

UC Irvine

UC Irvine Previously Published Works

Title

Toward Improved Hydrologic Prediction with Reduced Uncertainty Using Sequential Multi-Model Combination

Permalink

<https://escholarship.org/uc/item/0sq1d9ff>

ISBN

9780784409763

Authors

Moradkhani, Hamid

Hsu, K

Sorooshian, S

Publication Date

2008-05-01

DOI

10.1061/40976(316)576

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Toward Improved Hydrologic Prediction with Reduced Uncertainty using Sequential Multi-Model Combination

Hamid Moradkhani^{1*} (PE., MASCE), K. Hsu², and S. Sorooshian²

¹Department of Civil and Environmental Engineering, Portland State University

**Corresponding Author hamidm@cecs.pdx.edu*

²Department of Civil and Environmental Engineering, University of California, Irvine

Abstract

The contemporary usage of hydrologic models has been to rely on a single model to perform the simulation and predictions. Despite the tremendous progress, efforts and investment put into developing more hydrologic models, there is no convincing claim that any particular model in existence is superior to other models for various applications and under all circumstances. This results to reducing the size of the plausible model space and often leads to predictions that may well-represent some phenomena or events at the expenses of others. Assessment of predictive uncertainty based on a single model is subject to statistical bias and most likely underestimation of uncertainty. This endorses the implementation of multi-model methods for more accurate estimation of uncertainty in hydrologic prediction.

In this study, we present two methods for the combination of multiple model predictors using Bayesian Model Averaging (BMA) and Sequential Bayesian Model Combination (SBMC). Both methods are statistical schemes to infer a combined probabilistic prediction that possess more reliability and skill than the original model members produced by several competing models. This paper discusses the features of both methods and explains how the limitation of BMA can be overcome by SBMC. Three hydrologic models are considered and it is shown that multi-model combination can result in higher prediction accuracy than individual models.

1. Introduction

Over the past two decades significant effort has gone into development of watershed models by hydrologists. These models have been used to simulate and predict the behavior of the underlying physical processes in the natural system. The reliability on these models to accurately and precisely predict the nonlinear and complex behavior of the hydrologic system is dependent on the perception of the modeler from the governing processes in the system, followed by model

conceptualization, mathematical and computer modeling. Hydrologic and water resources engineers often use these models for calculating the watershed runoff, water resources planning and management including reservoir operation, storm water management, water distribution system and groundwater protection. Prediction with these models are often deterministic, relying on the most probable forecast without explicitly accounting for the uncertainty in the incomplete system representation (model structural uncertainty) or the uncertainty in the system initial condition, in the observation of system input and output and also in the parameters that identify the system. It is obvious that quantifying these uncertainties is necessary to evaluate model quality and predictive competence.

Owing to availability of massive data from various sources including ground-based and remotely-sensed observations, hydrologists have developed various types of models including data-driven or black box models, conceptual models and also physically-based models for both lump and distributed representation of a hydrologic system. The general practice by hydrologist and water resources engineers is to use a single model by complete reliance that the model can perform the simulation to their best advantage ignoring the fact that there is no such a perfect model in existence that fully represent the processes in all conditions (*Beven 2006; Smith et al., 2004*). Most of the predictive uncertainty analysis techniques developed by far are implemented on single models which are believed to result in underestimation of uncertainty and overconfidence in the model predictive capability. Multi-model combination methods have recently been advocated to benefit from the strength of various models in predicting the hydrologic variable of interest (*Neuman, 2003; Duan et al., 2007; Ajami et al., 2007; Vrugt and Robinson, 2007*). These studies have been motivated by the Bayesian Model Averaging (BMA) development by *Raftery et al., (1993, 2003, 2005)* and *Hoeting et al., (1999)*. BMA prediction is essentially the weighted average of the individual model predictions which has gained popularity in recent years. The BMA aims at providing the unconditional mean and variance of the predictant on the basis of several model forecasts. The main characteristic of the BMA is to rely on a set of time-invariant weighting parameters that are assumed to be normally distributed. Experience with modeling and simulation have shown that various models may perform differently at different periods (e.g. wet season vs. dry season or dry soil vs. wet soil initial condition) and reliance on just one model for simulating the processes in all conditions would be overconfidence on the model. Assuming the fixed parameter weight, as part of BMA technique,

does not provide the flexibility to models participate dynamically according to the conditions. Therefore, in this study we present a sequential Bayesian model combination method to overcome the fixed parameter assumption in the BMA method and the results are interpreted.

2. Multi-model Methods in Predictive Uncertainty Analysis

2.1. Bayesian Model Averaging (BMA)

Bayesian Model Averaging (BMA) is a statistical method which was originally developed to combine inferences and prediction from several statistical models (*Leamer 1978; Kass and Raftery 1995*). In other word, BMA was designed to postprocess the forecast model ensembles to deduce a predictive probability density function (PDF) of combined prediction that is more skillful and reliable than that of the original model members (model ensemble). *Raftery et al. 2005*) extended the BMA application to the ensemble of dynamic models (mainly weather forecasting models). The BMA predictive PDF of a quantity of interest is a weighted average of PDFs providing that the individual forecasts are unbiased or bias-corrected.

If y is the variable of interest to be forecasted (predictant), $D = [d_1, d_2, \dots, d_n]$ is the vector of observations (calibration data) and $M = [M_1, M_2, \dots, M_k]$ denotes the ensemble of individual model predictions, then the posterior distribution of y is represented as follows.

$$p(y | D) = \sum_{k=1}^K p(y | M_k) p(M_k | D) \quad (1)$$

Where, $p(y | M_k)$ is the forecast PDF according to model M_k and $w_k = p(M_k | D)$ is the posterior probability of prediction from model M_k given in below:

$$p(M_k | D) = \frac{p(D | M_k) p(M_k)}{\sum_{i=1}^k p(D | M_i) p(M_i)} \quad (2)$$

Where, $p(M_k)$ is the prior probability of model M_k being the true model and $p(D | M_k)$ is the likelihood of model M_k . The model posterior probability or the model weights should sum up to unity, that is, $\sum_{k=1}^K w_k = 1$. In the absence of prior knowledge in selecting the models at the beginning of prediction, all models are treated equally, i.e., $p(M_k) = 1/k$. This assumption simplifies the posterior probability presented by (2) to $p(M_k | D) = p(D | M_k)$ meaning that the

posterior of model M_k can be regarded as likelihood of calibration model M_k in predicting y . The likelihood of model M_k , in fact, quantifies the probability of success of model M_k to closely fit the observation D .

One of the challenges in BMA application lies in estimating the parameters, weights w_k and variance. One method proposed by *Raftery et al. (2005)* is to employ the Expectation-Maximization (EM) algorithm as an iterative procedure. BMA is essentially a post-processing of existing retrospective model simulations or hindcastings and the $P(M_k | Y)$ does not change with time meaning that once they are obtained, they remain fixed and used for for rest of the prediction. In the next section, we explain a procedure on sequentially estimating the $P(M_k | Y)$ enabling the procedure be applied in real time forecasting.

2.2. Sequential Bayesian Model Combination (SBMC)

The posterior distribution presented in eq. (2) as the original form of Bayes law is in the batch form where the available historical data is taken for the uncertainty estimation through that conditional probability. However, this form makes no attempt to include information from new observations when becoming available. The flexibility required to use the new information is provided by a sequential Bayesian scheme. *Moradkhani et al. (2005a &b)* showed that the methods based on sequential Bayesian estimation seem better able to benefit from the temporal organization and structure of information achieving better conformity of the model output with observations. In the multi-model selection process, it is intuitive that if certain models give better predictions than others for a specific portion of the process (time period), those models should be given higher level of participation in predicting the quantity of interest while still not ignoring the level (probability) of success of other models in prediction, i.e., using them may generate a better solution than using a fixed weighting factor as is done in BMA method. Therefore, implementing the model recursively provides a flexible framework to update models' posterior probability using newly-available observations (*Hsu et al., 2008*).

Let y_t denote the observation of predictive variable at time t and $p_{t-1}(M_j | D_{t-1})$ be the model prior distribution of the j^{th} model at time t , then the model posterior distribution in sequential form is written as:

$$p_t(M_j | D_t) = \frac{p(d_t | M_j) \cdot p_{t-1}(M_j | D_{t-1})}{\sum_{i=1}^k p(d_t | M_i) \cdot p_{t-1}(M_i | D_{t-1})} \quad (3)$$

For simplicity, the likelihood $p(d_t | M_j)$ was chosen to be normally distributed as

$$p_j(d_t | M_j) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left(-\frac{(y_t - \hat{y}_t(M_j))^2}{2\sigma_j^2}\right) \quad (4)$$

where $\hat{y}_t(M_j)$ is the model M_j estimate and σ_j is the standard deviation.

The Bayesian combination of predictors is expressed as an expected value (eq. 5 below) which is essentially the weighted prediction where the weights are posteriori probability of each model calculated by eq. (3):

$$E(y_t | M_1, M_2, \dots, M_k, D_t) = \sum_{j=1}^k E(y_t | M_j, D_t) p_t(M_j | D_t) \quad (5)$$

where $p_t(M_j | D_t)$ is the probability of the j -th model; $E(y_t | M_j, D_t)$ is the expected value of the model M_j estimate.

3. Hydrologic Models

To test the multi-model methodologies explained in previous section, four hydrologic models were used in this study including two conceptual models, mainly Sacramento Soil Moisture Accounting Model (SAC-SMA) and Hydrologic MODdel (HyMOD), a time series model, Autoregressive with eXogenous inputs (ARX) and a black box, Artificial Neural Networks (ANN) model. SAC-SMA model (Figure 1)(Burnash *et al.* 1973) is still the most widely used operational hydrologic model in the National Weather Service (NWS).

HyMoD is a simple hydrologic model which has its origin in the probability distributed soil moisture model (Boyle *et al.*, 2001) (Figure 2). ARX time-series model is simply a linear dynamic function that is used in many disciplines including hydrology (Wood *et al.*, 1980).

The ARX (n_1, n_2) time-series model is defined in a linear function as

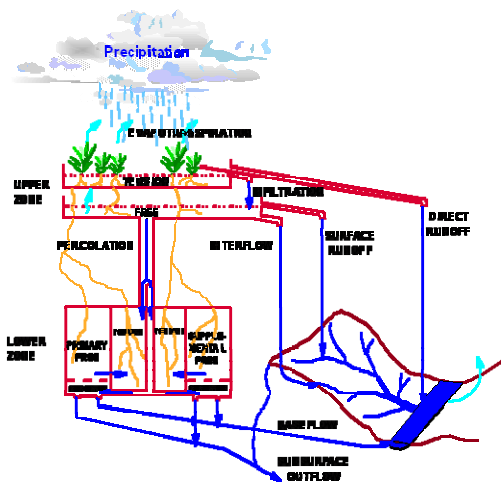


Figure 1. Schematic of SAC-SMA model

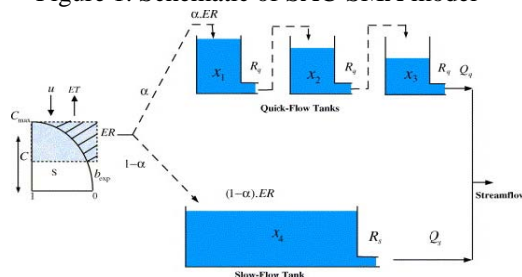


Figure 2. Schematic of SAC-SMA model

$$y_{t+1} = \sum_{i=0}^{n_1} a_i y_{t-i} + \sum_{j=0}^{n_2} b_j r_{t-j} + \varepsilon_{t+1} \quad (6)$$

where a_i and b_j are parameters, and $y(t)$ and $r(t)$ are the observed streamflow and rainfall sequences, respectively. The time unit t is one day, and ε_{t+1} is the error of streamflow estimation. The case study uses three previous time intervals of rainfall and streamflow observations as the inputs to the model (i.e., $n_1 = n_2 = 2$).

ANN models are black box or data-driven models that have been found suitable in many hydrology studies including streamflow prediction and precipitation estimation (Hsu et al., 1995; Moradkhani et al., 2004; Hong et al., 2005). In this study we used the Self Organizing Radial Basis (SORB) model developed by Moradkhani et al. (2004) (Figure 3).

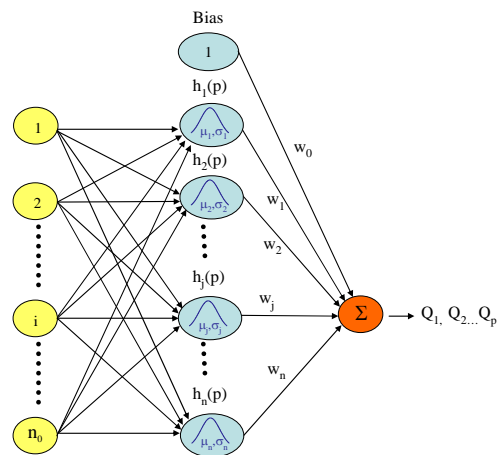


Figure 3. Architecture of SORB-ANN model

4. Study Test Basin

The Leaf River Basin with 1949 km² area is located north of Collins, Mississippi and was chosen as the test basin in this study. We used 36 years (1953~1988) of daily rainfall and streamflow data for this basin. For the models to be used in one-time-increment forward forecasting, their parameters must be calibrated from a set of historical data. For the two conceptual models, SAC-SMA and HyMOD, we used 11 years of data just for the calibration period as suggested by Yapo et al. (1996) to obtain the stable and reliable parameters. We used the same number of years for the ARX and ANN models in order for their parameters to be calibrated.

5. Analysis and Discussion

As mentioned in previous section, all models needed to be calibrated. The calibration of SAC-SMA and HyMOD were done using the Shuffled Complex Evolution-UA (SCE-UA) developed by Duan et al. (1992). The SCE-UA is known as an effective and efficient global optimization technique which has been tested and used in many other disciplines. The calibration of ARX model and SORB-ANN were done by means of simple least square (SLS) method. The common purpose in hydrological prediction is to maximize the predictive accuracy, precision

and reliability. Although there are many ways to measure the model performance, in this study we used, Nash Sutcliffe (NSE), Root Mean Square Error (RMSE), Correlation coefficient (CC), and Percent Bias (PBIAS) as follows:

$$NSE = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - y_m)^2} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - 1}} \quad (8)$$

$$CC = \frac{\sum_{t=1}^n (y_t - y_m)(\hat{y}_t - \hat{y}_m)}{\sqrt{\sum_{t=1}^n (y_t - y_m)^2 \sum_{t=1}^n (\hat{y}_t - \hat{y}_m)^2}} \quad (9)$$

$$PBIAS = 100 \frac{\sum_{t=1}^n |\hat{y}_t - y_t|}{\sum_{t=1}^n y_t} \quad (10)$$

We used these measures as different objective functions to calibrate each model. Therefore, by having four models and four objective functions, we create a sixteen-member ensemble of individual model predictions for the Leaf River basin. As a result, the model combination using BMA yields sixteen weights which are fixed for the whole period of simulation (calibration or evaluation). For the SBMC method, at each time of simulation, there exists sixteen-member ensemble which are changing throughout the simulation as explained in section 2.2. The verification statistics are seen for the individual models and also the model combination schemes in Figure 4.

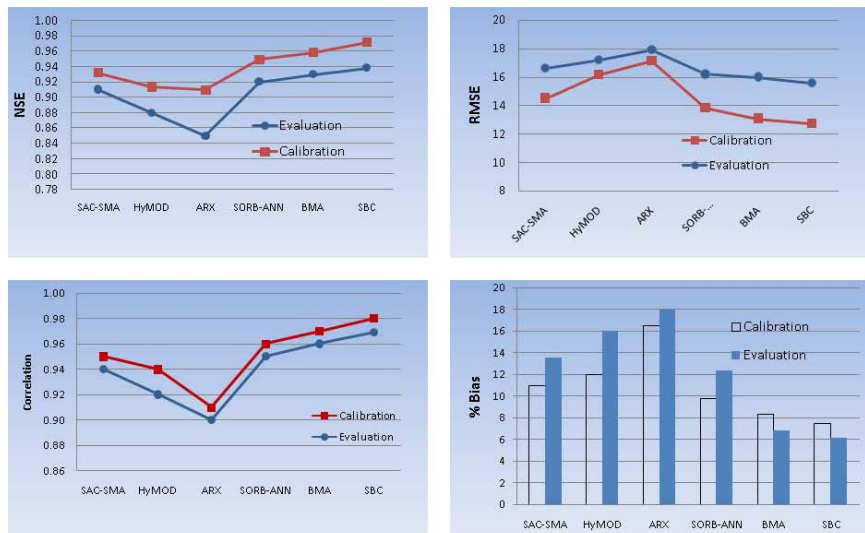


Figure 4. Verification statistics for both Calibration and evaluation periods

As seen in figure 4, the multi-model combination prediction resulted to better performance measures no matter what objective function used for the deterministic predictions resulted from the individual models. In other words, It is also seen that SBC outperforms BMA which is

believed to be due to the flexibility given to the time-varying weights which provides better adaptability of the model with real-time observation and showing that the posterior distribution of the models (weights) are subject to change due to climate variation.

We also extended our verification to probabilistic form. Probabilistic verification methods have been used extensively in the evaluation of meteorological and climate forecasts. One of the measures that is useful in our verification scheme is Ranked Probability Score (RPS) which is essentially the mean-squared error of the probability forecasts averaged over multiple events. In other words, RPS is the mean square error of probabilistic multi-category forecasts where observations are 1 (occurrence) for the observed category and 0 for all other categories. To generate forecast probability from the ensemble of models, the cumulative distribution function of all available historical observations is used to determine threshold streamflow values for non-exceedence probability categories. The pre-specified categories we used in our study are 10%, 35%, 70%, 90% and 100% nonexceedence. For example, the observed value for a given forecast category takes on the value of 1 if the observed flow value is less than the threshold for that category, otherwise the value is 0. The mathematical expression of RPS is given by:

$$RPS_t = \sum_{i=1}^J (F_t^i - O_t^i)^2 = \sum_{i=1}^J [P(\text{forecast} < \text{thresh}_i) - P(\text{observed} < \text{thresh}_i)]^2 \quad (11)$$

where F_t^i is the forecast probability and O_t^i is the observed value at each threshold category, $i=1, \dots, J$. The streamflow value for each day is treated as an event. The average RPS for a group of n evaluation period is given by:

$$RPS_m = \frac{1}{n} \sum_{t=1}^n RPS_t \quad (12)$$

When RPS is viewed in absolute value, it may not be that meaningful, therefore, we use the Rank Probability Skill Score (RPSS) which is a skill score based on RPS values given as:

$$RPSS = \left(1 - \frac{RPS}{RPS_{ref}} \right) \times 100\% \quad (13)$$

RPS_{ref} is the reference RPS value which here we computed from the original model ensemble. For the present study we found the two values of RPSS for each of BMA and SBC methods. $RPSS_{BMA} = 43.8\%$ and $RPSS_{SBC} = 51.2\%$ suggesting that in case of Bayesian Model Averaging,

we get 43.8% better prediction than that of the original ensemble predictions, and for Sequential Bayesian Combination about 51.2% better than the original ensemble predictions is resulted

6. Summary and Conclusion

The motivation for this study stems from the fact that there is no such a perfect hydrologic model available to date that performs best in all conditions. Therefore, we conducted an experiment to compare a well-known technique for multi-model combination, Bayesian Model Averaging (BMA), with Sequential Bayesian Model Combination (SBMC) for the one day ahead streamflow forecasting. Four hydrologic models were employed with different natures from times series type to black box and finally to conceptual models to provide deterministic forecast for the Leaf River Basin. Both deterministic and probabilistic verifications were made. The results show that in its current implementation, SBMC achieves a slightly better performance in terms of forecast accuracy, precision and skill score. It appears that sequential estimation in general can take advantage of structural organization of information content in the data to find the flexible weights that can adapt to hydrometeorological condition better than when the whole process is treated in batch and the weighting parameters are estimated once and for all.

In the present study we used four hydrologic models for one basin; however, more models with different development philosophy can be employed where the procedure is applied over more basins with various climate conditions.

Acknowledgment

The first author would like to acknowledge the partial financial support by NOAA-CPPA funding for this research.

References

- Ajami, N. K., et al., (2007): An integrated hydrologic Bayesian multi-model combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resources Research*, 43, W01403, doi:10.1029/2005WR004745.
- Beven K., (2006): A manifesto for the equifinality thesis, *J Hydrology*; 320(1-2):18–36.
- Boyle, D. P., H. Gupta, S. Sorooshian, et al., (2001): Toward improved streamflow forecasts: Value of semidistributed modeling, *Water Resources Research*, 37, 2749-2759.
- Burnash R. J. C., R. L. Ferreal and R. A. McGuire (1973), A generalized streamflow Simulation System: Conceptual Modeling for Digital Computers, U.S. Department of Commerce, National Weather Service and Department of Water Resources.

- Duan, Q., N.K. Ajami, X., Gao, S. Sorooshian, (2007): Multi-model ensemble hydrologic prediction using Bayesian model Averaging, *Adv. in Water Res.*, 30, 1371-1386.
- Duan, Q., S. Sorooshian, and H. Gupta (1992): Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, *Water Resources Research*, 28, 1015-1031.
- Hong, Y., et al. (2005), Self-organizing nonlinear output (SONO): A neural network suitable for cloud patch-based rainfall estimation at small scales, *Water Resources Research*, 41, W03008, doi:10.1029/2004WR003142.
- Hsu, K., H. Moradkhani, and S. Sorooshian, (2008), A Sequential Bayesian Approach for Hydrologic Model Selection and Prediction, *Water Resources Research*, in Review.
- Kass, R.E., and A. E. Raftery (1995): Bayes factors, *J. Amer. Stat. Assoc.*, 90, 773-795.
- Leamer, E. E., (1978): Specification searches, *Wiley*, 370 pp.
- Moradkhani, H., K. Hsu, H. Gupta, and S. Sorooshian (2004): Improved streamflow forecasting using self-organizing radial basis function artificial neural networks, *J. Hydrol*, 295, 246-262.
- Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. R. Houser (2005a), Dual stateparameter estimation of hydrological models using ensemble Kalman filter, *Adv Water Resour.*, 28(2), 135-147.
- Moradkhani H., K.-L. Hsu, H. Gupta, S. Sorooshian (2005b), Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604.
- Neuman S.P., (2003), Maximum Likelihood Bayesian Averaging of uncertain model predictions, *Stochast Environ Res Risk Assess*, 17, 291-305.
- Raftery A.E., (1993): Bayesian model selection in structural equation models. In K. A. Bollen and J. S. Long (Eds), *Testing Structural Equation Models*, PP. 163-180. Newbury Park, Calif.
- Raftery A.E., Balabdaoui F, Gneiting T, Polakowski M., (2003): Using bayesian model averaging to calibrate forecast ensembles. Technical Report no. 440, Department of Statistics, University of Washington.
- Raftery A.E., Gneiting T, Balabdaoui F, Polakowski M., (2005): Using Bayesian model averaging to calibrate forecast ensembles. *Mon Weather Rev*, 113:1155-74.
- Smith MB, Seo D-J, Koren VI, Reed SM, Zhang Z, Duan Q., (2004): The distributed model intercomparison project (DMIP): motivation and experiment design. *J Hydrology*, 298:4-26.
- Vrugt, J. A., and B. A. Robinson (2007): Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging, *Water Resources Research*, 43, W01411, doi:10.1029/2005WR004838.
- Wood, E. F., et al. (1980): Real-time forecasting/control of water resource systems : selected papers from an IIASA Workshop, October 18-21, 1976 edited by Eric F. Wood ; with the assistance of András Szölösi-Nagy, 1st ed., ix, 330 pp., Pergamon Press, Oxford ; New York.
- Yapo, P. O., et al. (1996): Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data, *J Hydrol*, 181, 23-48.