Advances in Water Resources 83 (2015) 148–159

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres

Bayesian experimental design for identification of model propositions and conceptual model uncertainty reduction

Hai V. Pham, Frank T.-C. Tsai*

Department of Civil and Environmental Engineering, Louisiana State University, 3418G Patrick F. Taylor Hall, Baton Rouge, Louisiana, 70803, USA

ARTICLE INFO

Article history:
Received 8 March 2014
Revised 10 April 2015
Accepted 28 May 2015
Available online 3 June 2015

Keywords:
Bayesian model averaging
Experimental design
Groundwater
Model discrimination
Observation network design
Uncertainty

ABSTRACT

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. This study introduces 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 a groundwater study in the Baton Rouge area, Louisiana. We design a new groundwater head observation network based on existing USGS observation wells. 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. Several design solutions are found to achieve an 80%-identifiable groundwater model in 5 years by using six or more existing USGS wells. The HBMA result shows that each highly probable proposition can be identified for each uncertain model component once the discrimination criterion is achieved. The variances of groundwater head predictions are significantly decreased by reducing posterior model probabilities of unimportant propositions.

1. Introduction

Identification of a reliable groundwater model for future applications is the ultimate goal of model development. This goal, however, remains challenging due to high uncertainty in groundwater systems. 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 [1,2]. Considering only one model may lead to statistical bias and underestimation of uncertainty because groundwater systems are often complex and have multiple interpretations [3–5].

In recent years, conceptual model uncertainty has received much attention in groundwater applications (e.g., [6–14]). Many studies have shown that contribution of conceptual model uncertainty to predictive uncertainty is significantly larger than that of model parameter uncertainty [3–5,7–10,14–17]. This motivates the research to consider multi-model methods that seek to obtain multi-model ensemble predictions and analyze uncertainty from a set of candidate groundwater models.

Bayesian model averaging (BMA) [18–21] 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 [22–24]. 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 [13,25,26].

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. For example, geological architecture can be one uncertain model component in groundwater modeling. Many hydrostratigraphy modeling techniques may be proposed to construct different geological architectures (propositions) potentially leading to

* Corresponding author. Tel.: 1 225 5784246
E-mail addresses: hpham28@lsu.edu (H.V. Pham), ftsai@lsu.edu (F.T.-C. Tsai).

http://dx.doi.org/10.1016/j.advwatres.2015.05.024
0309-1708/

an overwhelming number of models with non-dominant posterior model probabilities. Conducting prediction and uncertainty analysis using a great deal of computationally intensive groundwater models can become intractable. By incorporating many conceptual models, prediction results using BMA can become useless when the prediction uncertainty is very high. This concern highlights the importance of conducting an experimental design to discriminate groundwater model propositions, identify highly probable models, and in turn, reduce conceptual model uncertainty (model choice uncertainty) in prediction.

In the field of groundwater hydrology, many studies have conducted experimental designs to improve parameter estimation (e.g., [13,27–32]), reduce model prediction uncertainty (e.g., [33–36]) and minimize decision errors in hypothesis tests [37]. Some studies have used experimental designs to discriminate candidate models and identify the "true" model [38–42]. Knopman and Voss [39] 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 [41]. Usunoff et al. [38] 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. [42] recently applied the Kulback-Leibler information to discriminate flow and transport models.

Significant efforts have been invested in the development of different discrimination criteria for experimental designs in different fields [42–54], but only a few criteria have been applied to groundwater modeling, e.g., the differences between model outputs [38,39,41,55] and the Kulback-Leibler information [42].

Model discrimination criteria should serve not only for model discrimination, but also for model identification. As discussed in Box and Hill [43], 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. In line with Box and Hill, this study introduces a new experimental design to identify the most probable model under the BMA framework. A discrimination criterion is developed based on maximizing the maximum posterior model probability across all candidate models such that a unique and highly probable groundwater model can be identified. Posterior model probability is a straightforward quantitative indicator for model discrimination and is suitable for the experimental design since it represents model importance based on data evidence. The experimental design is implemented to a groundwater study at Baton Rouge, southeastern Louisiana. The purpose of the experimental design in this case study is to find the optimal number and location of groundwater observation wells and the number of sampling rounds such that a groundwater model meets the desired value of the discrimination criterion.

2. Discrimination criterion

Unlike prior studies that use differences between model outputs [38,39,41,55] or Kulback-Leibler information [42] to set up a discrimination criterion, this study uses the posterior model probability to discriminate one model from other candidate models. Let $\boldsymbol{M} = \{M_p;\ p = 1, 2, \ldots, P\}$ be a set of P candidate models developed by propositions from different sources of model structure uncertainty. The posterior model probability for a model M_p is $Pr(M_p|\boldsymbol{\Delta}^{\text{obs}})$, where $\boldsymbol{\Delta}^{\text{obs}}$ is existing observation data (e.g., groundwater heads, concentration, fluxes, etc. in aquifers). From Bayes' theorem, the sum of the posterior model probabilities is unity, i.e., $\sum_{p=1}^{P} Pr(M_p|\boldsymbol{\Delta}^{\text{obs}}) = 1$. A model is said to be γ-identifiable from the rest of the candidate models if its posterior model probability is superior to others and is over a probability threshold, γ. The probability threshold, γ, should be much larger than 50% (e.g., 80%) to ensure a unique and highly probable model. To find the maximum posterior model probability and check if the γ-identifiable model can be identified, we define the discrimination criterion as:

$$\max\{Pr(M_p|\boldsymbol{\Delta}^{\text{obs}}),\ p = 1, 2, \cdots, P\} \geq \gamma, \tag{1}$$

where the γ-identifiable model is the best model among the candidate models in terms of the highest posterior model probability.

3. Experimental design using BMA prediction

If none of the candidate models meets the probability threshold, the existing observation data $\boldsymbol{\Delta}^{\text{obs}}$ 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 γ-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:

$$D^* = \arg\limits_{D}\left[\max\max\left\{Pr\left(M_p|\boldsymbol{\Delta}^{\text{obs}}, \boldsymbol{\Delta}_D^{\text{new}}\right),\ p = 1, 2, \cdots, P\right\} \geq \gamma\right], \tag{2}$$

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

Since we do not know the future observation data $\boldsymbol{\Delta}_D^{\text{new}}$, 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, we take the deterministic approach and propose BMA mean predictions as the future observation data in the experimental design because the BMA mean prediction presents unbiased estimation.

3.1. 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, $\boldsymbol{\Delta}^{\text{new}}$, using all candidate models, BMA [18–21] 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_k, the future observation data presented by the BMA mean predictions are

$$E\left[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}\right] = \sum_{p=1}^{P} E\left[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}, M_p\right] Pr(M_p|\boldsymbol{\Delta}^{\text{obs}}), \tag{3}$$

where $E[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}, M_p]$ are the means of the future observation data predicted by model p at time t_1 given the existing observation data $\boldsymbol{\Delta}^{\text{obs}}$ and a design D, $E[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}]$ are the BMA means at time t_1, and $Pr(M_p|\boldsymbol{\Delta}^{\text{obs}})$ is the posterior model probability of M_p calculated using the existing observation data. This study adopts the large sample assumption of Draper [18], that $E[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}, M_p]$ is approximated by $E[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}, \hat{\boldsymbol{\beta}}_p, M_p]$, where $\hat{\boldsymbol{\beta}}_p$ is the maximum likelihood estimate of model parameters $\boldsymbol{\beta}_p$ of model M_p. The total data $\boldsymbol{\Delta}_1 = \{\boldsymbol{\Delta}^{\text{obs}}, E[\boldsymbol{\Delta}_D^{\text{new}}(t_1)|\boldsymbol{\Delta}^{\text{obs}}]\}$ are used to update the posterior model probabilities $Pr(M_p|\boldsymbol{\Delta}_1)$, which will be explained in the