Application of a Full Hierarchical Bayesian Model in Assessing Streamflow Response to a Climate Change Scenario at the Coweeta Basin, NC, USA

Author(s): Wu Wei, James S. Clark and James M. Vose
Source: Journal of Resources and Ecology, 3(2):118-128. 2012.
Published By: Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences
DOI: http://dx.doi.org/10.5814/j.issn.1674-764x.2012.02.003
URL: http://www.bioone.org/doi/full/10.5814/j.issn.1674-764x.2012.02.003

J. Resour. Ecol. 2012 3 (2) 118-128
DOI:10.5814/j.issn.1674-764x.2012.02.003
www.jorae.cn

Article

Application of a Full Hierarchical Bayesian Model in Assessing Streamflow Response to a Climate Change Scenario at the Coweeta Basin, NC, USA

WU Wei[1]*, James S. CLARK[1] and James M. VOSE[2]

1 Nicholas School of the Environment, Duke University, Durham, NC 27708, USA;
2 USDA-Forest Service, Coweeta Hydrologic Laboratory, Otto, NC 28763, USA

Abstract: We have applied a full hierarchical Baysian (HB) model to simulate streamflow at the Coweeta Basin that drains western North Carolina, USA under a doubled CO_2 climate scenario. The full HB model coherently assimilated multiple data sources and accounted for uncertainties from data, parameters and model structures. Full predictive distributions for streamflow from the Bayesian analysis indicate not only increasing drought, with substantial decrease in fall and summer flows, and soil moisture content, but also increase in the frequency of flood events when they were fit with Generalized Extreme Value (GEV) distribution and Generalized Pareto Distribution (GPD) under this doubled CO_2 climate scenario compared to the current climate scenario. Full predictive distributions based on the hierarchical Bayesian model, compared to deterministic point estimates, is capable of providing richer information to facilitate development of adaptation strategy to changing climate for a sustainable water resource management.

Key words: hierarchical Bayes; hydrological modeling; climate change; uncertainty; hydrological extremes

1 Introduction

Potential vulnerability of freshwater supplies to climate change is underscored by the recent widespread droughts and floods over the world. Understanding how the hydrologic cycle responds and contributes to global warming is needed to anticipate security of fresh water supply and assess risks of the resilience of ecosystem services, natural capital, human health, economic prosperity, and social stability from increasing flooding or water shortages under climate change (Vörösmarty *et al*. 2000; Jackson *et al*. 2001; Meybeck and Vörösmarty 2004; Barnett *et al*. 2008; Palmer *et al*. 2008). While global warming will likely intensify hydrological cycles, causing precipitation, and frequency and intensity of hydrological extremes such as floods and droughts to increase globally (Jackson *et al*. 2001; Oki and Kanae 2006; Gedney *et al*. 2006), spatio-temporal variation will determine dependability of fresh water supplies and risk of hydrological extremes (Groisman and Knight 2008) at regional and local scales which are more heterogeneous and more difficult to predict. Many regions, especially temperate ones, would experience increased summer drying from greater evapotranspiration, and maybe lower summer precipitation (Neilson and Marks 1994; Gleick 1987; Jackson *et al*. 2001; Loukas *et al*. 2002; Kilsby *et al*. 2007). In contrast, tropical regions may experience relatively smaller warming-induced changes in the hydrological cycle (IPCC 1996; Jackson *et al*. 2001). The predicted impact of precipitation change on hydrology is more uncertain compared to temperature impact due to the larger uncertainty involved in precipitation predictions under climate change.

Uncertainties for predictions of hydrological cycles under climate change are illustrated in different aspects in hydrological cycles, future climate predictions, and observations used for calibrating hydrological models. With the uncertainties, many variables that control runoff are typically crudely parameterized (Jakeman and Hornberger 1993; Kuczera and Mroczkowski 1998). Bayesian inferences show a promising tool since it can coherently assimilate data from multiple sources and account for the uncertainties in the hydrological system (e.g., Krzysztofowicz 1999; Vrugt and Robinson 2007; Ajami

Received: 2012-02-17 **Accepted:** 2012-05-09
*** Corresponding author:** WU Wei. **Current address**: Department of Coastal Sciences, Gulf Coast Research Laboratory, The University of Southern Mississippi, 703 East Beach Drive, Ocean Springs, MS 39564, USA. Email: wei.wu@usm.edu.

et al. 2007). It combines the advantages of consistency with simplicity of interpretation and capacity for learning from previous knowledge. Hierarchical Bayes represents a hierarchical modeling structure that can deal with the complexity and stochasticity using Bayesian inference. It decomposes the high-dimensional problem into levels in a fully consistent framework (Clark 2005). Three levels are usually decomposed to derive the joint distributions of the "unobservables" (parameters, latent variables to describe unobserved states, observation errors, and process errors): data level (Equation 1a); process level (Equation 1b); and parameter level (Equation 1c) (Clark *et al*. 2001).

$$\begin{aligned} &p(\text{parameters, processes} \mid \text{data, priors}) \\ &\quad \propto p(\text{data} \mid \text{process, data parameters}) \quad \text{-- Equation 1a} \\ &\quad \times p(\text{process} \mid \text{process parameters}) \quad \text{-- Equation 1b} \\ &\quad \times p(\text{all parameters} \mid \text{priors}) \quad \text{-- Equation 1c} \end{aligned} \tag{1}$$

Recent advances in computation, i.e., Markov chain Monte Carlo (MCMC), facilitate the implementation of Hierarchical Bayesian (HB) model and Bayesian inference in hydrological studies (Kuczera and Parent 1998; Campbell *et al*. 1999; Bates and Campbell 2001; Marshall *et al*. 2004). MCMC simulation generates a sample of "unobservables" to produce a Markov chain that converges to a stationary distribution. The algorithm included Gibbs sampling (Gelfand *et al*. 1990) and Metropolis-Hastings sampling (Harrio *et al*. 2001; Marshall *et al*. 2004) etc. The simulated posterior distribution can be summarized, in terms of uncertainties on parameters and latent states (Bates and Campbell 2001) and used to construct predictive distributions (Clark *et al*. 2001).

The prediction of flood and drought frequency will provide useful information on water resource management under changing climate. Failing to do so may have serious effects on environment and economy. Estimation of high flood and drought frequencies requires extrapolation beyond the range of observations, in which parametric probability distributions based on extreme value theory (Leadbetter *et al*. 1983; Embrects *et al*. 1999; Coles 2001) are commonly applied. Fisher-Tippet theorem (Fisher and Tippet 1928) states that the block maxima of a sequence of identically, independently distributed (iid) random variables follows a Generalized Extreme Value (GEV) distribution, while Pickands's study (Pickands 1975) shows that the excesses over a high threshold are Generalized Pareto (GP) distributed, a specialized model for peak over threshold (POT) type data. Other distributions have been used to estimate flood frequency, consisting of log normal, Gumbel extreme value type I (Gumbel EVI), double exponential model, log Pearson type III (LP3), members of the gamma distribution family, the Weibull, Escalante-Sandoval and Raynal-Vilasenor, and the logistic distribution family. For a comprehensive list of the models, please refer to Kidson and Richards (2005).The LP3 has been the official models in the US since 1967, but there are numerous studies concerning the model on underestimating flood events (Kidson and Richards 2005). Droughts in terms of deficit volumes or durations are less studied than floods and only a few studies apply block maxima and peak over threshold models for droughts derived from daily series, mainly due to the presence of a few big and many small drought events that are difficult to fit to a single parametric distribution (Engeland *et al*. 2004). Different studies have developed to solve this problem (Woo and Tarhule 1994; Madsen and Rosbjerg 1998; Kjeldsen *et al*. 2000) by removing the biggest and smallest drought events and fitting multiple distributions.

In this study, we demonstrate the potential consequences of temperature and precipitation changes under a doubled CO_2 climate scenario on hydrology, with a coherent assimilation of the crude specification of processes in hydrological models and multiple data sources using a hierarchical Bayesian analysis.

2 Methods

2.1 Site description

We studied two control watersheds (Watersheds 18 and 27) at the Coweeta Basin in the Nantahala Mountain Range of western North Carolina within the Blue Ridge Physiographic Province (35°03′N, 83°25′W) in the USA. Since 1934, the Coweeta Basin (1626 ha) has been a center of forest hydrological research in the mountains-piedmont of Georgia, South Carolina, North Carolina, and Virginia, and it has been a National Science Foundation (NSF) Long Term Ecological Research Site since 1980 for studies focusing on the impacts of disturbances and environmental gradients on biogeochemical cycles (Swank and Crossely Jr. 1988). Climate at Coweeta Basin is marine humid temperate and characterized by cool summers, mild winters and abundant rainfall in all seasons (Swift *et al*. 1988). Average annual precipitation varies from 1700 mm at low elevations (680 m) to 2500 mm on upper slopes (>1400 m). The hydrology is dominated by rain events, snow usually comprises less than 5% of the precipitation. The underlying bedrock is the Coweeta group (Hatcher Jr. 1979), which consists of quartz diorite gneiss, metasandstone and peltic schist, and quartzose metasandstone (Hatcher Jr. 1988). The regolith of the Coweeta basin is deeply weathered and averages about 7 m in depth.

Watershed 18 (WS18 12.5 ha) and watershed 27 (WS27 38.8 ha) serve as reference watersheds and have been unmanaged since being selectively logged in the early 1900's. The vegetation in both watersheds is mixed hardwoods. The elevation of Watershed 18 (low-elevation watershed) ranges from 726 to 993 m.a.s.l with an average slope of 52 and aspect of north-east. Watershed 27 (high-elevation watershed) has elevation from 1061 to 1454 m.a.s.l with an average slope of 55 and aspect of north-north-east. It was partially defoliated by fall crankerworm infestation from 1975 to 1979.

2.2 Climate change scenario

Our analysis is based on a climate scenario for the region under doubled atmospheric CO_2 concentrations predicted with the regional climate model (RegCM2, spatial resolution of 50 km) (Mearns *et al*. 2003), downscaled from CSIRO general circulation model (GCM) to the southeastern USA. In the RegCM2, temperature increases are predicted throughout the year, from 4 to 5℃ in winter and fall, from 4 to 7℃ in spring, and from 3 to 5℃ in summer compared to the current climate scenario. Precipitation predictions involve larger uncertainties compared to temperature predictions. For our region RegCM2 predicts change from –10% to 10% in winter, from 20% to 40% in spring, from –30% to –10% in summer, and from –10% to 0% in fall. The changes in temperature and precipitation will be sampled using uniform distribution over the ranges above under the climate change scenario in our HB model simulations.

2.3 Process model

We used a parsimonious daily lumped rainfall-runoff model with quick and slow flow components "GR4J" (Modele du Genie Rural a 4 parametres Journalier) (Perrin *et al*. 2003), but allowed for error at this process level (Wu *et al*. 2010). The parsimonious model has four parameters: the maximum capacity of soil moisture storage, a ground water exchange coefficient, the maximum capacity of routing storage, and a time base of a unit hydrograph (i.e. time of concentration of a watershed, defined as time required for water to travel from the most hydraulically remote point in the basin to the basin outlet). The model contains four sub-models: a soil moisture sub-model, an effective precipitation (the proportion of precipitation that could contribute to streamflow) sub-model, a non-linear routing slow streamflow sub-model, and a quick streamflow sub-model. These four sub-models are closely linked to each other. The effective precipitation sub-model uses soil moisture derived from the soil moisture sub-model as input to calculate percolation and then effective precipitation, then the effective precipitation is divided into two flow components: 90% is routed by a unit hydrograph and then a non-linear routing storage and becomes slow flow in the non-linear routing slow streamflow sub-model, and the remaining 10% is routed by a single unit hydrograph and becomes quick flow in the quick streamflow sub-model (Perrin *et al*. 2003; Wu *et al*. 2010).

The details on the GR4J model are contained by Perrin *et al*. (2003) and Wu *et al*. (2010).

2.4 Hierarchical Bayesian Model

Our hierarchical model structure was designed to estimate the components of streamflow generation, including the parameters, latent states of soil moisture and streamflow, and uncertainties of the inputs, parameters, and model structures (Wu *et al*. 2010). We assumed the major uncertainties of the model under current climate scenario include change in soil moisture and streamflow (termed "model misspecification" or "process error"). Thus, the sub-models for soil moisture, slowflow, and quickflow were stochastic, while the sub-model of effective precipitation was treated deterministically. A state space model, which accounted for temporal dependence, was implemented for simulating soil moisture. We also considered the effects of sampling or observation errors for the measurements of precipitation, streamflow and soil moisture. For convenience, we used priors that were conjugate with the likelihood (Calder *et al*. 2003), so that prior and posterior distribution had the same form.

Combining the data, process, and parameter models, we had the joint posterior of the "unobservables" (Equation 2, Wu *et al*. 2010):

$$
\begin{aligned}
&p(k_1,k_2,k_3,k_4,s,q,\sigma_1^2,\sigma_2^2,\tau_1^2,\tau_2^2,\tau_p^2 \mid p^o, temp, y, z, priors)\\
&\propto \prod_{t=1}^{T} N(y_t \mid q_t,\tau_1^2)\prod_{t=1}^{T} N(z_t \mid s_t,\tau_2^2)\\
&\prod_{t=1}^{T} N(q_t \mid f_1(s_t,k_2,k_3,k_4),\sigma_1^2)\\
&\prod_{t=2}^{T} N(s_t \mid f_2(s_{t-1},p_t,temp_t,k_1),\sigma_2^2)\\
&\prod_{t=1}^{T} N(\log(p_t^o) \mid \log(p_t),\tau_p^2)\\
&N(\log(k_1,k_2,k_3,k_4) \mid \log(B),V_B)\\
&IG(\sigma_1^2 \mid \alpha_{\sigma_1^2},\beta_{\sigma_1^2}) \quad IG(\sigma_2^2 \mid \alpha_{\sigma_2^2},\beta_{\sigma_2^2})\\
&IG(\tau_1^2 \mid \alpha_{\tau_1^2},\beta_{\tau_1^2}) \quad IG(\tau_2^2 \mid \alpha_{\tau_2^2},\beta_{\tau_2^2})\\
&IG(\tau_p^2 \mid \alpha_{\tau_p^2},\beta_{\tau_p^2})
\end{aligned}
\tag{2}
$$

where N represents normal distribution, log represents natural logarithm, IG represents inverse gamma distribution, k_1–k_4 are the four parameters in the GR4J model: k_1 denotes the maximum capacity of soil moisture storage, k_2 denotes a ground water exchange coefficient, k_3 denotes the maximum capacity of routing storage, and k_4 denotes a time base of a unit hydrograph, s denotes true soil moisture content, q denotes true log streamflow, σ_1^2 denotes lognormal error in combined slow flow and quick flow submodels, σ_2^2 denotes normal error in soil moisture submodel, τ_1^2 denotes lognormal observation error for streamflow measurements, τ_2^2 denotes normal observation error for soil moisture measurements, τ_p^2 denotes lognormal observation errors for precipitation measurements, p^o denotes observed precipitation, temp denotes temperature, y denotes observed log streamflow, z denotes observed soil moisture content, t denotes time, B is the mean of the priors of (k_1, k_2, k_3, k_4) and it is (350 mm, 0.0001, 90 mm, 1.7 days) based on Perrin *et al*. (2003), V_B is the variance-covaraince matrix of the logarithm of the priors of (k_1, k_2, k_3, k_4) and we assume there is no covariance between the parameters (0 at off-diagonal positions in the matrix) while the variances (diagonal positions in the matrix) are 0.4, 2.0, 0.1 and 0.2 which are weakly informative, and α and β are

the two parameters for the inverse gamma distribution for the observation and process errors.

We implemented Markov Chain Monte Carlo simulation in R (R Development Core Team 2008) to derive the joint posterior (Wu *et al*. 2010). With the estimated joint posterior, we verified the model could simulate streamflow well at other time (Wu *et al*. 2010). The more detailed description of the process and hierarchical model above, how the joint posterior was derived, and how the model was implemented can also be found in Wu *et al*. (2010).

For doubled CO_2 climate scenario, we randomly selected 3000 sets from the derived joint posterior of the unobservables, including the parameters and observation and process errors, to simulate soil moisture and streamflow. For each simulation, the changes in temperature and precipitation were sampled uniformly over the changing ranges predicted by RegCM2 for each season (Mearns *et al*. 2003). All these simulations composed of posteriors of soil moisture and streamflow predictions.

2.5 Flood and drought frequency estimation

The flood extremes are defined to be the highest daily average streamflow values. We applied GEV distribution (Equation 3 showing the cumulative distribution function of GEV) (Engeland *et al*. 2004) to block maxima data and GP distribution (Equation 4 showing the cumulative distribution function of GP) to data over a given threshold to analyze flood frequency using ExtRemes package in R software (Gilleland *et al*. 2004). Block maxima approach is to group the daily streamflow data into blocks of equal length and fit the distribution to the maximums of each block. We have tried different block length: season (3 months), half-a-year, year and found that the medians of yearly maxima from the full distribution of streamflow do not correlate with each other according to the autocorrelation analysis. Considering that the basin has not experienced disturbance since 1927, the medians of yearly maxima can be viewed as i.i.d. and can be fit using GEV distribution. Sometimes using only block maxima data can be wasteful since it ignores much of the data. A more useful way is to fit GP distribution to the exceedances over a given threshold. By using the tools of "mean residual life plot" and "fit threshold ranges (GDP)" provided in ExtRemes package, we chose 5 mm d^{-1} as the threshold.

$$F_{\xi,\mu,\sigma}(x)=\begin{cases}\exp\{-[1+\xi(\dfrac{x-\mu}{\sigma})]^{-\frac{1}{\xi}} & \text{if } \xi\neq 0\\ \exp\{-\exp(-\dfrac{x-\mu}{\sigma})\} & \text{if } \xi=0\end{cases} \quad (3)$$

With support $1+\xi(x-\mu)/\sigma>0$ where ξ is a shape parameter ($\xi>0$: Fréchet distribution; $\xi=0$: Gumbel distribution; and $\xi<0$: Weibull distribution), μ a location parameter and σ a scale parameter.

$$F_{\xi,\beta}(x)=\begin{cases}1-(1+\xi\dfrac{x}{\beta})^{-\frac{1}{\xi}} & \text{if } \xi\neq 0\\ 1-\exp(-\dfrac{x}{\beta}) & \text{if } \xi=0\end{cases} \quad (4)$$

where x is the excess, β is a scale parameter, and again ξ is a shape parameter. It can be shown that the GP distribution is related to the GEV distribution.

A drought event is based on choosing a truncation level q_0, below which the streamflow is considered as a drought occurrence (Engeland *et al*. 2004). Here we chose a

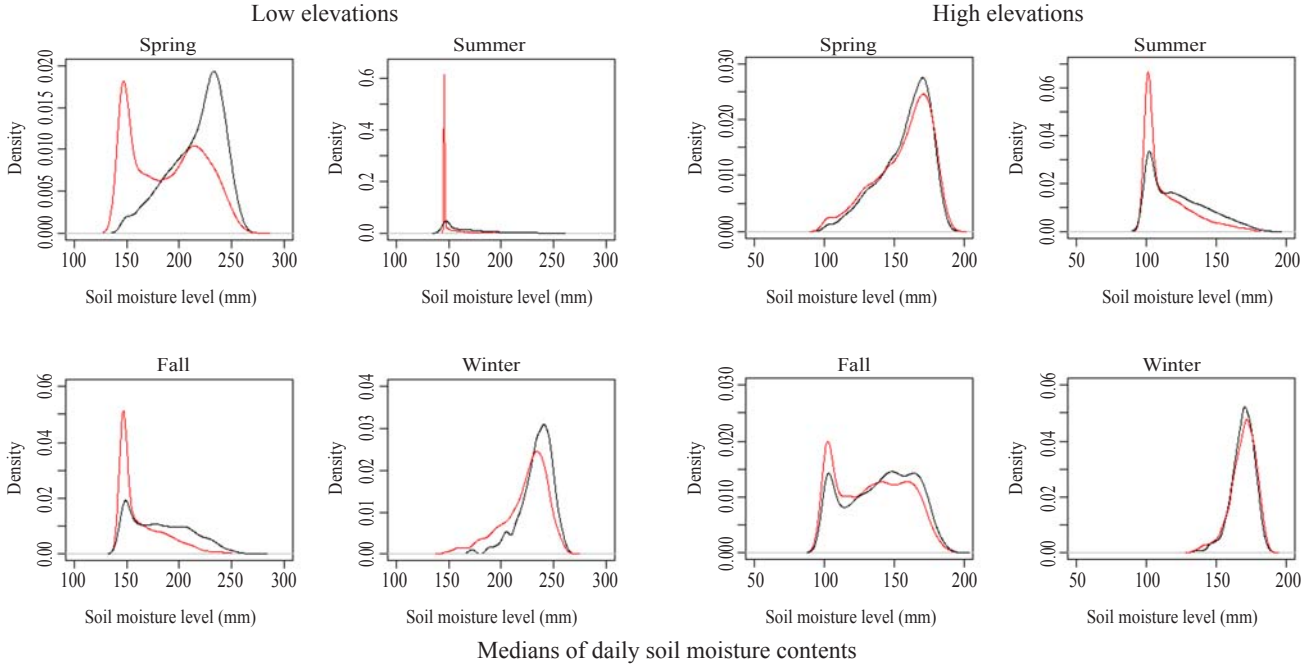

Fig. 1 Density plots of the daily medians of soil moisture contents over the simulation period for each season at the current climate scenario (black lines) and the doubled CO_2 scenario (red lines) at low (the left four panels) and high elevations (the right four panels).

threshold (such as 30-quantile, from Engeland *et al*. 2004) of the streamflow series at the current climate scenario and compare it to the proportion when the streamflow is below this threshold q_0 under the changing climate scenarios.

3 Results and discussion

We have successfully calibrated the HB model to simulate streamflow and soil moisture at the study area under the current climate scenario (Wu *et al*. 2010). Here we focus on reporting and discussing the results of the predictions under the doubled CO_2 climate scenario with full accounting of uncertainties of inputs, parameters, model structures, and climate predictions, deemed critical in light of problems in measurement of key variables.

From predictive distributions, we summarize the medians of daily soil moisture for i) baseline / current climate (from 1960 to 2004 for high elevations, from 1945 to 2004 for low elevations), for ii) the doubled CO_2 scenario, and for iii) differences between high and low elevations. The density of the medians of soil moisture over the 60 / 45 years for each season (spring: March to May; summer: June to August; fall: September to November; winter: December to February) generally shifts to lower values under the changing climate, and the most pronounced decreases occur in summer and fall, and at low elevations where soil moisture is more abundant compared to high elevations (Fig. 1). Predicted changes in daily median soil moisture from baseline climate to the doubled CO_2 scenario range from –29% to 5% at low elevation and from –22% to 7% at high elevation. Under the doubled CO_2 scenario, soil moisture is lower for most of the year with the exception of some days in winter when evapotranspiration is especially low. At baseline climate, soil moisture level for our low-elevation catchment is near the wilting point (11% volume, from SSURGO soil data) 14.1% of time (the 95% credible interval is 13.5%–14.1% with a median of 14.1%). This increases to 35.4% (the 95% credible interval is 34.0%–37.2% with a median of 35.4%) for the doubled CO_2 scenario. At high elevation, the percentage of time when soil moisture is near the wilting point (9% in volume) increases from 8.6% (the 95% credible interval is 8.3%–9.0% with a median of 8.6%) at baseline climate to 15.1% (95% credible interval is 14.0%–16.0% with a median of 15.1%) for the doubled CO_2 scenario.

We summarize the medians, lower and upper 2.5% quantiles in the absolute changes of streamflow for each season under the doubled CO_2 climate scenario compared to the current climate scenario over the simulation period (Fig. 2). The patterns of the changes are different at different seasons. Generally, the streamflow will increase in the spring, decrease in the summer and fall at low and high elevations. In the winter, the streamflow will decrease at low elevations, but increase at high elevations according to the contrast of the probability of percent changes larger than 0 and less than 0. The lower 2.5% quantiles show less uncertainty from their more concentrated probability density plots compared to the medians and upper 2.5% quantiles. Since the large values of streamflow usually involve higher measurement and process errors, the upper 2.5% quantiles show higher uncertainty from their flatter distributions.

Seasonal predictive distributions of relative change in streamflow based on daily medians (Fig. 3) show decreases in summer and fall, increases in spring, and no clear trends in winter, as in the predictive distributions of the absolute change (Table 1) at both low and high elevations. The relative change in streamflow based on the predicted daily lower and upper 2.5% quantiles shows the similar trend

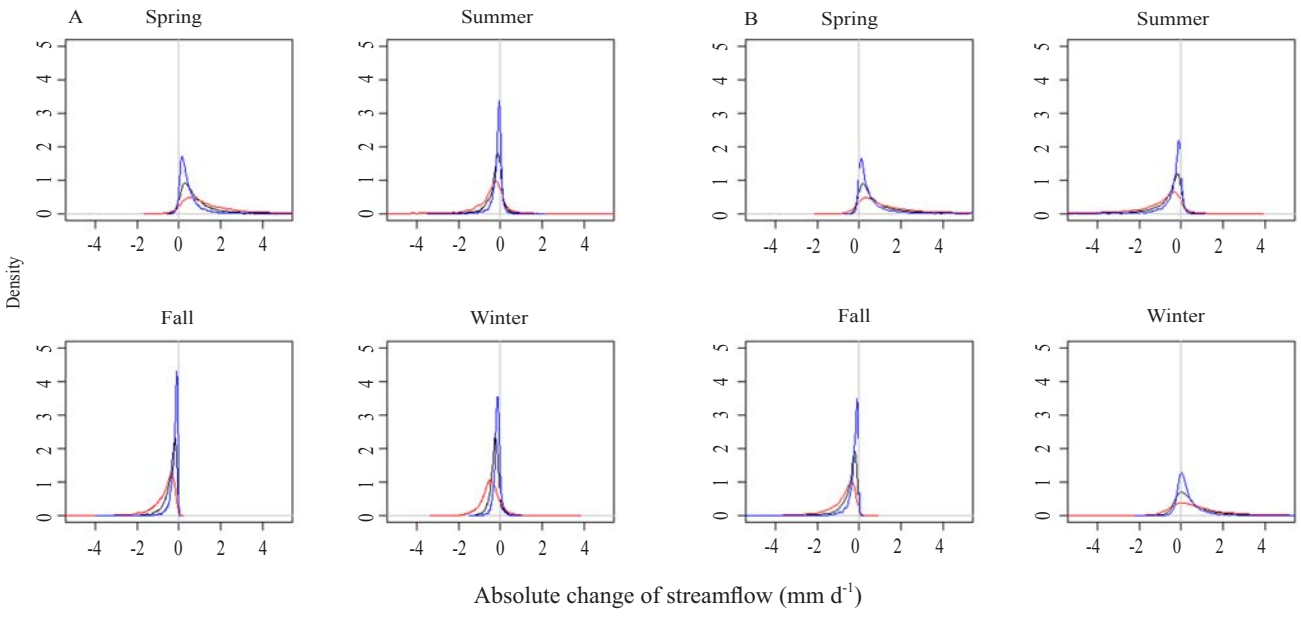

Fig. 2 Density plots of absolute change of streamflow (black: medians; blue: lower 2.5% quantiles; and red: upper 2.5% quantiles) over the simulation period for each season at the doubled CO_2 scenario compared to the current climate scenario at low (A) and high elevations (B).

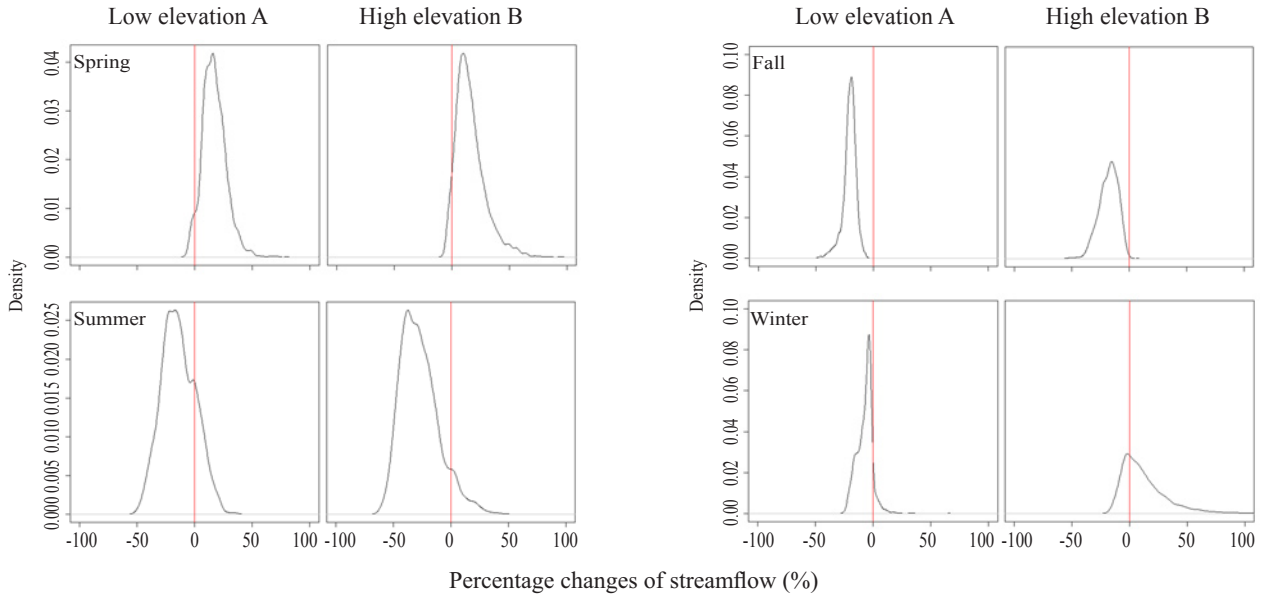

Fig. 3 Probability plots of the percent changes of streamflow based on the predicted daily medians over the simulation period at different seasons at low elevations (A) and high elevations (B) under the double CO2 scenario compared to the current climate scenario.

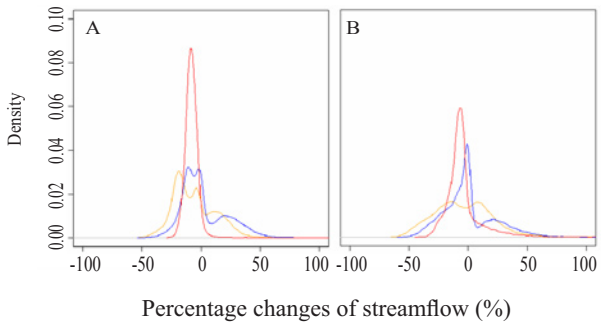

Fig. 4 Probability density plots of percent changes of streamflow based on the predicated daily medians over the simulation period under the doubled CO_2 scenario (orange line), compared with that for scenarios of only temperature change (red line) and only precipitation change (blue line), both shown for low elevations (A) and high elevations (B).

Table 1 Statistics of the percent change of streamflow based on the predictive daily medians over the simulation period under the double CO_2 scenario compared to the current climate scenario (unit: %).

Percent change		Lower 2.5% quantile	Median	Upper 2.5% quantile	Mean
Low elevations	Winter	–20.6	–5.84	5.47	–7.08
	Spring	–1.88	16.1	39.8	17
	Summer	–40.9	–15	15.8	–13.9
	Fall	–35.6	–20	–10.9	–20.7
High elevations	Winter	–11.9	7.64	68.7	13.1
	Spring	–2.19	13.7	49.9	16.3
	Summer	–52.7	–29.3	13.1	–27.1
	Fall	–35.5	–17.1	–4.34	–18

as the daily medians so is not reported here. In fact, Each day has full predictive distributions of soil moisture and streamflow and it is very difficult to present very clearly the full distributions each day with so many days, so we chose to focus on the medians from the posteriors and compare them under doubled CO_2 climate scenario and current climate scenario. Due to a nonlinear relationship between precipitation and streamflow, the decreases in precipitation have more impact than the increases at Coweeta Basin. In fall, the percent decrease of streamflow could be three times as large as the decrease in precipitation, while the increased spring precipitation is expected to have a small benefit because of higher demand by plants.

To understand the relative contributions of changes in precipitation vs. temperature, we compared streamflow for scenarios with changes in temperature or precipitation alone to the baseline climate scenario. Temperature change was found to have a substantially greater impact on median changes, whereas precipitation change had important consequences for extremes for low and high elevations (Fig. 4). Precipitation is responsible for more complex changes, being multi-modal. With precipitation change alone streamflow shows an overall decline with an over-time median decrease rate of 2.7% and 2.0% at low and high elevations respectively. With temperature change along, streamflow shows a larger decrease in over-time median rates of 8.3% and 6.9% at low and high elevations, respectively. Despite relatively lower sensitivity to precipitation change, in terms of these median changes of streamflow, extreme flows are most affected by precipitation change. At low elevations the 95% credible interval of percent streamflow change is (–26.6%, 44.3%) with precipitation changes alone, larger than the 95% credible interval of (–17.4%, 0.559%) predicted with the temperature change alone. At the high elevations, the

95% credible interval is (–36.9%, 51.6%) when there is only precipitation change and (–26.1%, 39.8%) when there is only temperature change. Whether streamflow is more sensitive to precipitation or to temperature depends on the soils, vegetation, and most importantly, climatic regime of the region. In a high-rainfall climate like Coweeta basin, small changes in precipitation may not have substantial impact beyond more frequent flooding. In cool regions increased temperature may have a large impact. For example, streamflow could be substantially more sensitive to temperature than to precipitation in Hudson Bay in Canada (Waggoner 1991). However streamflow is expected to be more sensitive to precipitation than temperature in some dry areas in Australia, China and Japan (Chiew *et al*. 1995; Guo *et al*. 2002; Tanakamaru and Kadoya 1993).

Reduced stream flows and more frequent drought have been predicted in many parts of the world (Fowler *et al*. 2007). In our study region, low flows associated with drought will be more frequent in summer and fall. At low elevations, streamflow is below 1 mm/day 13% of the time (median) for the baseline climate scenario, and 21% of the time for the doubled CO_2 scenario; at high elevations, the percentage of time when streamflow is less than 1 mm/day will increase from 14% at baseline climate scenario to 20% under the doubled CO_2 climate scenario. Moreover, the 30-quantile of the medians of streamflow under the baseline climate scenario for low elevations becomes 37-quantile under the changing climate scenario, indicating that the streamflow will be below a certain value 30% of time under the baseline scenario, and it will be below that same value 37% of time under the changing climate scenario. The same is for high elevations, 30-quantile of the medians of the streamflow under the baseline climate scenario becomes 36-

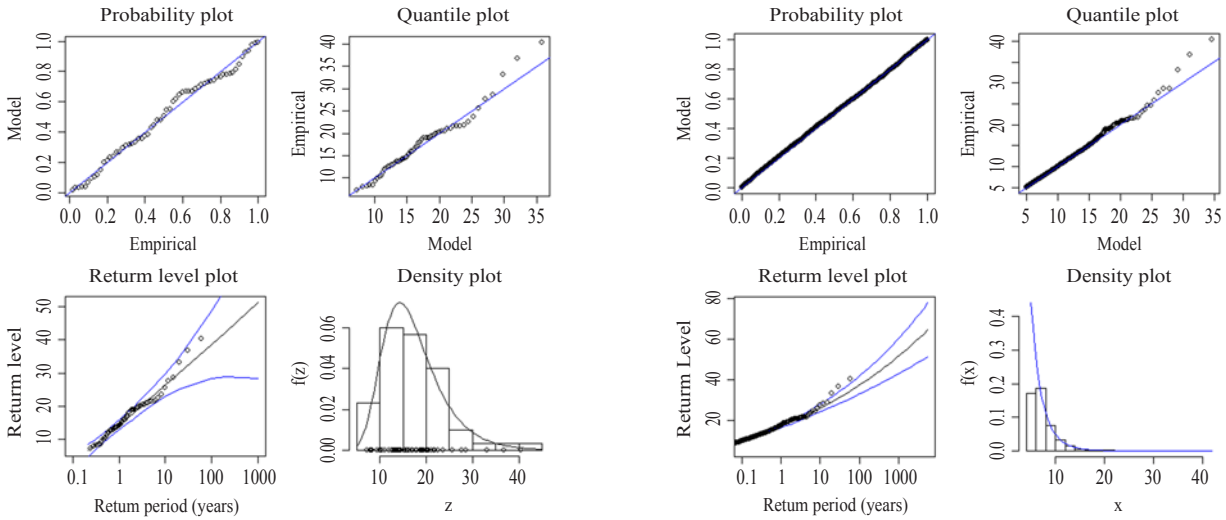

Fig. 5A The diagnostic plots of the GEV distribution (left four) and GP distribution (right four) for low elevations at current climate scenario, showing a good fit.

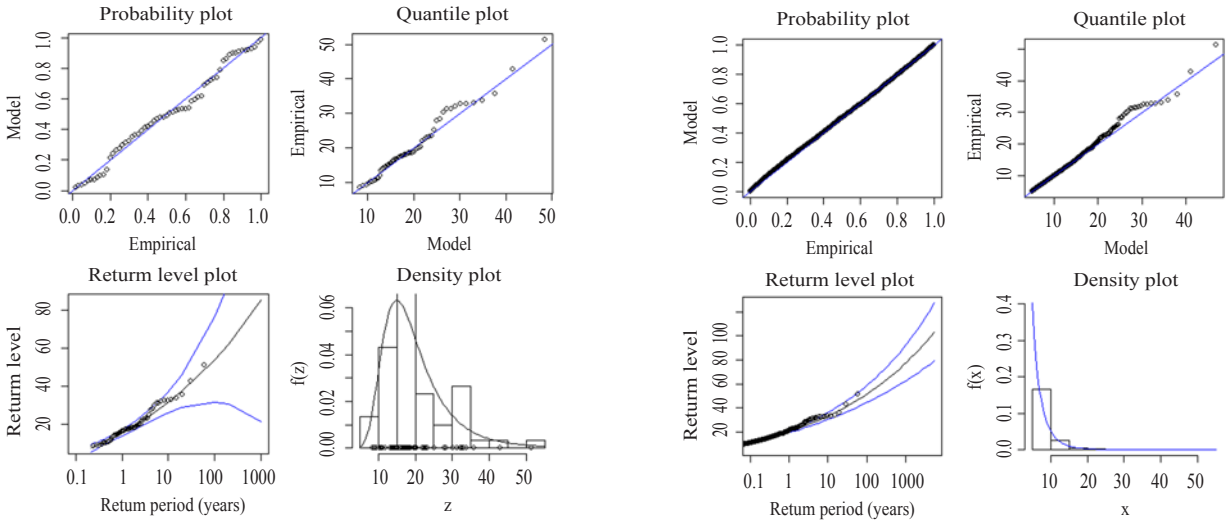

Fig. 5B The diagnostic plots of the GEV distribution (left four) and GP distribution (right four) for low elevations at the doubled CO_2 scenario, showing a good fit.

quantile under the changing climate.

On the other hand, analysis of synoptic flood generation suggests that any future increase in storm frequency and intensity might result in a higher frequency of extreme floods (Longfield and Macklin 1999). Even though precipitation has decreased over the last decade in some areas, there has been no concurrent decrease in the frequency in the extreme flows (Yu and Neil 1993). At the Coweeta Basin, the frequency of flood events will increase under the doubled CO_2 scenario at both low and high elevations. Specifically, exceedance probabilities of peak events are substantially larger under the doubled CO_2 scenario, and streamflow is higher under doubled CO_2 at the same exceedance probabilities, suggesting more frequent flood events in the doubled CO_2 scenario. GEV distribution and GP distribution are generally good fit to block maxima (medians of block maxima from the full distribution of streamflow) and data over a threshold respectively (from probability and quantile plots in Figs. 5 and 6) using maximum likelihood estimator. The following two tables (Tables 2 and 3) summarized the parameters and return level estimates for different return periods from both fitted distributions at low and high elevations.

The lower the return period, the more precise the estimate of the return level. GP distribution shows more precise estimate (smaller 95% CI) than the GEV distribution since more data are used to fit the distribution. Maximum likelihood estimates for the return level for a particular return period is higher under doubled CO_2 scenario than the baseline climate scenario, and there is no overlapping between the 95% confidence intervals for the estimates of the return level under the two climate

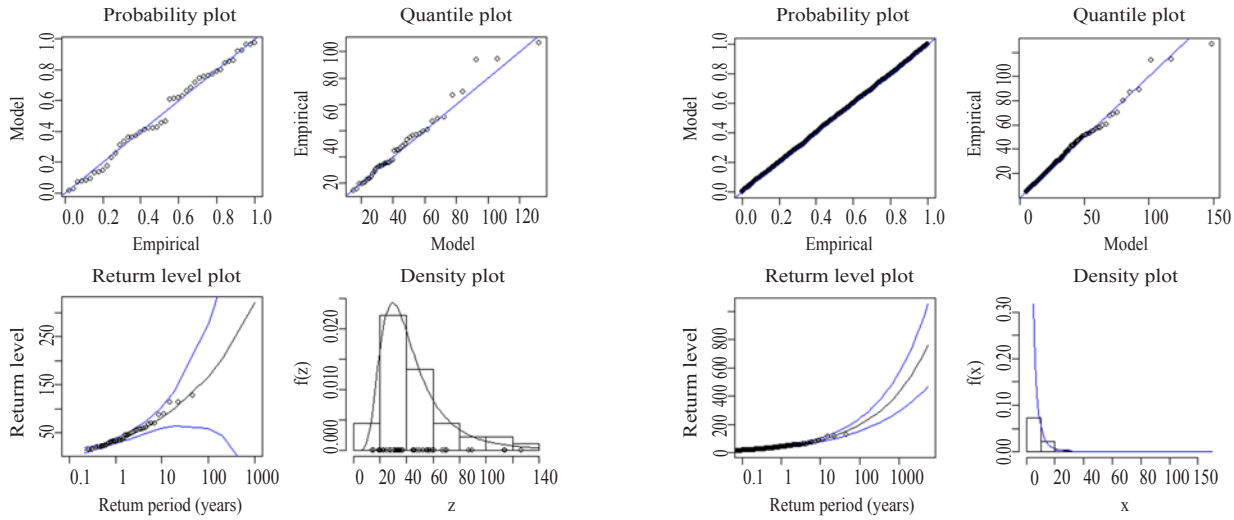

Fig. 6A The diagnostic plots of the GEV distribution (left four) and GP distribution (right four) for high elevations at current climate scenario, showing a good fit.

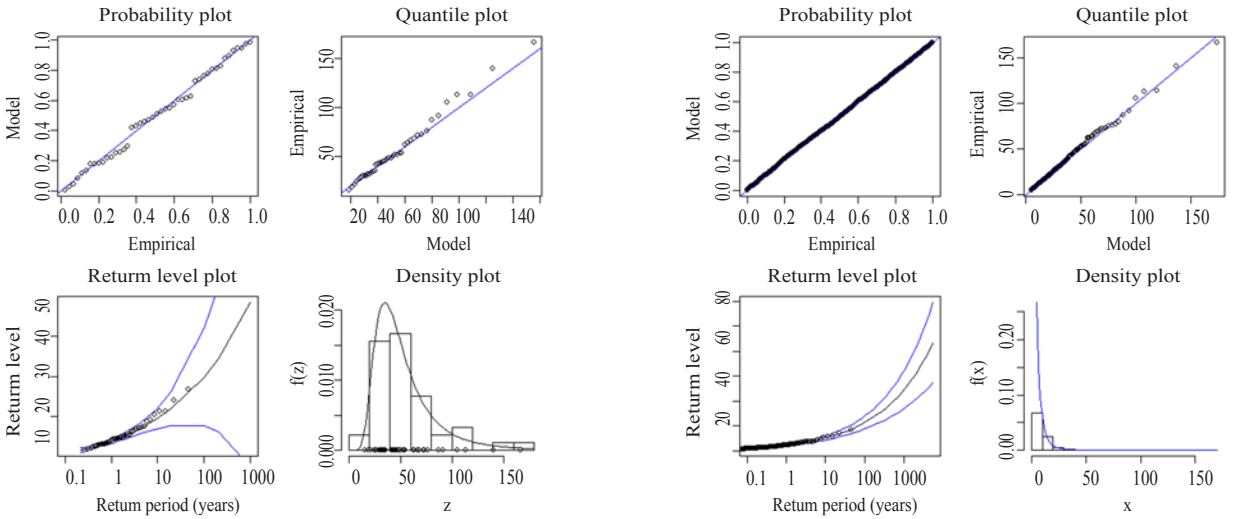

Fig. 6B The diagnostic plots of the GEV distribution (left four) and GP distribution (right four) for high elevations at the doubled CO_2 scenario, showing a good fit.

Table 2 The statistics of the fitting distributions at low elevations for flood frequency analysis (Note CI here represents confidence interval, not credible interval, since it is not Bayesian although the analysis is based on the Bayesian posteriors; 95% CI is from 2.5% quantile to 97.5% quantile. The same applies to Table 3).

Return period (years)	Model	Current / Doubled CO_2	Estimate of return level (mm d^{-1})	95% CI of the estimate (mm d^{-1})	Shape parameter	95% CI of the shape parameter
50	GEV	D*	46.48	(36.77, 69.38)	0.144	(–0.0613, 0.419)
50	GEV	C**	34.77	(29.71, 45.96)	0.014	(–0.143, 0.833)
50	GP	D	45.29	(42.67, 47.90)	0.149	(0.118, 0.183)
50	GP	C	33.72	(31.84, 35.61)	0.099	(0.0685, 0.133)
25	GEV	D	39.59	(32.78, 54.52)	0.144	(–0.0613, 0.419)
25	GEV	C	31.01	(27.10, 38.99)	0.014	(–0.143, 0.833)
25	GP	D	39.69	(37.43, 41.94)	0.149	(0.118, 0.183)
25	GP	C	30.29	(28.62, 31.95)	0.099	(0.0685, 0.133)
10	GEV	D	31.34	(27.16, 39.33)	0.144	(–0.0613, 0.419)
10	GEV	C	26	(23.24, 30.64)	0.014	(–0.143, 0.833)
10	GP	D	33.13	(31.30, 34.95)	0.149	(0.118, 0.183)
10	GP	C	26.09	(24.70, 27.47)	0.099	(0.0685, 0.133)
5	GEV	D	25.57	(22.62, 29.80)	0.144	(–0.0613, 0.419)
5	GEV	C	22.08	(19.92, 24.89)	0.014	(–0.143, 0.833)
5	GP	D	28.72	(27.18, 30.27)	0.149	(0.118, 0.183)
5	GP	C	23.16	(21.97, 24.35)	0.099	(0.0685, 0.133)

* D represents doubled CO_2 scenario; ** C represents the current climate scenario.

Table 3 The statistics of the fitting distributions at high elevations for flood frequency analysis.

Return period (years)	Model	Current / Doubled CO_2	Estimate of return level (mm d^{-1})	95% CI of the estimate (mm d^{-1})	Shape parameter	95% CI of the shape parameter
50	GEV	D	160.09	(112.67, 277.21)	0.258	(–0.119, 0.635)
50	GEV	C	135.58	(94.80, 240.09)	0.25	(–0.218, 0.985)
50	GP	D	180.41	(167.87, 192.96)	0.329	(0.291, 0.369)
50	GP	C	154.63	(143.60, 165.67)	0.334	(0.295, 0.376)
25	GEV	D	128.43	(96.72, 201.38)	0.258	(–0.119, 0.635)
25	GEV	C	109	(81.86, 173.33)	0.25	(–0.218, 0.985)
25	GP	D	142.34	(132.51, 152.16)	0.329	(0.291, 0.369)
25	GP	C	121.76	(113.15, 130.37)	0.334	(0.295, 0.376)
10	GEV	D	93.9	(75.85, 130.40)	0.258	(–0.119, 0.635)
10	GEV	C	79.81	(64.40, 111.28)	0.25	(–0.218, 0.985)
10	GP	D	103.65	(96.59, 110.70)	0.329	(0.291, 0.369)
10	GP	C	88.48	(82.33, 94.64)	0.334	(0.295, 0.376)
5	GEV	D	71.97	(60.21, 91.34)	0.258	(–0.119, 0.635)
5	GEV	C	61.13	(50.99, 77.90)	0.25	(–0.218, 0.985)
5	GP	D	81.22	(75.77, 86.67)	0.329	(0.291, 0.369)
5	GP	C	69.28	(64.54, 74.02)	0.334	(0.295, 0.376)

scenarios by fitting GP distribution, indicating return level is significantly larger under the doubled CO_2 scenario than the baseline climate scenario, and therefore the flood frequency is larger. It is not possible to determine the block maxima follow a Fréchet (ξ>0), Gumbel (ξ=0), or Weibull (ξ<0) distribution since the shape parameter (ξ) ranges from negative to positive. Since we have only 45 years' streamflow data for high elevations, and 60 years' data for low elevations, and the streamflow at higher elevations is larger than at low elevations (also see the results by Wu *et al*. 2010), the estimate of the return level are involved with more uncertainties (larger 95% CI range) at high elevations, especially for a return period of 50 years.

The Little Tennessee Basin (1797 square mile) where our study areas drain has experienced significant population growth and increased demand of freshwater resources for municipal- industrial- recreational- agricultural uses. Changes in streamflow generation in headwater catchments can have potentially devastating implications downstream. The prediction of both increased drought and flood events, together with lower streamflow in the summer and fall as early as 2100 when CO_2 concentration in the atmosphere is likely to double (Watson and Core Writing Team 2001) suggests that water resource managers will be challenged to meet the demands of a rapidly growing population for drinking water and flood

protection. According to the USGS (available at http://ga.water.usgs.gov/edu/qa-home-percapita.html, last access on April 26, 2012), the average American generally uses from 80–100 gallons of water each day. 79 493 population lived in Little Tennessee Basin from 2000 Census and it is expected to be more now. The study areas only occupy 0.011% of the whole basin, but the summer median flows (assuming the flow can totally be used for water supply) can support 1.2% to 3.1% of the population of the whole basin in water use, and the fall median flows can support 1.7 to 4.2% of the population. 15% of the streamflow reduction at low elevations and 29% reduction at high elevations in the summer under the doubled CO_2 climate scenario will reduce the population the watersheds can support to 0.9% to 2.2%, a reduction of 25% to 29% from the current climate scenario. 23% reduction in streamflow at low elevations and 18% reduction at high elevations in the fall under the doubled CO_2 climate scenario will reduce the population the watersheds can support to 1.4%–3.5%, a reduction of 17%–18% from the current climate scenario. The impact of possible climate change will be likely to couple with land use change in the basin to affect water supply even further.

4 Conclusions

Analysis shows that doubled CO_2 in atmosphere may exaggerates extremes in both soil moisture and runoff, with more frequent, intense droughts and floods from the analysis at both low and high-elevation watersheds. Informed management decisions will require accurate and reliable predictions of future conditions. Hierarchical Bayesian models integrating multiple long term data sets with scenarios of future change provide full predictive distributions with uncertainties from different sources accounted for. Compared to deterministic point estimates, the richer information from full distributions can be reliable basis for a more informed sustainable water resource management, helping resource managers anticipate hydrological extremes under climate change and adapt more effectively to climate change to ensure water security.

Acknowledgements

This research was made possible by grants from the National Science Foundation (NSF) through Coweeta Long Term Ecological Research (LTER). We thank USDA-Forest Service Coweeta Hydrologic Laboratory for the support of the data. We especially thank Drs. Chelcy FORD and Katherine ELLIOTT at Coweeta Hydrological Laboratory for the insightful discussion on hydrological cycles at Coweeta basin. We appreciate the critical comments from the anonymous reviewer and the editor, which helped improve the manuscript.

References

Ajami N K, Q Duan and S Sorooshian. 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research*, 43: W01403. doi: 464 1029/2005WR004745

Barnett T P, D W Pierce, H G Hidalgo, *et al*. 2008. Human-induced changes in the hydrology of the western United States. *Science*, 319: 1080-1083.

Bates B C and E P Campbell. 2001. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Water Resources Research*, 37 (4): 937-947.

Calder K, M Lavine, P Mueller and J S Clark. 2003. Incorporating multiple sources of stochasticity in population dynamic models. *Ecology*, 84: 1395-1402.

Campbell E P, D R Fox and B C Bates. 1999. A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Water Resources Research*, 35 (1): 211-220.

Chiew F H S, P H Whetton, T A McMahon and A B Pittock. 1995. Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments. *Journal of Hydrology*, 167: 121-147.

Clark J S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8: 2-14. doi: 10.1111/j.1461-0248.2004.00702.x

Clark J S, S R Carpenter, M Barbar, *et al*. 2001. Ecological forecasts: An emerging imperative. *Science*, 293: 657-660.

Coles S G. 2001. An Introduction to Statistical Modelling of Extreme Values. London: Springer.

Embrects P, C Klüppenberg and T Mikosch. 1999. Modelling Extremal Events for Insurance and Finance. Berlin: Springer.

Engeland K, H Hisdal and A Frigessi. 2004. Practical extreme value modeling of hydrological floods and droughts: A case study. *Extremes*, 7: 5-30.

Fisher R A and L H C Tippet. 1928. Limiting forms of the frequency distributions of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24: 180-190.

Fowler H J, C G Kilsby and J Stunell. 2007. Modelling the impacts of projected future climate change on water resources in north-west England. *Hydrology and Earth System Sciences*, 11: 1115-1126.

Gedney N, P M Cox, R A Betts, *et al*. 2006. Detection of a direct carbon dioxide effect in continental river runoff records. *Nature*, 439: 835-838.

Gelfand A E, S E Hills, A Racine-Poon and A F M Smith. 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, 85: 972-985.

Gilleland E, R Katz, G Young. 2004. extRemes: Extreme value toolkit, R package version 1.59. http://www.assessment.ucar.edu/toolkit/

Gleick P H. 1987. Regional hydrologic consequences of increases in atmospheric CO_2 and other trace gases. *Climatic Change*, 10: 137-160.

Groisman P Y and R W Knight. 2008. Prolonged dry episodes over the conterminous United States: new tendencies emerging during the last 40 years. *Journal of Climate*, 21: 1850-1862.

Guo S, Wang J, Xiong L, Ying A and Li D. 2002. A macro-scale and semi-distributed monthly water balance model to predict climate change impacts in China. *Journal of Hydrology*, 268: 1-15.

Haario H, E Saksman and J Tamminem. 2001. An adaptive metropolis algorithm. *Bernoulli*, 7(2): 223-242.

Hatcher Jr. R D. 1979. The Coweeta Group and Coweeta syncline: Major features of the North Carolina - Georgia Blue Ridge. *Southeastern Geology*, 21(1): 17-29.

Hatcher Jr. R D. 1988. Bedrock geology and regional geologic setting of Coweeta Hydrologic Laboratory in the eastern Blue Ridge. In: Swank W T and D A Crossley Jr. (ed.). Ecological Studies. New York: Springer-Verlag, 81-92.

IPCC (Intergovernmental Panel on Climate Change). 1996. Climate change 1995: The science of climate change. Cambridge, UK: Cambridge University Press, 529.

Jackson R B, S R Carpenter, C N Dahm, *et al*. 2001. Water in a changing world. *Ecological Applications,* 11(4): 1027-1045.

Jakeman A J and G M Hornberger. 1993. How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, 29: 2637-2649.

Kidson R and K S Richards. 2005. Flood frequency analysis: assumptions and alternatives. *Progress in Physical Geography*, 29(3): 392-410.

Kilsby C G, S S Tellier, H J Fowler and T R Howels. 2007. Hydrological impacts of climate change on the Tejo and Guadiana Rivers. *Hydrology and Earth System Sciences*, 11(3): 1175-1189.

Kjeldsen T R, A Lundorf and D Rosbjerg. 2000. Use of a two-component exponential distribution in partial duration modelling of hydrological droughts in Zimbabwean Rivers. *Hydrological Sciences Journal*, 45(2): 285-298.

Krzysztofowicz R. 1999. Bayesian theory of probabilistic forecasting via deterministic hydrological model. *Water Resources Research*, 35(9): 2739-2750.

Kuczera G and E Parent. 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The metropolis algorithm. *Journal of Hydrology*, 211: 69-85.

Kuczera G and M Mroczkowski. 1998. Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research*, 34, 1481-1489.

Leadbetter M R, G Lindgren and H Rootzén. 1983. Extremes and Related Properties of Random Sequences and Processes. New York: Springer.

Longfield S A and M G Macklin. 1999. The influence of recent environmental change on flooding and sediment fluxes in the Yorkshire Ouse basin. *Hydrological Processes*, 13: 1051-1066.

Loukas A, L Vasiliades and N R Dalezios. 2002. Climatic impacts on the runoff generation processes in British Columbia, Canada. *Hydrology and Earth System Sciences*, 6(2): 211-228.

Madsen H and D Rosbjerg. 1998. A regional Bayesian method for estimation of extreme streamflow droughts. In: Parent E, P Hubert, B Bobé and J Miquel (ed.). Statistical and Bayesian Methods in Hydrological Sciences, Technical documents in hydrology, series 20, Paris, France, 327-340.

Marshall L, D Nott and A Sharma. 2004. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research*, 40. doi: 55910.1029/2003WR002378

Mearns L O, F Giorgi, L McDaniel and C Shields. 2003. Climate scenarios for the Southeast US based on GCM and regional modeling simulations. *Climatic Change*, 60: 7-25.

Meybeck M, C J Vörösmarty. 2004. Part D. IGBP Global Change Series. In: Kabat P, et al. (ed.). Vegetation, Water, Humans and the Climate: A New Perspective on an Interactive System. Berlin: Spring-Verlag, 294-480.

Neilson R P and D Marks. 1994. A global perspective of regional vegetation and hydrologic sensitivities from climatic change. *Journal of Vegetable Science*, 5: 715-730.

Oki T and S Kanae. 2006. Global hydrological cycles and world water resources. *Science*, 313: 1068-1072.

Palmer M A, C A R Liermann, C Nilsson, *et al*. 2008. Climate change and the world's river basins: anticipating management options. *Frontiers in Ecology and the Environment*, 6(2): 81-89. doi: 10.1890/060148.

Perrin C, C Michel and V Andréassian. 2003. Improvement of a parsimonious model for streamflow simulation, *Journal of Hydrology*, 279: 275-289.

Pickands J. 1975. Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1): 119-131.

R Development Core Team. 2008. R: A language and environment for statistical computing, R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org

Swank W T and D A Crossley Jr. 1988. Forest hydrology and ecology at Coweeta. In: Swank W T and D A Crossley Jr. (ed.). Ecological Studies. New York: Springer-Verlag, 3-16.

Swift Jr. L W, G B Cunningham and L E Douglass. 1988. Climatology and hydrology. In: Swank W T and D A Crossley Jr. (ed.). Ecological Studies. New York: Springer-Verlag, 35-55.

Tanakamaru H and M Kadoya. 1993. Effects of climate change on the regional hydrological cycle in Japan. In: Bolle H J, R A Feddes and J D Kalma (ed.). Exchange processes at land surface for a range of space and time scales: proceedings of an international symposium held at Yokohama, Japan, 13-16 July, IAHS, 535-542.

Vörösmarty C J, P Green, J Salisbury, R B Lammers. 2000. Global water resources: Vulnerability from climate change and population growth. *Science*, 289: 284-288.

Vrugt J A and 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

Waggoner P E. 1991. U.S. water resources versus an announced but uncertain climate change. *Science*, 251: 1002.

Watson R T and Core Writing Team. 2001. IPCC Third Assessment Report - Climate change 2001: Synthesis report. Cambridge UK: Cambridge University Press.

Woo M K and A Tarhule. 1994. Streamflow droughts of northern Nigerian rivers. *Hydrological Sciences Journal*, 39(1): 19-34.

Wu W, J S Clark, J M Vose. 2010. Assimilating multi-source uncertainties of a parsimonious conceptual hydrological model using hierarchical Bayesian modeling. *Journal of Hydrology*, 394: 436-446.

Yu B and D T Neil. 1993. Long-term variations in regional rainfall in the south-west of Western Australia and the difference between average and high intensity rainfalls. *International Journal of Climatology*, 13: 77-88.

应用层次贝斯模型研究气候变化对威塔流域流量的影响

吴 蔚[1], James S. CLARK[1], James M. VOSE[2]

1 美国杜克大学环境学院，达勒姆，NC 27708，美国；
2 美国农业部林务局威塔水文实验室，奥托，NC 28763，美国

摘 要：我们应用层次贝叶斯模型模拟大气二氧化碳浓度加倍可能对美国北卡罗来纳州西部威塔（Coweeta）流域水文的影响。这个模型整合了多重数据来源并且同时考虑了数据，参数和模型结构的不确定性。贝叶斯分析的预测分布显示流量和土壤含水量在秋季和夏季将明显下降，这将造成这两个季节更严重的干旱。同时我们用通用极值分布（Generalized Extreme Value distribution）和通用普拉托分布（Generalized Pareto distribution）分析预测流量，结果显示洪水频率也会增减。层次贝叶斯模型，和许多只能得到最佳参数估计的水文模型相比，能提供更丰富的信息，包括预测的不确定性。这将有助于可持续水资源管理的大前提下发展应对气候变化的措施。

关键词：层次贝叶斯模型；水文模型；气候变化；不确定性；水文极值