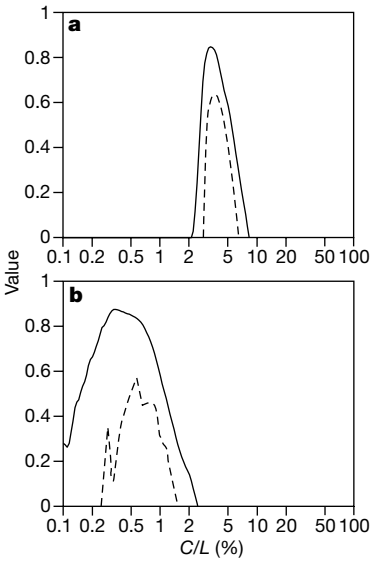

Figure 3 The potential economic value of probabilistic projections of climate change based on a generic binary decision model (see text for details). *C* and *L* are possible financial losses, relative to some reference index, associated with two alternative investments. Solid lines, value based on the CMIP2 greenhouse ensemble probability of E_n over the next 80 years. Dashed lines, value based on an estimate of E_n from a single member of the CMIP2 greenhouse ensemble. When value is zero or less, the corresponding estimate of the probability of E_n is no better than an estimate based on the control ensemble. A value of one would correspond to a perfect decision strategy. **a**, $n = 2$; **b**, $n = 3$.

globe. This value is illustrated in terms of the non-dimensional quantity $V_i = (M_i - M_l)/(M_u - M_l)$, where M_l and M_u are the lower and upper reference values respectively. Hence if $V_i = 0$, then the estimate of $\bar{p}$ of subset i is no better than an estimate which assumes that the climate of the twenty-first century will be the same as that of the twentieth century. On the other hand, perfect investment decisions have a value $V_i = 1$. In calculating V_i we average over all choices of verifying model. For V_2 we also average over all choices of forecast model.

Figure 3 shows V_1 and V_2 for $n = 2$ and $n = 3$, as a function of C/L. We do not show V_3 because it is never greater than zero. The reason for this is straightforward: the number of occurrences of extreme climate events is severely underestimated in the consensus projection. This arises because ensemble averaging tends to produce smooth fields which are unrealistically biased towards the climate-mean state.

If C/L is either sufficiently small, or sufficiently large, then accurate ensemble forecasts are clearly not needed to make good investment decisions (buy B or A, respectively). Hence, the greenhouse ensemble only has positive economic value (compared with the baseline value M_l) for a subset of possible C/L. This ensemble tends to be most valuable (compared with M_l) when C/L is comparable with the control ensemble estimate of $\bar{p}$. As $\bar{p}$ is smaller for E_3 than for E_2, then the greenhouse ensemble provides value for smaller C/L, when decisions are based on the occurrences of E_3 rather than E_2. It can be seen that the value of the greenhouse ensemble is never less than the value of a single deterministic projection, and that, overall, the difference in value between the ensemble and the single deterministic projection is larger for E_3 than for E_2. This is because the relative unreliability of the single deterministic projection is larger for E_3 than for E_2.

In fact, much larger ensembles are needed to provide reliable estimates of the probability of E_3 on a regional basis (for example, specific to the UK or for Bangladesh). Moreover, more research is needed to develop sound methodologies for representing model uncertainty in climate-prediction ensembles[10]. In addition, such extreme-event risk analysis would benefit from an increase in the resolution of present-day climate models (with grid sizes of hundreds of kilometres). For example, it would clearly be of more direct relevance for the decision problem discussed above if E was defined by some specific river bursting its banks during a particular season. However, estimating the future risk of this type of event would require feeding climate model output into a basin-specific hydrological model. In general, climate model grid sizes of the order of ten kilometres would be necessary to simulate precipitation statistics adequately over a typical catchment area (though for rivers with large catchment basins, such as the Brahmaputra, coarser climate model grids may be adequate). An ability to estimate reliably the probability of extreme climate events is an important factor in defining future computational requirements for the climate change problem. □

Received 27 September; accepted 18 December 2001.

1. Houghton, J. T. *et al.* (eds) *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, Cambridge, 2001).
2. Palmer, T. N. Predicting uncertainty in forecasts of weather and climate. *Rep. Prog. Phys.* **63,** 71–116 (2000).
3. Richardson, D. S. Skill and relative economic value of the ECMWF ensemble prediction system. *Q. J. R. Meteorol. Soc.* **126,** 649–668 (2000).
4. Palmer, T. N., Brankovic, C. & Richardson, D. S. A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Q. J. R. Meteorol. Soc.* **126,** 2013–2034 (2000).
5. Marsh, T. J. The 2000/01 floods in the UK—a brief overview. *Weather* **56,** 343–345 (2001).
6. Räisänen, J. & Palmer, T. N. A probability and decision-model analysis of a multi-model ensemble of climate change simulations. *J. Clim.* **14,** 3212–3226 (2001).
7. Meehl, G. A., Boer, G. J., Covey, C., Latif, M. & Stouffer, R. J. The coupled model intercomparison project. *Bull. Am. Meteorol. Soc.* **81,** 313–318 (2000).
8. Palmer, T. N. A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrisation in weather and climate prediction models. *Q. J. R. Meteorol. Soc.* **127,** 279–304 (2001).
9. Kharin, V. V. & Zwiers, F. W. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *J. Clim.* **13,** 3760–3788 (2000).
10. Allen, M. R. Do-it-yourself climate prediction. *Nature* **401,** 642 (1999).

Acknowledgements

We thank M. Blackburn and P. J. Webster for comments on an earlier draft of this Letter.

Increasing risk of great floods in a changing climate

P. C. D. Milly*, R. T. Wetherald†, K. A. Dunne* & T. L. Delworth†

** US Geological Survey, GFDL/NOAA; and † Geophysical Fluid Dynamics Laboratory/NOAA, P.O. Box 308, Princeton, New Jersey 08542, USA*

Radiative effects of anthropogenic changes in atmospheric composition are expected to cause climate changes, in particular an intensification of the global water cycle[1] with a consequent increase in flood risk[2]. But the detection of anthropogenically forced changes in flooding is difficult because of the substantial natural variability[3]; the dependence of streamflow trends on flow regime[4,5] further complicates the issue. Here we investigate the changes in risk of great floods—that is, floods with discharges exceeding 100-year levels from basins larger than 200,000 km^2—using both streamflow measurements and numerical simulations of the anthropogenic climate change associated with greenhouse gases and direct radiative effects of sulphate aerosols[6]. We find that the frequency of great floods increased substantially during the twentieth century. The recent emergence of a statistically

significant positive trend in risk of great floods is consistent with results from the climate model, and the model suggests that the trend will continue.

A focus on large basins facilitates the initiation of quantitative analyses of flood risk with numerical climate models; basins commonly used in trend analyses[4,5] are too small to be resolved by current models. Estimation of extreme-event risk requires multi-decadal observational records. Here, we consider 29 basins larger than 200,000 km^2 in area for which discharge observations span at least 30 yr. We analyse annual maximum monthly-mean flows, rather than annual maximum instantaneous flows; these two are strongly correlated in large basins. In contrast with earlier studies[4,5], this investigation has a global scope and focuses on extreme events; we analyse the 100-yr flood (that is, the river discharge that has a probability of 0.01 of being exceeded in any given year), which is commonly used in flood-risk assessment for river-basin planning and design of major structures. Choosing such a large-magnitude threshold probably reduces any distortion of our analysis by non-climatic factors such as land-use changes and river development.

For each basin, we fitted observed annual maximum monthly-mean discharges to Pearson's type III distribution[7] by the method of moments, and determined the 100-yr flood magnitude from the fitted distribution. The 100-yr flood was exceeded 21 times in our observational record of 2066 station-years. Flood events were concentrated in the latter half of the record; half of the observations were made after 1953, and 16 of the flood events occurred after 1953. Under the assumption that flood events were independent outcomes of a stationary process, we used binomial probability theory to determine a probability of 1.3% of having 16 or more of 21 events during the second part of the record. For observations from an extratropical subset of the basins (see below), the corresponding probability is 3.5%, for 7 out of 8 flood events in the second half of the record. Supplementary analyses for shorter return periods (2–50 yr) did not reveal significant trends, but 200-yr flood frequency increased significantly.

We now refine the simple analysis above to address certain shortcomings: the assumption of independence among flood events, the use of a crude index of the trend based on simple bisection of the historical sequence, and the presence of sampling errors in our estimates of 100-yr flood magnitudes. We first introduce a more robust measure, Z, of the flood-frequency trend; Z is the slope of the least-squares linear relation between annual flood frequency (number of flood events divided by number of operating stations) and time, with annual data values weighted by number of operating stations. To estimate the probability density function of Z under constant climate, we used output from a 900-yr

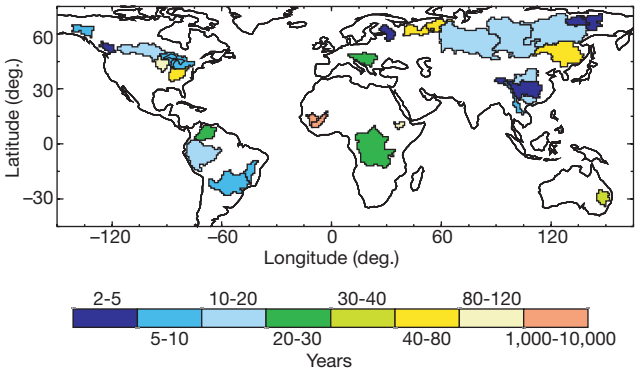

Figure 2 Map showing the gauged drainage areas and flood-risk sensitivities of the 29 river basins in this study. Colour indicates the modelled return period, under idealized quadrupling of atmospheric CO_2 concentrations, of the flood magnitude associated with a 100-yr return period in the control experiment. Although results for low-latitude basins are provided, the poor performance of the model in low latitudes should be kept in mind.

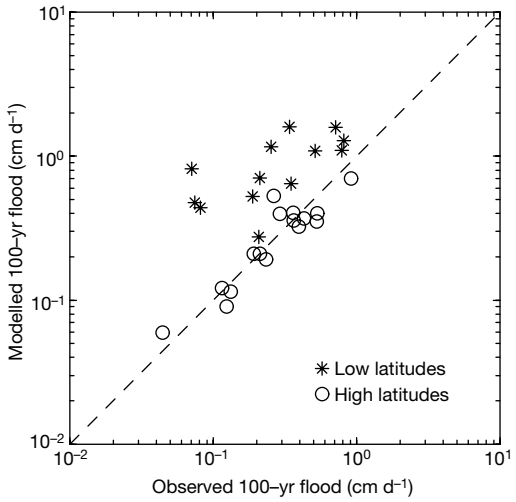

Figure 1 Estimates of 100-yr flood discharges based on model (control experiment years 201–900) and observations (full period of record). Discharges are expressed as monthly-mean discharges per unit basin area. The 16 high-latitude basins are those in North America, Europe, and northern Asia (Fig. 2), and are listed in Table 1. The starting point for basin selection was a previous synthesis and analysis of available streamflow records[13], in which rivers and gauging sites were chosen to minimize non-climatic sources of non-stationarity. That analysis yielded 34 non-nested basins that satisfy the present constraints on basin area (>200,000 km^2) and record length (at least 30 yr). Following more extensive tests for non-climatic non-stationarity in those records, we eliminated four basins on the basis of identified effects of dam construction and one basin on the basis of a change in the gauging site. Collectively, the 2066 station-years of observations on the remaining 29 basins span the 135-yr period 1865–1999; the average record length is 71 yr. For the observations and the model, annual maximum monthly-mean discharges were fitted to Pearson's type III distribution[7] by the method of moments, and the 100-yr flood was determined from the fitted distribution. To obtain discharge in the climate model, runoff was aggregated over all model cells upstream of the stream gauge of interest and routed through a reservoir whose storage is proportional to its outflow (the modelled river discharge at the gauge site). The constant of proportionality is a residence time whose value was assigned *a priori* on the basis of an analysis of observed basin-mean precipitation and discharge power spectra[14].

Table 1 Effect on river discharge of quadrupling atmospheric CO_2

River	Station	dq (%)	dQ (%)	*P* (%)
Yukon	Eagle, USA	54	15	11
Nelson	Bladder Rapids, Canada	76	36	8.8
Fraser	Hope, Canada	44	17	27
St Lawrence	Cornwall, Canada	20	15	12
Mississippi	Keokuk, USA	14	–1	0.90
Ohio	Metropolis, USA	–12	9	2.3
Danube	Orsova, Romania	–3	16	4.6
Neva	Novosaratovka, Russia	36	22	31
N. Dvina	Ust-Pinega, Russia	39	2	1.4
Pechora	Ust-Tsilma, Russia	25	5	2.2
Ob'	Salekhard, Russia	41	14	7.7
Yenisei	Igarka, Russia	23	13	5.9
Lena	Kusur, Russia	25	25	9.9
Yana	Dzanghky, Russia	42	43	32
Indigirka	Vorontsovo, Russia	42	30	46
Amur	Khabarovsk, Russia	4	10	2.4

Changes in discharge of extratropical rivers associated with idealized quadrupling of atmospheric CO_2 concentration in the model. dq is the relative change in annual mean discharge, dQ is the relative change in 100-yr annual maximum monthly discharge, and P is the annual probability of 100-yr flood (defined with respect to the control experiment), after quadrupling of atmospheric CO_2. We used the 100 yr of model output that begins 60 yr after stabilization of CO_2 concentration at the quadrupled level. Post-quadrupling distributions of annual maximum monthly flows were fitted to Pearson's type III distribution, which was then used to determine the probability of the control 100-yr flood.

'control' (constant radiative forcing) experiment with a coupled ocean–atmosphere–land model[6]. The model simulates well 100-yr flood thresholds (and annual discharge statistics) for basins far outside the tropics, but systematically overestimates flood magnitudes in the lower latitudes (Fig. 1); accordingly, we performed significance analyses both for the full set of basins and for the subset of 16 higher-latitude ('extratropical') basins, and we confined much of the subsequent analysis to the extratropical domain. For the historical record, $Z = 1.99 \times 10^{-4}\,\mathrm{yr}^{-1}$ with all basins ($1.41 \times 10^{-4}\,\mathrm{yr}^{-1}$ extratropical).

To evaluate the significance of these values of Z, we extracted from the control experiment 500 overlapping (hence, non-independent) 135-yr sequences (years 177–311, 178–312, ... , 676–810) of flows, mapped each to the time period 1865–1999, and sampled these sequences for the river-specific periods of observations in the historical record. For each sequence, we estimated the 100-yr flood level for each basin (thereby simulating the sampling error inherent in the observational analysis), determined flood occurrences, and calculated Z. The observed value of Z was exceeded in none of the 500 sequences when all basins were considered, and was exceeded in 3.2% of the sequences when only extratropical basins were considered. (Overall trends in flood frequency over the analysed part of the control experiment were negligible.) Thus, the model-based significance analysis, which implicitly uses the space–time correlation structure of floods in the model, essentially confirms and reinforces the simpler binomial analysis.

The apparent increase in flood risk might be associated with radiatively forced climate change. To assess flood-risk sensitivity to radiative forcing, we used a 300-yr 'idealized CO_2 quadrupling' experiment with a 1%-per-year growth (for 140 yr) of atmospheric CO_2 concentration from the control level to a stable, quadrupled level[8] (maintained for 160 yr). Modelled changes in annual mean discharge (relative to the control experiment) provide a measure of intensification of the water cycle as a result of idealized CO_2 quadrupling (Table 1). For the extratropical basins, these range from −12% to +76%, with a median value of +30%. Relative changes in the 100-yr monthly maximum discharge generally are smaller and less variable, with a median of +15%. In all but one of the basins, the control 100-yr flood is exceeded more frequently as a result of idealized CO_2 quadrupling (Fig. 2). The probability of exceeding this control flood changes by a factor that ranges from 0.90 to 46; in half of the basins, the factor exceeds 8 (implying a decrease in return period from 100 yr to shorter than 12.5 yr).

Given the substantial modelled sensitivity of flood risk to radiative forcing, we framed the hypothesis that historical changes in radiative forcing may explain the significant observed increase in flood risk. We examined the detectability of flood-risk change in five transient 'scenario' climate experiments (225 yr, 1865–2089) that shared common estimates of historical and projected future changes in radiative forcing by greenhouse gases and direct effects of sulphate aerosols, each with a distinct initial condition[6]. These experiments show an increase in extratropical flood frequency that generally is apparent early in the twenty-first century (Fig. 3). Thereafter, the flood rate is 2 to 8 times greater than its value during the historical period of observations. Values of extratropical Z were computed for each scenario with exactly the same gauging schedule as in the observations. Four of the experiments (all except scenario 4) had positive values of Z; the largest of these ($1.32 \times 10^{-4}\,\mathrm{yr}^{-1}$, in scenario 3) was slightly smaller than the observed value.

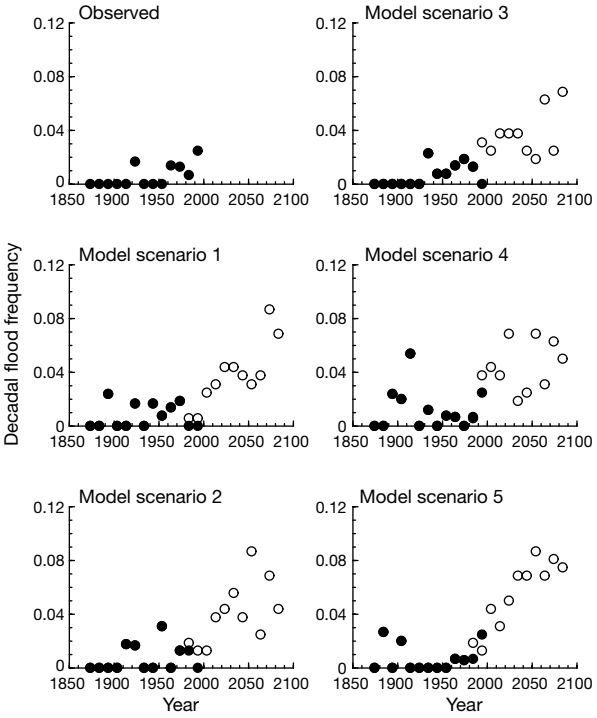

Figure 3 Decadal extratropical flood frequencies for observations and for scenario experiments. Flood frequency is the number of events exceeding the 100-yr discharge, divided by the number of station-years of observations. In plots of model outputs, filled circles are obtained with each station starting and ending as in the historical record; open circles are obtained with all stations continuing operation once begun. The symbols coincide before the 1980s, because no station record ends until 1986. Distinct 100-yr flood levels were defined for each of the six data sets with respect to the historical schedule of observations over that data set.

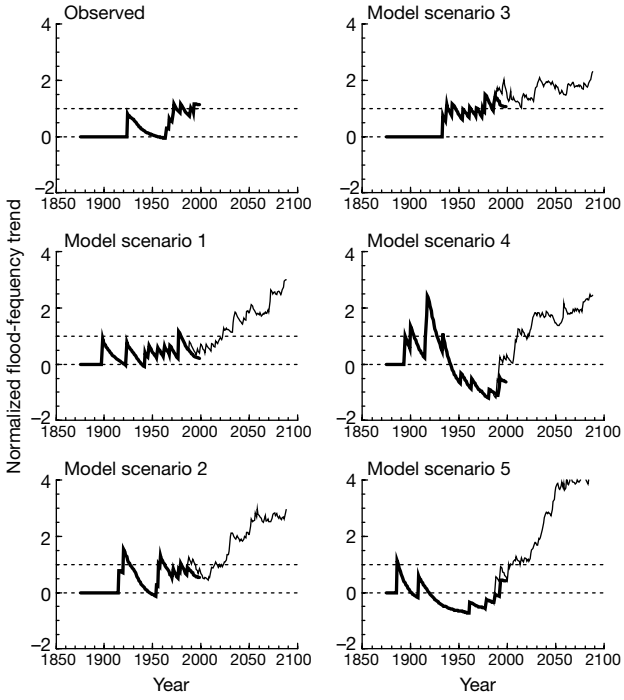

Figure 4 Normalized trend $Z(t)/Z_{95}(t)$ of extratropical flood frequency for observations and for scenario experiments. $Z_{95}(t)$ is the value of the flood-frequency trend statistic $Z(t)$ that equals or exceeds the values for 475 (95%) of the 500 control-experiment segments (see text). Normalized trends greater than 1 imply significance with respect to this measure. Results indicated by bold curves are obtained with each station starting and ending as in the historical record; results shown by light curves are obtained with all stations continuing operation once begun. These curves are identical before 1986.

Because trends and their statistical significance are random functions of time, an analysis of the time variation of detectability of flood-frequency change may be informative. Therefore, we generalized Z to $Z(t)$, where t is the hypothetical last year of available records, and evaluated $Z(t)$ for the observations and the scenario experiments (Fig. 4). The observed flood trend $Z(t)$ has been significantly (at 95% level, as evaluated from the control experiment) different from zero continuously since the flooding of the upper Mississippi River in 1993 and intermittently since 1972. Uninterrupted periods of statistically significant flood-frequency trends in scenarios 1 to 5 begin in years 2023, 2023, 1986, 2021 and 2006, respectively, and premonitory multidecadal periods of intermittent significance begin much earlier in scenarios 2 and 3 (1956 and 1937, respectively). Thus, the recent history of the observed trend index is generally consistent with the range of results from the scenario experiments.

Our detection of an increase in great-flood frequency and its attribution to radiatively induced climate change are tentative. The frequency of floods having return periods shorter than 100 yr did not increase significantly. Potentially significant effects of measurement non-stationarity are not easily assessed. The forced signal and unforced variability in the model contain errors of unknown magnitude. Absent from the model are forcings such as solar variability, volcanic activity, land-cover change[9], and water-resource development[10], and potential biospheric feedbacks such as CO_2-induced stomatal closure[11] and water-stress-induced root extension[12]. Especially evident from our study are the needs for improvements in simulation of tropical hydroclimate and continued commitment to stream-gauging programmes worldwide. □

Received 2 October; accepted 17 December 2001.

1. Cubasch, U. in *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. *et al.*) Ch. 9 (Cambridge Univ. Press, Cambridge, 2001).
2. White, K. S. *et al.* Technical Summary in *Climate Change 2001: Impacts, Adaptation and Vulnerability* (eds McCarthy, J. J. *et al.*) 19–73 (Cambridge Univ. Press, Cambridge, 2001).
3. McCabe, G. J. Jr & Wolock, D. M. Climate change and the detection of trends in annual runoff. *Clim. Res.* **8,** 129–134 (1997).
4. Lins, H. F. & Slack, J. R. Streamflow trends in the United States. *Geophys. Res. Lett.* **26,** 227–230 (1999).
5. Groisman, P. Ya., Knight, R. W. & Karl, T. R. Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century. *Bull. Am. Meteorol. Soc.* **82,** 219–246 (2001).
6. Knutson, T. R., Delworth, T. L., Dixon, K. W. & Stouffer, R. J. Model assessment of regional surface temperature trends (1947-1997). *J. Geophys. Res.* **104,** 30981–30996 (1999).
7. Pearson, K. *Tables for Statisticians and Biometricians* 3rd edn (Cambridge Univ. Press, Cambridge, 1930).
8. Delworth, T. L. *et al.* Simulation of climate variability and change by the GFDL R30 coupled climate model. *Clim. Dyn.* (submitted).
9. Chase, T. N., Pielke, R. A. Sr, Kittel, T. G. F., Nemani, R. R. & Running, S. W. Simulated impacts of historical land cover changes on global climate in northern winter. *Clim. Dyn.* **16,** 93–105 (2000).
10. Milly, P. C. D. & Dunne, K. A. Trends in evaporation and surface cooling in the Mississippi River basin. *Geophys. Res. Lett.* **28,** 1219–1222 (2001).
11. Sellers, P. J. *et al.* Comparison of radiative and physiological effects of doubled atmospheric CO_2 on climate. *Science* **271,** 1402–1406 (1996).
12. Milly, P. C. D. Sensitivity of greenhouse summer dryness to changes in plant rooting characteristics. *Geophys. Res. Lett.* **24,** 269–271 (1997).
13. Milly, P. C. D. & Dunne, K. A. Macroscale water fluxes: 1. Quantifying errors in the estimation of basin-mean precipitation. *Wat. Resour. Res.* (submitted).
14. Milly, P. C. D. & Wetherald, R. T. Macroscale water fluxes: 3. Effects of land processes on variability of monthly river discharge. *Wat. Resour. Res.* (submitted).

Acknowledgements

Streamflow data were collected and/or provided to us by the Global Runoff Data Centre (Germany), the US Geological Survey, Environment Canada, EarthInfo, Manaus Harbor Ltd/Portobras and Departmento Nacional de Aguas e Energia Electrica (Brazil), National Institute of Meteorology and Hydrology (Romania), J. Cordova, N. Garcia and J. Richey. We thank K. L. Findell, D. P. Lettenmaier and R. J. Stouffer for comments and suggestions.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.C.D.M. (e-mail: cmilly@usgs.gov).

Antarctic climate cooling and terrestrial ecosystem response

Peter T. Doran*, John C. Priscu†, W. Berry Lyons‡, John E. Walsh§, Andrew G. Fountain‖, Diane M. McKnight¶, Daryl L. Moorhead#, Ross A. Virginia☆, Diana H. Wall, Gary D. Clow††, Christian H. Fritsen‡‡, Christopher P. McKay§§ & Andrew N. Parsons****

** Department of Earth and Environmental Sciences, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607, USA*
† Land Resources and Environmental Sciences, 334 Leon Johnson Hall, Montana State University, Bozeman, Montana 59717, USA
‡ Byrd Polar Research Center, Ohio State University, 1090 Carmack Road, Scott Hall, Columbus, Ohio 43210, USA
§ Department of Atmospheric Sciences, University of Illinois, 105 South Gregory Street, Urbana, Illinois 61801, USA
‖ Department of Geology, Portland State University, Portland, Oregon 97207, USA
¶ Institute of Arctic and Alpine Research, 1560 30th Street, Campus Box 450, Boulder, Colorado 80309, USA
 # Department of Earth, Ecological and Environmental Sciences, 2801 W. Bancroft Street, University of Toledo, Toledo, Ohio 43606, USA
☆ Environmental Studies Program, Dartmouth College, 6182 Steele Hall, Hanover, New Hampshire 03755, USA
*** Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523, USA*
†† USGS—Climate Program, Box 25046, MS 980, Denver Federal Center, Denver, Colorado 80225, USA
‡‡ Division of Earth and Ecosystem Sciences, Desert Research Institute, 2215 Raggio Parkway, Reno, Nevada 89512, USA
§§ Space Science Division, NASA Ames Research Center, Moffet Field, California 94035, USA

The average air temperature at the Earth's surface has increased by 0.06 °C per decade during the 20th century[1], and by 0.19 °C per decade from 1979 to 1998[2]. Climate models generally predict amplified warming in polar regions[3,4], as observed in Antarctica's peninsula region over the second half of the 20th century[5–9]. Although previous reports suggest slight recent continental warming[9,10], our spatial analysis of Antarctic meteorological data demonstrates a net cooling on the Antarctic continent between 1966 and 2000, particularly during summer and autumn. The McMurdo Dry Valleys have cooled by 0.7 °C per decade between 1986 and 2000, with similar pronounced seasonal trends. Summer cooling is particularly important to Antarctic terrestrial ecosystems that are poised at the interface of ice and water. Here we present data from the dry valleys representing evidence of rapid terrestrial ecosystem response to climate cooling in Antarctica, including decreased primary productivity of lakes (6–9% per year) and declining numbers of soil invertebrates (more than 10% per year). Continental Antarctic cooling, especially the seasonality of cooling, poses challenges to models of climate and ecosystem change.

Terrestrial ecosystem research in the Antarctic is restricted to a few ice-free areas of the coast, including the McMurdo Dry Valleys (77–78° S, 160–164° E). The dry valleys region is the largest ice-free area on the Antarctic continent. It is a cold desert, comprising a mosaic of perennially ice-covered lakes, ephemeral streams, arid soils, exposed bedrock, and alpine glaciers. Published historical weather observations in the dry valleys are limited[11–14]. Biological activity is microbially dominated and diversity is low. The largest animals are soil invertebrates, of which soil nematodes are the most widely distributed[15].

Our 14-year, continuous automatic weather station record from the shore of Lake Hoare reveals that seasonally averaged surface air temperature has decreased by 0.7 °C per decade ($P = 0.21$) from 1986 to 1999 (Fig. 1a). The temperature decrease is most pronounced in