Bayesian Nonstationary Flood Frequency Analysis Using Climate Information

Moon, Young-Il* and Kwon, Hyun-Han**

···

Abstracts

It is now widely acknowledged that climate variability modifies the frequency spectrum of hydrological extreme events. Traditional hydrological frequency analysis methodologies are not devised to account for nonstationarity that arises due to variation in exogenous factors of the causal structure. We use Hierarchical Bayesian Analysis to consider the exogenous factors that can influence on the frequency of extreme floods. The sea surface temperatures, predicted GCM precipitation, climate indices and snow pack are considered as potential predictors of flood risk. The parameters of the model are estimated using a Markov Chain Monte Carlo (MCMC) algorithm. The predictors are compared in terms of the resulting posterior distributions of the parameters associated with estimated flood frequency distributions.

Keyword: Nonstationarity, Flood frequency analysis, climate information

···

1. Introduction

Climate is continuously changing and the signature of change is evident in records of hydrologic cycle and rainfall patterns. The most prominent example of interannual climate variability, El Niño-Southern Oscillation (ENSO), has been extensively investigated for its apparent effects on hydrologic variables. Mounting evidence demonstrates that climate variability also modifies the frequency of extreme hydrologic events (Kwon et al., 2007; Pizaro and Lall, 2002; Sankarasubramanian and Lall, 2003). It was found that large floods at a given site may be related to large scale atmospheric circulation anomalies and studied the use of mixture models for estimating flood frequency. More specifically, large scale circulation anomalies such as the Pacific Decadal Oscillation (PDO) and the Northern Atlantic Oscillation (NAO) which exhibit low frequency modes, may result in decadal scale variability in flood risk.

In recent decades, water resources managers are challenged by the recognition of nonstationarity in the frequency distribution of extreme hydrologic events. While the impact of climate variability on flood risk is acknowledged, promising use of this information, such as seasonal forecasts of flood risk and incorporation in flood management methodology remains rare. The existing techniques for flood frequency analysis regard the time series of annual maximum floods as stationary, an implicit assumption that the distribution of flood flows is

---
* Member·University of Seoul, Professor·E-mail : ymoon@uos.ac.kr
* Member·Columbia University Associate Research Scientist·E-mail : hk2273@columbia.edu

not significantly affected by climatic conditions (Sankarasubramanian and Lall, 2003). However, observed flood records that exhibit cycles, secular trends and low frequency variability provide evidence that the stationarity assumption is not valid in many cases. Furthermore, cyclical modes of variability in the frequency spectrum of hydrologic variables invalidate the common practice of treating hydrologic variables independently from the frequency of their generating mechanisms. In fact, they introduce the potential for prediction, for example of periods of enhanced flood risk.

Previous studies have modeled the impact of exogenous (to the flow record) variables on flood risk. A parametric quantile regression approach and a semiparametric local likelihood approach using climate indices were compared using synthetic data sets and for data from a streamflow gauging station in Montana (Sankarasubramanian and Lall, 2003). This study builds from that approach, comparing available surface and climate data as predictors of annual flood risk. In this study, a Bayesian modeling is used to estimate the mean values and distributions of the parameters of the flood frequency distribution.

2. Nonstationary Flood Frequency Analysis

A Bayesian based Nonstationary Flood Frequency Analysis model is proposed to incorporate nonstationarity using time-dependent climate predictors. The objective of Bayesian inference is to compute the posterior distribution of the desired variables, in this case the parameters of the annual maximum flood distribution. The posterior distribution $p(\theta \mid x)$ is given by Bayes Theorem as follow:

$$p(\theta \mid x) = \frac{p(\theta) \times p(x \mid \theta)}{p(x)} = \frac{p(\theta) \times p(x \mid \theta)}{\int_{\Theta} p(\theta) \times p(x \mid \theta) d\theta} \tag{1}$$

where θ is the vector of parameters of the distribution to be fitted, Θ is the space parameter, $p(\theta \mid x)$ is the likelihood function, x is the vector of observations and $p(\theta)$ is the prior distribution. Here, we present a method for incorporating climate information into updated estimates of the parameters for the extreme value distribution used to represent the annual maximum flood. The Bayesian based Nonstationary Flood Frequency Analysis is expressed in terms of both a location (mean) parameter $\mu(t)$ and a scale (standard deviation) parameter $\sigma(t)$, which both change with time, t. The parameters are hypothesized to be functions of climate indicators, such as SST, snowpack and GCM output that are developed here and others, such as ENSO and PDO, that are generally recognized climate phenomena with expected local impacts. Using the Gumbel extreme type distribution, the distribution of annual peak flood $Z(t)$ can be modeled as follows:

$$Z(t) \, Gumber(\mu(t), \sigma(t)) \tag{2}$$

In the present application, the standard deviation, $\sigma(t)$, was not found to vary

significantly as a function of time. As a result, Equation (2) is changed as follows:

$$Z(t) \ Gumber(\mu(t),\sigma) \tag{3}$$

3. Applications and discussion

The largest flow volumes in Montana streams usually occur during the spring and early summer months with the melting of the winter snowpack. The mechanisms for peak values are heavy rains falling during the spring thaw of a large snowpack. Here we evaluate the influence of the preceding snowpack near the streamflow station. The preceding snowpack (January-February-March-April) are averaged over eleven stations, and then used as a predictor in the model.

The final predictor used for flood risk estimation is output from the ECHAM 4.5 GCM. The ability to use GCM output to reproduce important hydrologic variables such as annual flood risk in present time may guide the use of such output for estimations of future flood risk under climate change scenarios. For this reason, we evaluate the ability of GCM output to produce skillful prediction of annual maximum streamflow. Winter precipitation is used because of the high skill achieved using observed snowpack. The nonstationarity of the annual maximum flood on the Clark Fork River is apparent in the time series depicted in Figure 1. In addition to the stochastic nature of extreme events, there is evidence that the flood time series is influenced by climate anomalies, that is, persistent anomalous conditions of climatologic variables such as SSTs at seasonal or longer timescales.
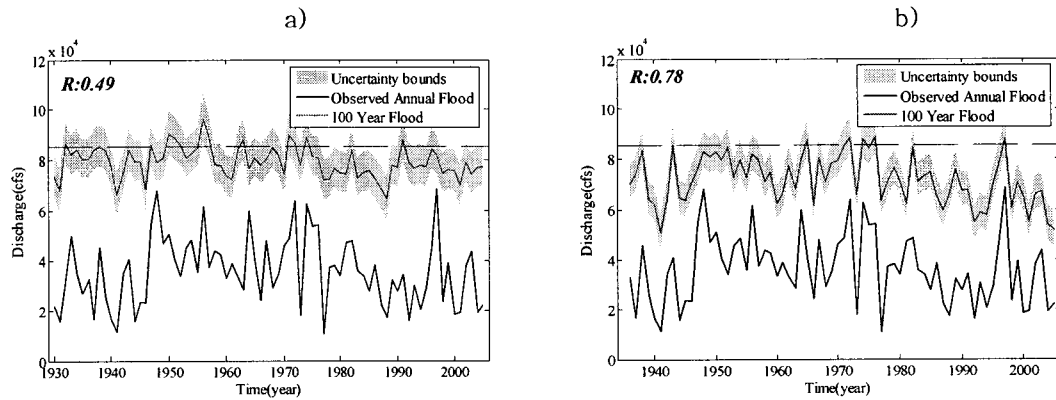

Figure 1. 100 year return period flood with a) SST PC-1 predictor and b) 100 year return period flood with snow predictor

Each of the climate indicators was then evaluated as a predictor of the annual peak flow using the Hierarchical Bayesian model to estimate the parameter values of the extreme value distribution function. Bayesian estimation of the parameter values allows the estimation of confidence intervals for each parameter. Note that uncertainty bands are smallest for the snowpack and SST predictors. Estimates of the annual flood value with 1% probability of occurring in a given year were produced using the Bayesian model based on each predictor. Each time series also includes the confidence interval based on the uncertainties associated with the parameter estimation. The signals show strong nonstationarity. In addition, they demonstrate the skill of the climate predictors in indicating the pronounced increases and decreases in flood risk inter-annually. Given the parameters of Gumbel distribution, Gumbel probability density function with different predictors is shown in Figure 2.
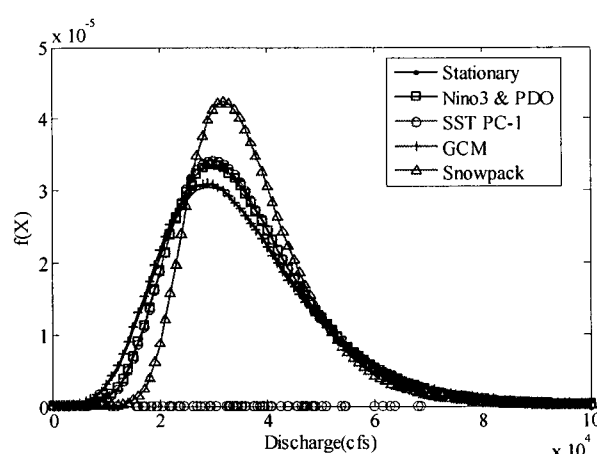

Figure 2. Probability Density Function with Different Predictors

There are two important implications of this work. First, forecasts of changes in the risk associated with the peak annual flood may provide value for reservoir managers and provides some impetus for managing flood dynamically, consistent with its nonstationary nature. Second, reduced uncertainty in the estimation of the 100 year flood, typically used as a design value in engineering, reduces the potential costs of over- and under-design of flood mitigation infrastructure and land use planning.

References

1. Kwon, H.-H., B. Casey, and U. Lall (2007), Nonstationary Flood Frequency Analysis in a changing climate condition: Application to Montana Annual Peak Discharge, *Geophysical Research Letter*, AGU, in the process of submission.
2. Pizaro, G., and U. Lall (2002), El Nino and Floods in the US West: What can we expect?, *Eos Trans.* AGU, 83, 349– 352.
3. Sankarasubramanian, A., and U. Lall (2003), Flood quantiles in a changing climate: Seasonal forecasts and causal relations, *Water Resour. Res.*, 39(5), 1134, doi:10.1029/2002WR001593.