aquifer structures using different methods is shown in Figure 5.4, where the IZ, GP, and IK models have the sand proportions of 36.05 %, 39.24 %, and 39.85 % with respect to the entire model domain, respectively.

![Figure 5.4: Sand distributions at layer 36 using (a) IZ, (b) GP and (c) IK methods, showing uncertainty in geological architecture construction.](image-url)
The second uncertainty source is from boundary condition conceptualization [Dean Oliver and Christakos, 1996; Feyen and Gorelick, 2004]. Due to limited groundwater level data, this study considers three sets of boundary values south of the BR fault. For the first set, denoted as B₀, the head boundary values for the “1,200-foot” sand south of the BR fault were determined 4.5 m higher than those measured at the USGS observation well EB-782A. The head boundary values for the “1,500-foot” sand and “1,700-foot” sand south of the BR fault were determined 1.5 m higher than those measured at the USGS observation well EB-780A. Another two alternatives that adjust B₀ by ±1.5 m are proposed and denoted as B⁻1.5 and B¹⁺1.5, respectively.

The third uncertainty source is from fault permeability architecture conceptualization. It has been recognized that geological faults are highly anisotropic and heterogeneous structures [e.g., Evans et al., 1997; Bense et al., 2003; Bense and Person, 2006; Boult et al., 2010, p.2; Ran et al., 2014] and strongly influence groundwater flow [Maslia and Prowell, 1990; Bredehoeft et al., 1992; Bense et al., 2003; Anderson and Bakker, 2008]. This study proposes two alternatives for the fault permeability architectures. The first alternative, denoted as HC², considers different homogeneous fault hydraulic characteristics for the BR fault and for the DSS fault. A fault hydraulic characteristic is defined as hydraulic conductivity per unit width of the fault. The second alternative, denoted as HC⁴, considers the fault hydraulic characteristic at the “1,200-foot” sand is different from that at the “1,500-foot” sand and “1,700-foot” sand.

Combinations of the alternatives from the three uncertainty sources result in 18 conceptual groundwater models. These models are named by the notations of the geological architectures, the boundary conditions and the fault permeability architectures (e.g., GP-B⁻¹.5-HC²).
5.3. Model Calibration Using the Parallel CMA-ES

The parallel computation of CMA-ES was carried using SuperMike-II, a supercomputer at Louisiana State University with 440 compute nodes and a peak performance of 146 TFlops (trillion floating-point operations per second). Each compute node is equipped with two 8-Core processors operating at a core frequency of 2.6 GHz. The performance of a parallel algorithm can be evaluated based on parallelization speedup that is the ratio of sequential execution time over parallel execution time and will be discussed as follows:

5.3.1. Optimal Population Size and Speedup of Parallel Runs

For this case study, the execution time for a single model simulation is around $0.93 \pm 0.31$ hours. Given minimal parallelization overhead, the speedup of the parallel CMA-ES is roughly equal to population size $\lambda$. For example, given $\lambda = 80$ and a stopping criterion 1.6 m fitting error, the parallel CMA-ES reached the stopping criterion in 37 iterations in about 34.4 hours. The sequential CMA-ES execution time would have been roughly $80 \text{ solutions} \times 34.4 \text{ hours} = 2,752$ hours. Note that calculated sequential execution time is slightly overestimated since the iteration time for parallel CMA-ES is the maximum of the running times of all the solutions in the iteration. No sequential runs are tested for this case study and the following results and discussion are for parallel runs.

Figure 5.5 demonstrates the speedup of the parallel CMA-ES with increasing the population size $\lambda$, which is equal to the number of processors. The optimal population size for the models is determined by performing calibration runs with different population sizes $\lambda = 16, 32, 48, 64$ and 80. For the best performance, Hansen and Ostermeier [2001] and Hansen et al. [2003] recommended $4 + \left\lceil 3\ln(n) \right\rceil \leq \lambda \leq 10n$. Thus, this study selected $\lambda = 80$ as the maximum
population size given \( n = 8 \). To maximize the efficiency of the parallel implementation, the number of processors is equal to the population size \( \lambda \).

![Graph](image)

**Figure 5.5:** The number of iterations for different population sizes required to reach several target fitting errors.

The number of iterations required to meet different stopping criteria of RMSE versus the number of processors is illustrated in Figure 5.5. More iterations are needed for smaller RMSE criteria. For example, using 32 processors, reaching RMSE of 1.57 m needs 80 iterations meanwhile reaching the smaller RMSE of 1.44 m needs 114 iterations. Small RMSE criteria may not be met using small population sizes as the search becomes less global. For example, the population size \( \lambda = 16 \) did not reach the RMSE 1.50 m criterion. Increasing the population size will always reduce the number of iterations. Thus, the optimal population size is \( \lambda = 10n = 80 \).

The optimal population size is at the upper limit \( \lambda = 10n \) given by Hansen et al. [2003]. As shown in Figure 5.5, given RMSE 1.57 m, speedup 2.05 is achieved by the optimal population size \( \lambda = 80 \) with respect to the default population size \( \lambda = 16 \).
Therefore, to take the advantage of the parallel CMA-ES, this study will use 80 processors for model calibration. This parallel CMA-ES is implemented to SuperMike-II, a high-performance computing system at Louisiana State University.

5.3.2. Model Calibration and Estimated Parameters

To calibrate a groundwater model and estimate model parameters, the parallel CMA-ES [Elshall et al., 2015] was adopted to minimize the root mean square error (RMSE) between the calculated and observed groundwater heads:

\[
\text{RMSE}_i = \left\{ \frac{1}{N_0 - 1} \sum_{j=1}^{N_0} \left[ h_j^{\text{cal}}(\beta_i) - h_j^{\text{obs}} \right]^2 \right\}^{1/2},
\]

where \( h_j^{\text{cal}}(\beta_i) \) is the calculated groundwater head \( j \) given model parameters \( \beta_i \), \( h_j^{\text{obs}} \) is the observed groundwater head \( j \) at a USGS well, and \( N_0 \) is the number of existing observation heads. In this study, \( N_0 = 2756 \) groundwater head data measured from 1975 to 2010 from 20 USGS wells in Figure 5.2.

The RMSE and estimated model parameters presented in Figure 5.6 and Table 5-1. The average number of iterations to each to reach these RMSEs is 77 iterations. The best model is GP-B_{1.5}-HC_{4} with RMSE = 1.47 m. However, the best model is not dominant from the second best model (IK-B_{+1.5}-HC_{4}, RMSE = 1.49 m) and the third best model (GP-B_{0}-HC_{4}, RMSE = 1.51 m) based on the RMSE values.

Results from the three best models showed that the estimated hydraulic conductivity of the “1,500-foot” sand and the “1,700-foot” sand is slightly higher than the “1,200-foot” sand. The BR fault and the DSS fault are found to be low-permeability faults that restrict horizontal flow. At the “1,200-foot” sand, the HC values of the two faults are in the same order of magnitude in which the BR fault is less permeable than the DSS fault. At the “1,500-foot” sand and the “1,700-foot” sand,
the estimated HC values of the Baton Rouge fault are over two orders of magnitude smaller than those of the Denham Springs-Scotlandville fault.

![Image](image_url)

**Figure 5.6:** A comparison of RMSE of 18 candidate groundwater models.

**Table 5-1:** RMSE and the estimated parameter values of 18 candidate models

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE (m)</th>
<th>Specific Storage (1/m)</th>
<th>Hydraulic Conductivity (m/d)</th>
<th>Hydraulic characteristic (1/m) of the DSS Fault</th>
<th>Hydraulic characteristic (1/m) of the BR Fault</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&quot;1,200-Foot&quot; Sand</td>
<td>&quot;1,500-1,700-Foot&quot; Sands</td>
<td>&quot;1,200-Foot&quot; Sand</td>
<td>&quot;1,500-1,700-Foot&quot; Sand</td>
</tr>
<tr>
<td>GP-B_{1.5}HC_4</td>
<td>1.47</td>
<td>1.55E-05</td>
<td>5.45E-06</td>
<td>21.16</td>
<td>0.0092</td>
</tr>
<tr>
<td>IK-B_{1.5}HC_4</td>
<td>1.49</td>
<td>4.45E-06</td>
<td>4.37E-06</td>
<td>18.29</td>
<td>0.0155</td>
</tr>
<tr>
<td>GP-B_{0}HC_4</td>
<td>1.51</td>
<td>1.69E-05</td>
<td>7.14E-06</td>
<td>20.23</td>
<td>0.0093</td>
</tr>
<tr>
<td>IK-B_{1.5}HC_3</td>
<td>1.53</td>
<td>2.93E-05</td>
<td>2.00E-06</td>
<td>17.23</td>
<td>0.0323</td>
</tr>
<tr>
<td>IZ-B_{1.5}HC_4</td>
<td>1.6</td>
<td>7.73E-06</td>
<td>2.65E-06</td>
<td>34.17</td>
<td>0.0081</td>
</tr>
<tr>
<td>IZ-B_{0}HC_4</td>
<td>1.62</td>
<td>7.17E-06</td>
<td>2.70E-06</td>
<td>32.10</td>
<td>0.0110</td>
</tr>
<tr>
<td>GP-B_{1.5}HC_2</td>
<td>1.64</td>
<td>2.38E-05</td>
<td>6.97E-06</td>
<td>17.23</td>
<td>0.0408</td>
</tr>
<tr>
<td>IK-B_{0}HC_2</td>
<td>1.64</td>
<td>1.01E-05</td>
<td>2.25E-06</td>
<td>16.69</td>
<td>0.0458</td>
</tr>
<tr>
<td>IK-B_{1.5}HC_2</td>
<td>1.64</td>
<td>2.78E-05</td>
<td>4.33E-06</td>
<td>16.43</td>
<td>0.0492</td>
</tr>
<tr>
<td>IZ-B_{1.5}HC_2</td>
<td>1.67</td>
<td>2.26E-05</td>
<td>2.46E-06</td>
<td>32.35</td>
<td>0.0118</td>
</tr>
<tr>
<td>GP-B_{0.5}HC_2</td>
<td>1.7</td>
<td>5.66E-06</td>
<td>3.87E-06</td>
<td>17.21</td>
<td>0.0369</td>
</tr>
<tr>
<td>IK-B_{1.5}HC_2</td>
<td>1.7</td>
<td>1.59E-05</td>
<td>5.48E-06</td>
<td>17.26</td>
<td>0.0443</td>
</tr>
<tr>
<td>IZ-B_{1.5}HC_2</td>
<td>1.72</td>
<td>8.24E-06</td>
<td>2.63E-06</td>
<td>27.37</td>
<td>0.0390</td>
</tr>
<tr>
<td>GP-B_{1.5}HC_2</td>
<td>1.78</td>
<td>6.36E-06</td>
<td>2.04E-06</td>
<td>16.53</td>
<td>0.0417</td>
</tr>
<tr>
<td>GP-B_{1.5}HC_4</td>
<td>1.78</td>
<td>5.45E-06</td>
<td>6.76E-06</td>
<td>23.13</td>
<td>0.0057</td>
</tr>
<tr>
<td>IZ-B_{0}HC_2</td>
<td>1.78</td>
<td>1.61E-05</td>
<td>2.04E-06</td>
<td>27.36</td>
<td>0.0380</td>
</tr>
<tr>
<td>IK-B_{0}HC_4</td>
<td>1.91</td>
<td>7.32E-05</td>
<td>4.66E-06</td>
<td>19.45</td>
<td>0.0127</td>
</tr>
<tr>
<td>IZ-B_{1.5}HC_2</td>
<td>1.97</td>
<td>1.18E-05</td>
<td>4.16E-06</td>
<td>27.19</td>
<td>0.0319</td>
</tr>
</tbody>
</table>
5.4. Results and Discussion

5.4.1. Simulated Heads and Groundwater Flow Fields

The simulated heads and one-standard deviation bounds of the best model, the second best model, the third best model, and other 15 models are shown Figure 5.7 and Figure 5.8. Only 12 USGS observation wells that have long observation head records are presented. The simulated heads of each model are calculated by running that model using the optimal parameters presented in Table 5-1. One-standard deviation bounds of simulated heads of each model are determined by using Monte Carlo simulation in which the mean and the covariance matrix of the estimated model parameters were determined after model calibration using the parallel CMA-ES.

We noted that conducting Monte Carlo simulations for this real case study is extremely computational burden. For example, using a MC simulation of 320 samples, it needs in average of $320 \times 0.93 \times 18 = 5357$ hours, using a personal computer. To reduce the MC simulation time, this study implemented “embarrassingly parallel” to the SuperMike-II supercomputer housed at LSU. Using 320 processors, the MC simulation is reduced to $0.93 \times 18 = 16.7$ hours.

The results showed that the simulated heads using the three best models are relatively similar and show good agreement to the observed heads. The best model and the second best model have small one-standard deviation bounds representing insignificant uncertainty in parameter estimations. These two models are equally good to represent the Baton Rouge aquifer system and can be used for future predictions and saltwater intrusion simulations. Other 16 candidate models can be eliminated as their model outputs were unable to fit well with the observation data.
Figure 5.7: Observed heads, simulated heads and one-standard deviation bounds at selected USGS observation wells in the “1,200-foot” sand.
Figure 5.8: Observed heads, simulated heads and one-standard deviation bounds at selected USGS observation wells in the “1,500-1,700-foot” sands.
We used the best model GP-B-1.5-HC4 to simulate groundwater heads at two largest pumping areas, which are the Industrial District pumping wells screened in the “1,200-foot” sand and the Lula wells and Government St. wells screened in the “1,500-1,700-foot” sands. The results showed significant groundwater level declines from 1975 to 2010 in both pumping areas as illustrated in Figure 5.9 and Figure 5.10. Given a sand unit, the water levels are very different in the horizontal direction but are relatively similar in the vertical direction. Also, the results showed different water levels between the “1,200-foot” sand and the “1,500-1,700-foot” sand in 2010.

Figure 5.9: Distribution of simulated groundwater heads for the Industrial District given by the best model GP-B-1.5-HC4 on (a) 1/1/1975; and (b) 12/31/2010.
Figure 5.10: Distribution of simulated groundwater heads for the Lula wells and Government St. wells pumping areas given by the best model GP-B-1.5-HC4 on (a) 1/1/1975; and (b) 12/31/2010.

Due to heavy pumping activities, two cones of depression have developed in the study area as shown in Figure 5.11. The first cone of depression is in the “1,200-foot” sand, at the Industrial District area as illustrated in Figure 5.11(a). The second cone of depression is in the “1,500-1,700-foot” sands, in the area close to the Lula Street pumping station and the Government Street pumping station as shown in Figure 5.11(b). These are important pumping stations that are providing industrial and municipal water supply for the Baton Rouge area.
The declination of groundwater level north of the Baton Rouge has changed naturally north-south flow direction. The groundwater flow now moves northward, across the BR fault, and potentially causing saltwater intrusion toward pumping stations.

Figure 5.11: Top view of simulated heads given by the best model on 12/31/2010 for (a) the “1,200-foot” sand; and (b) the “1,500-1,700-foot” sands.
5.4.2. Flow Budgets

This study used simulation results in December 2010 of the best model GP-B.1.5-HC4 to analyze the water budget of the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand in between two faults. The results showed that total inflow is 117,806 m$^3$/d of which flow from north of the DSS fault is 47,542 (40.4%); flow from east of the study area is 47,557 m$^3$/d (40.3%); flow from south of the BR fault is 16,726 (14.2%); and flow from the injection well is 2,589 (2.2%) as shown in Figure 5.12(a). Most of the outflow is due to the pumpage as illustrated in Figure 5.12(b). The pumping rate is 112,556 m$^3$/d that accounts for 95.5% of total outflow.

Figure 5.12: Groundwater budget for the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand in between the two faults in December 2010: (a) inflow; and (b) outflow.

The results also revealed strong groundwater flow interactions between the "1,200-foot" sand and the "1,500-1700- foot" sands in between two faults. The model estimated downward flow rate 22,278 m$^3$/day from the “1,200-foot” sand to the “1,500-1,700-foot” sands. This accounted for 38.7% of total inflow to the “1,500-1,700-foot” sands as shown in Figure 5.13(a). It estimated upward flow rate 7,896 m$^3$/day from the “1,500-1,700-foot” sands to the “1,200-foot” sand. This accounted for 13.7% of total inflow to the “1,200-foot” sand as shown in Figure 5.13(b). Moreover, the model estimated northward flow rate 8,851 m$^3$/day passing the BR fault to the “1,500-1,700-foot” sands. This accounted for 15.4% of total inflow to the “1,500-1,700-foot” sands. Most of the
outflow of the “1,500-1,700-foot” sands is due to the pumpage as shown in Figure 5.12(b). The pumping rate is 48,399 m$^3$/d that accounts for 84.2% of total outflow.

Figure 5.13: Groundwater budget for the “1,500-1,700-foot” sands in between the two faults in December 2010: (a) inflow; and (b) outflow.

5.4.3. Head Differences across the Two Faults

Figure 5.14(a) presents the vertical distribution of groundwater heads in a North-South cross-section on 12/31/2010. To calculate head differences across the two faults, this study used simulated heads from eight locations that close to the BR fault and the DSS fault. Those locations are marked as black dots Figure 5.14(a). The results of head differences across the two faults are demonstrated in Figure 5.14(b). This study found high head differences across the BR fault, and the head differences across the BR fault are much larger than that across the DSS fault. Across the BR fault, the head differences in 2010 were estimated 18.3 m and 27 m in the “1,200-foot” sand and in the “1,500-1,700-foot” sands, respectively. Across the DSS fault, the head differences in 2010 were estimated 5.83 m and 8.1 m in the “1,200-foot” sand and in the “1,500-1,700-foot” sands, correspondingly. In addition, simulated results showed increasing trends of the head differences across the BR fault and the DSS fault in the “1,500-1,700-foot” sands. The head difference across the BR fault increases from 19.8 m on 1/1/1975 to 26.2 m on 12/31/2010. Across the DSS fault, the head difference increases from 3.34 m to 7.62 m in the same period.
Figure 5.14: Head differences across the two faults: (a) a cross-section showing vertical distribution of groundwater heads at the end of model calibration. The locations with black dots were used to calculate the head differences; and (b) simulated head differences across the BR fault and the DSS fault.
5.4.4. Groundwater Head Predictions

Using the optimal estimated parameters, prediction models were executed to simulate groundwater heads at 19 USGS observation wells from 1/1/2011 to 12/31/2015. The assumptions for the prediction models are: (1) the future groundwater heads for boundaries are determined based on the trend of boundary heads over the last three years in the model calibration period (2008 to 2010); and (2) the future pumping rates are equal to the average monthly pumping rates of the last 5 years (2006 – 2010).

Figure 5.15 presents the predicted heads and one-standard deviation bounds using the best model, the second best model and other models. The prediction results showed dramatically water level declines for the first year of prediction (2011) at all USGS wells. After that, the water levels are likely to reach the steady state as the pumping rates are kept as constant. In addition, head predictions given by the best model and the second best model are comparatively different at the USGS observation wells. As the best model is not clearly superior to other candidate models represented by its agreement to the observation data, and predictions using the best model and the second best model are relatively different, their mean predictions are preferentially used to avoid biased predictions from using a single conceptual model [Li and Tsai, 2009; Chitsazan et al., 2014]. However, considering too many models may result in high head prediction uncertainty and may lose the purpose of model development [Bredhoeft, 2005; Højberg and Refsgaard, 2005]. In such situations, model discrimination and identification should be conducted to identify the best model or at least to reduce the number of models by collecting additional data through experimental designs.
Figure 5.15: Predicted heads and one-standard deviation bounds of the best model, the second best model and other 16 groundwater models at 19 USGS observation wells.
5.5. Conclusions

This study developed a three-dimensional groundwater flow model to better understand the impact of groundwater withdrawals that caused groundwater level decline and saltwater water intrusion in the “1,200-foot” sand, “1,500-foot” sand, and “1,700-foot” sand of the Baton Rouge aquifer system, Louisiana. A high vertically resolution grid was introduced to handle complex facies geometries of Baton Rouge fluvial aquifer system, including two geological faults. To accelerate the model calibration, the parallel CMA-ES has successfully implemented and applied to calibrate 18 groundwater models.

The parallel CMA-ES was found helpful in model calibration as it significantly reduces model calibration time. Moreover, the covariance matrix of the estimated parameters are valuable in analyzing parametric estimation uncertainty and head prediction uncertainty. The calibration result identified the best model represented by its lowest RMSE, but the best model is not much different from the second best and the third best model. Additional data are needed in order to reduce the number of competing models and in turn, identify the best model for saltwater intrusion prediction. The calibration results also reveal that the BR fault, and the DSS fault are low-permeability faults that restrict horizontal groundwater flow.

Using the best model for analyses, the study found strong groundwater flow interactions between the "1,200-foot" sand and the "1,500-1700- foot" sands for the area between two faults. Heavy pumpage has caused declination of water level north of the BR Fault, resulted in two cones of depression and high head differences across the BR fault and the DSS fault. The flow now moves northward, across the BR fault, causing saltwater intrusion toward pumping stations.
6. Bayesian Experimental Design for Identification of Model Propositions and Uncertainty Reduction

The methodology is tested using a real case study in the Baton Rouge area, southeastern Louisiana. This study developed 18 candidate groundwater models to simulate groundwater flow in the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand due to the uncertainty in model conceptualization. These models are different in boundary conditions, geological architectures, and fault permeability architectures as discussed in Section 5.2. The USGS observation wells were used to illustrate the experimental design as shown in Figure 6.1.

Figure 6.1: Map of the study area in the UTM (m) coordinate system, which includes location of the pumping wells, the USGS observation wells, the BR fault, and the DSS fault.

6.1. Model Calibration and Posterior Model Probabilities

To demonstrate the methodology, this study only used N=539 groundwater head values from seven USGS wells, EB-327, EB-413, EB-652, EB-782B, EB-807A, EB-996 and WBR-102A

---

2 This chapter is reproduced with modifications from Pham and Tsai [2015]. It is reprinted by permission from the publisher.
shown in Figure 6.1. The results of model calibration and posterior model probabilities are shown in Table 6-1. The RMSE of the 18 groundwater models ranges from 1.06 m to 1.71 m. The mean RMSE is 1.18 m and the standard deviation of the RMSE is 0.15 m. Using the estimated parameters and the full covariance matrix obtained from the CMA-ES for each model, realizations of the parameters are generated by Monte Carlo simulation and input into groundwater models to generate groundwater heads. Then, the error variances of groundwater heads are calculated using equation (9). Finally, the $Q_p$ values, the BIC values, the $\Delta$BIC values and the posterior model probabilities using a scaling factor $\alpha = 0.046$ were obtained and listed in Table 6-1.

Table 6-1: RMSE, weighted sum of squared errors, BIC values, and posterior model probabilities for the 18 groundwater models using variance window with 5% significance levels and $4\sigma_D$.

| Model         | RMSE (m) | $Q$   | BIC  | $\Delta$BIC | $\text{Pr}(M_i | \Delta_{obs})$ (%) |
|---------------|----------|-------|------|-------------|-------------------------------|
| GP-B-$1.5$-HC$_4$ | 1.19 | 316.3 | 1357 | 0           | 17.5                          |
| GP-B-$1.5$-HC$_2$  | 1.1   | 330.7 | 1359 | 2           | 16.8                          |
| GP-B$_0$-HC$_2$    | 1.11  | 334.1 | 1362 | 5           | 15.6                          |
| IK-B-$1.5$-HC$_4$  | 1.11  | 341.6 | 1383 | 25          | 9.8                           |
| GP-B$_0$-HC$_4$    | 1.34  | 348.9 | 1390 | 33          | 8.3                           |
| IK-B$_0$-HC$_2$    | 1.71  | 367.3 | 1396 | 38          | 7.3                           |
| IZ-B$_0$-HC$_2$    | 1.09  | 378.2 | 1407 | 49          | 5.7                           |
| IK-B+$1.5$-HC$_4$  | 1.19  | 366.8 | 1408 | 50          | 5.5                           |
| IZ-B+$1.5$-HC$_2$  | 1.06  | 382.5 | 1411 | 54          | 5.1                           |
| GP-B+$1.5$-HC$_2$  | 1.07  | 395.1 | 1423 | 66          | 3.9                           |
| IK-B$_0$-HC$_4$    | 1.14  | 435.5 | 1476 | 119         | 1.2                           |
| IK-B+$1.5$-HC$_2$  | 1.1   | 445.2 | 1474 | 116         | 1.2                           |
| IZ-B+$1.5$-HC$_4$  | 1.22  | 451.6 | 1492 | 135         | 0.8                           |
| IK-B-$1.5$-HC$_2$  | 1.06  | 494.9 | 1523 | 166         | 0.4                           |
| IZ-B$_0$-HC$_4$    | 1.12  | 479   | 1520 | 163         | 0.4                           |
| IZ-B+$1.5$-HC$_2$  | 1.23  | 511   | 1539 | 182         | 0.3                           |
| IZ-B-$1.5$-HC$_4$  | 1.17  | 504.9 | 1546 | 189         | 0.2                           |
| GP-B+$1.5$-HC$_4$  | 1.28  | 854.9 | 1896 | 539         | 0                             |
It is noted that the scaling factor is calculated based on the variance window of Tsai and Li [2008] using 5% significance level and $4\sigma_D$ window size. According to the posterior model probabilities, the top three groundwater models (GP-B_{1.5}-HC_4, GP-B_{1.5}-HC_2 and GP-B_0-HC_2) are deemed equally important, presented by comparable posterior model probabilities of 16.8%, 17.5%, and 15.6%, respectively. Six models have posterior model probability larger than 5% and have an influence on BMA model prediction and uncertainty.

Figure 6.2 shows the BMA tree of the posterior model probabilities and conditional posterior model probabilities for the three uncertain model components. The geological architecture obtained by the GP has the posterior model probability 62.0%, which implies a better geological architecture proposition than two other geological architecture propositions. The B_{1.5} proposition has the posterior model probability 49.9% that is better than two other boundary propositions. The HC_2 fault permeability architecture has the posterior model probability 55.6%, which is better than the HC_4 fault permeability architecture. Nevertheless, no proposition dominates others in each source of model structure uncertainty.

Figure 6.2: The BMA tree of posterior model probabilities and conditional posterior model probabilities before the observation network design. Posterior model probabilities and conditional posterior model probabilities less than 1% are not shown in the figure.
Having so many influential groundwater models indicates that the groundwater head data used in model calibration may not be sufficient. In what follows, the proposed experimental design is implemented on the USGS water wells to see if a single best model with high posterior model probability can be achieved.

6.2. Experimental Design Results

To illustrate the experimental design, this study only considers measurement locations belong to the USGS groundwater observation network that planned to monitor groundwater level declinations and saltwater intrusions in the Baton Rouge area. The 20 USGS water wells in Figure 6.1 are potential locations to be used in the experimental design.

The TSS approach was applied to the USGS observation wells shown in Figure 6.1. In the experimental design, monthly groundwater head data from 2011 to 2015 are gradually collected to see if the 80%-identifiable groundwater model can be determined, i.e., $\gamma = 80\%$. WBR-102A was not considered in the design since the location is too close to the west boundary. Following restrictions were made to the design: (1) observation wells are selected at the beginning of the observation, which is the first day of 2011; and (2) observation wells once selected remain active until the probability threshold is met or the maximum observation period is reached.

The BMA is applied to obtain the averaged head values of the 18 groundwater models, which will be used as “future observation” in experimental design. First, the prediction models were executed using the optimal estimated parameters to predict groundwater heads at 19 USGS observation wells. The assumptions for model prediction are: (1) the groundwater heads at four boundaries in the next 5 years (2011 to 2015) have the same trend as groundwater heads at the last three years in the model calibration period (2008 to 2010); and (2) the monthly pumping rates equal to the average pumping rates of the last 3 years (2008 – 2010). Then, equation (7) was used
to calculated groundwater heads at selected USGS water wells of a given experimental design. The posterior model probability for each groundwater model was recalculated using equation (12). It is noted that the purpose of the experimental design is to guide decision makers the potential observation wells to collect data before actual data are collected. Therefore, model parameters will remain the same in the design phase. Models should be re-calibrated once new real data become available, which is not the scope of this study.

There are 524,287 possible experimental design solutions given 19 USGS wells. Using an optimization approach may not guarantee to find the best design. Instead, this study developed a code to implement the “embarrassing parallel” algorithm to the SuperMike-II supercomputer to enumerate all possible design solutions. The total possible design solutions were spilt equally to 320 processors to be evaluated. Evaluation of one design took 82 seconds. Therefore, the computation time was reduced from 1.36 years (estimated runtime by using one processor) to 1.55 days (actual runtime by using 320 processors). Using the parallel computing, the result from the enumeration approach guarantees the goal of reaching the optimal solution.

6.2.1. Evolving Posterior Model Probabilities by Adding New Data

Figure 6.3(a) shows the evolution of the posterior model probabilities if the seven USGS observation wells used in model calibration continues to be used to collect monthly groundwater head data from 2011 to 2015. By the end of 2015, there are six groundwater models with posterior model probability larger than 1%. Three groundwater models have influential posterior model probabilities. The best model (GP-B.1.5-HC2) has posterior model probability 45%, which is not dominant. The second best model (GP-B.1.5-HC4) and the third-best model (IK-B0-HC2) have posterior model probability 33.6% and 14.1%, respectively. Using the seven USGS wells cannot meet the design criterion.
Figure 6.3: Change of posterior model probabilities of 18 groundwater models after 60 months of continuous groundwater head observation: (a) using the eight USGS observation wells used in model calibration; (b) using one USGS observation well at different locations; and (c) using a different number of best USGS observation wells.

To understand the influence of individual observation wells, Figure 6.3(b) shows the comparison of changes of posterior model probabilities if using only one USGS well to collect monthly groundwater head data from 2011 to 2015. All individual USGS wells support GP-B-1.5-HC2 model to be the best model except for EB-780A and EB-789B that support GP-B-1.5-HC4 model to be the best model. EB-782A, EB-168 and EB-807A show high posterior model probability differences that are larger than 10% between the best model and the second best model.
WBR-101, EB-996, EB-652, EB-946, EB-291, EB-327, and EB-301 show low posterior model probability differences that are less than 2% between the best model and the second best model. It is for sure that using one USGS well cannot meet the design criterion.

Given a number of observation wells, the enumeration method selects the best USGS water wells that maximize the posterior model probability of the best model. Figure 6.3(c) presents the result of posterior model probabilities after 60-month groundwater head observation. All design solutions support GP-B-1.5-HC_2 model to be the best groundwater model. Using only one USGS observation well, the best observation well location is at EB-782A, which is located close to the BR fault. The best model (GP-B-1.5-HC_2) has posterior model probability of 35.8%. The second best model (IK-B_0-HC_2) has posterior model probability 16.2%. Using two observation wells, EB-782A and EB-168 are the best two well locations that increase the highest posterior model probability to 44.7%. The second highest posterior model probability decreases to 15.9%. Figure 6.3(c) shows that it needs at least six observation wells in order to achieve the 80%-identifiable groundwater model in five years of groundwater head observation. These six USGS wells are EB-146, EB-413, EB-652, EB-789B, EB-946 and EB-996. In other words, using less than 6 USGS wells needs more than five years of observation or never meets the design criterion.

6.2.2. Non-unique Experimental Design Solutions

The least time to reach the design criterion 80% of posterior model probability is shown in Figure 6.4 given the best design solutions under different number of observation wells. Fourteen experimental designs satisfy the design criterion within 60 months. The least number of wells to meet the design criterion is 6. The shortest time to meet the design criterion is 41 months by using 13 USGS wells. It is interesting to see that the least observation time will increase if 18 or 19 USGS wells are used.
Figure 6.4: The least months to reach the design criterion of 80% posterior model probability given a number of best selected USGS observation wells.

The frequency of the USGS wells selected for the 14 experimental design solutions that meet the design criterion is shown in Figure 6.5. The most popular well is EB-168, which is selected 13 times, following by EB-146. The least selected well is EB-327, which has only six times.

Figure 6.5: The frequency of the USGS observation wells to be selected for the best observation network designs given the criterion of 80% posterior model probability.
6.2.3. Using Least Number of USGS Wells (6 wells)

The temporal change of posterior model probabilities using the best 6 USGS wells is shown in Figure 6.6. After first 12 months of observation, the posterior model probability for the best model GP-B-1.5-HC4 before the design is increased from 17.5% to 26.2%. However, the posterior model probability of GP-B-1.5-HC2 model is increased from 16.8% to 31.8% and GP-B-1.5-HC2 model becomes the best model. At the end of 60-month observation, GP-B-1.5-HC2 model has the posterior model probability 80.8%, which meets the design criterion. The second best model GP-B-1.5-HC4 has the posterior model probability 7.3%.

Figure 6.6: Change of posterior model probabilities of 18 groundwater models over time based on the best selected 6 USGS observation wells, EB-146, EB-413, EB-652, EB-789B, EB-946, and EB-996.

The BMA tree of the updated posterior model probabilities and conditional posterior model probabilities is shown in Figure 6.7, using the best 6 USGS wells for observation. The 80%-identifiable propositions for all levels are achieved. The result implies that the geological
architecture obtained by the GP, which has conditional posterior model probability 89.9%, should be kept. Two other geological architectures may be discarded. Under the GP geological architecture proposition, the B_{1.5} proposition has conditional posterior model probability 96%. Two other boundary propositions may be discarded. Under the GP geological architecture proposition and the B_{1.5} proposition, the HC_{2} fault permeability architecture has conditional posterior model probability 87.7%. Therefore, the HC_{4} fault permeability architecture can be discarded. As discussed previously, the posterior model probability of the best model is 80.8% shown at the base level, which is much greater than 7.3% of the second best groundwater model.

Figure 6.7: The BMA tree of posterior model probabilities and conditional posterior model probabilities in December 2015 after the observation network design with 6 best selected USGS observation wells. Posterior model probabilities and conditional posterior model probabilities less than 1% are not shown in the figure.

The variance reduction of predicted groundwater heads at the 19 USGS wells at the end of design period (12/31/2015) using the 6-well experimental design are shown in Figure 6.8 for the B_{1.5} model, the GP model and the hierarch model. Both between-model variances and within-model variances are significantly reduced by discriminating unimportant propositions through the experimental design.
Figure 6.8: Comparisons of the total variances of predicted groundwater heads at 19 USGS observation wells on 12/31/2015 without and with the observation network design of 6 best selected USGS observation wells after 60 months of observation.
6.2.4. Discussions

The fundamental idea of the experimental design is to collect data in order to reduce conceptual model uncertainty. Usually, observation data are not commonly available on many observation times. For example, in our study area, most of the USGS wells have less than four head observation data in a year because it is expensive to collect the head data for deep aquifers like ours. There are only a few wells being sampled frequently. Minimizing the number of observation time is one logical way to constraint the budget and to achieve the goal of the experimental design within a reasonable time. This study realizes that many factors will affect the design results. Believing that the methodology is valid, real future data should be collected according to the recommendations in order to carefully assess the design results. New real data should be used to re-calibrate the models and re-do the experimental design again. However, this post-analysis is beyond the scope of this study because (1) USGS do not have monthly groundwater head data for the selected observation wells. For most of the wells, the USGS only measure groundwater level one or two times per year and (2) actual pumping data for 2011-2014 are not available.

Head data (as well as other types of data, e.g. fluxes) at different locations can have varying impact in discriminating models. This study only considers groundwater head data in the illustrative example since groundwater head is the most achievable/economical data in the study area to be used to discriminate among groundwater models. In addition, the new observations are being treated as uncorrelated. In fact, these are some spatial correlation among the different locations that may cause issue of correlated head observations. However, the full covariance matrix of head observations is very tedious to obtain in the experimental design. This study assumed that data collected from different wells and from different sampling rounds are independent.
The predictive uncertainty in the BMA mean prediction will definitely affect the posterior model probability calculations and experimental design results. The future data $\Delta^{new}_D$ in equation (2) are not deterministic and have a PDF. The posterior model probability for each model should have a PDF. To make the problem more tractable, this study used the deterministic approach and considered future data $\Delta^\text{new}_D$ to be determined by the BMA mean prediction. Monte Carlo simulation approach can be used to assess the impacts of this type of uncertainty on the design results. However, this approach needs a much extensive computational resource than what have been incurred in this study, and was not investigated in this study.

The improvement of predictive performance depends on future information in the simulation models. For example, this study forecasted the pumping data and boundary head values for 2011-2015 as these data are unavailable. Therefore, the individual model predictions and the BMA mean prediction for 2011-2015 may not be close to the future real data because of incorrect future information that constructs the forecast models. However, this issue is unrelated to the proposed methodology. Adding new data will in some extent reduce uncertainty since the new data always bring some additional information about the system. However, what types of uncertainty and how effective the uncertainty can be reduced are really dependent on what types of data to collect and how you use the data, which is beyond the scope of this work. Our study is to collect spatiotemporal groundwater head data to reduce conceptual model uncertainty.

The hierarchical structure is subjected to the analyst's preference and prioritization of the sources of uncertainty. If one is more interested in one specific source of uncertainty than others, that specific source of uncertainty should be placed at the first level. However, the optimal design will be the same regardless the order of the sources of uncertainty because this study maximizes the posterior probability across all the models (base models), not the BMA models at the first level.
The results showed that the best model based on the existing data is not necessarily the best model after the experimental design using the BMA prediction data. The future data collected from different locations and sampling rounds will decide which model has highest potential to be the best model based on their posterior model probabilities. There is no guarantee which model will be the best model based on existing data without the experimental design.

6.3. Conclusions

Using posterior model probability as a discrimination criterion is a straightforward approach to identify a unique and highly probable model from a pool of candidate models. Since representing model importance based on data evidence, the posterior model probability is a direct quantitative indicator for the experimental design to guide data collection through an observation network. The time-sequential sampling approach presents a practical strategy for data collection, by which future data are gradually collected and used along with historical data to achieve the unique and highly probable model.

The HBMA method is found to be useful to illustrate the results of the experimental design for proposition discrimination and prediction uncertainty reduction. First, using the BMA mean prediction as future observation data to guide the time sequential sampling data collection strategy is an unbiased approach since all candidate models contribute to prediction according to their up-to-date posterior model probabilities. Second, once the design goal of achieving the \( \gamma \)-identifiable model is met, the unique and highly probable proposition for each uncertain model component is identified through the BMA tree of posterior model probabilities. By reducing posterior model probabilities of unimportant propositions, the prediction variance owing to different sources of uncertainty is significantly reduced.
The experimental design illustrated by the real groundwater case study is practical since this study only uses the USGS observation wells. Implementing the enumeration approach to a multi-core supercomputer is viable and assures finding the best number and location of the USGS wells for the design. This study finds several experimental designs that can achieve the 80%-identifiable groundwater model for modeling the “1,200-foot” sand, the “1,500-foot” sand, and the “1,700-foot” sand of the Baton Rouge area, Louisiana. There is a tradeoff between the number of the USGS wells sampled and the number of sampling rounds. Through the HBMA framework, the highly probable propositions are identified for the geological architecture, the head boundary values south of the BR fault, and the fault permeability. The HBMA shows significant reduction in head prediction variance at all USGS observation wells in the study area.

Using BMA mean prediction to re-calibrate simulation models in the time-sequential sampling approach is not suggested since the BMA predicted quantities are not real data. In addition, the posterior model probabilities calculated using BMA mean predictions only present a potential for a model to be the best model. Which model actually wins is only visible upon real data collection and model re-calibration. The experimental design might be repeated after some time when real future data are collected, the models are re-calibrated, and the posterior model probabilities are updated.
7. Optimal Observation Network Design for Conceptual Model Discrimination and Uncertainty Reduction

7.1. Model Calibration and Posterior Model Probability

To demonstrate the methodology this study only used 33 groundwater head measurements from 9 USGS observation wells in 2010 to calibrate the groundwater models. The simulation period is from January 1, 2010 to December 31, 2015 in which the calibration period is from January 1, 2010 to December 31, 2010, and the prediction period is from January 1, 2011 to December 31, 2015.

After model calibration, the RMSE of the 18 groundwater models ranges from 1.31 m to 2.07 m. The mean RMSE is 1.68 m and the standard deviation of the RMSE is 0.27 m. Equal prior model probabilities are considered for all models. The posterior model probabilities of the 18 groundwater models based on existing groundwater data are listed in Figure 7.1. The top four groundwater models (IK-B-1.5-HC4, IK-B+1.5-HC2, IZ-B0-HC4, and IZ-B+1.5-HC4) are indistinguishable, as indicated by comparable posterior model probabilities of 23.9%, 21.0%, 17.2%, and 16.2%, respectively.

![Figure 7.1: Posterior probabilities of 18 models given the existing observation data in 2010.](image)
In the observation network design, future groundwater heads are collected every three months from the USGS observation wells from 2011 to 2015. The BMA mean predictions are calculated using Equation (33), and will be used as “future observation” in the observation network design. The prediction models were executed using the optimal estimated parameters to predict groundwater heads at 19 USGS observation wells. The assumptions for model prediction are: (1) the groundwater heads at four boundaries in the next 5 years (2011 to 2015) have the same trend as groundwater heads at the last three years in the model calibration period (2008 to 2010); and (2) the monthly pumping rates equal to the average pumping rates of the last 3 years (2008–2010).

7.2. Pre-Analyses

Before conducting an optimal observation network design, the discrimination function value (F value) was calculated for the model domain for the first sampling round on March 31, 2011. Each computational cell has 18 predicted groundwater heads on March 31, 2011 by 18 groundwater models. Averaged F values over layers for a location is computed for the “1,200-foot” sand and the “1,500-1,700-foot” sands. Figure 7.2 shows the averaged F values over layers larger than 0.6 in the study area. Some USGS observation wells are not in high F zones. Considering homoscedasticity (equal variances) or heteroscedasticity (unequal variances) for the future groundwater heads produces dramatic results. Using the Box-Hill approach (homoscedasticity), the constant variance is calculated using the future groundwater heads predicted by 18 groundwater models at 19 USGS observation wells. Potential observation locations (high F values) are mostly found in the areas close to the two geological faults for the “1,200-foot” sand and the “1,500-1,700-foot” sands. Using our approach (heteroscedasticity), the potential observation locations are found south of the BR fault for these sands. The potential observation locations are also found north of the DSS fault for the “1,500-1,700-foot” sands.
Figure 7.2: Distributions of discrimination function values averaged over layers for the first sampling round considering homoscedastic and heteroscedastic future observations. F values smaller than 0.6 are blanked. The black dots are the locations of the USGS observation wells.

7.3. Impacts of Sources of Uncertainty

Impacts of specific sources of model structure uncertainty on F values for the “1,500-1,700-foot” sands are discussed. Figure 7.3(a) presents the distribution of averaged F values over layers (larger than 0.1) calculated by two groundwater models IK-B+1.5-HC2 (model 11) and IK-B+1.5-HC4 (model 12) that are different in the fault permeability architectures. Potential observation areas are found north-west of the DSS fault and the BR fault. Figure 7.3(b) presents averaged F value over layers (larger than 0.6) distribution using three groundwater models GP-B+1.5-HC4 (models 6), IK-B+1.5-HC4 (model 12) and IZ-B+1.5-HC4 (model 18) that differ in geological architectures. Potential observation areas are north-west of the DSS fault. In comparison to Figure
7.3(a), the potential observation areas are much wider, caused by large differences in groundwater head predictions and estimated variances due to uncertain geological architectures.

Figure 7.3: Distributions of discrimination function values averaged over layers after the first sampling round under different sources of model structure uncertainty.
Figure 7.3(c) shows the averaged F values over layers (larger than 0.6) distribution using three groundwater models IK-B-1.5-HC4 (models 8), IK-B0-HC4 (model 10), and IK-B+1.5-HC4 (model 12) that are different in the boundary condition south of the BR fault. Potential observation areas are south of the BR fault, where predicted groundwater heads are significantly different due to the uncertain boundary condition south of the BR fault. It indicates that the prediction uncertainty due to the uncertain boundary condition south of the BR fault is much higher than that due to uncertain fault permeability architecture and geological architecture.

Our analyses show that the discriminate function developed by Box and Hill is able to identify potential observation areas given different sources of model structure uncertainty. In addition, one could use the discrimination function value to prioritize sources of uncertainty and conduct observation network design to discriminate alternatives under each source of uncertainty.

7.4. Observation Network Design Using USGS Wells

From the previous analyses, it is found that considering heteroscedasticity for future groundwater heads is important. In addition, observation at the 19 USGS observation well locations may not produce high F values. Nevertheless, due to physical limitation, in what follows, an optimal observation network design based on the 19 USGS wells is conducted to achieve the 99% probability threshold while maximizing the F value.

7.4.1. Analyses of Individual USGS Wells

The F values for the first sampling round (March 31, 2011) calculated at the USGS wells are shown in Figure 7.4(a). EB-291, EB-146, and WBR-5 are the best three sites for data collection. Figure 7.4(b) shows the expected posterior model probabilities and standard deviations calculated using data collected at EB-291 after the first sampling round. The standard deviations of the
posterior model probabilities are very small. Model 18 (IZ-B_{1.5}+HC_4) is likely to become the best model with posterior model probability with mean 20.1% and standard deviation 2.47%. The maximum posterior model probability is significantly smaller than the 99% probability threshold. Observation data collected from one best USGS observation well after the first sampling round are unable to discriminate among the groundwater models.

Figure 7.4: Designs using one USGS well after the first sampling round: (a) discrimination function value at the USGS wells; and (b) expected posterior model probabilities and one-standard deviation bounds using data from EB-291.
Considering one USGS well at a time, Figure 7.5(a) shows the F values for every three-month sampling round up to 5 years. Using EB-291 at the first sampling round and EB-652 for the rest of the sampling rounds provides the maximum F values, which results in the three best models, IZ-B_{1.5}-HC_4, IZ-B_{1.5}-HC_4, and GP-B_{1.5}-HC_4. Their expected posterior model probabilities and standard deviations over the sampling rounds are shown Figure 7.5(b). The standard deviations are very small.

![Figure 7.5: Design using one best USGS well at different sampling rounds: (a) evolution of discrimination function values using data from EB-291 or from EB-652; and (b) expected posterior model probabilities and one-standard deviation bounds of three best models found at the end of the design.](image)

After 60 months of data collection, the model IZ-B_{1.5}-HC_4 has expected posterior model probability 74.9%, which is smaller than the 99% probability threshold. The second best model (IZ-B_{1.5}-HC_4) and the third best model (GP-B_{1.5}-HC_4) have expected posterior model probabilities of 11.9%, and 6.73%, respectively. In conclusion, using observation data collected from one USGS
well is not enough to discriminate among candidate models after 5 years of quarterly data collection. Multiple wells are needed.

### 7.4.2. Optimal Observation Network Design

This study enumerates all possible combinations from the 19 USGS observation wells to obtain the global optimal design solutions. The “embarrassingly parallel” algorithm is implemented on SuperMike-II, a supercomputer. All 524,287 possible design solutions were split equally among 320 processors to be evaluated. Evaluation of one design took 15 seconds. Therefore, the computation time required for all designs was reduced from 91 days (estimated runtime by using one processor) to 6.83 hours (actual runtime by using 320 processors).

By maximizing the F value with the enumerative approach, Figure 7.6 presents the best USGS wells for different numbers of observation locations up to the 4th sampling round, after which the best observation locations are the same as that of the 4th sampling round. Figure 7.6 shows an interesting rule that if a USGS well is selected in a sampling round, the same USGS well will continue to be selected for the design in the same sampling round as the number of wells increases. This is because the F value calculated using equation (25) is additive over observation data. This mathematical property significantly reduces location search difficulty by searching for one USGS well at a time within a sampling round. For example, in the first sampling round (Figure 7.6a), EB-291 is the best observation location if one well is considered. EB-146 is the new best observation location along with EB-291 if two wells are considered. WBR-5 is the new best observation location along with EB-291 and EB-146 if three wells are considered, and so forth. However, given the same number of wells to be considered, a USGS well selected in one sampling round may not be selected in another sampling round.
Figure 7.6: Best observation locations given a desired number of USGS wells for the first four sampling rounds.

The optimal observation network design found that using 12 USGS wells results in the highest F value and identifies groundwater model GP-B_{1.5}-HC_{4} that passes the 99% probability
threshold at 60 months (20 sampling rounds). As shown in Figure 7.7, using 1, 2, 4, 7, 8, 15, 16, and 17 USGS wells could not find a 99%-identifiable groundwater model within 60 months. Using a different number of USGS wells may find three other 99%-identifiable groundwater models within 60 months. They are IZ-B-1.5-HC₄, IZ-B+1.5-HC₄, and IK-B-1.5-HC₂. However, these solutions are suboptimal: the enumerative approach with high-performance computing is able to find the global optimum solution.

The expected posterior model probabilities and standard deviations of the four identified groundwater models in Figure 7.7 are shown in Figure 7.8. The small standard deviations indicate low posterior model probability uncertainty arising from model parameter uncertainty. Between 3-24 months, the posterior model probability of the 99%-identifiable groundwater model GP-B.
$1.5$-$HC_4$ decreases as more future data are added. The non-monotonic behavior indicates that the future data in this period support other groundwater models. Future data after 24 months start to support $GP-B_{1.5}$-$HC_4$ and discriminate it from other groundwater models after 33 months.

![Figure 7.8: Expected posterior model probabilities and one-standard deviation bounds of the four identified groundwater models in Figure 7.7.](image)

Before the optimal observation network design, large prediction uncertainty of groundwater heads at the 19 USGS wells is shown in Figure 7.8 for December 31, 2015. High prediction variances are due to using 18 groundwater models for prediction. After the observation network design, the total prediction variances according to equation (14) are significantly reduced due to using only one groundwater model, $GP-B_{1.5}$-$HC_4$, for head prediction.
7.5. Conclusions

The Box-Hill discrimination function is expanded to account for multiple independent spatiotemporal future observations. Maximizing the Box-Hill discrimination function in an optimal observation network design needs a probability threshold in order to distinctly identify one model for a system. Using posterior model probability as a discrimination criterion to identify the best model for a system provides direct data evidence since posterior model probability describes the importance of a model based on data and its prior probability.

The BMA method is found to be useful in handling various sources of uncertainty in the optimal observation network design. The BMA mean predictions are less biased by future observation data than those of individual models. Moreover, the BMA method is able to incorporate existing observation data and quantify future observation uncertainty arising from conceptual and parametric uncertainties in the discrimination function.

The optimal observation network design is implemented in a groundwater study in the Baton Rouge, Louisiana area to collect new groundwater levels from USGS observation wells.
This study found that considering homoscedastic or heteroscedastic future groundwater heads drastically impact potential observation locations. Heteroscedasticity of future groundwater heads should be considered in optimal observation network design to account for various sources of future observation uncertainty. Using the USGS observation wells may not be efficient according to the Box-Hill discrimination function value. Nevertheless, they are able to provide data to meet the probability threshold. After the optimal observation network design, variances of head predictions are significantly reduced due to eliminating models that have insignificant posterior model probabilities.
8. Concluding Remarks

This study aims at developing a reliable groundwater model for future head predictions and saltwater intrusion simulations under parametric and conceptual model uncertainty. To do so, the study first developed a grid generation technique that automatically creates MODFLOW grids from well logs. Then, multiple candidate models were proposed and carefully calibrated using the parallel CMA-ES implemented in a high-performance computing system. Final, experimental designs were introduced to discriminate candidate models and identify the highly probable groundwater model and model propositions. The key findings for each research step are summarized as follows:

With respect to the grid generation, this study presented a technique for automatic generation of structured and unstructured MODFLOW grids from well logs. The technique can preserve facies geometries of complex hydrostratigraphic architecture by using fine vertical discretization and regional geological dip. In addition, the technique avoids a possible overwhelming number of computational cells by introducing upscaling approaches.

The technique was applied to generate MODFLOW grids for a real-world fluvial aquifer system of the “1,200-foot” sand, the “1,500-foot” sand, the “1,700-foot” sand, and the “2,000-foot” sand in the Baton Rouge area, including two faults. The technique could correlate a large number of well logs and preserves geological information and depositional environment. In addition, the derived MODFLOW grids preserved the geometries of the faulted hydrostratigraphy, including pinch-outs and narrow hydraulic connections through the faults. The result of the case study demonstrates the capability of the technique that exactly converts complex hydrostratigraphic architectures including faults into MODFLOW grid. As a result, model structure errors are reduced in future flow and transport simulations. A MATLAB code was
developed in which new grids can be generated given different input parameters (e.g., criteria for eliminating thin sand and thin clay, number of layers) or when additional well log data become available.

With respect to modeling of groundwater flow in the “1,200-foot” sand, the “1,500-foot” sand and the “1,700-foot” sand, the results showed the BR fault and the DSS fault are low permeability faults that limit horizontal flow. The model shows strong groundwater flow interactions between the "1,200-foot" sand and the "1,500-1700- foot" sands for the area between two faults. In addition, heavy pumping in the area in between the two faults has caused declination of groundwater level north of the BR fault, resulted in two cones of depression and high head differences across the BR fault and the DSS fault. The declination of groundwater level north of the Baton Rouge has changed flow direction in which the flow moves northward, across the BR fault, and potentially causing saltwater intrusion toward pumping stations. The parallel CMA-ES was found helpful in model calibration as it significantly reduces the model calibration time. Moreover, the covariance matrix of the estimated parameters are valuable in analyzing parametric estimation uncertainty and head prediction uncertainty.

With respect to the experimental design for identification of model propositions, this study found that using posterior model probability as a discrimination criterion is a straightforward approach to identify a unique and highly probable model from a pool of candidate models. As posterior model probability describes model importance by taking into account data evidence and its prior probability, it is a direct quantitative indicator to guide future data collection through an optimal observation network in the experimental design.

The time-sequential sampling approach presents a practical strategy for data collection, by which future data are gradually collected and used along with historical data to achieve the unique
and highly probable model. Using the BMA mean predictions as future observation data is an unbiased approach since the BMA mean predictions are less biased by future observation data than those of individual models. In addition, all candidate models contribute to prediction according to their up-to-date posterior model probabilities based on both historical data and future data. The HBMA method is found to be useful to illustrate the results of the experimental design for proposition discrimination and prediction uncertainty reduction.

The experimental design illustrated by the real groundwater case study is practical since this study only uses the USGS observation wells. Implementing the enumeration approach to a multi-core supercomputer is viable and assures finding the best number and location of the USGS wells for the design. Through the HBMA framework, the highly probable propositions are identified for the geological architecture, the head boundary values south of the BR fault, and the fault permeability. The HBMA shows significant reduction in head prediction variance at all USGS observation wells in the study area.

With respect to the optimal observation network design for model discrimination, this study adopts the Box-Hill discrimination function and posterior model probabilities to develop an optimal observation network design to discriminate conceptual models. The discrimination criterion is posterior model probability. The Box and Hill discrimination function is expanded to account for multiple independent spatiotemporal observations. The BMA method is used to predict future observation data and to quantify future observation uncertainty in the discrimination function.

The BMA method is found to be useful in handling various sources of uncertainty in the optimal observation network design. First, the BMA mean prediction is introduced as future observation data, which is less biased and more reliable than prediction using individual models.
Second, the BMA method is able to incorporate existing observation data and quantify future observation uncertainty arising from conceptual and parametric uncertainties in the discrimination function.

The optimal observation network design is implemented in a groundwater study in the Baton Rouge, Louisiana area to collect new groundwater levels from USGS observation wells. Heteroscedasticity of future groundwater heads should be considered in optimal observation network design to account for various sources of future observation uncertainty. Using the USGS observation wells may not be efficient according to the Box-Hill discrimination function value. Nevertheless, they are able to provide data to meet the probability threshold. After the optimal observation network design, variances of head predictions are significantly reduced due to eliminating models that have insignificant posterior model probabilities.
References


Aquaveo (2014), *Groundwater Modeling System (GMS)*, Aquaveo, LLC.


Carle, S. F. (1999), *T-PROGS: Transition probability geostatistical software*, University of California, Davis.


Chitsazan, N. (2014), Bayesian Saltwater Intrusion Prediction and Remediation Design under Uncertainty, Louisiana State University, Baton Rouge, LA.


Pham, H. V., and F. T.-C. Tsai (submitted), Construction of structured and unstructured MODFLOW grids from well logs. (Under review)

Pham, H. V., and F. T.-C. Tsai (submitted), Optimal Observation Network Design for Conceptual Model Discrimination and Uncertainty Reduction. (Under review)


RockWare (2014), *RockWorks 16*, RockWare Incorporated, Golden, CO 80401 USA.


Appendix A: Derivation of a Multi-Data Discrimination Function

To derive a multi-data discrimination function from Shannon entropy, the study follows Box and Hill [1967] and uses Kullback’s inequality [[Kullback, 1997], page 15, corollary 3.1]:

\[
\sum_{j=1}^{P} \Pr(M_j | \Delta_{obs}^{\text{new}}) \Pr(\Delta_{D}^{\text{new}} | M_j) \ln \frac{\Pr(\Delta_{D}^{\text{new}} | M_j)}{\Pr(\Delta_{D}^{\text{new}} | M_j)} \geq \Pr(\Delta_{D}^{\text{new}} | M_j) \ln \frac{\Pr(\Delta_{D}^{\text{new}} | M_j)}{q(\Delta_{D}^{\text{new}})}
\]  

(A1)

Integrating equation (A1) with new data \( \Delta_{D}^{\text{new}} \) and then taking an average over models with the posterior model probabilities \( \Pr(M_j | \Delta_{obs}) \), the right side of the inequality becomes the expected entropy change, \( R \), in equation (24). By doing so, the left side of the inequality is defined as the discrimination function, \( F \):

\[
F = \sum_{i=1}^{P} \sum_{j=1}^{P} \Pr(M_i | \Delta_{obs}) \Pr(M_j | \Delta_{obs}) \int \Pr(\Delta_{D}^{\text{new}} | M_i) \ln \frac{\Pr(\Delta_{D}^{\text{new}} | M_i)}{\Pr(\Delta_{D}^{\text{new}} | M_j)} d\Delta_{D}^{\text{new}}
\]  

(A2)

which can be further derived as

\[
F = \sum_{i=1}^{P} \sum_{j=1}^{P} \Pr(M_i | \Delta_{obs}) \Pr(M_j | \Delta_{obs}) \left( \int \Pr(\Delta_{D}^{\text{new}} | M_i) \ln \Pr(\Delta_{D}^{\text{new}} | M_i) d\Delta_{D}^{\text{new}} - \int \Pr(\Delta_{D}^{\text{new}} | M_j) \ln \Pr(\Delta_{D}^{\text{new}} | M_j) d\Delta_{D}^{\text{new}} \right) \\
+ \int \Pr(\Delta_{D}^{\text{new}} | M_j) \ln \Pr(\Delta_{D}^{\text{new}} | M_j) d\Delta_{D}^{\text{new}} - \int \Pr(\Delta_{D}^{\text{new}} | M_i) \ln \Pr(\Delta_{D}^{\text{new}} | M_i) d\Delta_{D}^{\text{new}}
\]  

(A3)

Kullback’s inequality indicates that the expected entropy change is a lower bound of the discrimination function value, i.e., \( R \leq F \).

The marginal likelihood function \( \Pr(\Delta_{D}^{\text{new}} | M_j) \) is usually built under the Gaussian assumption and data independence on future observation data \( \Delta_{D}^{\text{new}} \) as follows:

\[
\Pr(\Delta_{D}^{\text{new}} | \eta) = (2\pi)^{-\frac{N}{2}} |\sigma|^{-1} \exp \left[ -\frac{1}{2} (\Delta_{D}^{\text{new}} - \eta)^{T} \sigma^{-2} (\Delta_{D}^{\text{new}} - \eta) \right]
\]  

(A4)
η and σ² are the expected values and known diagonal covariance matrix of future observation data \( \Delta_D^{\text{new}} \). |σ| is the determinant. N is the number of future observation data \( \Delta_D^{\text{new}} \). The randomness of the expected values η can be characterized by the outputs of a model as follows:

\[
\text{Pr}(\eta|M_j) = (2\pi)^{-\frac{N}{2}} |\sigma_j|^{-1} \exp\left[-\frac{1}{2}(\eta - \Lambda_j)^T \sigma_j^{-2} (\eta - \Lambda_j)\right]
\]

where \( \Lambda_j \) and \( \sigma_j^{-2} \) are the predictions and variances of future observation data by model \( M_j \) at the future data domain.

By considering all possible η, the marginal likelihood function can be obtained by

\[
\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) = \int \text{Pr}\left(\Delta_D^{\text{new}} | \eta\right) \text{Pr}(\eta|M_j) d\eta
\]

Substituting equations (A4) and (A5) into (A6), the marginal likelihood function \( \text{Pr}(\Delta_D^{\text{new}} | M_j) \) becomes

\[
\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) = (2\pi)^{-\frac{N}{2}} |\Sigma_j|^{-1} \exp\left[-\frac{1}{2}(\Delta_D^{\text{new}} - \Lambda_j)^T \Sigma_j^{-2} (\Delta_D^{\text{new}} - \Lambda_j)\right]
\]

where \( \Sigma_j^{-2} = \sigma^2 + \sigma_j^{-2} \) is the diagonal covariance matrix. Substituting (A7) to (A3), the integral terms have analytical solutions as follows.

\[
\int \text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) \ln\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) d\Delta_D^{\text{new}} = \ln\left[(2\pi)^{\frac{N}{2}} |\Sigma_j|^{-1}\right] - \frac{N}{2}
\]

(A8)

\[
\int \text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) \ln\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) d\Delta_D^{\text{new}} = \ln\left[(2\pi)^{\frac{N}{2}} |\Sigma_j|^{-1}\right] - \frac{N}{2}
\]

(A9)

\[
\int \text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) \ln\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) d\Delta_D^{\text{new}} = \ln\left[(2\pi)^{\frac{N}{2}} |\Sigma_j|^{-1}\right] - \frac{N}{2} \text{tr}(\Sigma_j^{-2} \Sigma_i^{-2})
\]

(A10)

\[
\int \text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) \ln\text{Pr}\left(\Delta_D^{\text{new}} | M_j\right) d\Delta_D^{\text{new}} = \ln\left[(2\pi)^{\frac{N}{2}} |\Sigma_j|^{-1}\right] - \frac{N}{2} \text{tr}(\Sigma_j^{-2} \Sigma_i^{-2})
\]

(A11)
Substituting (A8)-(A11) into (A3), the discrimination function becomes

\[ F = \frac{1}{2} \sum_{i=1}^{P} \sum_{j=i+1}^{P} \text{Pr}(M_j | \Delta_{\text{obs}}) \text{Pr}(M_j | \Delta_{\text{obs}}) \]

\[ \sum_{n=1}^{N} \left[ \frac{\left( \sigma_{n,i}^2 - \sigma_{n,j}^2 \right)^2}{\left( \sigma_n^2 + \sigma_{n,i}^2 \right)\left( \sigma_n^2 + \sigma_{n,j}^2 \right)} + \left( \Delta_{n,i} - \Delta_{n,j} \right)^2 \left( \frac{1}{\sigma_n^2 + \sigma_{n,i}^2} + \frac{1}{\sigma_n^2 + \sigma_{n,j}^2} \right) \right] \]  

(A12)

where \( \{ \Delta_{n,i} \} \) and \( \{ \Delta_{n,j} \} \) are the elements of \( \Delta_i \) and \( \Delta_j \), respectively; \( \{ \sigma_n^2 \} \) are the diagonal elements of \( \sigma^2 \); and \( \{ \sigma_{n,i}^2 \} \) are the diagonal elements of \( \sigma_i^2 \).
Appendix B: Open Access Permission

Open access permission for the Advances in Water Resources paper:

This is a License Agreement between Hai V Pham ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier
Elsevier Limited
The Boulevard, Langford Lane
Kidlington, Oxford, OX5 1GB, UK

Registered Company Number 1982084

Customer name
Hai V Pham

Customer address
2000 Brightside Dr.
Baton Rouge, LA 70820

License number 3658870646935
License date Jun 30, 2015

Licensed content publisher Elsevier

Licensed content publication Advances in Water Resources
Licensed content title Bayesian experimental design for identification of model propositions and conceptual model uncertainty reduction
Licensed content author Hai V Pham, Frank T.-C. Tsai
Licensed content date September 2015
Licensed content volume number 83
Licensed content issue number n/a

Number of pages 12
Start Page 148
End Page 159
Type of Use reuse in a thesis/dissertation

Portion full article
Format electronic
Are you the author of this Elsevier article? Yes
Will you be translating? No
Title of your thesis/dissertation EXPERIMENTAL DESIGNS FOR MODEL DISCRIMINATION AND UNCERTAINTY REDUCTION IN GROUNDWATER MODELING
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows.

"Reprinted from Publication title, Vol / edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment
terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE
The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the website must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at [http://www.sciencedirect.com/science/journal/xxxxxx](http://www.sciencedirect.com/science/journal/xxxxxx) or the Elsevier homepage for books at [http://www.elsevier.com](http://www.elsevier.com), Central Storage. This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

   Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at [http://www.elsevier.com](http://www.elsevier.com). All content posted to the website must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve**: In addition to the above the following clauses are applicable: The website must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors**: the following clauses are applicable in addition to the above:

**Preprints**:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts**: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.
Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement

- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (JPA) is the definitive final record of published research that appears or will appear in the journal and embodies all value adding publishing activities including peer review or ordination, copy editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs)

https://r100.copyright.com/App/PrintableLicenseFrame.jsp?publisherID=70&publisherName=ELS&publicationID=0309-1726&publicationID=90676&rightID=1&... 57
and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation:** If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for ProQuest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PIs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

**Elsevier Open Access Terms and Conditions**

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

**Terms & Conditions applicable to all Open Access articles published with Elsevier:**

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

**Additional Terms & Conditions applicable to each Creative Commons user license:**

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant
DOIs), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

**20. Other Conditions:**

v1.7

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.
Vita

Hai Pham was born in Hai Phong, Vietnam in October, 1983. He received his bachelor’s degree in Civil and Environmental Engineering from Vietnam Water Resources University in 2006. Then he worked three years as a researcher at the Vietnam Institute of Meteorology, Hydrology and Environment (IMHEN). In 2011, he moved to Seoul, Korea where he obtained a master’s degree in Civil Engineering at the Dongguk University. In August 2011, he moved to Baton Rouge, Louisiana to start his doctoral program in the Department of Civil and Environmental Engineering at Louisiana State University. He is currently a candidate for the degree of Doctor of Philosophy in Civil Engineering.