Hydrological cycle in the upper Mississippi River basin: 20th century simulations by multiple GCMs

Eugene S. Takle,[1,2] Manoj Jha,[3] and Christopher J. Anderson[1]

Received 26 May 2005; revised 24 July 2005; accepted 18 August 2005; published 28 September 2005.

[1] We used 20th century simulations by nine global climate models (GCMs) to provide input for a streamflow model to simulate baseline hydrologic conditions in the Upper Mississippi River Basin (UMRB). Statistical tests revealed that streamflow data produced by members of the GCM multi-model ensemble were serially uncorrelated at all lags and formed unimodal distributions and that GCM multi-model results may be used to assess annual streamflow in the UMRB. Although all low-resolution GCMs produced large differences from observations of streamflow and hydrological components simulated by the streamflow model, the nine-member ensemble performed quite well. Results of statistical tests indicate that, of all models used, the high-resolution GCM – the only high-resolution model tested – gives simulated streamflows much closer to observed values, despite the fact that its low-resolution sister model has no advantage over the other seven low-resolution models. **Citation:** Takle, E. S., M. Jha, and C. J. Anderson (2005), Hydrological cycle in the upper Mississippi River basin: 20th century simulations by multiple GCMs, *Geophys. Res. Lett.*, *32*, L18407, doi:10.1029/2005GL023630.

1. Introduction

[2] A key question underlying the Global Energy and Water Cycle Experiment (GEWEX, http://www.gewex.org/gewex_overview.html) is whether the hydrological cycle is changing. Recent observations and modeling suggests acceleration of the hydrological cycle at high latitudes in the Northern Hemisphere [*Stocker and Raible*, 2005; *Wu et al.*, 2005]. Assessments of local and regional impacts of changes in the hydrological cycle in future climates call for improved capabilities for modeling the hydrological cycle and its individual components at the subwatershed level.

[3] Determination of impacts of climate change on streamflow requires regional or local representations of meteorological variables derived from global models. These higher resolution datasets can be acquired by (1) extracting grid-point values directly from GCM datasets and linearly interpolating values from global grid points to domain points of interest, (2) using regional climate models (RCMs) to dynamically downscale GCM results, and (3) using statistical models to determine point or regional values from large-scale fields from GCMs. Streamflow models, such as the Soil and Water Assessment Tool (SWAT) [*Arnold and Fohrer*, 2005], accept a wide range of meteorological datasets and use internal weather generators to fill in missing values and create refined details, such as the partitioning of daily precipitation between rain and snow. Therefore, it is not clear whether spatial or temporal refinement of GCM results is warranted when such results are used as input to SWAT. Coupled atmosphere-ocean GCMs have improved physical process models and resolution since the last assessment report of the *Intergovernmental Panel on Climate Change* [2001], and advances in computing capabilities now permit the use of multi-model ensembles, which may reduce biases. Surely, if method (1) for deriving regional/local values gives good results there might be little incentive to perform (2) or (3).

[4] We previously [*Jha et al.*, 2004] reported use of RCM output to drive SWAT for the UMRB where the RCM was driven by reanalysis and a single GCM. We report herein some implications of using multiple GCMs for input to SWAT to estimate annual streamflow and hydrological budget components. We use a subset of 20th century (20C) results of nine GCMs being made available for the IPCC 4th Assessment Report (http://www-pcmdi.llnl.gov/).

2. Domain

[5] The UMRB has a drainage area of 447,500 km^2 up to the point just before the confluence of the Missouri and Mississippi Rivers (Grafton, IL) (Figure 1). Land cover in the basin is diverse and includes agricultural lands, forests, wetlands, lakes, prairies, and urban areas.

[6] For modeling with SWAT, the basin is divided into 119 subwatersheds, each of which is subdivided into hydrological response units (HRUs) such that the basin consists of 474 HRUs. Observed climate data used as input to the hydrological model are provided by 111 weather stations distributed relatively uniformly across the basin. *Jha et al.* [2004] give details of land use, soils, and topography data for the UMRB.

[7] Surface elevations in the UMRB range from 85 m to 640 m ASL with no locations having abrupt changes over this range. Hence, our study domain lacks fine-scale orographic features that otherwise would surely compromise the ability of GCMs to describe the spatial distribution of hydrological processes over a region containing only a few GCM grid points. We ask whether under these conditions GCMs can deliver climatic variables that, when downscaled by simple linear interpolation to provide local values, can allow a calibrated hydrological model to reproduce measured annual mean and interannual variability of streamflow.

3. Models

3.1. SWAT Model

[8] SWAT [*Arnold and Fohrer*, 2005] is a continuous time, long-term, watershed scale hydrologic and water

[1]Department of Agronomy, Iowa State University, Ames, Iowa, USA.
[2]Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa, USA.
[3]Department of Economics, Iowa State University, Ames, Iowa, USA.

0094-8276/05/2005GL023630$05.00

Figure 1. The Upper Mississippi River Basin (UMRB) and delineated subwatersheds. See color version of this figure in the HTML.

quality model. The model was developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in large complex watersheds with varying soils, land use, and management conditions over long periods of time. It is a physically based model that operates on daily time steps and uses readily available inputs.

[9] Subdivision of the watershed into HRUs enables SWAT to reflect differences in evapotranspiration for various crops and soils. Flow amounts estimated for all HRUs are summed and routed through channels, ponds, and/or reservoirs to the watershed outlet. Upland components include hydrology, weather, soil temperature, plant growth, and land and water management. Stream processes include channel flood routing, and ponds and reservoirs contain water balance and routing.

[10] Meteorological input to SWAT includes daily values of maximum and minimum temperature, total precipitation, mean wind speed, total solar radiation, and mean relative humidity. The hydrologic cycle simulated at the HRU level is based on the balance of precipitation, surface runoff, percolation, evapotranspiration, and soil water storage. SWAT partitions total daily precipitation into rain or snow using the mean daily temperature. Snow cover is allowed to be non-uniform due to shading, drifting, topography, and land cover and is allowed to decline non-linearly based on an areal depletion curve. Snowmelt, a critical factor in partitioning between runoff and base flow, is controlled by air and snow pack temperatures, melting rate, and areal coverage of snow. On days when the maximum temperature exceeds 0°C, snow melts according to a linear relationship of the difference between the average snow pack maximum temperature and the threshold temperature for snowmelt. The melt factor varies seasonally, and melted snow is treated the same as rainfall for estimating runoff and percolation. SWAT simulates surface runoff volumes for each HRU using a modified SCS curve number method [*Mishra and Singh*, 2003]. Further details can also be found in the SWAT User's manual [*Neitsch et al.*, 2002]. The version of SWAT used to produce results reported herein is the same model calibrated for the UMRB baseline conditions that was reported by *Jha et al.* [2004].

3.2. Global Climate Models

[11] GCM results were available from nine models (see Table 1) in the IPCC Data Archive (http://www-pcmdi.llnl.gov/) at the time of this writing, including two versions of models from three of the laboratories. While not spanning the full range of model variability (since only a single realization was used for each model) and giving disproportionate weight to models from these three laboratories, results derived therefrom give a preliminary view of streamflow resulting from direct use of data generated by multiple GCMs. We use a subset (i.e., 1961–2000) of model output from the runs simulating the 20C because we have streamflow for this period for comparison with model results. Grid point values from the GCMs were linearly interpolated to domain points of interest.

4. Results

[12] Distributions of annual streamflow are shown in Figure 2, together with the observed gage data (labeled GAGE) at Grafton, IL and results of SWAT driven by observed weather conditions from stations in the basin (labeled OBS). Comparison of GAGE and OBS reveals that SWAT introduces a slight positive bias to the annual streamflow but gives quite good representations of the distribution (inter-annual variability) and extremes. The GCM/SWAT multi-model mean annual streamflow is 282 mm, which is 29 mm (11%) larger than the gage data. Both GAGE and OBS distributions have mode of 300 mm within narrow peaks, but the fraction of annual streamflow $\geq$300 mm is

Table 1. Global Models Used in the SWAT-UMRB Simulations

Institution	Model Name	Lon × Lat Resolution
NOAA Geophysical Fluid Dynamics Laboratory (USA)	GFDL-CM 2.0	2.5° × 2.0°
NOAA Geophysical Fluid Dynamics Laboratory (USA)	GFDL-CM 2.1	2.5° × 2.0°
Center for Climate System Research (Japan)	MIROC3.2 (medres)	2.8° × 2.8°
Center for Climate System Research (Japan)	MIROC3.2 (hires)	1.125° × 1.125°
Meteorological Research Institute (Japan)	MRI	2.8° × 2.8°
NASA Goddard Institute for Space Studies (USA)	GISS_AOM	4° × 3°
NASA Goddard Institute for Space Studies (USA)	GISS_ER	5° × 4°
Institut Pierre Simon Laplace (France)	IPSL-CM4.0	3.75° × 2.5°
Canadian Centre for Climate Modeling and Analysis (Canada)	CGCM3.1(T47)	3.8° × 3.8°

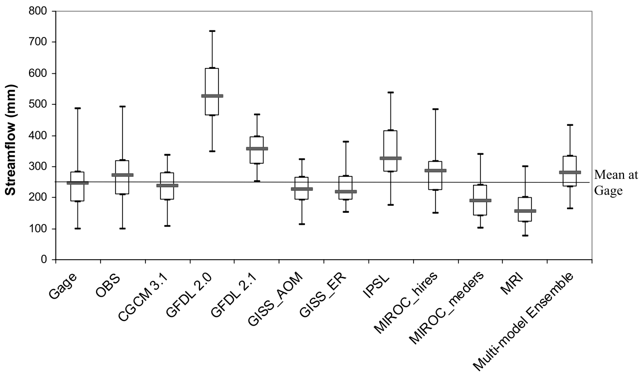

Figure 2. Variability of annual values of GCM/SWAT simulations for a sub-period of the 20C. Measured data at Grafton, IL are labeled as Gage, and SWAT run driven by observed climate is labeled as OBS. Plotted values give median (bold line), quartiles (box values), and lowest and highest values (extremes of whiskers). Dotted line gives mean of the data reported by the gage at Grafton, IL. See color version of this figure in the HTML.

47.3% in GAGE and 61.3% in OBS, thus giving the OBS mean a positive bias (8.7% larger).

[13] We found that each GCM/SWAT streamflow time series was serially uncorrelated at all lags, suggesting that each GCM/SWAT collection of annual streamflow values could be represented as an independent sample from a population rather than as a time series. All data form unimodal distributions (though with varying spread about the peak), and therefore may be modeled by normal distributions. We examined whether each GCM/SWAT might form distributions indistinguishable from OBS/SWAT. Evaluation by use of T-tests of the hypotheses of zero difference between the means of annual streamflow produced by OBS/SWAT and the individual GCM/SWAT simulations revealed that all pair-wise comparisons except MIROC3.2(hires)/SWAT could be rejected at the 2% or higher level (Table 2). The T-test for the MIROC3.2(hires) had a p-value of 0.8312, whereas the p-value for MIROC3.2(medres) was 4.1×10^{-5}, giving strong support to the conclusion that high resolution for the MIROC3.2 model substantially improves simulation of UMRB streamflow. Lack of high-resolution simulations with other models precludes testing the generality of this result. However, the fact that the only high-resolution model reproduced the

Table 2. P-Values of T-Test of Individual GCM/SWAT Streamflow, Pooled GCM/SWAT Streamflow (Labeled GCM POOL), and Measured Streamflow (Labeled Gage) Compared to OBS/SWAT

GCMs	P-Value
GFDL-CM 2.0	4.8303E-17
GFDL-CM 2.1	3.3774E-5
MIROC3.2(medres)	4.1050E-5
MIROC3.2(hires)	0.8312
MRI	0.3963E-8
GISS_AOM	0.0098
GISS_ER	0.0124
IPSL-CM4.0	0.0050
CGCM3.1(T47)	0.0229
GCM POOL	0.5979
Gage	0.1667

Table 3. Hydrological Components Simulated by SWAT[a]

Hydrological Components	OBS/SWAT (1968–1997)	Measured Data	HadCM2/ RegCM2 ~1990	GFDL-CM 2.0	GFDL-CM 2.1	MIROC 3.2, medres	MIROC 3.2, hires	MRI	GISS_AOM	GISS-ER	IPSL-CM 4.0	CGCM 3.1
Precipitation	846	846	900	1032	910	736	821	707	746	746	793	859
Snowfall	118	-	244	213	196	110	104	134	125	95	202	140
Snowmelt	116	-	241	211	193	107	100	130	120	94	200	138
Surface runoff	100	-	148	215	140	55	75	58	63	51	147	80
Baseflow	181	-	213	330	223	145	213	109	170	182	196	161
Potential ET	967	-	788	759	854	1054	984	1011	744	729	692	970
Evapotranspiration (ET)	557	-	533	484	540	531	527	532	505	506	445	611
Total water yield	275	253	350	531	353	194	279	162	227	227	336	232
				Standard Deviation of Annual Values								
Precipitation	113	-	101	110	78	78	88	66	63	56	87	86
Streamflow	84	81	95	108	55	58	69	50	52	53	90	57

[a]Measured streamflow data are at Grafton, IL (USGS gage # 05587450). All values are average annual values (in mm) averaged over 1963–2000 (unless otherwise specified); Years 1961 and 1962 are simulated as initialization period. HadCM2/RegCM2 SWAT simulations [*Jha et al.*, 2004] are average over 10-year period.

Table 4. Results for the Multi-Model Ensemble Mean of SWAT Driven by GCMs and Observed Meteorological Conditions for Sub-Periods of the 20C[a]

Hydrological Components	OBS/SWAT	Measured Data	GCM/SWAT		MIROC 3.2, hires		HadCM2/RegCM2/SWAT	
			Mean	% Diff.	Amount	% Diff.	Amount	% Diff.
Precipitation	846	846	817	−03	821	−03	900	+06
Snowfall	118	-	147	+25	104	−12	244	+206
Snowmelt	116	-	144	+24	100	−13	241	+208
Surface runoff	100	-	98	−02	75	−25	148	+48
Base flow	181	-	192	+06	213	+18	213	+18
Potential ET	967	-	866	−10	984	+02	788	−15
ET	557	-	520	−07	527	−05	533	−04
Total water yield	275	253	282	+11	279	+10	350	+38

[a]Percent differences are calculated from measured data when available and otherwise from results of SWAT driven by observed meteorology. Different averaging periods were used as follows: OBS/SWAT: 1968–1997; GCM/SWAT: 1963–2000; MIROC3.2 (hires)/SWAT: 1963–2000; and HadCM2/RegCM2/SWAT: 1990–1999. Results in the last two columns are from *Jha et al.* [2004].

mean of the record and none of the lower resolution models did adds incentive to further explore the resolution issue.

[14] The coarse GCM/SWAT distributions with smallest p-values (CGCM3.0, GISS_AOM, GISS_ER) each were slightly skewed in the sense opposite to that of OBS/SWAT for which the distribution mean is composed of relatively large weighting on annual streamflow values ≥300 mm. Despite the overlap of interquartile ranges in Figure 2, the variance was insufficient to accommodate the difference of mean values.

[15] T-tests indicate statistically significant differences between GCM-simulated and observed annual streamflow. We examined potential for creating an ensemble distribution composed of all GCM/SWAT results. We computed pair-wise correlation for all GCM combinations and found none statistically different from zero, indicating the individual time series of GCM/SWAT annual streamflow are uncorrelated to one another. Furthermore, the GCM data, by definition, come from different sources. One potential advantage of creating an ensemble of GCMs is that model errors may be "averaged-out" if errors are uncorrelated. Since the individual time series of GCM/SWAT annual streamflow are uncorrelated to one another, we may hypothesize that there is a population from which all GCM/SWAT results represent independent samples. We tested the hypothesis of zero difference between mean annual streamflow of the pooled GCM/SWAT and OBS/SWAT results and found a p-value of 0.5979, suggesting that a GCM/SWAT multi-model ensemble may provide valid assessments of annual streamflow in the UMRB. However, it should be pointed out that physical processes may be poorly represented or completely absent in GCM simulations, which may preclude detailed process analysis, such as water cycling between terrestrial and atmospheric reservoirs.

[16] GCM/SWAT values of standard deviation generally are smaller than OBS/SWAT values for all GCMs. The average of the individual GCM/SWAT standard deviations is 71 mm compared to 79 mm for the gage data and 84 mm for the OBS/SWAT data. The average ratio of standard deviation to mean for the GCMs is 0.27 compared to 0.33 for GAGE and 0.31 for OBS. The average of the standard deviation is less skillful than individual model standard deviations.

[17] SWAT calculates components of the hydrological budget from the meteorological data supplied by each model. Rainfall gages in the UMRB provide measurements of precipitation, and gage data at Grafton provide measurements of streamflow. We estimate other (unmeasured) hydrological components with SWAT using weather-station input. Precipitation amounts (Table 3) derived directly from GCMs vary from −16 to +22% of observed. SWAT estimates that 14% of the observed precipitation in the basin comes in the form of snow, while the GCM-derived estimates put this percentage at 13–22%. Runoff varies from −49% to +115% of that calculated for observed climate inputs to SWAT. Evapotranspiration (ET) and potential ET span a more narrow range of −23% to +9%, and total water yield (i.e., surface runoff + base flow − transmission losses, the latter being a minor factor) is from −35% to 110% of the gage-measured streamflow. Table 3 also presents standard deviations of precipitation and corresponding simulated streamflows. It was found that interannual precipitation variability is correlated with interannual streamflow variability, as indicated by the coefficient of determination value of 0.71. This suggests that model skill in simulating precipitation is crucially important for skillful simulations of annual streamflow variability.

[18] GCM/SWAT model means of hydrological components show quite good agreement with gage data and observations-driven simulations (Table 4). Snowfall (and resulting snowmelt) presents the dominant challenge among the hydrological components, with the mean of all global models giving about 25% more snow than simulated by SWAT from observed weather, possibly due to a seasonal positive bias in precipitation or negative bias in temperature. However, the high-resolution MIROC3.2 results also agree with observations.

[19] Climate models generally produce too many light rain events and too few intense events [*Gutowski et al.*, 2003] even if rainfall totals are accurate. The impact of this bias, compared to the true intensity spectrum, is to reduce runoff and increase ET and/or base flow. Low bias on rainfall likely would lead to low runoff, base flow, ET, and hence water yield, while excess rain would have the opposite effect.

[20] We previously reported results of using a regional climate model (RegCM2) to dynamically downscale results of a global model (HadCM2) to the UMRB [*Jha et al.*, 2004]. The HadCM2/RegCM2/SWAT results (Table 4) show large differences from the GCMs in partitioning precipitation to snowfall (27%), which can be traced to a

1–2 mm/day positive bias in precipitation and low temperature bias in HadCM2/RegCM2 in winter and spring.

[21] From Tables 3 and 4, we conclude that: (1) use of a GCM drawn at random to drive SWAT could lead to sizable errors in streamflow and hydrological cycle components, (2) use of the mean streamflow from a multi-model ensemble of GCM/SWAT simulations, by contrast, performs quite well for this task, (3) the lone high-resolution GCM does as well as the multi-model ensemble mean despite large errors in its lower-resolution sister model, and (4) the downscaled results of a global model by a regional model (models chosen on the basis of availability) used to drive SWAT are inferior to those resulting from the GCM model mean and the high-resolution GCM.

5. Conclusions

[22] We found that GCM/SWAT values for annual streamflow were serially uncorrelated at all lags and form unimodal distributions, suggesting that the data may be modeled as independent samples from an identical normal distribution and that GCM/SWAT multi-model ensemble results may provide a valid approach for assessing annual streamflow in the UMRB. The multi-model ensemble mean of GCM/SWAT simulations demonstrated good performance in reproducing observed precipitation (3% error) and streamflow (11% error) despite large differences among individual models. MIROC3.2(hires) – the only high resolution model tested – simulated observed streamflow with a p-value 36 times larger than the next best model, suggesting a benefit of grid refinement of GCMs.

[23] **Acknowledgments.** We acknowledge the international modeling groups for providing their data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the model data, the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy.

References

Arnold, J. G., and N. Fohrer (2005), Current capabilities and research opportunities in applied watershed modeling, *Hydrol. Processes*, *19*, 563–572.

Gutowski, W. J., Jr., S. G. Decker, R. A. Donavon, Z. Pan, R. W. Arritt, and E. S. Takle (2003), Temporal scale of precipitation errors in central US climate simulation, *J. Clim.*, *16*, 3841–3847.

Intergovernmental Panel on Climate Change (2001), *Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., 881 pp., Cambridge Univ. Press, New York.

Jha, M., Z. Pan, E. S. Takle, and R. Gu (2004), Impact of climate change on stream flow in the Upper Mississippi River Basin: A regional climate model perspective, *J. Geophys. Res.*, *109*, D09105, doi:10.1029/2003JD003686.

Mishra, S. K., and V. P. Singh (2003), *Soil Conservation Service Curve Number (SCS-CN) Methodology*, *Water Sci. Technol. Libr.*, vol. 42, Springer, New York.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, R. Srinivasan, and J. R. Williams (2002), Soil and Water Assessment Tool: User manual, version 2000, TWRI Rep. TR-192, 455 pp., Tex. Water Resour. Inst., College Station.

Stocker, T. F., and C. C. Raible (2005), Water cycle shifts gear, *Nature*, *434*, 830–833.

Wu, P., R. Wood, and P. Stott (2005), Human influence on increasing Arctic river discharges, *Geophys. Res. Lett.*, *32*, L02703, doi:10.1029/2004GL021570.

C. J. Anderson and E. S. Takle, Department of Agronomy, Iowa State University, Ames, IA 50011, USA. (gstakle@iastate.edu)

M. Jha, Department of Economics, Iowa State University, Ames, IA 50011, USA.