Derivation of Design Flood Using Multisite Rainfall Simulation Technique and Continuous Rainfall-Runoff Model

Hyun-Han Kwon[1]
Korea Institute of Construction Technology, Goyang-Si, Gyeonggi-Do, 411712, Korea

Abstract: Hydrologic pattern under climate change has been paid attention to as one of the most important issues in hydrologic science group. Rainfall and runoff is a key element in the Earth's hydrological cycle, and associated with many different aspects such as water supply, flood prevention and river restoration. In this regard, a main objective of this study is to evaluate design flood using simulation techniques which can consider a full spectrum of uncertainty. Here we utilize a weather state based stochastic multivariate model as conditional probability model for simulating the rainfall field. A major premise of this study is that large scale climatic patterns are a major driver of such persistent year to year changes in rainfall probabilities. Uncertainty analysis in estimating design flood is inevitably needed to examine reliability for the estimated results. With regard to this point, this study applies a Bayesian Markov Chain Monte Carlo scheme to the NWS-PC rainfall-runoff model that has been widely used, and a case study is performed in Soyang Dam watershed in Korea. A comprehensive discussion on design flood under climate change is provided.

Key Words: Weather state model, downscaling, GCM, large-scale climate pattern

1. Introduction

Estimates of design flood frequencies are routinely required for engineering purposes. Design floods are required for planning and operation measures, structures design, as well as for safety and risk analysis of existing structures. Depending on the characteristics of the system, design floods are frequently provided as a peak discharge value or as a flood hydrograph corresponding to a prespecified return period (*Kwon et al.,* 2007).

The streamflow data to support design purposes in South Korea are very limited. In addition, streamflow data from gauged regions compared to rainfall data is relatively insufficient to analyze flood frequency. Therefore, it is particularly important to utilize stochastic simulation for the floods and rain events in order to quantify the inherent uncertainty and to provide reliable estimates of the characteristic frequencies (*Kwon et al.,* 2007).

In this study we propose to derive the flood frequency curves based on weather state based multisite rainfall generation scheme and Bayesian continuous rainfall-runoff. Two novel approaches allow us to consider a full spectrum of uncertainties from input data to parameter estimation in the models.

The proposed methodology could be viewed as an alternative or supplement of the traditional approaches in order to confidently provide flood-frequency estimates. The advantage of derived distribution technique is that one can account all the aspects of uncertainties in the hydrologic variables. By applying this approach to the rainfall-runoff models, the return periods of simulated floods can be empirically derived.

2. Methodology

2.1 Weather State Based Multisite Rainfall Generation Model (WSMR)

Precipitation is an important component for water resources systems, and daily rainfall series are being used as a main input in hydrologic models. Stochastically generated daily rainfall is usually used to assess water resources systems. Although numerous studies literature (*Rajagopalan et al*., 1996; *Sharma and Lall*, 1999) for daily rainfall generation at a single site were researched in the hydrological and climate literature, keen attention has been paid to the spatial dependence (*Kwon et al*., 2008). The spatial dependence at different sites in a watershed needs to be considered, and this issue is especially critical in simulating rainfall, which is represented by the variability in space and time.

Markov chains have been a popular method for modeling daily precipitation occurrence. Typically a two-state (wet or dry), one-step model is used, and the state transition probabilities (e.g., transition from a wet day to a wet day, a wet day to a dry day) are estimated from the data. One problem with such a description is that the transition probabilities may vary over the year, i.e., the process of precipitation occurrence is nonstationally. There is an implicit assumption that the occurrence process is stationary over the period. This assumption may not be tenable.

In this paper, we developed a multivariate weather state model which is relatively complex model considering weather pattern, nonstationrity and Markov statistics with the intention of providing a practical tool for the simulation of daily rainfall in Korea for use with hydrologic model.

WSMR relate broad scale atmospheric circulation patterns to local rainfall by postulating weather states to act as a link between the two disparate scales. For instance, let $\mathbf{R} = \{R_t^1, \ldots, R_t^n\}$ be a multivariate random vector giving precipitation amounts at a network of n sites. Let $\mathbf{S}_t$ be the weather state at time t and $\mathbf{X}_t \in R^D$ be the vector of atmospheric measures at time t for $1 \le t \le T$.

The two main assumptions on WSMR are: first, $P(\mathbf{R}_t \mid S_1^T, \mathbf{R}_1^{t-1}, \mathbf{X}_1^T) = P(\mathbf{R}_t \mid S_t)$ where $\mathbf{X}_1^T$ indicate the sequence of atmospheric data from time 1 to T (i.e., the length of sequence) and similarly for S_1^T and $\mathbf{R}_1^{t-1}$, second, $P(S_t \mid S_1^{t-1}, \mathbf{X}_1^T) = P(S_t \mid S_{t-1}, \mathbf{X}_t)$.

2.2 Bayesian Rainfall-Runoff Model

Rainfall-runoff models are widely used in understanding and quantifying inflows with the predicted precipitation from large scale climate models, and to provide information that can be used in water resources management. Calibration of rainfall-runoff models with respect to local observational data is used to improve model predictability. In the present study, a Shuffled Complex Evolution Metropolis (SCEM-UA) global optimization algorithm is used for optimization and uncertainty assessment of hydrologic model parameters in rainfall-runoff model. [*Vrugt, et al*., 2003c] combined the strengths of the Monte Carlo Markov Chain (MCMC) sampler with the concept of complex shuffling from SCE-UA to form an algorithm. The SCEM-UA has been successfully used for hydrologic applications such as rainfall-runoff model optimization, large scale streamflow simulation and streamflow forecasting [*Feyen*, *et al.*, 2007; *Vrugt, et al.*, 2003c; *Vrugt, et al.*, 2006a]. The SCEM-UA algorithm is similar to the SCE-UA global optimization method [*Duan, et a*l., 1992], but employs the Metropolis Hastings scheme [*Metropolis, et al.*, 1953] [*Hastings*, 1970] instead of the Downhill Simplex method for population evolution. Hence, The SCEM-UA is able to simultaneously infer both the most likely parameter set and its underlying posterior probability distribution with a single optimization run [Vrugt, et al., 2003a; Vrugt, et al., 2006a; Vrugt, et al., 2006b].

Among many rainfall-runoff models, this study employs the NWS-PC model as a hydrologic model. The NWS-PC, a PC version of NWSRFS (National Weather Service River Forecasting System) developed by National Weather Service, consist of soil moisture and flow routing components. The soil moisture component used the SAC-SMA (Sacramento Soil Moisture Account) model and flow routing component uses the kinematic wave and Muskingum methods in the HEC-1 mode (*Tabios Ⅲ et al*, 1986).

The SCEM-UA algorithm starts with an initial population of parameter sets randomly distributed throughout the given maximum and minimum parameter space. Then the model is run for each parameter set. Therefore, the posterior density $p(\theta \mid \mathbf{x})$ being the correct parameter set given knowledge from measurements $\mathbf{x}$ for each parameter set is estimated from the model output and the measurements using a Bayesian inference scheme. An objective of Bayesian methods is to compute the posterior distribution of the desired variables, in this case the parameters of the annual maximum flood distribution. The posterior distribution $p(\theta \mid \mathbf{x})$ is given by Bayes Theorem as follow:

$$p(\theta|\mathbf{x}) = \frac{p(\theta) \times p(\mathbf{x} \mid \theta)}{p(\mathbf{x})} = \frac{p(\theta) \times p(\mathbf{x} \mid \theta)}{\int_{\Theta} p(\theta) \times p(\mathbf{x} \mid \theta) d\theta} \propto p(\theta) \times p(\mathbf{x} \mid \theta) \tag{3}$$

where θ is the vector of parameters of the distribution to be fitted, Θ is the space parameter, $p(\theta \mid \mathbf{x})$ is the likelihood function, x is the vector of observations and $p(\theta)$ is the prior distribution.

The Gelman and Rubin convergence statistic [*Gelman*, et al., 2003] is calculated on the generated posterior densities to check whether convergence to a stationary target distribution has been achieved. For a more detailed explanation of the SCEM-UA algorithm, including a description of the statistical formulas, please refer to the article by [*Vrugt*, et al., 2003c]. The posterior distributions of thirteen parameters are updated and optimized in the Bayesian framework.

RESULTS AND DISCUSSION

The primary application we consider in this paper is the simulation of the daily rainfall of the Soyang Dam watershed from its 1987-2008 using observed daily rainfall. $BIC_k = 2LL(\Theta_k) - p \log T$ where $LL(\Theta_k)$ is the log- likelihood of the model with k weather states, p is the number of parameters and T is the number of days of observed data used to train the model. The number of weather states in the model has a considerable influence on the performance of the model. A typical approach to the identification of the appropriate number of states is to minimize the Bayesian Information Criterion (BIC). The BIC is used here to select the number of weather states. Five weather states are selected for the each month in WSDM model. Figure 1 presents the frequency and amount of precipitation associated with each state for May.

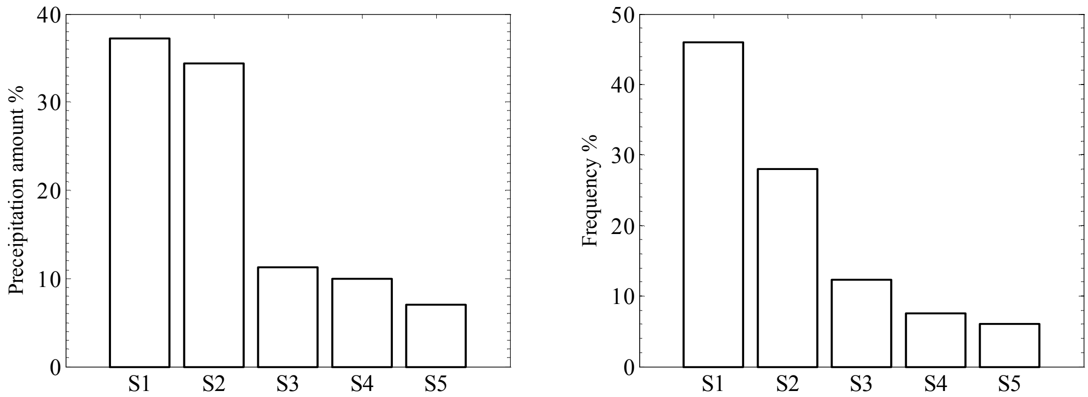

Figure 1. The estimated percentage of amount and frequency of seasonal rainfall according to each state.

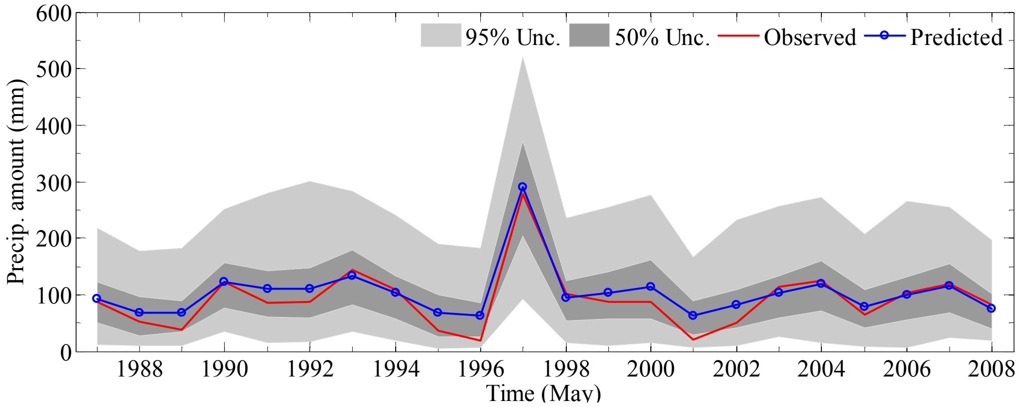

Figure 2. Model performance of the simulated daily rainfall

Selected statistics of rainfall amount as simulated by the WSDM for the May are compared, and the performance in reproducing the amounts at 31-day sequences for May is plotted in Figures 2. For the May season, the correlations with the observed statistic are around 0.9. The model performance is quite good in simulating daily rainfall in terms of reproducing both high- and low-frequency.

This study applies a Bayesian Markov Chain Monte Carlo scheme to the NWS-PC rainfall-runoff model that has been widely used. The NWS-PC model is calibrated against observed daily runoff, and thirteen parameters in the model are optimized as well as posterior distributions associated with each parameter are derived. The Bayesian Markov Chain Monte Carlo shows an improved result in terms of statistical performance measures and graphical examination. Figure 3 shows model calibration results using Bayesian Markov Chain Monte Carlo simulation from 1989 to 1999. Finally, the simulated daily rainfall series are fed into a NWC-PC rainfall-runoff model to generate discharge scenarios. Figure 4 shows the flood frequency curve with the theoretical line of log-normal distribution with probability paper. The proposed approach showed that the overall trend of design flood was similar to the design flood from traditional flood frequency analysis.

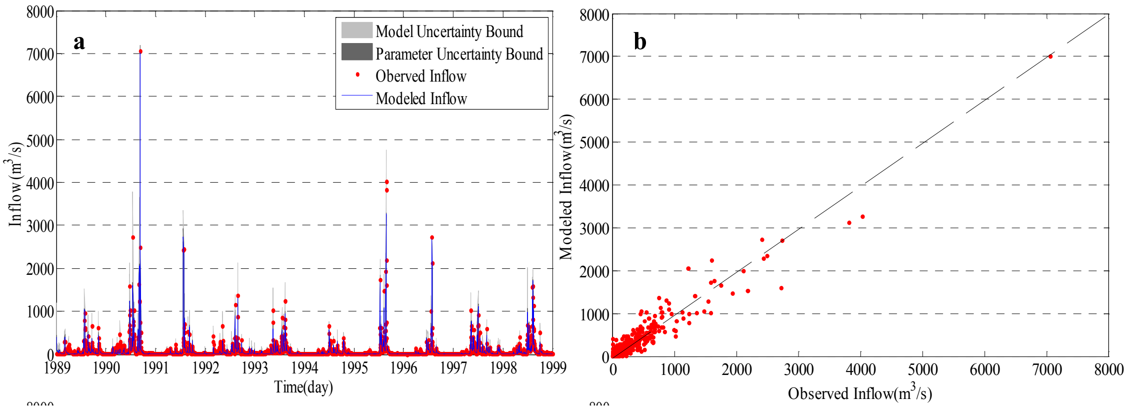

Figure 3. Model calibration results using Bayesian Markov Chain Monte Carlo simulation from 1989 to 1999. a) Comparison between observed flow and modeled flow, b) scatter plot.

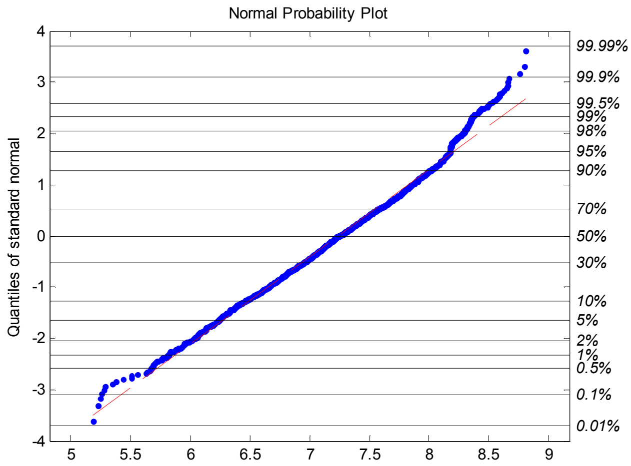

Figure 4. Derived flood frequency curve fitted by lognormal distribution

References

Duan, Q., V.K. Gupta and S. Sorooshian, Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res. 28 (1992) (4), pp. 1015–1031.

Feyen, L., et al. (2007), Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model, J Hydrol, 332, 276-289.

Hastings, W. K. (1970), Monte-Carlo Sampling Methods Using Markov Chains and Their Applications, Biometrika, 57, 97-&.

Kwon, H.-H., Y-I Moon, and A. F. Khalil, 2007. Nonparametric Monte Carlo Simulation for Flood Frequency Curve Derivation: An Application to a Korean Watershed. *Journal of the American Water Resources Association* (JAWRA) 43(5):1316-1328. DOI: 10.1111/j.1752-1688.2007.00115.x

Kwon, H.-H., U. Lall. and J. Obeysekera 2008, Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida, Stoch Environ Res Risk Assess, DOI 10.1007/s00477-008-0270-2.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equations of state calculations by fast computing machines, J. Chem. Phys., 21, 1087–- 1091, 1953.

Rajagopalan. B., Lall, U., and Tarboton, D.G. (1996) Nonhomogeneous markov model for daily precipitation, Journal of hydrologic engineering, Vol 1, 33-40.

Sharma, A. and Lall, U. (1999) A nonparametric approach to daily rainfall simulation. Mathematics and Computers in Simulation 48:367-371.

Tabios Ⅲ, G., Obeysekera, J. T., and Salas. J.D. (1986) Hydrology and water Resource Program, Colorado State Univ., Ft Collins, Colorado.

Vrugt, J. A., et al. (2003a), Identification of rainfall interception model parameters from measurements of throughfall and forest canopy storage, Water Resour Res, 39, -.

Vrugt, J. A., et al. (2003b), Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resour Res, 39, -.

Vrugt, J. A., et al. (2003c), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resour Res, 39, -.

Vrugt, J. A., et al. (2006a), Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting Model, J Hydrol, 325, 288-307.

Vrugt, J. A., et al. (2006b), Real-time data assimilation for operational ensemble streamflow forecasting, J Hydrometeorol, 7, 548-565.