LETTERS

Global pattern of trends in streamflow and water availability in a changing climate

P. C. D. Milly[1], K. A. Dunne[1] & A. V. Vecchia[2]

Water availability on the continents is important for human health[1,2], economic activity[3], ecosystem function[4] and geophysical processes[5]. Because the saturation vapour pressure of water in air is highly sensitive to temperature, perturbations in the global water cycle are expected to accompany climate warming[6]. Regional patterns of warming-induced changes in surface hydroclimate are complex and less certain than those in temperature, however, with both regional increases and decreases expected in precipitation and runoff. Here we show that an ensemble of 12 climate models exhibits qualitative and statistically significant skill in simulating observed regional patterns of twentieth-century multidecadal changes in streamflow. These models project 10–40% increases in runoff in eastern equatorial Africa, the La Plata basin and high-latitude North America and Eurasia, and 10–30% decreases in runoff in southern Africa, southern Europe, the Middle East and mid-latitude western North America by the year 2050. Such changes in sustainable water availability would have considerable regional-scale consequences for economies as well as ecosystems.

Streamflow is a temporally lagged, spatial integral of runoff over a river basin. Averaged over many years, runoff generally is equal to the difference between precipitation and evapotranspiration and, hence, to the convergence of horizontal atmospheric water flux. From a resource perspective, runoff is a measure of sustainable water availability. However, streamflow can be affected by anthropogenic disturbances, which may generate spurious (that is, nonclimatic) changes; at the spatial scale of basins to be considered here, the most significant of these disturbances is associated with the diversion of water for the irrigation of cropland.

In support of an assessment of forced climate change conducted by the Intergovernmental Panel on Climate Change (IPCC), many climate-modelling centres recently performed '20C3M' simulations of climate with prescribed external forcing (variations in atmospheric composition and solar irradiance) for the late nineteenth century and the whole of the twentieth century. Forcing was not identical across models, but generally included estimated historical variations of radiatively active atmospheric gases and aerosols (including volcanic emissions) and solar irradiance. Control simulations with temporally invariant preindustrial forcing ('PICNTRL') were also performed, as were integrations into the future with an assumed forcing model ('SRESA1B').

The annual runoff fields from a total of 62 runs of the 20C3M experiment on 21 different models (one to nine runs per model) were integrated spatially over 165 river basins with long-term (28–99 years, median 59 years) streamflow measurements (see Methods). Climate-model runoff commonly does not reflect the time lag associated with storage in river basins. Although the missing time lag may not affect computed climates, it can strongly affect the temporal variability of streamflow, which is important for our analysis. Accordingly, the climate-model basin runoff time series were converted to equivalent streamflow time series by routing them through a model of a linear reservoir in such a way as to match the observed serial correlation (see Methods).

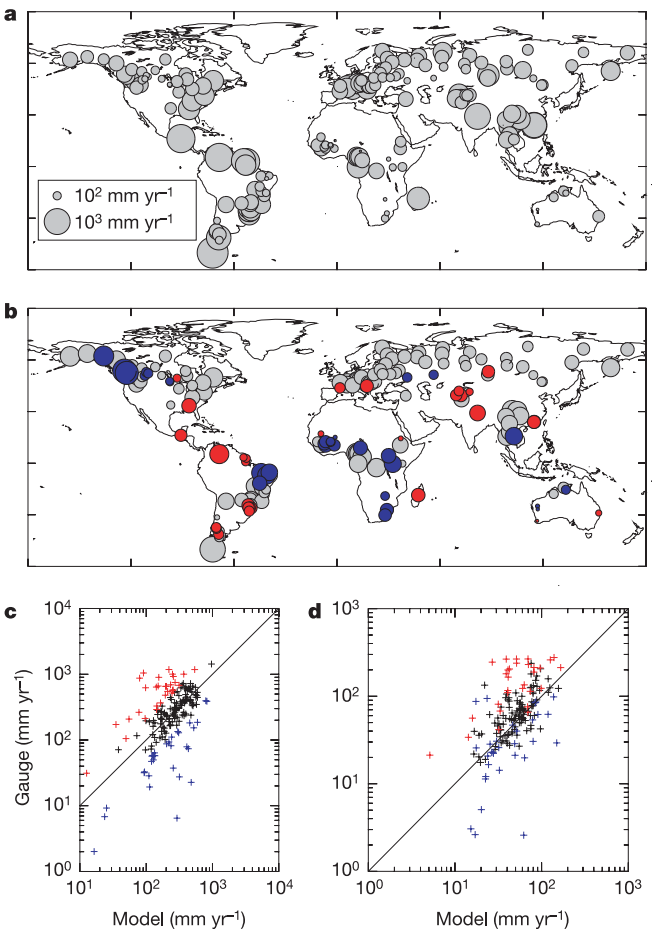

Figure 1 | Annual runoff rate (streamflow per unit basin area, in mm yr^{-1}). **a**, Global distribution of mean values from stream-gauge observations; runoff is proportional to the area of each circle, and each circle is centred at a gauged basin centroid. **b**, Global distribution of ensemble (geometric) means from 35 model runs. Here and in **c** and **d**, gauges with a mean greater than double (smaller than half) the observations are shown in blue (red). (Geometric mean was used because runoff estimates varied greatly across models for a given gauge; results for arithmetic means are similar.) **c**, Observed versus model ensemble means. **d**, Observed versus model ensemble standard deviations.

[1]US Geological Survey, Geophysical Fluid Dynamics Laboratory/NOAA, PO Box 308, Princeton, New Jersey 08542, USA. [2]US Geological Survey, 821 E. Interstate Ave., Bismarck, North Dakota 58503-1199, USA.

To maximize the signal-to-noise ratio in our analysis, we chose to work with an ensemble of models. The realism of hydroclimatic simulations varies across models, so we elected to form the ensemble from a subset of the models, with the selection based on performance. We ranked the models with respect to root-mean-square (r.m.s.) error (over the 165 basins and all runs) of the logarithm of long-term mean discharge per unit area; the logarithmic transform is commonly used in hydrology because flows can range over orders of magnitude. We retained the 12 models (35 runs of 20C3M) with the lowest error for use in the ensemble analyses presented here (see Methods).

The mean and standard deviation of annual streamflows (expressed per unit basin area) for the ensemble model output generally range from about one-half to double the observed values, except in a few regions (Fig. 1). In relation to the observed values, the ensemble model mean values tend to be large in much of Africa, the Nordeste region of South America (northeastern Brazil), and northwest North America, and small in northern low latitudes of the Americas and southern South America. Despite the presence of these large local-scale differences between the model ensemble and observed values, the global-scale relation is strong.

To characterize twentieth-century changes in streamflow at gauges, we used the difference, D, between the average annual streamflows for 1971–98 (based on available sample size m) and 1900–70 (sample size n) (see Methods). Let s and r be the sample standard deviation and lag-1 autocorrelation of the pooled annual streamflow time series. Under the null hypothesis of a stationary climate, the normalized difference, defined as $Z = D/\{s(1/m + 1/n)^{1/2}[(1+r)/(1-r)]^{1/2}\}$ and henceforth termed the 'trend', is approximately standard-normally distributed[7]. Thus, normalizing D in this way accounts for local differences in record lengths and streamflow variability and persistence. Although Z can be used to measure local significance of change, our objective is not to assess local significance but rather to determine whether the observed Z values for the 165 basins are correlated significantly with the values predicted by the 12-model ensemble. If so, we can conclude that external forcing explains a significant part of global streamflow change for the twentieth century.

The pattern of hydroclimatic change indicated by the ensemble mean of the model trends qualitatively resembles the pattern in the observations (Fig. 2). The observed tendency towards less runoff in sub-Saharan Africa, southern Europe, southernmost South America, southern Australia and western mid-latitude North America generally is seen in the ensemble. The ensemble reproduces the observed increases in runoff in the La Plata basin of southern South America, southeastern through central North America, the southeastern quadrant of Africa, and northern Australia. In northern Eurasia and far northwestern North America, the ensemble shows a strong upward trend of runoff, which is consistent with, though more robust than, a general upward trend in the observations in this zone. (Patterns of change in model precipitation (not shown) are qualitatively similar to those of model runoff, as shown previously for similar experiments[8].)

Differences between ensemble trends and observed trends are also apparent. Outside the tropics, the observed trends show less coherence in space than the ensemble trends. Consistent disagreement in sign of the trend is most apparent in Central America and northern South America (where a marked bias in average runoff has already been noted), northeastern Europe, and central and southeast Asia.

The correlation between ensemble and observed trends in streamflow across basins is +0.34. The correlations between trends computed from individual models and the observed trends are all positive, ranging from +0.05 to +0.28, with a mean value of +0.16. A linear-regression slope of 1.51 for observed versus ensemble

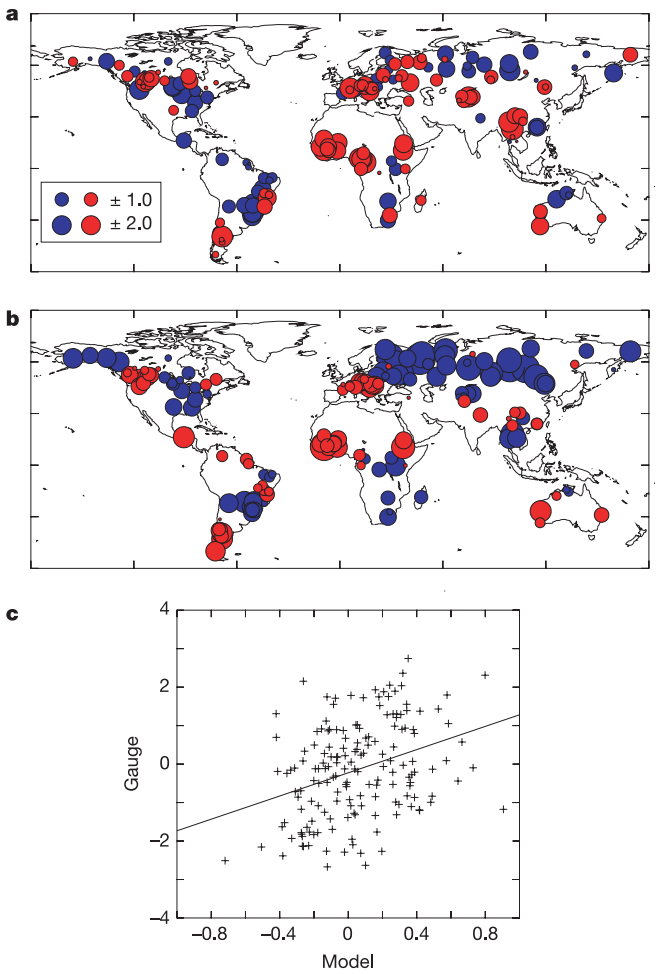

Figure 2 | Global distributions of trend (Z) in streamflow from 1900–70 to 1971–98. **a**, Stream-gauge observations. **b**, Ensemble (arithmetic) means of 35 model-run Z values, multiplied by $35^{1/2}$ to account for the reduction in variance caused by averaging. **c**, Plot of observations against means of 35 model-run Z values. The ordinary least-squares regression line shown has the equation gauge data = 1.51 × (model ensemble) − 0.23.

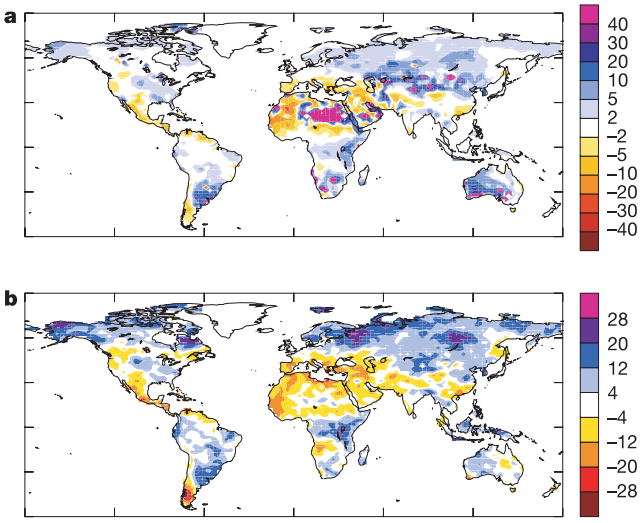

Figure 3 | Relative change in runoff during the twentieth century. **a**, Ensemble (arithmetic) mean of relative change (percentage) in runoff for the period 1971–98, computed as 100 times the difference between 1971–98 and 1900–70 runoff in the 20C3M experiments, divided by 1900–70 runoff. **b**, Number of runs (out of a total of 35) showing a positive change minus the number showing a negative change.

trends indicates that observed trends are larger on average than modelled trends, but not significantly so.

Could the positive correlation between the ensemble and observed trends have arisen, by chance, as a result of internal (that is, unforced, natural) variability in the climate system? To address this question of statistical significance, we needed to estimate the sampling distribution of the correlation coefficient between the ensemble trends and the trends computed from repeated realizations of an unforced climate system. To do this, we formed as many distinct 99-year segments of output from the PICNTRL experiment as were available from the 12 models in the ensemble. This sampling yielded 49 synthetic sets of observations, which were mapped to the period 1900–98, masked to allow the use only of values from years and gauges when and where real observations were made, and then used to compute the trend statistics. We determined the correlation of trends in each of these synthetic observation sets with the ensemble average of those from the 20C3M time series. The 49 correlation values ranged from −0.32 to +0.33, with a mean of 0.01; the 49 regression slopes ranged from −1.12 to +1.25, with a mean of 0.05. We assume that the 49 correlation values can be used to approximate the distribution from which the value +0.34 would have been drawn under the null hypothesis of a stationary hydroclimate. Because none of the 49 values are as large as +0.34, we infer that the correlation between the forced-model ensemble trends and the observed trends is statistically significant. This inference relies on the assumption that the models faithfully represent interbasin correlation of internal variations of runoff in the models.

Figures 3 and 4 show twentieth-century and twenty-first-century percentage changes in runoff estimated by the model ensemble, with indications of the degree of agreement among models on the direction of change. The model projections for the twenty-first century are dependent on various assumptions, for example those connected with future greenhouse-gas emissions, volcanic activity and solar variability. Quantitative projections by the model ensemble also are affected by large model errors in some basins (Fig. 1), but the demonstrated retrospective skill suggests qualitative validity of the projections. The ensemble-average change in runoff by the period 2041–60 shows a pattern generally consistent with that of twentieth-century change, although amplified and with important qualitative differences. In general, areas of increased runoff shrink over time (that is, from the late twentieth century to the mid twenty-first century), whereas areas of decreased runoff grow. Initial increases of runoff in the twentieth century are projected to reverse in the twenty-first century in eastern equatorial South America, southern Africa and the western central plains of North America. Modelled drying of the Mediterranean region extends farther north into Europe in the twenty-first-century runs than in the twentieth-century runs.

Almost all model runs agree on the direction of twenty-first-century change in certain regions (Fig. 4). These agreements include increases (typically 10–40% by 2050) in the high latitudes of North America and Eurasia, in the La Plata basin of South America, in eastern equatorial Africa and in some major islands of the equatorial eastern Pacific Ocean. Prominent regions of agreement on decreasing (typically 10–30%) runoff include southern Europe, the Middle East, mid-latitude western North America, and southern Africa.

On the basis of this analysis, it seems that a significant part of twentieth-century hydroclimatic change was externally forced, that larger changes can be expected in the coming decades, and that climate models can help now to characterize future changes. Henceforth it may be prudent to include projections of forced hydroclimatic change as factors in assessments of water availability, thereby facilitating their consideration not only in water management but also in economic and ecological assessment and planning.

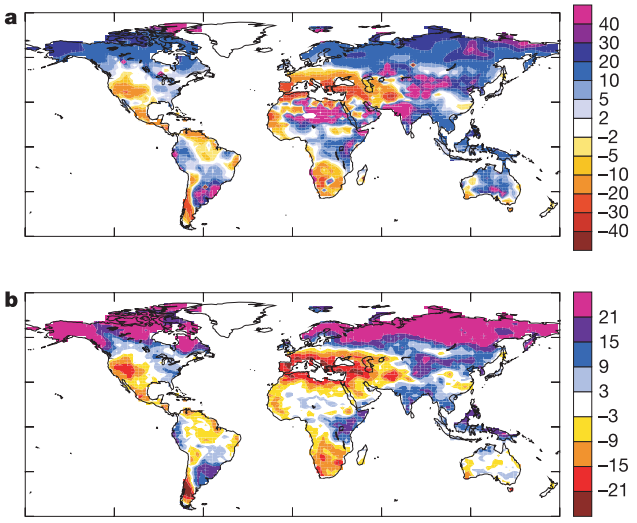

Figure 4 | Relative change in runoff in the twenty-first century. **a**, Ensemble (arithmetic) mean of relative change (percentage) in runoff for the period 2041–60, computed as 100 times the difference between 2041–60 runoff in the SRESA1B experiments and 1900–70 runoff in the 20C3M experiments, divided by 1900–70 runoff. **b**, Number of pairs of runs (out of an available total of 24 pairs) showing a positive change minus the number showing a negative change.

METHODS

From a previously defined set of 663 gauged river basins[9], we selected the 165 gauges judged most suitable for analysis of hydroclimatic change. To be included, a basin was required to have at least 28 years of data and no more than 10% of values missing during the period of record. To avoid overweighting of relatively gauge-rich Europe and North America in the analyses, only basins with drainage area greater than 50,000 km^2 were included for those continents. Net diversions for irrigation (diversions minus return flows, estimated as the product of irrigated area in the basin[10] and the excess, if any, of mean potential evaporation[11] over mean precipitation[12]) were required to be less than 10% of mean flow. Results reported throughout this paper were only slightly sensitive to these subjective numerical constraints on record length, missing values, basin area, and irrigation. The monthly time series of observed discharge were obtained from the Global Runoff Data Centre and averaged to annual values for all analyses reported here. Delineation of the drainage basin associated with each gauge was determined by use of the Simulated Topological Network (STN-30p)[13].

For year i, the conversion from basin-average model annual runoff $y(i)$ to model streamflow $q(i)$ was given by $q(i) = rq(i-1) + (1-r)y(i)$. For each model and each gauge, the value of r was assigned the difference between the lag-1 autocorrelation of annual values from observed streamflow and the lag-1 autocorrelation of annual model runoff; in the rare cases in which this difference exceeded the largest of all observed values of autocorrelation in the observed discharge (0.90), it was set to the latter value instead. The initial value of q was set equal to the time-average value of y. For comparability, only years with observations were sampled from the models when making comparisons with observations.

The r.m.s. difference between the observed and modelled natural logarithm of discharge ranged from 0.98 to 3.5. Ensembles were formed from the 12 models for which the r.m.s. difference was less than 1.3. In terms of institutional designations used by the Program for Climate Model Diagnosis and Intercomparison (PCMDI), these 12 models are CCSM3, CGCM3.1(T63), ECHAM5/MPI-OM, ECHO-G, FGOALS-g1.0, GFDL-CM2.0, GFDL-CM2.1, GISS-AOM, MIROC3.2(hires), MRI-CGCM2.3.2, UKMO-HadCM3 and UKMO-HadGEM1.

The 1900–98 time range was selected for our analysis because it was the longest for which all models provided output. The choice of 1970 for the break in analyses of change was based partly on our observation from previous model investigations that global-mean measures of hydroclimatic change became noticeable at about this time; additionally, this choice maximized the number of basins with data both before and after the break. We investigated the sensitivity of our results to changes in the break year and found that the results for a 1980 break were similar to those for a 1970 break. A 1960 break generally resulted in smaller (and less significant) trends, and a 1990 break resulted in more variable (and less significant) trends because of the small sampling period thereafter. For break years of 1960, 1970, 1980 and 1990, we found that 12, 0, 0

and 5, respectively, of the 49 synthetic observations yielded correlations exceeding those of the real observations. If all 21 models (62 runs of 20C3M and 81 segments of PICNTRL) are used instead of the selected 12 models, the numbers of synthetic observations with correlations exceeding those of the real observations are 20, 6, 1 and 7 (out of 81) when using a break year of 1960, 1970, 1980 and 1990.

Received 18 May; accepted 12 October 2005.

1. Reiter, L., Falk, H., Groat, C. & Coussens, C. M. (eds) *From Source Water to Drinking Water: Workshop Summary* (National Academies Press, Washington DC, 2004).
2. United Nations Educational Scientific and Cultural Organization. *Water for People—Water for Life, The United Nations World Water Development Report* (Berghahn Books, Oxford, 2003).
3. Shiklomanov, I. A. & Rodda, J. C. (eds) *World Water Resources at the Beginning of the 21st Century* (Cambridge Univ. Press, Cambridge, 2003).
4. Mooney, H., Cropper, A. & Reid, W. Confronting the human dilemma. *Nature* **434,** 561–562 (2005).
5. Cazenave, A. *et al.* Space techniques used to measure change in terrestrial waters. *Eos* **48,** 59 (2004).
6. Allen, M. & Ingram, W. J. Constraints on future changes in climate and the hydrologic cycle. *Nature* **419,** 224–232 (2002).
7. Brockwell, P. J. & Davis, R. A. *Time Series: Theory and Methods* Ch. 7 (Springer, New York, 1987).
8. Manabe, S., Wetherald, R. T., Milly, P. C. D., Delworth, T. L. & Stouffer, R. J. Century-scale change in water availability: CO_2-quadrupling experiment. *Clim. Change* **64,** 59–76 (2004).
9. Fekete, B. M., Vörösmarty, C. J. & Grabs, W. High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Glob. Biogeochem. Cycles* **16** (2002); published online 7 August 2002 (doi:10.1029/1999GB001254).
10. Siebert, S., Döll, P., Feick, S. & Hoogeveen, J. *Global Map of Irrigated Areas version 2.2* (Johann Wolfgang Goethe University, Frankfurt am Main, Germany; Food and Agriculture Organization of the United Nations, Rome, 2005).
11. Darnell, W. L. *et al. Surface Radiation Budget: A Long-term Global Dataset of Shortwave and Longwave Fluxes* [online] ⟨http://www.agu.org/eos_elec/95206e.html⟩ (1996).
12. Adler, R. F. *et al.* The Version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4,** 1147–1167 (2003).
13. Vörösmarty, C. J., Fekete, B. M., Meybeck, M. & Lammers, R. B. Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. *Glob. Biogeochem. Cycles* **14,** 599–621 (2000).

Acknowledgements We thank T. Cohn, T. Delworth, I. Held, G. Hodgkins, H. Lins and R. Stouffer for advice. Streamflow data were provided by national hydrometric services through the Global Runoff Data Centre, Koblenz, Germany. Climate-model outputs were provided by modelling centres through PCMDI. We acknowledge the JSC/CLIVAR Working Group on Coupled Modelling and their Coupled Model Intercomparison Project and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. This investigation was performed within the IPCC Global Hydroclimates Analysis Project.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.C.D.M. (cmilly@usgs.gov).