1. Introduction

The services provided by aquatic systems are fundamentally important to humans. In addition to providing clean water for consumption and agriculture, aquatic ecosystems sustain biodiversity and provide support for basic ecological processes as well as important economic activities, including fisheries and recreation. Nevertheless, aquatic systems are heavily impacted by human activities including land use changes associated with agriculture and urbanization, as well as physical modification to river channels which result in altered flow regimes (Miltner et al., 2004; Paul et al., 2006; Sullivan et al., 2006). In addition, the Earth’s climate is predicted to exhibit significant changes in temperature and precipitation during this century due to human activities (Hansen et al., 2006). These expected climatic changes have been detected and already have resulted in measurable impacts on the physical environment (IPCC, 2007).

Increased temperature is the most commonly identified issue regarding predicted changes in climate during the coming century, and the potential impacts of this warming have received the majority of attention (IPCC, 2007). Changes in precipitation patterns are anticipated to be a significant component of climate change as well. Modifications of precipitation patterns including the changes in the magnitude and temporal variability of annual precipitation may result in relatively intense rainfall concentrated during particular times of the year (Kattenberg et al., 1996). Changes in precipitation, in combination with increases in temperature, can have dramatic effects on the...
hydrology of aquatic systems, subsequently impacting water resources as well as the aquatic taxa which are adapted to particular flow regimes (Poff et al., 1997).

Accurate information on the spatial variation in streamflow and the assessment of the potential impacts of climate change on future streamflow regimes are critical for water resource management, particularly in the context of water quantity, quality, and aquatic ecosystem sustainability. The coupling of hydrologic models with global climate models (GCMs) makes the assessment of climate change impacts on water resources possible. Previous studies have examined the impacts of future climate projections downscaled from GCM simulations on water resources (Cherkauer and Sinha, 2010; Hay et al., 2011; Jha et al., 2006, 2004; Kang and Ramirez, 2007; Lettenmaier et al., 1999; Nijssen et al., 2001; Takele et al., 2005). However, most of these studies have focused on the change in the overall water budget rather than the spatial and temporal changes in streamflow variability. Streamflow magnitude and variability are both essential variables influencing the survival, growth, and reproduction of aquatic species, and the directional alteration of these variables can impact local community structure and cause populations to decline (Bain et al., 1988).

Hydrologic models used to predict future water resources under projected warming should accurately reproduce observed streamflow through calibration (Duan et al., 1993; Gupta et al., 1998; Sivapalan et al., 2003; Wagener et al., 2007). A significant challenge in calibration is the identification of appropriate model parameters for distributed hydrologic models. In contrast to lumped models, distributed models account for watershed spatial heterogeneity by using a relatively larger number of parameters. However, not all parameters are measurable because the scale of measurement is usually smaller than the effective scale at which the parameters are applied (Beven, 2001b).

When models are comprised of a relatively large number of parameters, the issue of equifinality is a major concern (Beven, 1993, 2001a; Lo et al., 2010, 2008). That is, multiple sets of parameter combinations can yield similar results. Moreover, distributed hydrologic models can potentially amplify the problems associated with parameter estimations if spatially distributed data are unavailable for calibration. In this case, model calibration usually relies on measured hydrologic responses at a single watershed outlet (Githui et al., 2009; Rouhani et al., 2007; Zhan et al., 2006), such that the phenomenon of “predicting the correct result for the wrong reasons” may occur (Jetten et al., 2003). Though distributed hydrologic models are widely used, there are still very few extensive calibration and validation studies conducted at distributed ground measurements in both water quantity and quality modeling (Beven, 2002). To reduce the possibility of apparently accurate simulations at the watershed outlet resulting from a combination of locally inaccurate simulations, multi-site calibration within a watershed is recommended (Gul and Rosbjerg, 2010; White and Chaubey, 2005; Zhang et al., 2010).

The goal of this study is to predict spatial variation of streamflow and assess the potential impact of climate change on streamflow in watersheds located primarily in Illinois in the Midwestern United States. A distributed hydrologic model, the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998), was calibrated and validated with measured streamflow from multiple gauged sites. The Sequential Uncertainty Fitting Algorithm (SUFI-2) (Abbaspour et al., 2004, 2007) was used for model calibration, validation, and uncertainty analysis. After the SWAT model was calibrated and validated, 26 biased-corrected and spatially downscaled future climate projections derived from nine GCMs were used to drive the validated SWAT model in order to assess the potential impacts of climate change on water resources in the studied watersheds.

2. Materials and methods

2.1. SWAT hydrologic model

SWAT is a physically-based and distributed hydrologic model developed to predict the impacts of changes in landscape management practices on water, sediment, and agricultural chemical yields (Arnold et al., 1998). In addition, SWAT is capable of assessing the impacts of climate change on hydrologic responses and agricultural activities by adjusting climatic variables based on future projections (Arnold and Fohrer, 2005; Neitsch et al., 2005a,b). SWAT typically operates on a daily time step for long-term simulations at large watershed scales. SWAT accounts for spatial heterogeneities by first dividing a large watershed into several sub-basins, and then further dividing the sub-basins into multiple hydrologic response units (HRUs). Each HRU is a combination of unique soil, land cover and management strategies. The simulated water quantity and quality from each sub-basin are routed by streamflow and distributed to the watershed outlet. For a more detailed description of SWAT, see Neitsch et al. (2005d).

2.2. Calibration and uncertainty analysis using SUFI-2

Due to the processes resulting in equifinality (Beven and Binley, 1992), it is difficult to manually calibrate a distributed model in which there are numerous parameters influencing the simulated hydrologic response. The SUFI-2 algorithm was used to assist model calibration, validation and uncertainty analysis (Abbaspour et al., 2004, 2007). Compared with similar techniques such as the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992), Parameter Solution (ParaSol) (van Griensven and Meixner, 2006), and Bayesian inference methods (Kuczera and Parent, 1998), SUFI-2 requires fewer simulations to achieve a similar level of performance (Yang et al., 2008). Instead of identifying absolute parameter values, the characterization of parameter ranges is more important (Bardossy and Singh, 2008). Starting with the initial parameter ranges, SUFI-2 is capable of generating different parameter combinations, comparing simulations with observations, and identifying the optimal parameter ranges. Moreover, instead of calibrating model parameters based on hydrologic responses from a single watershed outlet, SUFI-2 is able to simultaneously calibrate parameters based on distributed data within a watershed. Hydrologic models cannot avoid uncertainties originating from input data, parameters, and model structures (Abbaspour et al., 2007; Dillah and Protopapas, 2006; Dubus and Brown, 2002; Leenhardt, 1995; Zhang et al., 1993). However, SUFI-2 maps more uncertainties onto the parameter ranges and quantifies overall uncertainty in the output of hydrologic response using a 95% prediction uncertainty (95PPU), which, in this study, was calculated at the 2.5% and 97.5% levels of the cumulative distribution of an output variable obtained through the Latin hypercube sampling technique (Abbaspour et al., 2007). Moreover, SUFI-2 quantifies the uncertainties using P-factor and R-factor statistics. P-factor is the percentage of measured data falling into the 95PPI confidence interval, whereas R-factor is the average breadth of the 95PPI band divided by the standard deviation of measured data. The goal of SUFI-2 is to include the majority of measured data with the smallest possible uncertainty bands.

2.3. Study area and data

The study area consists of four watersheds: the Rock River watershed (RRW), Illinois River watershed (IRW), Kaskaskia River watershed (KRW), and Wabash River watershed (WRW) (Fig. 1). All four watersheds are located east of the Mississippi River, primarily
cover portions of Illinois, Wisconsin, and Indiana, and drain a total area of 206,928 km$^2$ (Table 1). The watershed outlets of RRW, IRW, and KRW are located at confluences with the Mississippi River, while the WRW outlet is located at the confluence of the Wabash River and the Ohio River. Dominant soils in Illinois typically have silt loam and silt clay loam textures (Eltahir and Yeh, 1999; Yeh et al., 1998). More information on the Illinois hydrometeorology can be found in Yeh and Famiglietti (2008, 2009).

Streamflow data measured from United States Geological Survey (USGS) gauging stations were used for SWAT calibration and validation. Average monthly streamflow data recorded from 1975 to 2009 were available from 100 USGS gauging stations distributed within these four watersheds (Fig. 1). Three Geographic Information System (GIS) data layers were used to parameterize SWAT: a gridded 30 m digital elevation model (DEM) (NED, 2000), a 1:250,000 digital soil dataset from the State Soil Geographic

### Table 1

| Watershed Information including the number of precipitation weather stations, temperature weather stations, calibration stream gauge stations, drainage area, and land use. |
|---------------------------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Acronym                        | Drainage area (km$^2$) | Number of Prec stations | Number of Temp stations | Number of Calibration stations | Land use above the watershed outlet (%) |
| Rock River watershed           | RRW            | 28401.02         | 33              | 28              | 14              | 75.41           | 8.42            | 9.64            | 5.31            | 1.22            |
| Illinois River watershed       | IRW            | 72985.53         | 78              | 55              | 41              | 71.33           | 12.14           | 12.49           | 2.73            | 1.31            |
| Kaskaskia River watershed      | KRW            | 15418.76         | 29              | 18              | 20              | 70.54           | 16.16           | 9.10            | 3.64            | 0.56            |
| Wabash River watershed         | WRW            | 90123.05         | 83              | 58              | 25              | 68.54           | 18.94           | 9.32            | 1.85            | 1.35            |

Prec: precipitation; Temp: temperature.
and GCM20c3m

is a multiplicative scaling of average monthly precipitation

\[ \text{GCM20c3m} \]

OBSVIC

\[ \frac{P_{\text{obs}}}{P_{\text{GCM20c3m}}} \]

\[ i = 1, 2, \ldots, 12 \]

\[ (1) \]

where \( P_i \) is a multiplicative scaling of average monthly precipitation for each month \( i \) from 1975 to 1999, \( P_{\text{obs}} \) and \( P_{\text{GCM20c3m}} \) are individual average monthly precipitation estimates for 1975–1999 from OBSVIC and GCM20c3m, respectively.

2.4. Climate data

SWAT requires daily climate data, including precipitation, maximum air temperature, minimum air temperature, solar radiation, wind speed, and relative humidity to drive the water balance model. When observed climate data are not available, the stochastic weather generator model within SWAT can simulate daily weather data or fill in gaps of missing data based on average monthly climate statistics from neighboring stations (Sharpley and Williams, 1990). In this study, climate data from four sources, (1) observed climate data from the National Weather Service (NWS) weather stations (OBSNWS) from 1975 to 2009 (http://lwf.ncdc.noaa.gov/oa/ncdc.html), (2) gridded observed climate data compiled by the Variable Infiltration Capacity (VIC) group (OBSVIC) from 1975 to 1999, (3) climate data from various GCMs from 1975 to 1999, and (4) future climate data from various GCMs from 2046 to 2065 and 2081 to 2100, were used to drive SWAT (Fig. 2). OBSNWS from 1975 to 2009, including daily precipitation and minimum and maximum air temperature, were obtained from 223 precipitation stations and 159 temperature stations from the NWS (Fig. 1), and were input into SWAT to drive the streamflow simulations for calibration and validation. The other necessary climate variables were simulated by the SWAT weather generator. Nine global climate projections obtained from the online archive “Bias Corrected and Downscaled WCRP (World Climate Research Programme’s) CMIP3 (Coupled Model Intercomparison Project phase 3) Climate Projections” (Maurer et al., 2010) were used for future SWAT model predictions (Appendix A). A gridded climate dataset (OBSVIC) representing 20th century surface climate conditions based on observations and complied by the Variable Infiltration Capacity group (Maurer et al., 2002) was used to correct the bias of a given GCM’s simulation of the 20th century climate (GCM20c3m). The basis for bias-correction was developed by focusing on datasets of OBSVIC and GCM20c3m at a 2° spatial resolution. The bias-correction was then applied to the GCM future climate projections datasets. Downscaling spatially translated bias-corrected GCM projections (1961–2000, 2046–2065, 2081–2100) data from climate models at a 2° spatial resolution to a 1/8° resolution is more relevant to watershed-scale hydrologic modeling (Hidalgo et al., 2008; Maurer and Hidalgo, 2008; Maurer et al., 2010), thereby providing approximately 184, 434, 94, and 529 grid points within the RRW, IRW, KRW, and WRW, respectively. Analysis of GCM predictions also indicated that a bias in precipitation estimates from 2046 to 2065 and from 2081 to 2100 occurred during the downscaling process, which resulted in an underestimation of precipitation in watersheds in this study. We corrected the precipitation bias introduced from spatial downscaling by calculating a simple multiplicative scaling of average monthly precipitation for each grid between OBSVIC and GCM20c3m, and applying the multiplicative scaling to each grid of downscaled GCM precipitation data from 2046 to 2065 and 2081 to 2100. A simple multiplicative scaling of average monthly precipitation was calculated as:

\[ P_i = \frac{P_{\text{obs}}}{P_{\text{GCM20c3m}}} \]

2.5. Model setup and statistical analysis

The SWAT ArcGIS (version 9.3) interface (Neitsch et al., 2005a,b) was used to write SWAT input files from the DEM, soil, and land use GIS data layers. There were 303, 405, 169, and 902 sub-basins delineated in RRW, IRW, KRW, and WRW, respectively. In order to reduce computational burden in large watersheds, the dominant soil type and land use were characterized in each sub-basin. The workflow of SWAT simulation runs with various past and future climate datasets are displayed in Fig. 2. Climate data from OBSNWS including daily precipitation, maximum air temperature, and

Fig. 2. Schematic diagram of SWAT simulation runs with various climate data sources including: (1) observed climate data from National Weather Service (NWS); (2) gridded observed climate data compiled by the variable infiltration capacity (VIC) group; (3) current climate data (20c3m) from each GCM; and (4) future climate data from each GCM. Measured streamflow from 1978 to 1999 (\( Q_{\text{obs1}} \)) and 2000–2009 (\( Q_{\text{obs2}} \)) were used for calibration and validation, respectively.
minimum air temperature from 1975 to 2009 were input to SWAT. The first three years of simulations were used as a spin-up period and excluded from the analysis. We chose a calibration period from 1978 to 1999 when the internal USGS gauge stations were simultaneously active. Streamflow data from 2000 to 2009 were used for validation. Daily streamflow simulations from SWAT were aggregated into monthly streamflow for calibration and validation. No artificial reservoirs and ponds were included due to the lack of discharge information. Reservoirs and ponds were treated as water or wetlands in land use.

In SWAT, streamflow is the sum of surface runoff, interflow, and base flow generation. SWAT was independently calibrated against measured streamflow from each GCM. The goodness-of-fit of model simulation and prediction capability was evaluated using the Nash–Sutcliffe coefficient (E) and the coefficient of determination (R²).

In the course of the iterative SUFI-2 calibration, the best set of parameters from the last iteration was obtained at the maximum average E of the calibration stations. This set of parameters was subsequently used in SWAT to assess the errors of a given projected climate model and to predict future streamflow in response to climate change. Daily precipitation and maximum and minimum air temperature in two 20-year periods (2046–2065 and 2081–2100) from nine GCMs (26 simulations) were input to SWAT. We quantified the predicted streamflow and its variability from two 10-year periods (2051 to 2060 and 2086 to 2095). When comparing past and future climate variables and streamflow, the 10-year period from 1990 to 1999 was used because both OBSVIC and GCM20C3M have the most recent climate in 1999. The change in future predicted streamflows were assessed based on simulated streamflow from each GCM. Inter-annual variability of predicted streamflow was quantified as the coefficient of variation (CV) and calculated as:

\[
CV_{\text{intra}} = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{\text{STD}(Q_{m})}{Q_{m}} \right)_i
\]

where \(Q_{m}\) is monthly streamflow, \(\text{STD}(Q_{m})\) and \(Q_{m}\) are the standard deviation and the mean of monthly streamflow in a year, and \(i\) is the year number. \(CV_{\text{intra}}\) is the N-year mean of CV for a sub-basin. 

For each sub-basin, inter-annual streamflow variability was calculated as:

\[
CV_{\text{inter}} = \frac{\text{STD}(Q_{a})}{\bar{Q}_{a}}
\]

where \(Q_{a}\) is annual streamflow, \(\text{STD}(Q_{a})\) and \(\bar{Q}_{a}\) are the standard deviation and the mean of annual streamflow. In this study, 10-year inter-annual streamflow variability for the periods 1990–1999, 2051–2060, and 2086–2095 were analyzed. Overall, the intra- and inter-annual streamflow variability is the mean of \(CV_{\text{intra}}\) and \(CV_{\text{inter}}\) from all sub-basins within the study watersheds.

Errors in the past and future streamflow predictions may originate from the SWAT model, observed climate data, and future climate datasets simulated by GCMs. Evaluation of these errors is
made by comparing SWAT simulations with measured streamflow records (Jha et al., 2004). The possible error sources from SWAT and the climate datasets are listed in Table 3. The magnitude of the errors was evaluated by calculating the following bias and root mean square error (RMSE):

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^{N} (Q_{i,i} - Q_{m,i})$$

(4)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Q_{i,i} - Q_{m,i})^2}$$

(5)

where $Q_{m,i}$ and $Q_{i,i}$ are the measured and simulated streamflow on month $i$, respectively, and $N$ is the number of years of streamflow data. The bias provides a measure of annual mean error between simulations and observations. The RMSE gives an estimate of the variability between model simulations and observations, which is used to assess the validity of the model in reproducing the seasonal cycle. Errors were evaluated using observed and simulated streamflow from the following stream gauges: Rock River near Joslin, IL (USGS 05586100), Illinois River at Valley City, IL (USGS 05586100), Kaskaskia River near Venedy station, IL (USGS 05594100), and Wabash River at Mt. Carmel, IL (USGS 03377500).

3. Results

3.1. Parameter ranges during calibration

For each watershed, although the initial parameter ranges are identical, the optimal parameter ranges from calibration are different (Table 2). For example, the initial range of CN2 (SCS runoff curve number for moisture condition II) is –0.2 to 0.2 for all watersheds, but the final range is –0.11 to –0.01, –0.01 to 0.10, –0.13 to –0.02 and –0.04 to 0.03 for RRW, IRW, KRW, and WRW, respectively (Table 2). Similar results are found for other parameters (Table 2). Not all of the 20 parameters are consistently important for the accurate simulation of streamflow among watersheds, only the following four parameters are found to be sensitive in all four watersheds: CN2, ESCO (soil evaporation compensation factor), RCHRG_DP (deep aquifer percolation fraction), and EPCO (plant uptake compensation factor).

3.2. SWAT calibration and validation

Simulated monthly total streamflow from SWAT is in reasonable agreement with the measurements from the 100 USGS gauge stations (Fig. 3). During the calibration, the $R^2$ of 90% of the stations, as well as the coefficient $E$ of 77% of the stations, exceeded 0.6. The number of stations with $R^2$ and $E$ over 0.6 decreases to 78% and 71%, respectively, during the validation period (Fig. 3). More than 60% of the measurements bracketed within the 95PPU are found at 80% of the stations during calibration and 47% during validation. The $R$-factor is below 0.8 at 87% of the stations in the calibration, while it decreases to 57% of the stations during validation (Fig. 4).

Table 3 The sources of error in simulated streamflow from SWAT and climate models.

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>Error source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q_{sim}$ versus $Q_{obs1}$</td>
<td>OBSNWS * SWAT</td>
</tr>
<tr>
<td>$Q_{sim}$ versus $Q_{obs2}$</td>
<td>OBSNWS * SWAT</td>
</tr>
<tr>
<td>$Q_{sim}$ versus $Q_{obs3}$</td>
<td>GCM</td>
</tr>
<tr>
<td>$Q_{sim}$ versus $Q_{obs4}$</td>
<td>Climate change</td>
</tr>
</tbody>
</table>

$Q_{sim}$ and $Q_{obs}$ are simulated and observed streamflows represented in Fig. 2.

Generally, SWAT accurately reproduces long-term mean seasonality and month-to-month variability of hydrologic responses as well as short-term dynamics of individual events, as indicated by the optimal $E$ of 0.53, 0.78, 0.75, and 0.76 during the calibration period for the RRW, IRW, KRW, and WRW, respectively. According to the parameter ranges obtained from the calibration process of SWAT, predicted streamflow during the validation period exhibits a reasonable reproduction of the dynamics of the observed streamflow during events and between events (Fig. 5). The optimal $E$ values in the validation period are 0.68, 0.71, 0.81, and 0.72 in the RRW, IRW, KRW, and WRW, respectively.

3.3. Predicted streamflow from 2051 to 2060 and 2086 to 2095

In order to evaluate the impact of climate change on the spatial variability of streamflow, we first compared 1990–1999 climate conditions from the OBSVIC and GCM20c3m data. Average annual temperature from the OBSVIC dataset is well reproduced by the nine GCM20c3m estimates. The maximum differences in mean temperature between OBSVIC and GCM20c3m are 0.34 °C, 0.26 °C, 0.14 °C and 0.26 °C in the RRW, IRW, KRW, and WRW, respectively. Annual precipitation of OBSVIC was well matched by the nine GCM20c3m estimates, with errors of approximately ±1%. The future mean temperature in 2051–2060 and 2086–2095 increases in comparison with the past mean temperature in 1990–1999. The mean temperature differences between 26 GCM projections for 2051–2060 (2086–2095) and the 1990–1999 temperature are +3.05 °C (+4.57 °C), +2.64 °C (+4.09 °C), +3.27 °C (+4.72 °C), and +3.22 °C (+4.68 °C) in the RRW, IRW, KRW, and WRW, respectively. The percentage changes of annual precipitation from 1990–1999 to 2051–2060 range from −17.98% to 7.25%, −22.98% to 6.91%, −27.58% to −5.31%, and −18.64% to 7.73%, in the RRW, IRW, KRW, and WRW, respectively. The percentage changes in annual precipitation from 1990–1999 to 2086–2095 range from −24.22% to 13.16%, −22.09% to 15.48%, −33.09% to −0.8%, and −21.02% to 16.17%, in the RRW, IRW, KRW, and WRW, respectively.

The SWAT-predicted monthly average streamflows produced by the 26 GCM projections indicates a wide range of potential flow regimes (Fig. 6). Generally, the simulated 2051–2060 and 2086–2095 streamflows from the 26 GCM projections converge during summer, but exhibit wider ranges during winter. For example, the range of the projected 2051–2060 streamflows among GCM projections is 13.71 mm/month in August (maximum 15.77 mm/month and minimum 2.07 mm/month), but increases to 42.28 mm/month in December (maximum 55.17 mm/month and minimum 12.89 mm/month) (Fig. 6a, KRW). However, the relative changes of projected streamflows during the summer do not represent smaller percent ranges compared with 1990–1999 simulations from each GCM20c3m. For example, the relative range of the projected 2051–2060 streamflow is from −80.28% to 57.69% in August and from −64.11% to 94.10% in December (Fig. 6a, KRW). Overall 17, 17, 20, and 18 of 26 predicted annual streamflows in the RRW, IRW, KRW, and WRW, respectively, are predicted to increase in 2051–2060 compared with the simulated streamflow from individual GCM20c3m estimates. During 2086–2095, 18, 15, 20, and 16 of 26 predicted annual streamflows in RRW, IRW, KRW, and WRW, respectively, are predicted to decrease relative to GCM20c3m estimates. The annual streamflows for 2051–2060 (2086–2095) are projected to decrease up to 45.2% (61.3%), 48.7% (49.8%), 48.7% (56.6%), and 41.1% (44.6%) in RRW, IRW, KRW, and WRW, respectively (Appendix B).

3.4. Intra-annual and inter-annual streamflow variability

The majority of simulations indicate that intra-annual and inter-annual streamflow variability will decrease in the future (Table 4).
Spatially averaged intra-annual CV in 1990–1999 is 0.99, which is higher than 21 of the 26 estimates of the intra-annual CV in 2086–2095. None of the intra-annual CV estimates in 2051–2060 are higher than the 1990–1999 measure. The spatially averaged intra-annual CV for the 1990–1999 period using OBSNWS data is 0.71, 0.97, 1.18, and 1.10 in RRW, IRW, KRW, and WRW, respectively. However, the CV for 2051–2060 (2086–2095) period decreases to 0.44 (0.48), 0.90 (0.97), 0.99 (1.07), and 0.86 (0.94) in RRW, IRW, KRW, and WRW, respectively (Appendix C). Approximately 51% of sub-basins have CV estimates greater than 1.0 during 1990–1999 (Fig. 7a). Under the highest CV from the 26 GCM projections, approximately 51% of the sub-basins have intra-annual CV greater than 1.0 during 2051–1999 (Fig. 7a). Under the highest CV from the 26 GCM projections, approximately 51% of the sub-basins have intra-annual CV greater than 1.0 during 2051–2060 (Fig. 7b). The percentage of sub-basins with intra-annual CV estimates greater than 1.0 increases to 85% in 2081–2096 (Fig. 7c). The spatially averaged inter-annual streamflow CV is 0.43 during 1990–1999. Only two of the 26 inter-annual streamflow CV estimates are larger than 0.43 during 2051–2060 and 2086–2095.

3.5. Uncertainty of model simulations

The levels of uncertainty of the SWAT streamflow simulations were calculated by comparing the SWAT simulated streamflow from OBSNWS, OBSVIC, and OBS20C3M (Table 5). A positive value of bias indicates an overestimation of SWAT streamflow simulations, while a negative value indicates an underestimation. The highest percentage of biases is found for OBSNWS (15.0%) and OBSVIC (15.0%) (Table 5). The GCM3 exhibited the highest percentage of biases with an average overprediction of contemporary streamflow of 13.4%, 12.3%, 11.9%, and 13.7%, in the RRW, IRW, KRW, and WRW, respectively (Table 5). The measure of biases in GCM-based future streamflow predictions is the difference between current and future streamflow simulations based on various GCM projection datasets (Jha et al., 2004). Simulated streamflows from various future GCM estimates generally exhibit a relatively higher percentage of bias compared to comparisons between observed contemporary climate data.
Predicted future streamflows based on various GCM3 projections (A2, A1B, B1) result in, on average, the highest percentage of absolute bias, ranging from 48% to 53% among the four watersheds. However, the percentage of highest absolute bias is approximately 61% from GCM6 in the RRW (Fig. 8). The majority of future streamflow predictions based on GCM projections during the 2051–2060 and 2086–2095 periods indicate overall reductions in streamflow. However, a predicted increase in streamflow is found when using the climate data projected by GCMs 1, 2 and 9 (Fig. 8).

4. Discussion

4.1. SWAT calibration and validation

The simulated streamflows produced by SWAT are generally consistent with observations from the four watersheds during the calibration and validation periods. Only three of 100 gauge stations have negative $E$ values. The low goodness-of-fit between observations and simulations could result from any combination of errors from observed and simulated streamflow, climate data, and/or geophysical GIS data including land use and soil maps. However, the influence of poor simulations from these three stations declines with the distance downstream. When a poorly simulated streamflow is routed downstream, the agreement between simulated and observed streamflow is expected to decrease due to error propagation (Abbaspour et al., 2007; Dillah and Protopapas, 2000; Dubus and Brown, 2002; Leenhardt, 1995; Zhang et al., 1993). However, most $E$ values in downstream stations remain between 0.6 and 1.0 in the four watersheds, suggesting error propagation is masked or the influence of error from upstream stations with low goodness-of-fit estimates is small. The results imply the existence of hydrologic error compensation (Cao et al., 2006), in which the combination of poorly simulated streamflow from upstream sub-basins may yield reasonable results at downstream sub-basin and watershed outlets.
With more than 20 years of streamflow calibration data from 14, 41, 20, and 25 USGS gauge stations at RRW, IRW, KRW, and WRW, respectively, streamflow simulations were validated at each station in the four watersheds, and a general agreement with the corresponding observations was found. The results suggest the importance of multi-site calibration and validation when hydrologic responses across large watersheds are desired from model simulations. Without the multi-site calibration and validation, spatial variation within the watershed would be masked (Takken et al., 1999), suggesting the importance of collecting spatially distributed data (when available) to conduct simultaneous multi-site calibration (Zhang et al., 2008).

Fig. 5. Time series of observed and simulated monthly streamflow data generated during the validation period (2000–2009) at the (a) USGS 05446500 Rock River near Joslin, Illinois stream gauge, (b) USGS 05586100 Illinois River at Valley City, Illinois stream gauge, (c) USGS 05594100 Kaskaskia River near Venedy, Illinois stream gauge, and (d) USGS 03377500 Wabash River at Mt. Carmel, Illinois stream gauge.
most sensitive parameters (CN2, ESCO, RCHRG_DP, and EPCO) identified in this study are physically based. As a proxy of hydrologic processes, the results imply that water fluxes in surface, groundwater, and evapotranspiration would be the primary priorities for data collection for conducting multi-site calibration and validation in relatively large watersheds.

4.2. Predicted streamflow

With 26 bias-corrected and downscaled climate projections for the periods 2051–2060 and 2086–2095, our results suggest that annual streamflow will likely decrease based on the majority of GCM-based SWAT model simulations. However, the monthly streamflow pattern is altered throughout the year. The results suggest that streamflow will tend to increase in winter but decrease in summer for both future time periods. The altered pattern of monthly streamflow is more obvious in KRW and WRW, which are located in the southern part of the study region. This altered streamflow pattern is potentially driven by changes in the water budget including alteration of evapotranspiration induced by changes in precipitation and temperature. Warmer temperatures may also play an additional role in the alteration of streamflow patterns. When the temperature is warmer in winter, precipitation more frequently falls as rain instead of snow, which could explain why reduced precipitation in January and February can still result in higher streamflow in the four watersheds.

![Fig. 6. Simulated streamflow using OBS in-c (solid line) and simulated streamflow for: (a) 2051–2060, and (b) 2086–2095 from GCMs with A2 (red), A1B (green) and B1 (blue) scenarios in absolute values (top) and relative to simulations (bottom) in Rock River (RRW), Illinois River (IRW), Kaskaskia River (KRW), and Wabash River (WRW) watersheds. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image)

<table>
<thead>
<tr>
<th>1990–1999</th>
<th>2051–2060</th>
<th>2086–2095</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>Highest.CV</td>
<td>Lowest.CV</td>
</tr>
<tr>
<td>Intra-annual</td>
<td>0.9969</td>
<td>0.9925</td>
</tr>
<tr>
<td>Inter-annual</td>
<td>0.4343</td>
<td>0.4943</td>
</tr>
</tbody>
</table>
Precipitation is predicted to be relatively high in July and August by most GCM projections; however, the streamflow is predicted to be relatively low in these two months. This result suggests that the annual streamflow may be dominated by precipitation, but warmer temperatures could change the rate of evapotranspiration, subsequently influencing monthly streamflow patterns (Islam et al., 2012). Previous studies have indicated similar relationships between rising air temperature and decreased streamflow (Chen et al., 2007; Hawkins and Austin, 2012; Najjar et al., 2010; Tang et al., 2012).

4.3. Streamflow variability

Our results indicate that the majority of simulated streamflow from the 26 GCM projections from 2051–2060 and from 2086 to 2095 exhibit lower variability in terms of intra- and inter-annual
CV compared with that during 1990–1999. Streamflow variability is primarily influenced by factors including frequency and magnitude of precipitation and land use and land cover characteristics (Chang et al., 2012; Nippgen et al., 2011). In this study, land use and land cover were not changed during SWAT simulations for periods 1990–1999, to 2051–2060 and 2086–2095, which limits changes in climate as the primary factor influencing changes in streamflow variability.

The analyses of streamflow variability (CV) suggests that different regions have different buffering capabilities in response to potential climate change scenarios (Hay et al., 2011; Roots, 1989). A plausible explanation is that the hydrologic response in a given watershed may be dominated by different hydrologic processes (Grayson and Bloschl, 2000; Sivakumar, 2004, 2008). Previous studies indicate that total rainfall, rainfall intensity, and soil water are dominant variables driving hydrologic processes in agricultural watersheds (Houser et al., 2000). Temperature is the dominant variable for stream discharge in middle- to high-latitude watersheds where snow tends to accumulate during the winter (Liu et al., 2007; Nijssen et al., 2001). With climate change, in terms of increased temperature and altered precipitation patterns, the response to climate change in snow-covered areas may be more severe than in warmer regions because of extreme hydrologic events associated with snowmelt. The increased temperature would result in relatively greater changes in snowmelt in winter or earlier spring and subsequently cause more soil moisture to evaporate. Even within the same watershed, different land-use types may differentially influence the hydrologic response to climate change. For example, forests (as opposed to agricultural landscapes) could better mitigate the impact of intense precipitation by limiting the amount of soil erosion from the land surface (Nearing et al., 2005; Schiettecatte et al., 2008). The results from this study suggest that different strategies are likely required in different regions in response to anticipated climate change.

4.4. Uncertainties in climate change impact

Climate data from OBSvic were compiled from the National Weather Service COOP database. Maurer et al. (2002) developed the OBSvic according to the daily precipitation and maximum and minimum temperature of OBSnws and gridded these data into a 1/8° spatial resolution using the synergraphic mapping system algorithm. The uncertainty of SWAT simulations combined with that of OBSnws or OBSvic provides a threshold that can be compared to uncertainties arising from using different GCM climate projection data. The bias percentage from SWAT combined with OBSnws or OBSvic is estimated to be between ±15%, which suggests uncertainties from GCM initial conditions (GCM20c3m) are reasonable. However, the bias percentage from most of the 26 GCM projections is greater than ±15% (Fig. 8). These results suggest dramatic changes in streamflow in the future, with most cases indicating reductions in streamflow. In addition, the variability among future streamflow predictions implies that the selection of future projections from climate model simulations critically affects the prediction of water yield under climate change (Stone et al., 2003).
5. Conclusions

We used the SWAT distributed hydrologic model in combination with the SUFI-2 multi-site calibration procedure to demonstrate that spatial and temporal variation in streamflow over large watersheds can be reasonably represented through multi-site calibration and validation. Without the availability of distributed hydrologic data for calibration and validation, distributed hydrologic models are best characterized as lumped models. Distributed hydrologic models are more complex than lumped models in terms of their structure and number of parameters. However, not all model parameters are sensitive to the simulations of hydrologic responses, thereby providing the opportunity to reduce the number of calibration parameters in the distributed models. Among the 20 parameters calibrated, only four parameters are found to be sensitive in all four watersheds, suggesting the importance of multi-site calibration and validation across relatively large spatial scales.

Many of the difficulties and limitations during calibration and validation were data related. We used multi-site streamflow data to calibrate and validate the SWAT. However, streamflow is only one of the components of the water cycle, and these data are widely available compared with other aspects of the water cycle, such as evapotranspiration and soil moisture. Soil moisture data should eventually be available across Illinois, thus offering the opportunity to refine model predictions.

The coupling of hydrologic models and GCM projections allows for the prediction of future streamflow conditions. Using climate model projections to drive the distributed hydrologic simulations, our results suggest the potential for dramatic changes in streamflow as well as streamflow variability during the coming century in most watersheds in Illinois and adjacent midwestern states. In particular, annual streamflows for 2051–2060 (2086–2095) could decrease up to 45.2% (61.3%), 48.7% (49.8%), 48.7% (56.6%), and 41.1% (44.6%) in the RRW, IRW, KRW, and WRW, respectively, while the intra-annual streamflow variability will likely decrease in all watersheds in the periods of 2051–2060 and 2086–2095 compared to that in 1990–1999. Such predictions of hydrologic response to the projected changes in future climate may provide the foundation for water management strategies focused on the mitigation of the impacts of climate change on aquatic resources and ecosystems.

Acknowledgements

We thank the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP’s Working Group on Coupled Modelling (WGCM) for providing the WCRP CMIP3 multi-model dataset. We also thank M. Anthony, C. Beachum, M. Chu, M. Michel, S. Niu, and two anonymous reviewers for providing extremely helpful comments on a previous version of this manuscript. Funding for this research was provided to JHK by the Environmental Protection Agency’s (EPAs) Science to Achieve Results (STARS) Consequences of Global Change for Water Quality program (EPA-G2008-STAR-D2) and from the National Science Foundation (DEB-0844644).

Appendix A.

The GCM model groups, model designation, and acronym. Each GCM has three SRES scenarios for future greenhouse gas emissions forcing global climate. The three SRES scenarios are A2 (higher emissions path), A1B (“middle” emissions path), and B1 (“lower” emissions path). GCM8 only has A2 and B1 scenarios. When the GCM has multiple simulations featuring unique initial conditions to simulate future climate projections, only the first run was used.

<table>
<thead>
<tr>
<th>Modeling group</th>
<th>IPCC model designation</th>
<th>Acronym</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Canadian Centre for Climate Modeling and Analysis</td>
<td>CGCM3.1(T47)</td>
<td>GCM1</td>
</tr>
<tr>
<td>2 Meteo-France/Centre National de Recherches Meteorologiques, France</td>
<td>CNRM-CM3</td>
<td>GCM2</td>
</tr>
<tr>
<td>3 US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory, USA</td>
<td>GFDL-CM2.0</td>
<td>GCM3</td>
</tr>
<tr>
<td>4 US Dept. of Commerce/NOAA/Geophysical Fluid Dynamics Laboratory, USA</td>
<td>GFDL-CM2.1</td>
<td>GCM4</td>
</tr>
<tr>
<td>5 Institute Pierre Simon Laplace, France</td>
<td>IPSL-CM4</td>
<td>GCM5</td>
</tr>
<tr>
<td>6 Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC), Japan</td>
<td>MIROC3.2 (medres)</td>
<td>GCM6</td>
</tr>
<tr>
<td>7 Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA</td>
<td>ECHO-G</td>
<td>GCM7</td>
</tr>
<tr>
<td>8 Max Planck Institute for Meteorology, Germany</td>
<td>ECHAM5/MPI-OM</td>
<td>GCM8</td>
</tr>
<tr>
<td>9 Meteorological Research Institute, Japan</td>
<td>MRI-CGCM2.3.2</td>
<td>GCM9</td>
</tr>
</tbody>
</table>

Appendix B.

Predicted annual streamflows for three periods 1990–1999, 2051–2060 and 2086–2095 in the RRW, IRW, KRW, and WRW based on climate inputs from nine GCMs and three emission scenarios for each GCM (Appendix A). "na" indicates cases where relevant emission scenarios were not available for the particular GCM.
Appendix C.

Intra-annual variability (Coefficient of variation (CV)) of predicted streamflow for three periods 1990–1999, 2051–2060 and 2086–2095 in the RRW, IRW, KRW, and WRW based on climate inputs from nine GCMs and three emission scenarios for each GCM (Appendix A). “na” indicates cases where relevant emission scenarios were not available for the particular GCM.
References


IPCC. 2007 Climate Change 2007: The physical science basis. Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L., (Eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 996.


