Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change

Scott Steinschneider,[1] Austin Polebitski,[1] Casey Brown,[1] and Benjamin H. Letcher[2]

[1]   The cascade of uncertainty that underscores climate impact assessments of regional hydrology undermines their value for long-term water resources planning and management. This study presents a statistical framework that quantifies and propagates the uncertainties of hydrologic model response through projections of future streamflow under climate change. Different sources of hydrologic model uncertainty are accounted for using Bayesian modeling. The distribution of model residuals is formally characterized to quantify predictive skill, and Markov chain Monte Carlo sampling is used to infer the posterior distributions of both hydrologic and error model parameters. Parameter and residual error uncertainties are integrated to develop reliable prediction intervals for streamflow estimates. The Bayesian hydrologic modeling framework is then extended to a climate change impact assessment. Ensembles of baseline and future climate are downscaled from global circulation models and are used to drive simulations of streamflow over parameters drawn from the posterior space. Time series of streamflow statistics are calculated from baseline and future ensembles of simulated flows. Uncertainties in hydrologic model response, sampling error, and the range of future climate projections are integrated to help determine the level of confidence associated with hydrologic alteration between baseline and future climate regimes. A case study is conducted on the White River in Vermont, USA. Results indicate that the framework can be used to present a reliable depiction of the range of hydrologic alterations that may occur in the future.

1.   Introduction

[2]   The threat of nonstationary hydrology has motivated significant research efforts investigating the potential impacts of climate change on regional hydrology and implications for local water resource systems. Despite these efforts, uncertainty in both future climate conditions and regional hydrologic response confounds the interpretation of results and diminishes their utility in water resources planning [*Lopez et al.*, 2009]. A systematic approach is required to account for the uncertainty in hydrologic impact assessments so that decision makers can consider adaptation strategies contextualized by the uncertainty in design statistics critical to the decision-making process. In this paper we propose a statistical framework that quantifies several sources of uncertainty in long-range projections of hydrologic alteration, including uncertainties in future climate, hydrologic model predictive skill, model parameterization, and sampling error of estimated hydrologic statistics. These uncertainties are integrated to develop a probabilistic description of potential alterations to regional hydrology useful for water resources planning.

[3]   In the vast majority of studies, hydrologic alteration under climate change is assessed using future climate scenarios, as simulated by global circulation models (GCMs), that are downscaled to a location of interest and used to force a regional hydrologic model. The simulated hydrologic response is then compared to a baseline response based on historic climate data, and measures of hydrologic alteration are computed [*Gleick*, 1986]. There are multiple sources of uncertainty that degrade this process, including those associated with the GCMs (i.e., inaccuracy at subcontinental scales, inconsistencies across models, parameterization, uncertain boundary conditions, difficulty in assessing predictive skill), the ambiguity between different downscaling techniques, and the hydrologic model (i.e., model structure, input and output data used for calibration, parameterization) [*Wood et al.*, 1997]. GCM accuracy and consistency, along with the choice of downscaling methodology, are considered to be the primary sources of uncertainty and have garnered significant research attention [*Räisänen and Palmer*, 2001; *Palmer and Räisänen*, 2002; *Piani et al.*, 2005; *Stainforth et al.*, 2005; *Fowler et al.*, 2007; *Stainforth et al.*, 2007a; *Lopez et al.*, 2009]. Errors associated with the hydrologic model, however, have received less emphasis in studies

[1]Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, Massachusetts, USA.
[2]S. O. Conte Anadromous Fish Research Center, Leetown Science Center, U.S. Geological Survey, Turners Falls, Massachusetts, USA.

Corresponding author: A. Polebitski, Department of Civil and Environmental Engineering, University of Massachusetts Amherst, 130 Natural Resources Rd., Amherst, MA 01002, USA. (polebitski@ecs.umass.edu)

considering hydrologic alteration under climate change. In the majority of climate change impact assessments, hydrologic simulations of future climate are treated largely as deterministic output that can be used to directly identify hydrologic alterations [*Chao*, 1999; *Hamlet and Lettenmaier*, 1999; *Lettenmaier et al.*, 1999; *Nijssen et al.*, 2001]. Some studies have explored the impacts of hydrologic model uncertainty on climate impact assessment results, but they often only investigate uncertainties in parameterization [*Arnell*, 1999; *Cameron et al.*, 2001; *Wilby*, 2005], model structure [*Boorman and Sefton*, 1997; *Jiang et al.*, 2007], or a combination of both [*Wilby and Harris*, 2006; *Kay et al.*, 2009; *Prudhomme and Davies*, 2009a, 2009b], and almost never formally account for prediction error, which can often dominate total model uncertainty [*Stedinger et al.*, 2008].

[4] While parameter and structural errors are important components of the total uncertainty in hydrologic model results, accounting for these uncertainties alone may not guarantee reliable predictive bounds for streamflow estimates. For a watershed exhibiting significant heterogeneity or unexplainable behavior, many types of hydrologic response may be challenging to simulate even with an ensemble of model structures or parameterizations. The assumption that a set of hydrologic models with multiple parameterizations is complete enough to reliably bound true hydrologic response is difficult to verify [*Renard et al.*, 2010]. This is especially true if the models struggle to reproduce certain aspects of the observed streamflow and exhibit errors that vary across the magnitude and timing of hydrologic responses. To generate reliable predictive bounds, a formal quantification of residual error is needed. If predictive uncertainty associated with the hydrologic model is not formally addressed and propagated through climate change impact analyses, claims of hydrologic alteration from such studies can be overstated and misguide water resources decision makers.

[5] In a related line of research, predictive uncertainty in hydrologic modeling has been extensively explored and mature methods for quantifying error have been developed. Early efforts focused on pseudo-Bayesian methods [*Beven and Binley*, 1992; *Beven and Freer*, 2001], and later more formal Bayesian techniques emerged to properly account for both residual and parameter uncertainties [*Bates and Campbell*, 2001; *Marshall et al.*, 2004; *Stedinger et al.*, 2008; *Schoups and Vrugt*, 2010]. Further studies have dissected model error into its component parts, investigating the impacts of uncertain input and response data on model predictions [*Kavetski et al.*, 2006a, 2006b; *Thyer et al.*, 2009; *Renard et al.*, 2010]. Other innovative approaches for assessing hydrologic model uncertainty include Bayesian recursive estimation [*Thiemann et al.*, 2001], Bayesian hierarchical mixture of experts [*Marshall et al.*, 2007], and simultaneous parameter optimization and data assimilation [*Vrugt et al.*, 2005; *Clark and Vrugt*, 2006]. These techniques can be extended to climate impact studies to quantify the total uncertainty in hydrologic models and demonstrate the extent to which it obscures the differences between future and baseline hydrologic conditions.

[6] To the authors' knowledge, only one study has attempted to simultaneously quantify hydrologic model prediction and parameterization error and then propagate that uncertainty through climate impact assessments of hydrologic alteration [*Khan and Coulibaly*, 2010]. This study employed a Bayesian neural network rainfall-runoff model to explore climate-impacted hydrology. In this study, the posterior distribution of model parameters and the final distribution of model predictions were assumed Gaussian to improve the tractability of Bayesian integrals, despite the availability of Markov chain Monte Carlo (MCMC) sampling procedures that allow for more complex and accurate distributional assumptions. More importantly, uncertainty bounds were only generated for the streamflow trace generated using the mean of ensemble climate members, rather than for each climate member individually. This approach artificially deflates the true uncertainty in future hydrologic model projections because hydrologic model error should be integrated with the range of uncertainties stemming from GCMs and downscaling techniques.

[7] The study presented here will contribute to the science of hydrologic uncertainty analysis under climate change by developing a framework in which hydrologic model error is formally characterized and appropriately integrated with other sources of future climate uncertainty to better quantify the total uncertainty of hydrologic alterations under future climates. This allows a comparison of the range of projected changes in streamflow due to climate change to be compared with the uncertainty due to hydrologic model error. Hydrologic model prediction error is formally characterized with an appropriate likelihood function and combined with prior distributions of model parameters using Bayes' Theorem. MCMC sampling is used to evaluate the posterior distributions of hydrologic and error model parameters. Reliable uncertainty bounds for streamflow estimates are constructed from the integration of parameter and residual uncertainties and evaluated over the historic record. The Bayesian hydrologic modeling framework is then extended to a climate change impact assessment. Ensembles of baseline and future climate data are downscaled from GCMs and used to drive simulations of streamflow over parameter samples drawn from the posterior space. While GCM projections do not fully capture climate change uncertainty, the range of climate projections can be described as an estimate of the irreducible range of climate uncertainty, a minimum bound [*Stainforth et al.*, 2007a; *Wilby and Dessai*, 2010]. Time series of streamflow statistics are generated from baseline and future ensembles of simulated flows. Appropriate probability distributions are then fit to these statistics, enabling the estimation of streamflow quantiles and their sampling error for the ensemble of baseline and future conditions. Quantile estimates are directly compared between baseline and future scenarios in the context of their cumulative uncertainties. The framework can be used to highlight the complex interactions between different sources of uncertainty and their effects on future estimates of design flow statistics used in decision making. An application of this framework is presented for the White River Basin in Vermont using a version of the monthly ABCD hydrology model [*Thomas*, 1981] with a snow component.

[8] The paper will proceed as follows. Section 2 provides background on Bayesian inference techniques in rainfall-runoff modeling and their potential use for error propagation in future hydrologic simulations. Section 3 delineates the methodology used to quantify the total uncertainty of hydrologic alteration under future climate change scenarios. The methodology is applied and results presented in section 4,

and the study concludes in section 5 with a discussion of future research needs.

2. Bayesian Methods in Hydrologic Modeling and Their Use in Climate Change Studies

[9] Bayesian methods provide a formal mechanism to characterize the error in hydrologic model predictions, along with uncertainties surrounding parameterization. In a Bayesian framework, previous knowledge about parameter values can be incorporated into model calibration through a probability density function (pdf) known as the prior distribution. A joint pdf is then used to summarize the distribution of model residuals, and MCMC sampling procedures can be used to characterize the posterior distributions of hydrologic and error model parameters. If the error model correctly represents the distribution of model residuals, parameter and residual uncertainties can be integrated to develop predictive bounds for streamflow estimates. A relatively simple Bayesian formulation for rainfall-runoff modeling is described below that can be employed to help propagate uncertainties in a climate change impacts analysis. The Bayesian formulation presented below can be used to emphasize the importance of prediction error in uncertainty analyses under climate change and highlight the complex interactions between different sources of modeling uncertainty. In section 5, we discuss other challenges (e.g., source separation of uncertainties, choice of error model, and model structural errors) facing a complete quantification of hydrologic modeling uncertainty and their implications for the framework presented in this work.

2.1. Bayesian Hydrologic Modeling

[10] Let a conceptual rainfall-runoff model be formulated as follows:

$$Q = M(\theta_M, X) + \varepsilon \qquad (1)$$

where Q equals the vector of observed streamflows of length n, θ_M equals the set of hydrologic model parameters, X equals the matrix of inputs, $\hat{Q} = M(\theta_M, X)$ represents the streamflow model predictions, and ε equals residual model errors. Model residuals are assumed to follow a probability distribution described by a hypothesized joint pdf with a set of residual error model parameters θ_ε. Initially, no assumptions are made regarding the functional form of the error model $\varepsilon(\theta_\varepsilon)$. That is, model residuals may be autocorrelated, non-Gaussian, or heteroskedastic. However, we assume that errors associated with input data measurements, response data measurements, and model structure are aggregated into the error term ε. The implications of this simplifying assumption are discussed in section 5.1.

[11] Before proceeding with calibration, all previous knowledge about the set of hydrologic and error model parameters, $\theta = \{\theta_M, \theta_\varepsilon\}$, is summarized in a prior distribution, denoted $P(\theta)$. If no prior information is available, vague priors can be used so that calibration is driven by observed data only. The likelihood function, $L(Q|\theta, X)$, is based on the error model and is essentially a measure of hydrologic model skill. For certain hydrologic models applied at coarse temporal resolutions the choice of error model may be relatively simple, while many other applications may require more care in the identification of an appropriate error model [*Kuczera*, 1983]. These issues are discussed further in section 5.2. With an error model and associated likelihood function chosen, Bayes' theorem can then provide the joint posterior distribution of all model parameters,

$$P(\theta|Q, X) = \frac{L(Q|\theta, X) \times P(\theta)}{\int L(Q|\theta, X) \times P(\theta) \times d\theta} \qquad (2)$$

[12] The integral in the denominator is a constant of proportionality required to ensure that the right hand side term is a well-defined probability density function. MCMC methods can be used to evaluate the joint posterior distribution by sampling parameter values that are consistent with the combined information of the data and prior knowledge.

[13] To calculate predictive bounds on simulated streamflow, uncertainties in both model parameters and predictive skill need to be integrated. A time series of predicted percentiles, Q_α, for the $1 - \alpha$ nonexceedance level can be constructed for the vector of true streamflows Q as follows [*Schoups and Vrugt*, 2010]:

$$P(Q \le Q_\alpha | X) = freq([M(\theta_{M,j}, X) + \varepsilon(\theta_{\varepsilon,j})]_{j=1,\dots,J} \le Q_\alpha | X) = \alpha \qquad (3)$$

where $j = 1, \dots, J$ is the number of parameter sets sampled from the posterior distributions of θ_M and θ_ε. That is, J samples of model estimates, $M(\theta_{M,j}, X)$, and model errors, $\varepsilon(\theta_{\varepsilon,j})$, are generated for each simulated time step to produce a pdf of predicted values from which the predicted percentile can be inferred. The notation $freq()$ is used to acknowledge that the probability of the true vector of streamflows Q falling below the vector of percentiles Q_α is approximated using the frequency with which the sum of model predictions and errors fall below those percentiles. A 95% predictive bound around the time series of true streamflows Q can be formed with the bounded region $[Q_{.025}, Q_{.975}]$. If $\varepsilon(\theta_\varepsilon)$ is set to zero, then model error associated with parameter uncertainty can be isolated.

2.2. Integrating Uncertainties From the Hydrologic Model and Future Climate Projections

[14] Uncertainty in future climate must be integrated with errors from the hydrologic model to develop an appropriate range of possible hydrologic alterations. These uncertainties arise primarily from errors inherent to GCM simulations, which have been shown to exhibit poor skill at predicting even mean climate conditions at subcontinental scales [*Wood et al.*, 1997; *Stainforth et al.*, 2005]. Additional uncertainty stems from the downscaling technique used to transfer coarse GCM climate fields into meaningful climate changes at the local scale [*Fowler et al.*, 2007]. The climate science literature is ripe with studies exploring different methods to quantify future climate uncertainty. This study does not aim to thoroughly review all of these approaches or examine the merit of each. Rather, a brief overview of common methods is presented and then one method is chosen to demonstrate how future climate uncertainties can be nested in a framework aimed at quantifying the total uncertainty in future hydrologic projections.

[15] The most common approach relies on an ensemble of future climate scenarios to bracket possible climate changes.

These scenarios are developed using climate simulations from multiple GCMs that have been forced with several emission scenarios and initiated with different starting conditions, often downscaled with only one technique. Some studies have attempted to address downscaling uncertainty by using multiple downscaling methods [*Wilby and Harris*, 2006]. Other studies have attempted to assign nonuniform probabilities to different projections, using measures of bias and convergence to inform the choice of probabilities [*Tebaldi et al.*, 2005]. No matter how they are used, however, direct use of downscaled, multimodel GCM output as forcing data can only generate a lower bound on the maximum range of future climate uncertainty [*Stainforth et al.*, 2007a]. Since GCM simulations over the historic record do not fully explore the multiple sources of uncertainty at play, it is difficult, if not impossible, to develop a satisfying error model and bracket the true uncertainty of future climate projections. The quantification of future climate uncertainty remains largely intractable at present, as expectations for future experiments is that the uncertainty will increase. This study considers the simplest and most common quantification of future climate uncertainty where an ensemble of Z projections of future climate developed from several GCMs and emissions scenarios are downscaled to the region of interest using one downscaling technique. This represents a minimum range of climate uncertainty but allows a comparison of the range of GCM projections to hydrologic modeling uncertainty. The framework presented in section 3 can easily be extended to accommodate more complex quantifications of future climate uncertainty.

3. Framework to Quantify Hydrologic Uncertainties Under Future Climate Scenarios

[16] The Bayesian hydrologic model described in section 2 can be used to help quantify the uncertainty of important streamflow statistics, Y, generated under baseline and future climate conditions. Here, Y is a time series of a statistic of interest (e.g., average annual flows, average monthly flows, annual peak flows, etc.) calculated from a simulated time series of streamflow. We present an approach that quantifies uncertainty in inferred quantiles of Y stemming from future climate projections, the hydrologic model, and sampling error.

[17] Assume that Z climate change projections are available to provide a model-based range of possible future climate changes. For each climate change projection $z \in Z$, streamflow simulations are generated from one of two sequences of climate drawn from z: (1) a baseline series, X_z^b, which is generated from a downscaled time series of historical (1950–1999) conditions, or (2) a future series, X_z^f, generated from a downscaled time series of future conditions (2050–2099). Hereafter, all baseline (b) and future (f) variables will be denoted with superscripts. After the hydrologic model is calibrated in the Bayesian framework to historic observations, Monte Carlo resampling is used to select K parameter sets from the posterior parameter space over which to simulate an ensemble of K streamflow traces for both baseline and future climates. These ensembles capture the parameter and residual uncertainties in the hydrologic model. The simulation procedure can be repeated for each climate sequence $z \in Z$, producing a total of K×Z streamflow simulations for both baseline and future

climates. Time series of streamflow statistics, $Y_{z,k}^b$ and $Y_{z,k}^f$, can then be developed from these K×Z baseline and K×Z future streamflow projections.

[18] To make an inference on the pth quantile, Y_p, of the statistic Y, sampling error in the estimation of Y_p must also be propagated through the analysis. If the climate projections are of limited length, then sampling error could contribute significantly to uncertainties in quantiles of projected hydrologic statistics and therefore need to be accounted. For each climate projection z and posterior parameter sample k, appropriate pdf's can be fit to the baseline $Y_{z,k}^b$ and future $Y_{z,k}^f$ time series. Since both time series are of limited length, the true parameter values of the fitted pdf's will be unknown, but their uncertainty can be described using their sampling distributions. D samples of the pth quantiles $(Y_{z,k}^b)_{p_{d=1,\ldots,D}}$ and $(Y_{z,k}^f)_{p_{d=1,\ldots,D}}$ can be estimated using D draws from the sampling distributions of the fitted probability model parameters. Predictive bounds for the quantiles $(Y^b)_p$ and $(Y^f)_p$ can then be estimated at some confidence level $(1 - \alpha)100\%$ using the $\frac{\alpha}{2}$ and $(1 - \frac{\alpha}{2})$ percentiles of the ensemble of Z×K×D estimates of the pth quantile under both baseline and future climate conditions. We note here that an alternative approach to splitting the climate into pseudostationary baseline and future time periods would be to fit a nonstationary probability model [*Khaliq et al.*, 2006] to a transient climate over the entire timeframe (1950–2099).

[19] The methodology proceeds as follows (Figure 1).

[20] 1. Calibrate the hydrologic model over a set of historic climate and streamflow observations as stated in section 2 to develop posterior distributions of hydrologic and error model parameters. Evaluate the model using a split-sampling testing procedure. If possible, conduct a differential split-sample test to determine the capacity of the model to adequately model changes in climate [*Klemes*, 1986].

[21] 2. Sample K hydrologic and error model parameter sets, $\mathbf{\Theta}_M = \{\boldsymbol{\theta}_{M,1}, \boldsymbol{\theta}_{M,2}, \ldots, \boldsymbol{\theta}_{M,K}\}$ and $\mathbf{\Theta}_\varepsilon = \{\boldsymbol{\theta}_{\varepsilon,1}, \boldsymbol{\theta}_{\varepsilon,2}, \ldots, \boldsymbol{\theta}_{\varepsilon,K}\}$ from their posterior distributions developed in step 1.

[22] 3. For the ith scenario of future climate, z_i, selected from an ensemble of projections $\mathbf{Z}$, develop a baseline climate sequence of length N, $X_{z_i}^b = \{x_{z_i,1}^b, x_{z_i,2}^b, \ldots, x_{z_i,N}^b\}$, that represents historic climate. Climate information in $X_{z_i}^b$ can be downscaled from a historic period (i.e., 1950–1999) simulated in z_i, and can include variables such as temperature, precipitation, potential evapotranspiration, etc.

[23] 4. Generate a future climate sequence, $X_{z_i}^f$, that is downscaled from the same climate projection but is representative of some future time period (i.e., 2050–2099).

[24] 5. To develop the kth time series of baseline streamflow, $Q_{z_i,k}^b$, sample N perturbations $\varepsilon^b(\boldsymbol{\theta}_{\varepsilon,k}) = \{\varepsilon_{k,1}^b, \varepsilon_{k,2}^b, \ldots, \varepsilon_{k,N}^b\}$ from the error model $\varepsilon(\boldsymbol{\theta}_{\varepsilon,k})$ using the kth error model parameter set $\boldsymbol{\theta}_{\varepsilon,k}$. Then drive the hydrologic model with the baseline climate sequence using the kth hydrologic model parameter set $\boldsymbol{\theta}_{M,k}$ and add the output to the error series $\varepsilon^b(\boldsymbol{\theta}_{\varepsilon,k})$:

$$Q_{z_i,k}^b = M(\boldsymbol{\theta}_{M,k}, X_{z_i}^b) + \varepsilon^b(\boldsymbol{\theta}_{\varepsilon,k})$$

[25] The kth time series of future streamflow $Q_{z_i,k}^f$ can be generated in the same fashion by substituting $X_{z_i}^b$ with $X_{z_i}^f$ and $\varepsilon^b(\boldsymbol{\theta}_{\varepsilon,k})$ with a new sequence of errors $\varepsilon^f(\boldsymbol{\theta}_{\varepsilon,k})$.
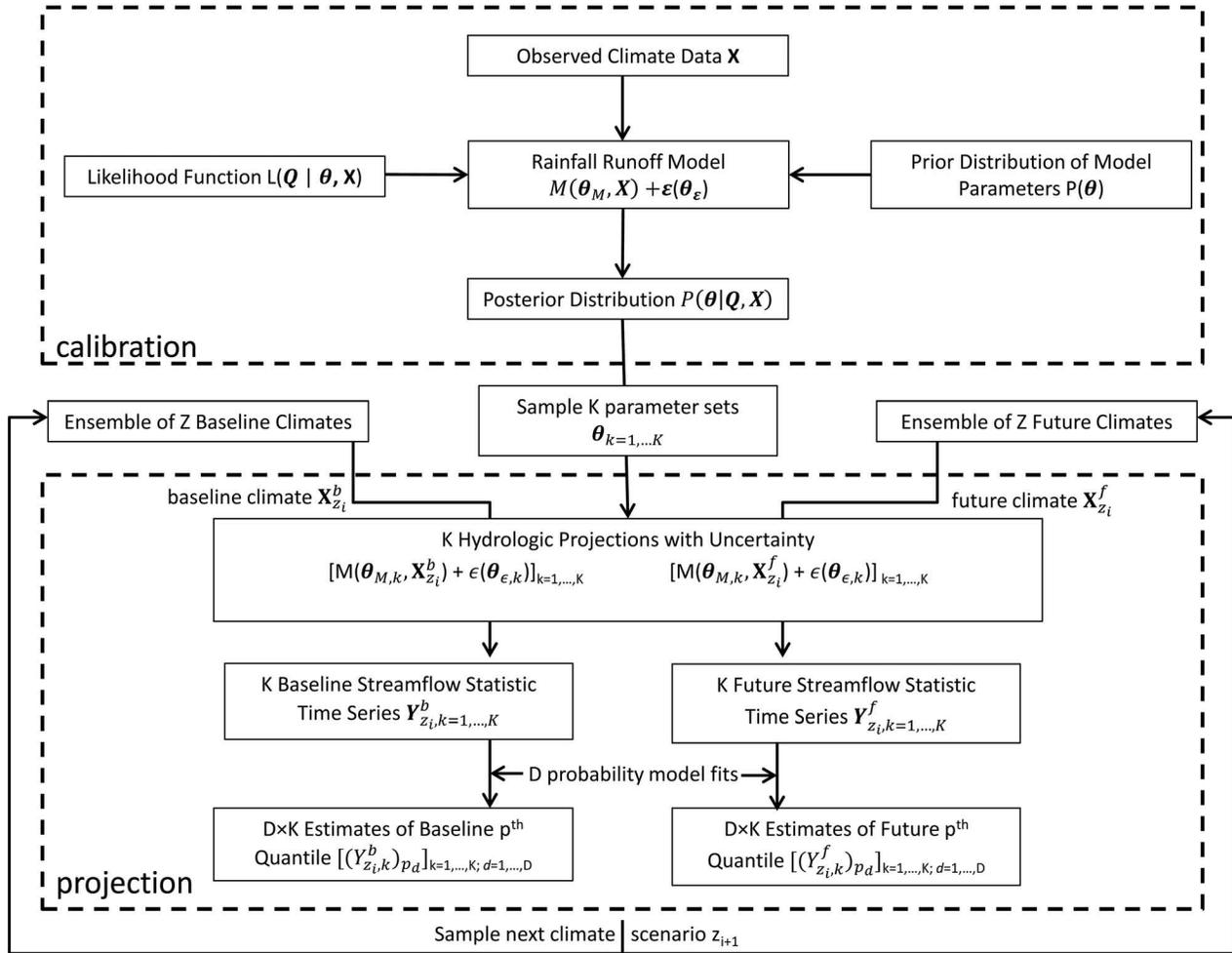
Figure 1. Flowchart of the statistical framework for a hydrologic uncertainty analysis under climate change.

[26] 6. Repeat step 5 K times to develop K time series of baseline and future streamflow.

[27] 7. Calculate the time series of streamflow statistics $Y^b_{z_i,k}$ and $Y^f_{z_i,k}$ for each of the K parameter samples for both baseline and future climate conditions.

[28] 8. Fit an appropriate probability model to each time series of streamflow statistics. An estimate of the pth quantile, $(Y^b_{z_i,k})_p$ and $(Y^f_{z_i,k})_p$, can be inferred from the fitted probability models for both baseline and future statistics. For instance, if the streamflow statistic (or its logarithms) are normally distributed, the pth quantile can be estimated as $Y_p = \mu_y + \xi_p \times \sigma_y$, where μ_y is the mean of the statistic, σ_y is its standard deviation, and ξ_p is the 100p percentile of the standard normal distribution. To account for sampling error, draw D estimates of probability model parameters (i.e., $\mu_{y,d}$ and $\sigma_{y,d}$ with d = 1,..., D) from their sampling distributions to produce D estimates of the pth quantile, $(Y^b_{z_i,k})_{p_{d=1,...,D}}$ and $(Y^f_{z_i,k})_{p_{d=1,...,D}}$. In this study, sampling distributions were taken as the posterior distributions of probability model parameters developed via a Bayesian fit of the probability model to the streamflow statistics $Y^b_{z_i,k}$ and $Y^f_{z_i,k}$. Vague distributions (e.g., uniform distributions) can be used as priors for probability model parameters in the Bayesian fit.

Alternatively, estimates of sampling distributions for different probability models are often available in the literature.

[29] 9. Repeat steps 3–8 for each climate projection $z \in Z$. This will produce K×Z×D different estimates of the pth quantile for both baseline and future climate conditions. The expected value of the pth quantile of Y for baseline and future conditions can be calculated by taking the mean across all K×Z×D quantile estimates, $\overline{(Y^b)_p}$ and $\overline{(Y^f)_p}$. Similarly, predictive intervals $[(Y^b)_{p,\frac{\alpha}{2}}, (Y^b)_{p,(1-\frac{\alpha}{2})}]$ and $[(Y^f)_{p,\frac{\alpha}{2}}, (Y^f)_{p,(1-\frac{\alpha}{2})}]$ for the pth quantile can be developed using the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$ percentiles of the ensemble of Z×K×D quantile estimates. These two intervals quantify the total considered uncertainty in estimates of the pth quantile of a streamflow statistic Y for both baseline and future conditions. They can be directly compared to provide a reliable depiction of how distinct the future hydrologic alteration for that statistic will be after accounting for all sources of uncertainty considered.

4. Application of Statistical Framework in a Climate Impact Assessment

[30] An application of the statistical framework described above is presented for the White River Basin, located in

central Vermont. Records of monthly precipitation, temperature, and potential evapotranspiration are used to drive a Bayesian calibration of a conceptual rainfall-runoff model of the basin. An adaptation of the ABCD conceptual hydrologic model that incorporates a new snow modeling scheme is chosen for this purpose. After calibration, posterior distributions of both hydrologic and error model parameters are examined for convergence, and a probabilistic evaluation of the error model is presented to ensure the distribution of model residuals is well characterized. After the model is evaluated, the framework for climate impact assessments is applied to an ensemble of transient GCM climate scenarios.

4.1. White River Basin

[31] The White River is a major tributary of the Connecticut River in New England, draining 1790 square kilometers in the east central portion of Vermont (Figure 2). Running 97.6 km from the Green Mountains to the Connecticut River Valley below, the White River is the largest gauged basin in the Connecticut River Watershed without significant regulation from upstream reservoirs or land use changes. Precipitation rates are relatively constant throughout the year, averaging approximately 100 mm m^{-1}. Regional estimates suggest about 70% of all winter precipitation falls as snow [*Huntington et al.*, 2004]. Seasonal variations in temperature drive snow accumulation and melt processes that dominate hydrologic response throughout the winter and spring months. Streamflow is lowest during the summer and early fall months when evapotranspiration rates reach their peak.

4.2. ABCD Hydrologic Model

[32] An altered version of the ABCD hydrologic model is considered to model monthly streamflow in the White River Basin. The original ABCD model is a four parameter (a,b,c,d), conceptual rainfall-runoff model designed through a control volume analysis on upper soil moisture zone storage [*Thomas*, 1981]. The model converts monthly averaged precipitation and potential evapotranspiration into estimates of monthly streamflow by diverting water between two soil storage zones, losses to evapotranspiration, and the stream. The model has been recommended as an effective parsimonious model with physically meaningful parameters capable of efficiently reproducing monthly water balance dynamics in both theory [*Vogel and Sankarasubramanian*, 2003] and practice [*Alley*, 1984; *Vandewiele et al.*, 1992]. A detailed review of the original ABCD model formulation can be found in [*Fernandez et al.*, 2000].

[33] A snow component similar to that of *Martinez and Gupta* [2010] was added to the ABCD model to simulate the snow accumulation/melt processes that dominate much of the hydrologic cycle in northern latitude watersheds. A snow storage zone is added that stores all incoming precipitation as snow water equivalent during times of year when the temperature falls below a threshold T_{Snow}. A second threshold, T_{rain}, delimits the temperature above which all precipitation falls as rain. When temperatures rise above T_{rain}, all water held in the snow storage zone melts and is added to incoming precipitation for that month. This threshold melt process is highly representative of springtime hydrology seen in northern New England rivers. When monthly temperatures fall between T_{rain} and T_{snow}, a fraction of the incoming precipitation for that month enters the snow storage component, and the remainder falls as rain. In addition, a fraction of the water held as snow is available for melt and is added to the effective rainfall for that month. The rate of melt is given by the parameter e. The total snow melt in time t is given by

$$Melt_t =$$

$$\begin{cases} S_{t-1} & T_t \geq T_{rain} \\ (S_{t-1} + frac_t \times P_{tot,t}) \times e \times (1 - frac_t) & T_{snow} < T_t < T_{rain} \\ 0 & T_t \leq T_{snow} \end{cases}$$

$$(4)$$
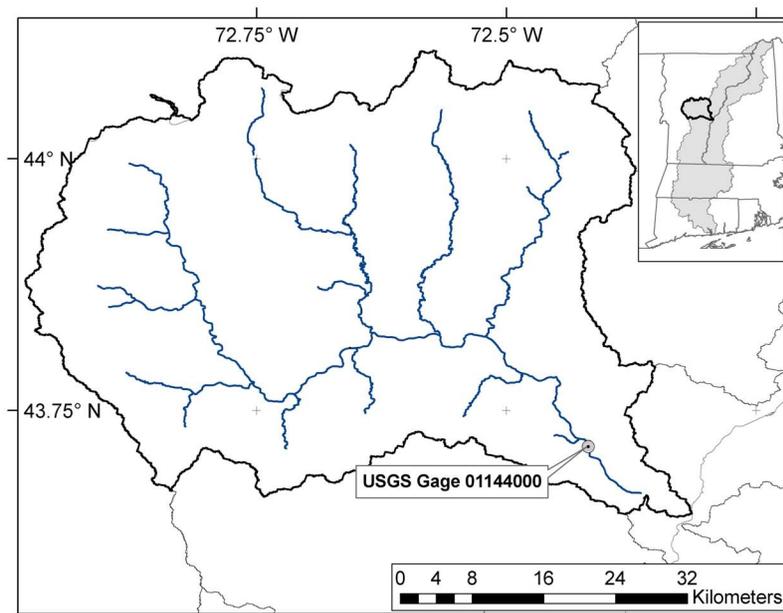
Figure 2. Schematic of the White River Basin, Vermont, USA.

where S_{t-1} is the water stored as snow in the previous month, $P_{tot,t}$ is the total precipitation, T_t is the mean monthly temperature, and $frac_t$ is the fraction of precipitation that falls as snow, equal to $\frac{T_{rain}-T_t}{T_{rain}-T_{snow}}$. The water stored as snow in month t is given by

$$S_t = \begin{cases} 0 & T_t \geq T_{rain} \\ (S_{t-1}+frac_t \times P_{tot,t}) - Melt_t & T_{snow} < T_t < T_{rain} \\ P_{tot,t} + S_{t-1} & T_t \leq T_{snow} \end{cases}$$
(5)

[34] The effective precipitation input to the model (precipitation available for runoff, soil zone storage, ET, etc.) is then given by

$$P_{eff,t} = \begin{cases} P_{tot,t} + Melt_t & T_t \geq T_{rain} \\ (1-frac_t) \times P_{tot,t} + Melt_t & T_{snow} < T_t < T_{rain} \\ 0 & T_t \leq T_{snow} \end{cases}$$
(6)

[35] In total, three parameters are used to represent snow accumulation/melt processes, bringing the total number of model parameters to seven (a, b, c, d, e, T_{rain}, T_{snow}). During calibration, the parameter T_{snow} is not directly calibrated because its prior distribution would have to be conditioned on the value of T_{rain} to ensure it took a smaller value. To circumvent this issue, a nonnegative parameter $dif = T_{rain} - T_{snow}$ is used, from which T_{snow} can be directly computed.

[36] *Martinez and Gupta* [2010] performed a thorough analysis on the suitability of a similar snow-augmented ABCD model structure for catchments throughout the United States, testing the model using several diagnostic statistics including Nash-Sutcliffe efficiency, bias, and variance error. That study found that the snow-augmented ABCD model structure significantly improves results for snow-dominated watersheds in New England and is a suitable structure for many catchments in the region, supporting its use in this study.

4.3. Bayesian Calibration and Evaluation

[37] Historic, monthly averages of precipitation and maximum, minimum, and mean daily temperatures were gathered for the basin over the period of January 1980 to December 2005 from the gridded observed meteorological data set produced by *Maurer et al.* [2002]. Average monthly streamflows were collected from the U.S. Geological Survey (USGS) West Hartford gauge (ID 01144000) located at the mouth of the White River. Monthly averages of maximum, minimum, and mean daily temperatures were combined with estimates of monthly extraterrestrial solar radiation to produce a time series of potential evapotranspiration using the Hargreaves method [*Hargreaves and Samani*, 1982]. Solar radiation was calculated using the method presented by *Allen et al.* [1998].

[38] Based on past hydrologic modeling experience for monthly flows in the New England region, a normal distribution with mean zero and standard deviation σ was initially chosen to characterize the sampling distribution of the residuals of the natural logarithms of observations and model predictions (hereafter referred to simply as model residuals)

$$f(\varepsilon_{\ln}) = \frac{1}{\sqrt{2 \times \pi \times \sigma^2}} \times \exp\left(-\frac{\varepsilon_{\ln}^2}{2 \times \sigma^2}\right)$$
(7)

where $\varepsilon_{\ln} = \ln(Q) - \ln(\hat{Q})$. The likelihood function for the observed streamflow values, $\boldsymbol{Q}$, is then given by

$$L(\boldsymbol{Q}|\boldsymbol{\theta},\boldsymbol{X}) = (2 \times \pi \times \sigma^2)^{\frac{-n}{2}} \prod_{t=1}^{n} \exp\left(-\frac{(\ln(Q_t) - \ln(\hat{Q}_t))^2}{2 \times \sigma^2}\right)$$
(8)

[39] The prior for the unknown parameter σ was set to a gamma distribution with known shape $\lambda = 1$ and scale $\zeta = 2.5$ parameters. The posterior of this parameter characterizes the level of uncertainty in hydrologic model estimates. A verification of the chosen sampling distribution for model residuals is described below.

[40] Past studies were used to inform prior distributions for the hydrologic model parameters a, b, c, and e [*Alley*, 1984; *Vandewiele et al.*, 1992; *Fernandez et al.*, 2000; *Martinez and Gupta*, 2010], and the remaining model parameters (d, T_{rain}, dif) were given vague priors in the form of uniform distributions or normal distributions with large variances. Initial states were also calibrated in the model to avoid any parameter biases from incorrect initial conditions. The slice sampler was chosen for the MCMC sampling and was implemented in the JAGS programming language (M. Plummer, rjags: Bayesian graphical models using MCMC, R package version 2.2.0-4, http://CRAN.R-project.org/package=rjags). Three chains were used in the sampling, and the Gelman and Rubin factor was used to test for convergence [*Gelman and Rubin*, 1992]. Calibration was implemented over the period between January 1980 and December 1999, leaving 6 years of data for evaluation. Table 1 summarizes the prior and posterior distributions for all parameters inferred in the MCMC sampling, as well as allowable ranges for each parameter. Figure 3a shows the history plots of parameter a for the three chains, and Figure 3b presents histograms of the prior and posterior distributions of parameter a. For all model parameters, the Gelman and Rubin convergence factor was within 0.005 of 1, suggesting that convergence was reached for all calibrated parameters.

[41] Figure 3c presents a normal probability plot of the model errors $\varepsilon_{\ln}$ generated from the hydrologic simulation under the median posterior parameter set over the evaluation period (January 2000 to December 2005), and Figure 3d shows their autocorrelation coefficients. Results from the Q-Q plot suggest that model residuals follow a normal distribution relatively well. Most autocorrelation coefficients in Figure 3d are insignificant, including that at lag 1. There are some coefficients that exhibit small but significant values, particular at seasonal lag times. An autocorrelation component could be added to the error model, but this would require additional parameters to be estimated in the calibration, creating a tradeoff between problem dimensionality and error model accuracy. The seasonal autocorrelation seen in Figure 3d is rather low and not considered worth the increased dimensionality needed to model its behavior.

Table 1. Summary of Prior and Posterior Distributions for All Model Parameters[a]

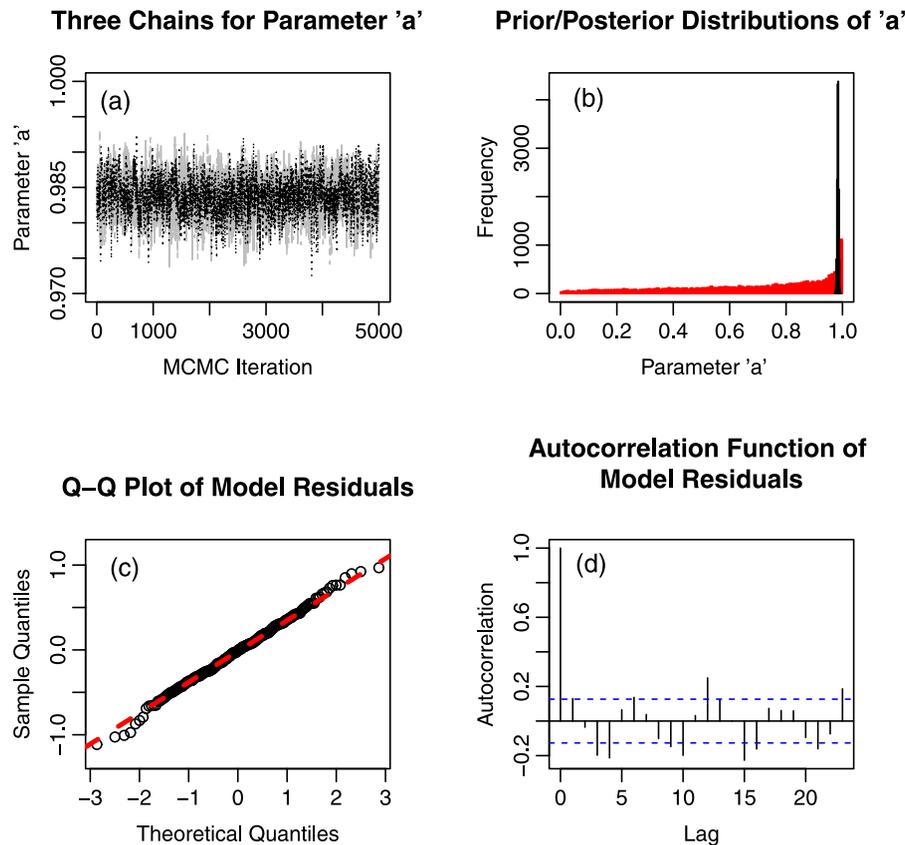| Parameter	Allowable Range	Prior Distributions	Posterior Distribution			
			First Quartile	Median	Mean	Third Quartile
a	$(0, 1)$	Beta (a = 1.2, b = 0.6)	0.982	0.984	0.984	0.986
b (mm)	$(0, \infty)$	Normal ($\mu = 300$, $\varphi = 100$)	303	310	310	316
c	$(0, 1)$	Beta (a = 0.6, b = 1.2)	0.14	0.18	0.18	0.22
d	$(0, 1)$	Uniform (a = 0, b = 1)	0.45	0.66	0.74	0.90
e	$(0, 1)$	Beta (a = 0.8, b = 1.8)	0.141	0.205	0.206	0.268
T_{rain} (°C)	$(-\infty, \infty)$	Normal ($\mu = 0$, $\varphi = 4$)	-1.65	-1.48	-1.47	-1.31
dif (°C)	$(0, \infty)$	Uniform (a = 0.01, b = 20)	12.9	13.9	14.0	15.0
σ(ln(mm))	$(0, \infty)$	Gamma ($\lambda = 1$, $\zeta = 2.5$)	0.13	0.14	0.14	0.15

[a]Normal priors are given mean (μ) and standard deviation (φ) hyperparameters. Gamma priors have shape (λ) and scale (ζ) hyperparameters.

Therefore, the original choice of a normal error model with no autocorrelation component for ε_{ln} was considered adequate for this modeling exercise.

[42] Figure 4 shows the observed monthly streamflow for the last 5 years of calibration and the entire evaluation period, as well as model estimates generated by the median values of the posteriors for hydrologic model parameters. The Nash-Sutcliffe efficiency (NSE), mean flow bias, and variance error for simulated streamflow using the median parameter set equals 0.82, -1.4%, and $+6.6\%$ for the calibration period and 0.67, -5.2%, and -15.1% for the evaluation period. The bias and variance errors are expressed as a percentage of observed values. These performance statistics are considered either "good" or "acceptable" in other hydrologic modeling studies [*Martinez and Gupta*, 2010]. Also shown in Figure 4 are error bounds consistent with the 2.5th and 97.5th percentiles of streamflow estimates, calculated according to equation 3. Observed data from the calibration and evaluation periods fell outside the 95% predictive interval 3.3% and 6.7% of the time, respectively, again suggesting that the error model adopted is appropriate for this application.

[43] An additional evaluation procedure was conducted to further evaluate the adequacy of the error model. The details of the procedure are given by *Laio and Tamea* [2007]. In brief, the procedure tests whether probabilistic

Three Chains for Parameter 'a'   **Prior/Posterior Distributions of 'a'**

**Q–Q Plot of Model Residuals**   **Autocorrelation Function of Model Residuals**

Figure 3. MCMC and model error diagnostics, including (a) the history plot for parameter a shown for the three MCMC chains, (b) a histogram of the prior (red) and posterior (black) distribution for parameter a, (c) a Q-Q plot showing sample quantiles of model error ε_{ln} against theoretical quantiles of a standard normal distribution, and (d) the autocorrelation function of model errors ε_{ln}.
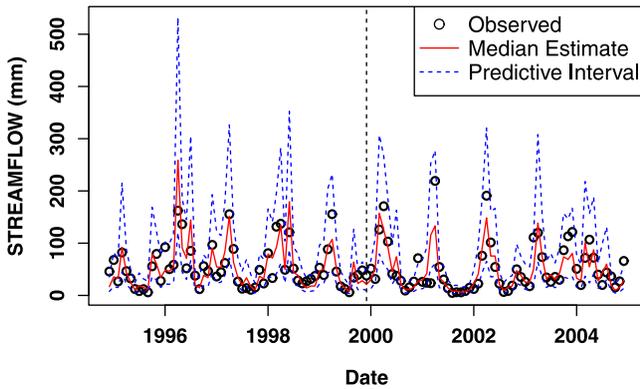
Figure 4. Time series of streamflow during calibration (left of vertical dashed line) and evaluation phases (right of vertical dashed line). Only a portion of the calibration time period is shown for clarity.

predictions for a set of streamflow observations are adequate in a statistical sense. To conduct the test, the cumulative distribution function of predicted streamflow at time t is evaluated with respect to the observation q_t at t via a probability integral transform, $v_t = P_t(q_t)$. If the probabilistic predictions of streamflow are suitable then the v_t values will be mutually independent and distributed uniformly between 0 and 1. To test uniformity, a probability plot can be employed to graphically examine how well the distribution of v_t values matches a U(0,1) distribution. The condition of mutual independence can be tested using the Kendall's tau test of independence.

[44] The probability plot of v_t values versus a theoretical uniform distribution are shown in Figure 5, along with Kolmogorov confidence bands at the 95% confidence level. The distribution of v_t values match that of a U(0,1) distribution very well, satisfying the first condition of the test. In addition, the condition of mutual independence was met under the Kendall's tau test of independence (p value of 0.81), satisfying the second condition of the test. These results provide further support for the error model chosen in this application.
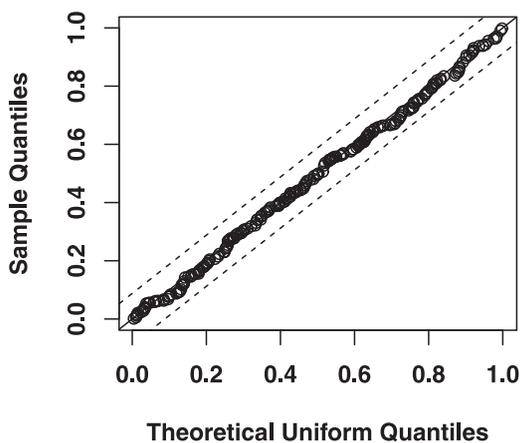

Figure 5. Q-Q plot of the sample quantiles of the v_t values versus those of a U(0,1) distribution. Kolmogorov confidence bands (dashed lines) at the 95% confidence level are also shown.

4.4. Future Climate Scenarios

[45] Seventy-three transient future climate simulations, running from 1950 to 2100 and sampled across the A1b, A2, and B1 emission scenarios, were gathered from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project Phase 3 (CMIP3) multimodel data set. GCM simulations were downscaled according to the bias correction and statistical downscaling (BCSD) approach described by *Maurer et al.* [2007]. For each GCM simulation, a baseline and future climate scenario (i.e., time series of mean monthly temperatures and total monthly precipitation) was taken from 50 year windows of downscaled climate data centered about the years 1975 and 2075, respectively. Figure 6 shows the absolute and percent difference between mean annual temperatures and mean annual precipitation, respectively, for these two periods across all 73 projections. We note here that maximum and minimum monthly temperatures are not provided in the downscaled CMIP3 data set but are required for calculations of potential evapotranspiration. To generate maximum and minimum monthly temperature fields for baseline and future scenarios, the average differences between maximum and mean monthly temperature and minimum and mean monthly temperature were calculated for each month over the historic record. These average differences were then added to each time series of mean monthly temperature for all projections from the CMIP3 data set to generate the maximum and minimum monthly temperature fields.

### 4.5. Projections of Hydrologic Response With Uncertainty

[46] The Z = 73 baseline (1950–1999) and future (2050–2099) climate scenarios taken from the CMIP3 data set were each used to drive an ensemble of K = 5000 hydrologic model simulations, each with different parameter sets drawn from the posterior distributions developed in section 4.3. Four different annual streamflow statistics (Y) were considered in the analysis, including average January, March, April, and October streamflows. These monthly statistics were chosen because they exhibit a wide range of changes under future climate and highlight the importance of including hydrologic model error in climate impact assessments. These statistics were assumed to follow a lognormal distribution, similar to the observed historic streamflow data. This assumption was validated for each of these
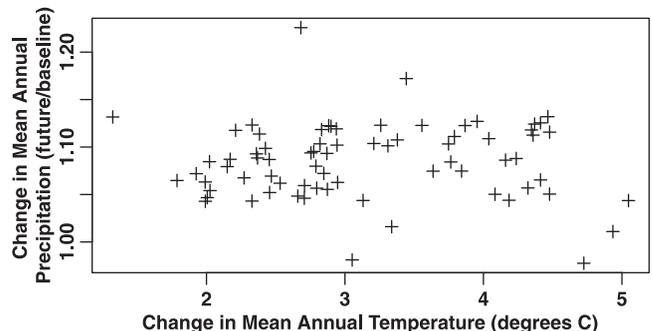
Figure 6. The change in mean annual precipitation and mean annual temperature between baseline and future time slices across all 73 climate scenarios.

statistics under a large sample of climate scenarios and parameter sets using probability plots. Sampling error in the quantiles of these statistics was estimated using D = 1000 different estimates for the mean and standard deviation of the fitted lognormal distributions drawn from their posterior distributions. Results are presented as follows. The isolated effects of hydrologic model residual error on the estimation of these statistics are considered first. The integration of uncertainties from the range of climate projections, model residual error, model parameterization, and sampling uncertainty are then addressed. An analysis of alteration in different monthly statistics is then presented in the context of their integrated uncertainty estimates.

[47] Figure 7 presents the pdf of a fitted lognormal distribution to January monthly streamflows developed from one GCM scenario over the baseline period forced with one sample of hydrologic and error model parameters. Two pdf's are shown, one developed from the original streamflow trace, and a second developed from the same trace after being perturbed with noise generated from the error model. The variability in both future climate and parameter estimates is omitted by considering only one climate trace and parameter set, therefore isolating the effects of residual error on the distribution of the January flow statistic. As expected, the addition of residual error to the simulated streamflow trace causes the spread in January flows to increase. Addition of residual uncertainty to the model output appropriately adjusts the data so that it better represents the actual precision with which we can estimate characteristics of the streamflow statistic. Since the error model is logarithmic, the spread increases more at higher streamflow values than it does at lower values, suggesting different levels of precision for different magnitudes of flow. Interestingly, this highlights one of the difficulties in the choice of error model. While a transformation might make the data more tractable for a given error model, the application of
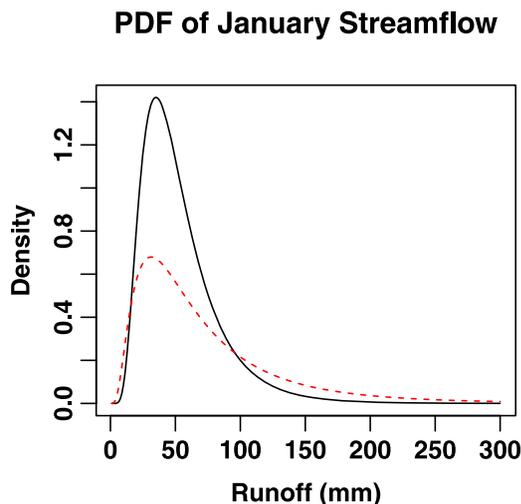
that error model may lead to asymmetric uncertainty estimates after the transformation is reversed.

[48] To develop comprehensive uncertainty bounds around future hydrologic statistics, the residual error of the hydrologic model needs to be integrated with uncertainties in model parameterization, future climate projections, and sampling error. Figure 8 shows 95% predictive intervals for quantile estimates of baseline period January streamflow plotted against nonexceedance probabilities for different considerations of uncertainty. Figures 8a–8c show the isolated contributions of climate uncertainty, hydrologic model parameter and residual error, and sampling error to the uncertainty of quantile estimates, respectively. The range of quantile estimates in Figure 8a stems from the ensemble of baseline climate scenarios run over the median hydrologic model parameter set without the addition of residual noise. The range in Figure 8b was developed for only one ensemble member of baseline climate, but both parameter and residual uncertainties from the hydrologic model were considered. The influence of hydrologic model parameter and residual errors are aggregated and presented together in Figure 8b in order to represent the total added uncertainty from the hydrologic model. In Figure 8c, one baseline climate scenario was used to drive the hydrologic model with the median parameter set and no additional noise, but sampling uncertainty was calculated for each quantile. We note that the ranges of uncertainty in Figures 8a–8c are dependent on the climate ensemble member or parameter set that was held constant during their development and are thus only used to illustrate the range of isolated uncertainty bounds. Figures 8d–8f show the predictive bounds for quantile estimates when climate, hydrologic model, and sampling uncertainties are integrated together. Figure 8d is the same as Figure 8a, but Figure 8e shows the uncertainty bounds for quantile estimates when climate uncertainty, parameter uncertainty, and residual uncertainty are considered simultaneously. Figure 8f shows the total integrated uncertainty with sampling error considered as well.

[49] When comparing isolated and integrated uncertainties, it immediately becomes clear that uncertainties from climate projections, hydrologic model parameter and residual error, and sampling error cannot be independently added to generate reliable predictive bounds for estimates of hydrologic statistics and their properties. This is seen in Figure 8e and 8f, in which the range of uncertainty for many quantiles, particularly the larger ones, is greater than the sum of the uncertainties of their component parts (Figures 8a–8c). This property highlights the dependence of uncertainty bounds on the interactions between the different sources of uncertainty.

[50] This is a particularly important point, so we present a simplified example to emphasize it here. Consider a normalized streamflow quantile, Y_p, with zero mean and a variance conditional on either isolated climate uncertainty ($\sigma^2_{Y_p,c}$) or hydrologic modeling uncertainty ($\sigma^2_{Y_p,h}$). Assuming Y_p is normally distributed, a $(1-\alpha)$ predictive interval under isolated climate uncertainty and isolated hydrologic modeling uncertainty could be respectively written as $[-\sigma_{Y_p,c} \times \xi_{\frac{\alpha}{2}}, \sigma_{Y_p,c} \times \xi_{\frac{\alpha}{2}}]$ and $[-\sigma_{Y_p,h} \times \xi_{\frac{\alpha}{2}}, \sigma_{Y_p,h} \times \xi_{\frac{\alpha}{2}}]$, where $\xi_{\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$ percentile of the standard normal distribution. Now assume that Y_p can be expressed under the simple additive model $Y_p = \varepsilon_c + \varepsilon_h$, where $\varepsilon_c \sim N(0, \sigma^2_{Y_p,c})$ and ε_h
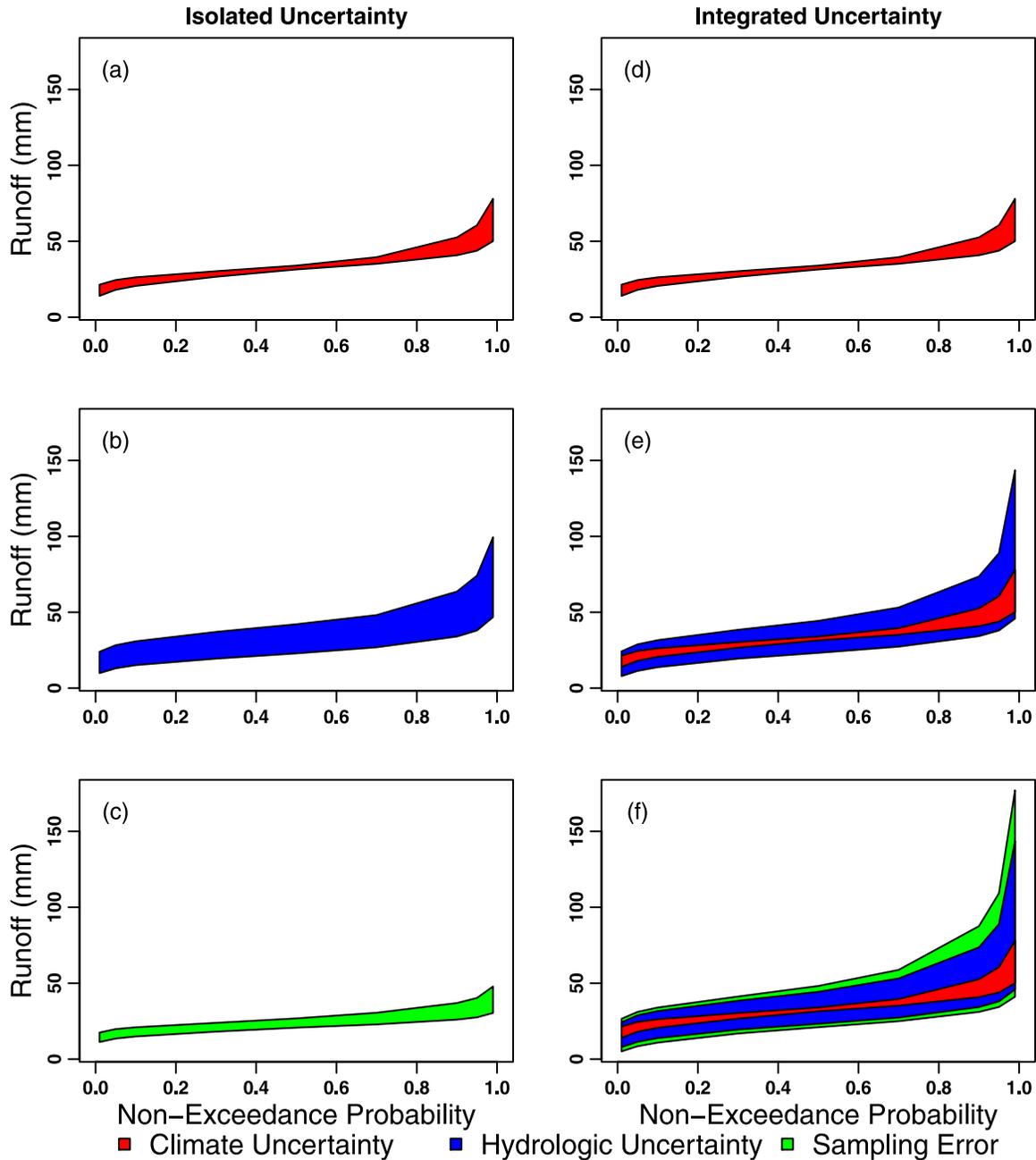
PDF of January Streamflow

Figure 7. Probability density functions of baseline January monthly streamflow with (red dashed line) and without (black solid line) a perturbation with noise generated from the error model. Only one GCM scenario (z = 1) and parameter set (k = 1) were used to generate the streamflow trace.

Figure 8. (a–c) Isolated and (d–f) integrated 95% predictive intervals for quantiles of January streamflow over the baseline period. Uncertainty originating from a range of climate scenarios, parameter and residual errors in the hydrologic model, and sampling error are shown in isolation in Figures 8a, 8b, and 8c, respectively. Climate uncertainty in Figure 8a is repeated in Figure 8d, the integration of climate, parameter, and residual uncertainties is presented in Figure 8e, and Figure 8f shows the cumulative uncertainty after sampling error is also considered.

$\sim N(0, \sigma^2_{Y_p,h})$. Assuming that variations in Y_p stemming from climate and hydrologic modeling uncertainty are independent, we would expect that the total variance of Y_p would equal the sum of the isolated variances, $\sigma^2_{Y_p} = \sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}$. However, the predictive interval for Y_p under integrated climate and hydrologic modeling uncertainty would be given as $\left[-\sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \times \xi_{\frac{\alpha}{2}}, \sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \times \xi_{\frac{\alpha}{2}}\right]$, which does not correspond to the sum of the two isolated intervals above

because $\sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \neq \sigma_{Y_p,c} + \sigma_{Y_p,h}$. Therefore, even under the simplifying assumption that variations in Y_p can be described by the simple additive model above, we would not expect uncertainty intervals to be additive. Thus, there is no reason to believe that uncertainty intervals would be additive given a more complex situation in which variations in Y_p can be influenced by the interactions of different sources of uncertainty within a hydrologic modeling framework.

[51] The dependence of variations in Y_p on interactions between different sources of uncertainty can be traced to several contributing factors. First, the hydrologic model being considered is nonlinear, so different parameterizations of that model will result in nonlinear responses to a given climate. When those various parameterizations are used to simulate hydrologic response over a range of climates, there is the potential that the combination of an extreme climate ensemble member and parameter set will lead to significantly different streamflow responses than that seen under just climate or parameter uncertainty alone. Another source of dependency arises from the interaction between the error model and the ensemble of climate members. Because the error model used in this application is based on a logarithmic transformation, the uncertainty of large quantile values becomes highly skewed to the right after residual uncertainty is accounted. If an ensemble climate member leads to slightly larger quantile values for the streamflow statistic being considered, the residual error estimated for those larger quantiles could lead to the significant expansion of their predictive bounds. Finally, there are significant interactions between sampling error estimation and both hydrologic and climate model uncertainties. Sampling error uncertainty bounds will grow with the uncertainty in the parameters of the distribution used to model the streamflow statistic. The sampling distributions of these parameters will likely change when climate and hydrologic model uncertainties are considered, causing the magnitude of sampling error to change with respect to its range when considered in isolation.

[52] After aggregating the uncertainties from climate scenarios, the hydrologic model, and sampling error, it becomes evident that some quantile values for certain streamflow statistics can only be estimated with limited precision. This is shown for the cumulative error under baseline climate conditions in Figure 8f. In the case of future climate conditions, the range of climate projections becomes far more significant. Figure 9 compares the cumulative uncertainty of January monthly flows evaluated over the historic and future climate conditions. Figure 9a is the same as in Figure 8f, but Figure 9b shows the uncertainty in future climate projections.

[53] Two primary differences arise between the baseline and future cumulative uncertainties for January flow quantiles. First, the underlying climate uncertainty is far greater under the future scenarios than those of the baseline. This is expected because the baseline climate projections are all directly mapped to the historical trace of temperature and precipitation via downscaling. Thus, the range of historical projections does not model climate uncertainty or even climate model uncertainty but rather is an artifact of the bias correction method. Consequently, the range of future projections also does not model the uncertainty of future climate or even the model uncertainty of future climate projections. Nonetheless, the range of climate projections is commonly used to provide some sense of the uncertainty in the projections that arise due to model error and internal variability and are used for that purpose. That range, albeit a minimum range of climate uncertainty, significantly increases the uncertainty in the quantile estimates relative to hydrologic modeling uncertainty as shown in the comparison between Figures 9a and 9b. Second, the sampling error for larger quantiles is vastly greater for the future scenarios than for the baseline. This is due to the greater spread of January flows under future conditions and its influence on sampling error estimates. Overall, it is clear that the cumulative uncertainty for quantile estimates of this statistic is much greater for the future than it is under baseline conditions.

[54] Quantile estimates can be directly compared between baseline and future scenarios in the context of their
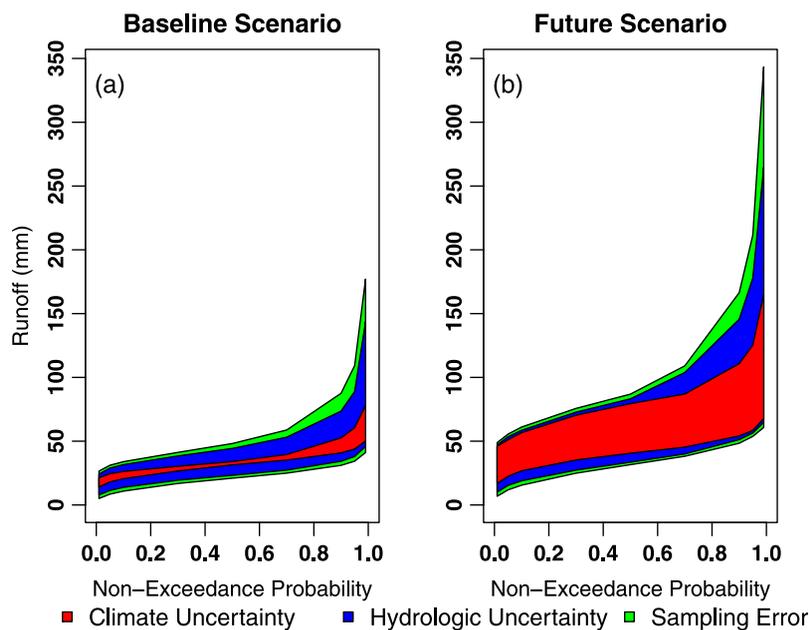
Figure 9. Integrated 95% predictive bounds for January flow quantiles under the (a) baseline and (b) future periods.

cumulative uncertainties to help determine the level of confidence that can be associated with their possible alteration under climate change. Figure 10 presents the cumulative uncertainty of quantile estimates of monthly streamflow statistics in the White River for future and baseline conditions. Here, no distinction is made between the different sources of uncertainty (e.g., climate, hydrologic, or sampling errors). Rather, the cumulative 95% predictive intervals for flow quantiles under baseline and future conditions are overlaid on each other to provide a representation of whether changes in streamflow under climate change exceed the range of uncertainty that arises during the modeling process. Less overlap between predictive intervals of flow quantiles under baseline and future conditions provides greater confidence that the flow quantile will actual differ under future climate conditions. Figure 10a shows that there are significant differences between the distributions of January flows in the baseline and future periods even after accounting for cumulative modeling uncertainties. Results suggest that climate projections of January flows are significantly higher in the future than in the present, likely due to a shift in the snowfall to precipitation ratio driven by increased wintertime temperatures. Over most January quantiles, approximately half of the bounded region for future conditions lies completely outside the range of baseline uncertainty. This suggests that this range of climate changes rises to a level that is well above the baseline uncertainty.

[55] Figures 10b and 10c show results for March and April average streamflows, respectively. The range of climate projections show March flows increasing in the future while April flows decrease. These changes are consistent with earlier snowmelt occurrences and decreases in snowpack storage that historically have persisted into the later spring. Interestingly,

the highest quantiles of March flows for the future period show minor departures from those of the baseline; differences become more noticeable for flows below the 95th percentile. This is not the case for April flows, which show more significant departures between baseline and future flows at the highest quantiles. This suggests that more confidence can be associated with shifts in the highest flows during April than in March. This is likely because snowpack, a driving factor of the largest spring flows, is consistently reduced in April under all future hydroclimatic projections, but is more variable across the projections in the month of March.

[56] Figure 10d shows results for the month of October. The range of climate projections exceeds only minutely the baseline uncertainty bounds for October quantiles. The spread in the future period for most quantiles extends both below and above that of the baseline period, although the changes are extremely small except for the higher quantiles. These results suggest that no real change in most October flow quantiles are projected in this set of CMIP3 climate changes.

5.   Discussion of Future Research Needs

[57] The framework in this study addresses many types of uncertainty in future hydrologic alterations and integrates them together to form a more comprehensive expression of the total uncertainty surrounding future hydrologic variables. Nevertheless, several simplifying assumptions were made regarding the quantification of hydrologic model uncertainty in this analysis. Various challenges still hinder a complete quantification of this uncertainty, including source separation of uncertainties, choice of error model, and model structural errors. A discussion of each of these issues
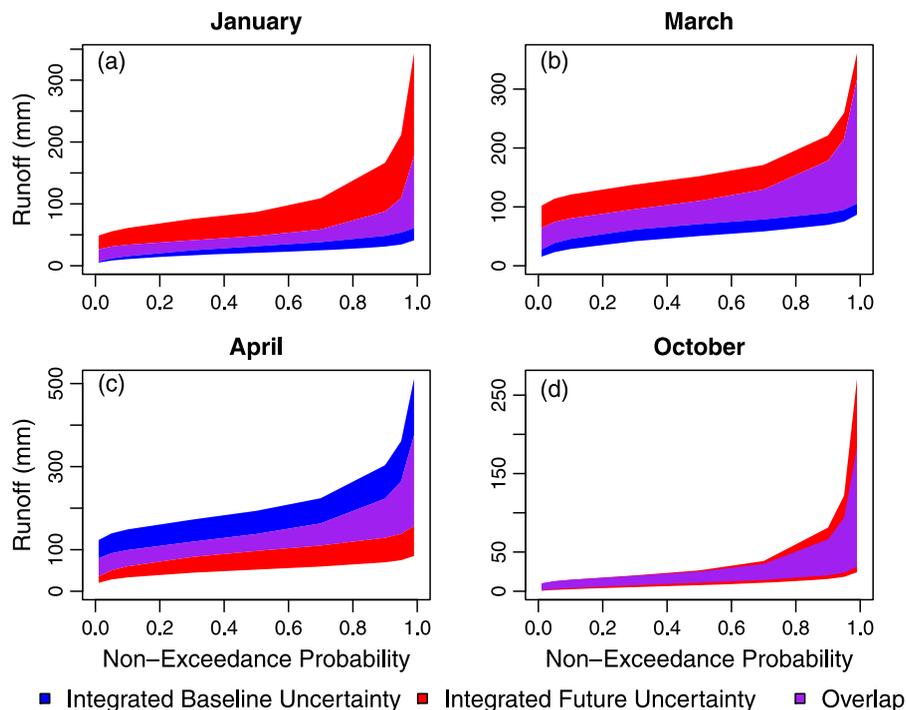
Figure 10.   Integrated 95% predictive bounds in flow quantiles for baseline and future periods for the months of (a) January, (b) March, (c) April, and (d) October.

follows to highlight further research needed to bolster the framework presented in this study.

5.1. Input and Response Data Uncertainties in Future Hydrologic Projections

[58] To simplify the modeling approach this study aggregated all errors associated with input data measurements, response data measurements, and model structure into one error term ε, but the aggregation of different types of error into one term can have significant implications for the quantification of uncertainties in future hydrologic projections [*Thyer et al.*, 2009]. Errors in forcing data sets (e.g., input precipitation data, temperature data, etc.) and observations (e.g., streamflow measurements) are particular to the historic record. Their influence on uncertainty estimates for streamflow predictions should be isolated to the historic period and removed from uncertainty estimates of future streamflow projections. Approaches have been proposed to quantify and separate different sources of uncertainty in hydrologic modeling through Bayesian methods [*Kavetski et al.*, 2006a, 2006b; *Huard and Mailhot*, 2008]. These approaches represent possible contributions of uncertainty from input and output measurement errors using prior distributions chosen by the modeler. Prior distributions for input and output data permit corruptions in those measurements to be filtered out of the calibration process, allowing for more robust and unbiased estimation of hydrologic and error model parameters, along with their associated uncertainties. While not employed in this study, methodologies for separating input and response data uncertainties from uncertainties in future hydrologic projections are promising tools that should be explored in future applications of the proposed framework.

5.2. Error Model Identification and Associated Challenges

[59] The choice of error model used to represent the probabilistic structure of model residuals also plays a critical role in accurately assessing uncertainties in future hydrologic projections. In the vast majority of hydrologic applications, model errors violate assumptions of normality, independence, and homoscedasticity [*Kuczera*, 1983]. If the error model is incapable of capturing these characteristics, parameter estimates can become biased and inferences of parameter and residual uncertainty can degrade [*Thyer et al.*, 2009]. This could significantly impede efforts to accurately propagate hydrologic modeling uncertainty through a climate change impacts analysis.

[60] Previous studies have proposed many alterations to the error model to capture different characteristics of residual error. Several studies have employed autoregressive moving average (ARMA) models and various transformations to model autocorrelated, non-Gaussian, and heteroskedastic errors [*Kuczera*, 1983; *Bates and Campbell*, 2001; *Thiemann et al.*, 2001]. Perhaps the most inclusive error model is proposed by *Schoups and Vrugt* [2010], in which residual errors were modeled using an autoregressive polynomial, a time-variant standard deviation linearly related with mean predicted flow, and a random noise component described by a skew exponential power distribution. The three components allowed the error model to simultaneously model residuals exhibiting autocorrelation, heteroskedasticity, and nonnormality, respectively, without the use of a transformation. A flexible parameterization, inferred through Bayesian techniques, allowed the structure of model errors to be determined during calibration, circumventing the difficulties of specifying error structure a priori. Overall, the advances in explicitly representing the stochastic nature of hydrologic model error are promising and suggest that Bayesian methods to quantify predictive uncertainty may be reliable for complex, high temporal resolution (e.g., daily) models often used in climate change impact analyses. Further research is needed to test this hypothesis.

5.3. Structural Errors in Hydrologic Modeling

[61] Structural errors in conceptual hydrologic modeling arise because spatially and temporally averaged representations of a catchment are often unable to simulate the true dynamics of a distributed and heterogeneous watershed. Structural errors may present one of the biggest challenges to the use of hydrologic models in predicting catchment response to climate change, especially when those responses fall outside the range of historic variability. Efforts to accurately characterize structural error in hydrologic models have met with only moderate success. Many studies assume input and output data are known and lump structural errors into a residual error term [*Bates and Campbell*, 2001; *Marshall et al.*, 2004; *Stedinger et al.*, 2008]. This was the approach taken in this study. Other approaches consider fluxes in rainfall-runoff models as stochastic, using state space approaches [*Vrugt et al.*, 2005] and time-varying parameter values [*Kuczera et al.*, 2006; *Reichert and Mieleitner*, 2009] to compensate for structural deficiencies stemming from spatial and temporal averaging. The use of several different model structures is also a popular choice [*Boorman and Sefton*, 1997; *Wilby and Harris*, 2006; *Jiang et al.*, 2007; *Kay et al.*, 2009; *Prudhomme and Davies*, 2009a, 2009b], and methods like Bayesian model averaging have recently been employed to help generate more reliable predictive intervals from these ensembles [*Duan et al.*, 2007; *Marshall et al.*, 2007]. However, it is often difficult to determine whether enough model structures are considered to develop a complete accounting of structural uncertainties. These different approaches and their underlying assumptions are summarized in more detail by *Renard et al.* [2010]. The formal characterization of structural model uncertainty remains a primary challenge to the hydrologic modeling community, especially as the need for insight about future hydrologic alterations under previously unseen climate forcings increases.

6. Conclusions

[62] There is a growing recognition that advancements in climate change alteration studies are required to inform water resource planners and managers of the magnitude and sources of uncertainty in future hydrologic projections. In particular, of interest is whether projected changes in streamflows are important relative to the baseline error of the hydrologic modeling process. A statistical framework for investigating this question was presented here. Our approach was able to propagate uncertainty from a hydrologic model into future streamflow projections and integrate that uncertainty with

other sources, producing a more complete uncertainty analysis of future hydrology under climate change.

[63] This study employed a very simple but common approach for quantifying future climate uncertainty based on an ensemble of future climate projections. More comprehensive approaches exist, including those that treat climatological uncertainty with formal probability distributions [*Tebaldi et al.*, 2005]. These approaches present an interesting possibility of recasting the entire cascade of model results in a probabilistic framework. However, GCM simulations are projections, not predictions, and therefore a limit likely exists for how useful direct GCM output will be in developing reliable bounds on future climate. It is difficult to compare the raw projections with observations in meaningful ways to assess skill and error, and current practices that rely on a comparison of the marginal distributions of GCM simulations against those of the observations provide "only a limited kind of confidence" [*Stainforth et al.*, 2007b]. In addition, the downscaling methods are often calibrated over the entire historic record, leaving cross-validation approaches impossible. Nonetheless, the framework presented here allows a comparison of the range of climate projections with hydrologic modeling uncertainty.

[64] The application to the White River Basin demonstrates how a comprehensive treatment of uncertainty can reveal varying levels of precision associated with hydrologic alterations across a spectrum of hydrologic responses. This information could be very valuable in assisting water resource managers with decisions regarding adaptation measures to possible climate changes. Depending on the projected direction and severity of climate change impacts on regional hydrology, water resources investments for adaptation can be quite expensive. The possible regret associated with those investments increases rapidly with the uncertainty surrounding future hydrologic alterations, particularly key design flow statistics. Since the minimization of regret is often used to govern decisions regarding large capital investments, a reliable quantification of future hydrologic uncertainty is critical for a robust application of decision theory to climate change adaptation investments in the water sector. This study provides a meaningful contribution toward that end. Future work will propagate future hydrologic uncertainties developed in this study through systems and environmental models to understand the impacts of integrated hydrologic and climate uncertainties on decision-making in fields like water resources and ecohydrology.

References

Allen, R., L. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration guidelines for computing crop water requirements, *FAO Irrig. Drain. Pap. 56*, Food and Agric. Organ., Rome.

Alley, W. M. (1984), On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models, *Water Resour. Res*, 20, 1137–1149.

Arnell, N. W. (1999), The effect of climate change on hydrologic regimes in Europe: A continental perspective, *Global Environ. Change*, 9, 5–23.

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 Resour. Res.*, 37(4), 937–947.

Beven, K. J., and A. M. Binley (1992), The future of distributed hydrologic models: Model calibration and uncertainty prediction, *Hydrol. Processes*, 6, 279–298.

Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems, *J. Hydrol.*, 249, 11–29.

Boorman, D. B., and C. E. M. Sefton (1997), Recognizing the uncertainty in the quantification of the effects of climate change on hydrologic response, *Clim. Change*, 35, 415–434.

Cameron, D., K. Beven, and P. Naden (2001), Flood frequency estimation under climate change (with uncertainty), *Hydrol. Earth Syst. Sci.*, 4(3), 393–405.

Chao, P. (1999), Great Lakes water resources: Climate change impact analysis with transient GCM scenarios, *J. Am. Water Resour. Assoc.*, 35(6), 1499–1507.

Clark, M. P., and J. A. Vrugt (2006), Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters, *Geophys. Res. Lett.*, 33, L06406, doi:10.1029/2005GL025604.

Duan, Q., N. K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Adv. Water Resour.*, 30(5), 1371–1386.

Fernandez, W., R. M. Vogel, and A. Sankarasubramanian (2000), Regional calibration of a watershed model, *Hydrol. Sci. J.*, 45(5), 689–707.

Fowler, H., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modeling, *J. Climatol.*, 27, 1547–1578.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, 7(4), 457–511.

Gleick, P. (1986), Methods for evaluating the regional hydrologic impacts of global climatic changes, *J. Hydrol.*, 88, 97–116.

Hamlet, A. F., and D. P. Lettenmaier (1999), Effects of climate change on hydrology and water resources in the Columbia River Basin, *J. Am. Water Resour. Assoc.*, 35(6), 1597–1623.

Hargreaves, G. H., and Z. A. Samani (1982), Estimating potential evapotranspiration, *J. Irrig. Drain. Eng.*, 108(3), 225–230.

Huard, D., and A. Mailhot (2008), Calibration of hydrologic model GR2M using Bayesian uncertainty analysis, *Water Resour. Res.*, 44, W02424, doi:10.1029/2007WR005949.

Huntington, T. G., G. A. Hodgkins, B. D. Keim, and R. W. Dudley (2004), Changes in the proportion of precipitation occurring as snow in New England (1949–2000), *J. Clim.*, 17, 2626–2636.

Jiang, T., Y. D. Chen, C. Xu, X. Chen, X. Chen, and V. P. Singh (2007), Comparison of hydrologic impacts on climate change simulated by six hydrologic models in the Dongjiang Basin, south China, *J. Hydrol.*, 336, 316–333.

Kavetski, D., G. Kuczera, and S. W. Franks (2006a), Bayesian analysis of input uncertainty in hydrologic modeling: 2. Application, *Water Resour. Res.*, 42, W03408, doi:10.1029/2005WR004376.

Kavetski, D., G. Kuczera, and S. W. Franks (2006b), Bayesian analysis of input uncertainty in hydrologic modeling: 1. Theory, *Water Resour. Res.*, 42, W03407, doi:10.1029/2005WR004368.

Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones (2009), Comparison of uncertainty sources for climate change impacts: Flood frequency in England, *Clim. Change*, 92, 41–63.

Khaliq, M. N., T. B. M. J. Ouarda, J. C. Ondo, P. Gachon, and B. Bobee (2006), Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review, *J. Hydrol.*, 329(3–4), 534–552.

Khan, M. S., and P. Coulibaly (2010), Assessing hydrologic impact of climate change with uncertainty estimates: Bayesian neural network approach, *J. Hydrometeorol.*, 11, 482–495.

Klemes, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, 31, 13–24.

Kuczera, G. (1983), Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty, *Water Resour. Res.*, 19(5), 1151–1162, doi:10.1029/WR019i005p01151.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterizing model error using storm-dependent parameters, *J. Hydrol.*, 331(1–2), 161–177, doi:10.1016/j.hydrol.2006.05.010.

Laio, F., and S. Tamea (2007), Verification tools for probabilistic forecasts of continuous hydrologic variables, *Hydrol. Earth Syst. Sci.*, 11, 1267–1277.

Lettenmaier, D. P., A. W. Wood, R. N. Palmer, E. F. Wood, and E. Z. Stakhiv (1999), Water resources implications of global warming: A U.S. regional perspective, *Clim. Change*, 43(3), 537–579.

Lopez, A., F. Fung, M. New, G. Watts, A. Weston, and R. L. Wilby (2009), From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England, *Water Resour. Res.*, *45*, W08419, doi:10.1029/2008WR007499.

Marshall, L., D. Nott, and A. Sharma (2004), A comparative study of Markov chain Monte Carlo method for conceptual rainfall-runoff modeling, *Water Resour. Res.*, *40*, W02501, doi:10.1029/2003WR002378.

Marshall, L., D. Nott, and A. Sharma (2007), Towards dynamic catchment modeling: A Bayesian hierarchical mixture of experts framework, *Hydrol. Processes*, *21*, 847–861.

Martinez, G. F., and H. V. Gupta (2010), Toward improved identification of hydrological models: A diagnostic evaluation of the "*abcd*" monthly water balance model for the conterminous United States, *Water Resour. Res.*, *46*, W08507, doi:10.1029/2009WR008294.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States, *J. Clim.*, *15*(22), 3237–3251.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy (2007), Fine-resolution climate projections enhance regional climate change impact studies, *Eos Trans. AGU*, *88*(47), 504.

Nijssen, B., G. M. O'Donnell, A. F. Hamlet, and D. P. Lettenmaier (2001), Hydrologic sensitivity of global rivers to climate change, *Clim. Change*, *50*(1), 143–175.

Palmer, T. N., and J. Räisänen (2002), Quantifying the risk of extreme seasonal precipitation events in a changing climate, *Nature*, *415*, 512–517.

Piani, C., D. J. Frame, D. A. Stainforth, and M. R. Allen (2005), Constraints on climate change from a multi-thousand member ensemble of simulations, *Geophys. Res. Lett.*, *32*, L23825, doi:10.1029/2005GL024452.

Prudhomme, C., and H. Davies (2009a), Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: Baseline climate, *Clim. Change*, *93*, 177–195.

Prudhomme, C., and H. Davies (2009b), Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: Future climate, *Clim. Change*, *93*, 197–222.

Räisänen, J., and T. N. Palmer (2001), A probability and decision-model analysis of a multimodel ensemble of climate change simulations, *J. Clim.*, *14*, 3212–3226.

Reichert, P., and J. Mieleitner (2009), Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, *Water Resour. Res.*, *45*, W10402, doi:10.1029/2009WR007814.

Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res.*, *46*, W05521, doi:10.1029/2009WR008328.

Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res.*, *46*, W10531, doi:10.1029/2009WR008933.

Stainforth, D. A., et al. (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, *433*, 403–406.

Stainforth, D. A., T. E. Downing, R. W. A. Lopez, and M. New (2007a), Issues in the interpretation of climate model ensembles to inform decisions, *Philos. Trans. R. Soc. A*, *365*, 2163–2177.

Stainforth, D. A., M. R. Allen, E. R. Tredger, and L. A. Smith (2007b), Confidence, uncertainty and decision-support relevance in climate predictions, *Philos. Trans. R. Soc. A*, *365*, 2145–2161.

Stedinger, J. R., R. M. Vogel, S. U. Lee, and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resour. Res.*, *44*, W00B06, doi:10.1029/2008WR006822.

Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns (2005), Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles, *J. Clim.*, *18*, 1524–1540.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, *37*(10), 2521–2535.

Thomas, H. A. (1981), Improved methods for national water assessment, report, *contract WR 15249270*, U.S. Water Resour. Counc., Washington, D. C.

Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrologic modeling: A case study using Bayesian total error analysis, *Water Resour. Res.*, *45*, W00B14, doi:10.1029/2008WR006825.

Vandewiele, G. L., C.-Y. Xu, and Ni-Lar-Win (1992), Methodology and comparative study of monthly water balance models in Belgium, China and Burma, *J. Hydrol.*, *134*, 315–347.

Vogel, R. M., and A. Sankarasubramanian (2003), Validation of a watershed model without calibration, *Water Resour. Res.*, *39*(10), 1292, doi:10.1029/2002WR001940.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten (2005), Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, *41*, W01017, doi:10.1029/2004WR003059.

Wilby, R. L. (2005), Uncertainty in water resource model parameters used for climate change impact assessment, *Hydrol. Processes*, *19*, 3201–3219.

Wilby, R. L., and S. Dessai (2010), Robust adaptation to climate change, *Weather*, *65*, 180–185, doi:10.1002/wea.543.

Wilby, R. L., and I. Harris (2006), A framework for assessing uncertainties in climate change scenarios: Low-flow scenarios for the River Thames, UK, *Water Resour. Res.*, *42*, W02419, doi:10.1029/2005WR004065.

Wood, A. W., D. P. Lettenmaier, and R. N. Palmer (1997), Assessing climate change implications for water resources planning, *Clim. Change*, *37*, 203–228.