Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: http://www.tandfonline.com/loi/thsj20

Has the magnitude of floods across the USA changed with global CO_2 levels?

R. M. Hirsch & K. R. Ryberg

To cite this article: R. M. Hirsch & K. R. Ryberg (2012) Has the magnitude of floods across the USA changed with global CO_2 levels?, Hydrological Sciences Journal, 57:1, 1-9, DOI: 10.1080/02626667.2011.621895

To link to this article: http://dx.doi.org/10.1080/02626667.2011.621895

Copyright This work was authored as part of the Contributor's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

View supplementary material

Published online: 24 Oct 2011.

Submit your article to this journal

Article views: 7046

View related articles

Citing articles: 48 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=thsj20

Download by: [University of North Texas] **Date:** 24 January 2017, At: 17:20

Has the magnitude of floods across the USA changed with global CO_2 levels?

R. M. Hirsch[1] and K. R. Ryberg[2]

[1] *US Geological Survey, 432 National Center, Reston, Virginia 20192, USA*
rhirsch@usgs.gov

[2] *US Geological Survey, 821 E. Interstate Ave., Bismarck, North Dakota 58503, USA*

Received 5 April 2011; accepted 28 July 2011; open for discussion until 1 July 2012

Citation Hirsch, R.M. and Ryberg, K.R., 2012. Has the magnitude of floods across the USA changed with global CO_2 levels? *Hydrolological Sciences Journal*, 57 (1), 1–9.

Abstract Statistical relationships between annual floods at 200 long-term (85–127 years of record) streamgauges in the coterminous United States and the global mean carbon dioxide concentration (GMCO2) record are explored. The streamgauge locations are limited to those with little or no regulation or urban development. The coterminous US is divided into four large regions and stationary bootstrapping is used to evaluate if the patterns of these statistical associations are significantly different from what would be expected under the null hypothesis that flood magnitudes are independent of GMCO2. In none of the four regions defined in this study is there strong statistical evidence for flood magnitudes increasing with increasing GMCO2. One region, the southwest, showed a statistically significant negative relationship between GMCO2 and flood magnitudes. The statistical methods applied compensate both for the inter-site correlation of flood magnitudes and the shorter-term (up to a few decades) serial correlation of floods.

Key words floods; trends; climate change; statistics; carbon dioxide

L'intensité des crues aux Etats-Unis a-t-elle changé avec les niveaux mondiaux de CO_2?

Résumé Nous avons étudié les relations statistiques entre les crues annuelles mesurées en 200 stations de jaugeage à long terme (85 à 127 ans d'enregistrement) de la partie continentale des Etats-Unis et la concentration moyenne globale du dioxyde de carbone (GMCO2). Nous nous sommes limités aux stations de jaugeage peu ou pas influencées par la réglementation ou la croissance urbaine. La partie continentale des Etats-Unis a été divisée en quatre grandes régions et on a utilisé la méthode du bootstrap stationnaire pour tester si les modéles de ces associations statistiques sont sensiblement différents de ce qui serait attendu sous l'hypothése nulle que l'intensité des crues est indépendante de GMCO2. Dans aucune des quatre régions définies dans cette étude n'existent de fortes preuves statistiques que l'intensité des crues augmente avec la croissance de GMCO2. Une région, au Sud-Ouest, présente une relation décroissante statistiquement significative entre GMCO2 et l'intensité des crues. Les méthodes statistiques appliquées compensent la corrélation de l'intensité des crues entre stations et l'autocorrélation à court terme (jusqu'à quelques dizaines d'années) des séries de crues.

Mots clefs crues; tendances; changement climatique; statistiques; dioxyde de carbone

INTRODUCTION

One of the anticipated hydrological impacts of increases in greenhouse gas concentrations in the atmosphere is an increase in the magnitude of floods (Trenberth 1999, IPCC 2007, Gutowski *et al*. 2008). Greenhouse gases change the energy balance of the atmosphere and lead to atmospheric warming, which increases the water-holding capacity of the atmosphere, which in turn, potentially changes the amounts of precipitable water. The resultant warming also changes the form of precipitation (more rain and less snow), changes the timing of snowmelt (Dettinger and Cayan 1995, Milly *et al.* 2002, Hodgkins and Dudley 2006), is projected to change storm tracks, and may change the frequency and intensity of large-scale ocean/climate conditions such as El Niño/Southern Oscillation; therefore, the idea that river flood characteristics have changed, or will change, as a result of increased greenhouse gas

ISSN 0262-6667 print/ISSN 2150-3435 online
This work was authored as part of the Contributor's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.
http://dx.doi.org/10.1080/02626667.2011.621895
http://www.tandfonline.com

concentrations is reasonable. Effective flood mitigation strategies depend on accurate assessments of flood risk. Land and water resource managers are asking questions about how to estimate future flood risks with increasing greenhouse gas concentrations. One approach to estimating these changes is through the use of atmospheric models coupled with hydrological models, for example Milly *et al*. (2002), but such approaches have serious limitations associated with the temporal and spatial scales of general circulation models (GCMs) and their physical fidelity to important flood-producing processes. An alternative approach is the statistical analysis of a collection of long-term, high-quality flood records to examine relationships between flood magnitudes and greenhouse gas concentrations. Using long flood records has the combined advantage of observing flooding behaviour over a substantial range of greenhouse gas concentrations and helps to limit the potential for confusing patterns of long-term persistence with secular trends. The importance of long-term persistence and the difficulties it presents to hydrologic data analysis is discussed by Koutsoyiannis (2003) and by Cohn and Lins (2005). Kundzewicz and Robson (2004) discuss several methodological issues regarding the detection of secular trends in hydrological records. They highlight the importance of running the analysis with high quality data sets with at least 50 years of record.

This study attempts to quantify the changes taking place in flood behaviour in the coterminous United States (USA) as a function of greenhouse gas concentrations. The influence of enhanced greenhouse forcing on floods is not thought to be a direct influence. Rather it likely to be a result of multiple factors that are themselves influenced by greenhouse forcing. For example, snow pack dynamics have already changed in many areas, the result being that flood potential may be increased due to storms producing more rain and less snow than they would have in the absence of this forcing. However, the amount of water in storage as snow pack may be decreased, resulting in potentially smaller flood volumes from rain-on-snow events. With enhanced greenhouse forcing, warmer air temperatures may result in depleted soil moisture and/or lesser amounts of frozen ground at the time when the most intense rainfall events occur. Thus, even increased volumes of precipitation may result in smaller flood peaks given the greater moisture-storage capacity of the soil. Across an area as large and diverse (in terms of climate and topography) as the coterminous USA it is reasonable to expect that the interplay of these and many other factors will be quite variable. Given the recognized shortcomings of climate models in terms of simulating the many variables that are significant to hydrology it is useful to supplement the more process-based studies with this, much simpler, empirical approach to the question and seek insight on the net effect of the possible linkages between greenhouse forcing and floods at national and regional scales. For simplicity, we have represented increased greenhouse forcing here by the annual time series of GMCO2. It is assumed that the relationship between GMCO2 and flood magnitudes may vary across the USA and may depend on the size of the watershed. The approach is predicated on the idea that the increase in GMCO2 over the past century is an unplanned "experiment" and that every watershed that has been monitored over that time can be viewed as an "experimental subject". The streamflow data set consists of annual flood series from 200 streamgauges operated by the US Geological Survey (USGS) in the coterminous USA, of at least 85 years length through water year 2008, from basins with little or no reservoir storage or urban development (less than 150 persons per square kilometre in 2000). Details of selection criteria, the streamgauge sites, and the peak streamflow data, are provided in the supplementary material (Appendix and Tables S1 and S2*).

The working hypothesis of this study is that change in GMCO2 is a dominant factor in driving changes in flood behaviour for these watersheds. However, human influences associated with large numbers of very small impoundments and changes in land use also could play a role in changing flood magnitude. Unfortunately, at time scales on the order of a century, it is difficult to make a quantitative assessment of the changes in these factors over time. This is a weakness in all such longitudinal studies of streamflow change, but we believe that the site selection process employed limits the risk of confusing land-based and atmospheric causes of hydrologic change. The selection of suitable streamgauges for studies such as this one is always a compromise. If the criteria had been significantly more restrictive in terms of the amount of reservoir storage, urban development, or other human activities in the watersheds, the selected sites would be limited almost entirely to very small watersheds, typically in remote and often mountainous areas, atypical of the types of watersheds that produce floods that result in significant amounts of economic loss. In addition, the

*Note: Tables S1 and S2 are available online only.

record lengths available would be much shorter than the ones used here. Our criteria were predicated on obtaining maximum record length and a wide range of drainage areas.

ASSUMPTIONS ABOUT RELATIONSHIPS BETWEEN FLOOD MAGNITUDE AND GMCO2

For each of the 200 streamgauges in the study a regression model was fit to the annual peak streamflow. The form of the model was:

$$\ln(Q_i) = \beta_0 + \beta_1 \times C_i + e_i \quad (1)$$

where Q_i is the annual peak streamflow for year i, in $m^3\ s^{-1}$; C_i is the GMCO2 for year i in parts per million (ppm; see auxiliary material for sources for all variables used in the analysis), and e_i is the unexplained variation for year i. Evaluation of the residuals from these regressions suggests an approximately symmetrical distribution in most cases. The form of the model was evaluated through consideration of model residuals, and through tests for constant variance and serial correlation (see Appendix).

There is a growing body of research on the question of trends in flooding in Europe and across the globe, for example: Mudelsee *et al.* 2003, Kundzewicz 2004, Lindström and Bergström 2004, Kundzewicz *et al.* 2005, Hannaford and Marsh 2008. A number of previous studies have explored temporal changes in annual peak discharges or annual maximum daily discharges in the USA (Lins and Slack 1999, McCabe and Wolock 2002, Villarini *et al.* 2009, Villarini and Smith 2010). This study contrasts with these other US studies in two respects. First, those studies conceptualized the change as being a temporal trend, so they used analyses such as the Mann-Kendall test to evaluate changes in the distribution of flood magnitudes over time. This study explicitly relates changes in flood magnitudes to CO_2. If we consider the hypothesis that there is such a relationship then we would expect a low rate of change in the earlier parts of the record and much more rapid rates of change in more recent years. The data analysis used here is designed to be more sensitive to such a pattern of change *versus* any arbitrary pattern of monotonic change. The other difference is that the studies mentioned above focus on the number and location of streamgauges where the trends are statistically significant. In contrast, the goal of this study is to search for patterns of relationship, drawing conclusions from the ensemble of streamgauges rather than making decisions to reject or not reject a null hypothesis that $\beta_1 = 0$ at any single site. This is similar to the approach used by Douglas *et al.* (2000), which explicitly considered spatial correlation, but that study used substantially shorter hydrologic records and did not consider greenhouse gas concentrations as an explanatory variable. This study is similar in philosophy to longitudinal studies in epidemiology in which the goal is to explore associations between possible disease-causing mechanisms and health outcomes of a population of individuals. Such studies are designed to provide empirical evidence about risk factors and to guide process-based research, but draw no conclusions about the mortality or morbidity of individuals in the study. Our emphasis is not on the "significance" of associations for any specific watershed, but rather to identify patterns of association across the USA.

RESULTS

Figure 1 shows estimates of β_1 at every streamgauge and the four regions used: Northeast (NE), Southeast (SE), Northwest (NW) and Southwest (SW). The

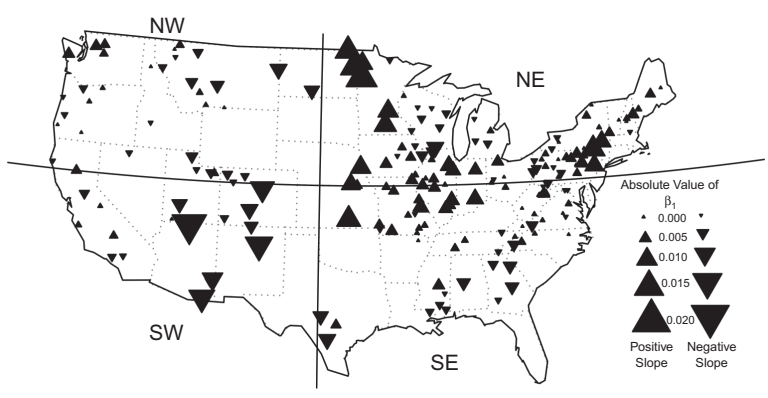

Fig. 1 Regression results for all 200 streamgauges. Triangles indicate the magnitude and sign of the estimate of β_1. The boundaries of the four regions are shown.

boundaries of the four regions are simple. The boundary between east and west is the 100th meridian, generally dividing the more-arid west and more-humid east. The boundary between north and south is the 40th parallel, generally separating colder and warmer climates. Table S2 of the auxiliary material lists the results for each streamgauge. Of the 200 streamgauges there are 48 for which the null hypothesis, $\beta_1 = 0$, would be rejected at $\alpha = 0.05$. Of these, the sign of the estimated value of β_1 was positive in 30 cases and negative in 18 cases. Under the null hypothesis the expected numbers would be five positive and five negative. The fact that the actual numbers are so much larger than the expected numbers could be a consequence of spatial and temporal correlations in the data and/or the presence of regionally specific causal relationships to GMCO2.

The overall pattern of the β_1 estimates shows roughly equal numbers of positive and negative values. The largest positive values are focused in a north–south band running from Minnesota and eastern North Dakota through eastern Kansas, with a strong focus in and near the watershed of the Red River of the North (located at the boundary between Minnesota and North Dakota). There is a smaller area of moderately high positive values in an area of New Jersey, eastern Pennsylvania, and eastern New York. The largest set of negative values is in the Rocky Mountains and arid southwest. There are some notable similarities and notable differences between the spatial pattern of change shown in Fig. 1, and the pattern of projected change in annual runoff from 1980–1999 to 2090–2099, as illustrated in maps published by the Intergovernmental Panel on Climate Change (IPCC), such as Figure 3.4 of Bates *et al.* (2008), recognizing that the former is focused on flood magnitude while the latter is on overall water availability. The similarity is strong in terms of the trend towards drying conditions in the Rocky Mountains and arid southwest. In addition, the relatively neutral results in the Southeast and Northwest quadrants of the USA also show a general agreement between this study and the IPCC projections. However, the highly focused area of very high β_1 values near and to the south of the Red River of the North shows up as an area of virtually no change in the runoff projection map. This study's analysis of floods in this particular area, as well as other analyses related to mean runoff conditions in the same area, are very much at odds with the climate change impact assessment reported by the IPCC.

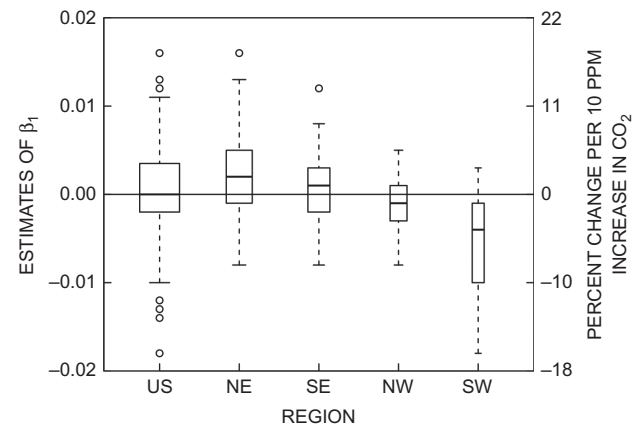

Fig. 2 Boxplots of estimates of β_1 (left-hand scale). On right-hand scale the estimates are re-expressed in units of percent change per 10 ppm increase in GMCO2. Box width is proportional to the square root of the sample size.

Figure 2 shows the β_1 estimates as a set of box plots for the USA as a whole and for each of the four "quadrants" of the USA. The box plots show distinct regional differences in the values of β_1 with the NE showing the strongest tendency towards positive values and the SW showing the strongest tendency toward negative values. In addition, for each region, we computed Kendall's tau (Kendall 1938) for the relationship between β_1 and drainage area, and between the absolute value of β_1 and the drainage area. In no case did we find a significant relationship.

STATIONARY BOOTSTRAP ANALYSIS

In order to evaluate if these regional patterns are substantially different from what might be expected by chance alone if there were no relationship between flood magnitudes and GMCO2, we performed a series of bootstrap analyses. Kundzewicz and Robson (2004) make particular note of the benefits of block bootstrap approaches to compensate for spatial and temporal correlation in the analysis of long-term hydrologic records. Bootstrap techniques have been applied in other regional studies of streamflow trends, for example Douglas *et al.* (2000) and Wilson *et al.* (2010). The specific method used was a variation of stationary (block) bootstrap (Politis and Romano 1994), in which the relationship between the annual floods and GMCO2 was randomized. The procedure preserves the spatial correlation of the annual flood data. It also approximately preserves the shorter lag component (up to about two decades) of the serial correlation structure of the flood data. The procedure

is implemented as follows: the GMCO2 record was kept fixed, that is, the annual values remained in order from earliest date (1882) to latest date (2008). Peak streamflow for a region was represented by stationary bootstrap replicates. For each replicate, the stationary bootstrap procedure selects a random sample, with replacement, of blocks of years across all of the streamgauges in the region. These blocks are of random length with block length geometrically distributed with a mean length of 20 years (mean lengths of 10 and 30 were also used but the effect of varying the mean block length on the results was very small). For each bootstrap replicate, the resampled time series of $\ln(Q)$ at each streamgauge are regressed against the GMCO2 values for the entire period of record. For iteration j, m_j is computed as the median estimate of β_1 across all of the streamgauges in the region. The test statistic is M, the median of the estimated values of β_1 for a region (with the observations in their original order). The null hypothesis is that the expected value of M is zero in the region (i.e. no relationship between flood magnitudes and GMCO2). Using 10 000 iterations of the bootstrap, we computed the attained, two-sided significance level for each region. The p-value is the fraction of the iterations in which $|m_j| \geq |M|$. Table 1lists the results of the procedure by region.

DISCUSSION

The only strong statistical result is the negative relationship between GMCO2 and flood magnitudes in the SW region. The results are suggestive of a positive relationship in the NE region. The other two regions were not suggestive of a relationship in one direction or the other. The NE results are strongly influenced by the cluster of positive values in the area surrounding the Red River of the North and all of the records in this cluster are highly correlated. It has long been recognized that this region of the USA (and adjoining areas of Canada) has experienced a series of highly persistent quasi-periodic oscillations in hydrological conditions over time spans of a century or more (Rannie 1998, Knox 2000, Vecchia 2008). It also is well known that this region has experienced an abrupt and large increase in precipitation amounts, particularly since about 1980. The temporal pattern of floods that is observed in the data for this region is consistent with one or more of three lines of argument: (1) greenhouse forcing as a cause of the change, (2) changes in land management (particularly land drainage and levee construction) as a cause of the change, or (3) persistent atmospheric conditions that have time scales of a century or more and which could reverse themselves at some unknown future time/area, as a cause of the change. Knowledge based on historical observations and/or palaeo-hydrological information subjectively points to the latter of these three possibilities as being more likely, although a combination of these influences is also supportable. The southern Rocky Mountains and southwestern desert areas present a similar puzzle. Palaeo-hydrological records and historical information suggest oscillations that reflect much wetter conditions in the late 19th century and very early 20th century than those experienced in recent decades (National Research Council 2007). Here again, the influence of greenhouse forcing (changing storm tracks and decreasing the size of winter snowpacks), or land-management, or the existence of some highly persistent quasi-periodic phenomena are all plausible explanations of the patterns observed. In particular, in the southwestern region, the warming that has taken place is likely to be causing decreased winter snowpacks in some of the watersheds and this decrease may be contributing to a decreased potential for flooding.

There are caveats about these analyses and their usefulness in relating future changes in flood magnitude to future increases in GMCO2, including the recognition that the response could be highly non-linear or may have important threshold levels. Another possibility is that flood behaviour is influenced by other global atmospheric variables in addition to GMCO2. To test this idea we

Table 1 The number of streamgauges, median estimate of β_1, percentage of streamgauges with positive β_1, and attained two-sided significance levels for the stationary bootstrap.

Region	NE	SE	NW	SW
Number of streamgauges	85	61	36	18
Median estimate of β_1	0.0016	0.0009	−0.0006	−0.0041
Percent of streamgauges with $\beta_1 > 0$	62%	61%	36%	22%
Attained two-sided significance	0.14	0.40	0.57	0.0019

regressed ln(Q) values against GMCO2 and indices of El Niño/Southern Oscillation, Pacific Decadal Oscillation, and Atlantic Multidecadal Oscillation (see Table S2). Our conclusion from this analysis is that, although knowledge of these atmospheric variables may be useful in estimating flood risk in some particular years in parts of the USA, they provide little additional insight on the question of the relationship of flood magnitudes to GMCO2. We found that the inclusion of these other variables leads to a modest decrease in evidence for a positive association of GMCO2 to flood magnitude. How increased greenhouse forcing will change the frequency and intensity of these phenomena is an important question to be explored because of the influence that they have on flood-producing atmospheric conditions. It is possible that changes in these climate modes may prove to be as significant to flooding as the direct impact of factors such as changes in air temperature or the water-holding capacity of the atmosphere.

CONCLUSIONS

The question of how floods are related to GMCO2 concentrations is an important one for adaptation to climate change. This study suggests positive associations in colder areas with moderate to high precipitation amounts and negative associations in some of the dryer parts of the USA. The results of this study do not mean that no strong relationship between flooding and GMCO2 will emerge in other areas in the future. It may be that the greenhouse forcing is not yet sufficiently large to produce changes in flood behaviour that rise above the "noise" in the flood-producing processes. What these results do indicate is that except for the decreased flood magnitudes observed in the SW there is no strong empirical evidence in any of the other 3 regions for increases or decreases in flood magnitudes in the face of the 32% increase in GMCO2 that has taken place over the study period. However, it is crucial that analysis of the empirical data be conducted repeatedly as greenhouse forcing changes over time because such empirical analyses are a valuable check on the results of theoretical or model-driven studies of this issue.

We believe that a wide range of empirical approaches to this issue need to be undertaken and some may reveal patterns that this analysis was unable to discern. In particular, much more detailed analyses of these questions focused on particular regions may be very fruitful. Changes in the intensity or frequency of certain storm tracks (such as "atmospheric rivers" discussed by Dettinger 2011), and changes in the relative importance of the role of snowpack and rain on snow events are crucial to the insights needed to project regional changes in flood hazards. The addition of new data over time should increase the power and usefulness of data-driven approaches. Milly *et al.* (2008) state "In a nonstationary world, continuity of observations is crucial" and we add that continuing analysis of those observations is also crucial (see National Research Council 2009, p.13).

Acknowledgements This work was supported by the US Geological Survey, National Streamflow Information Program and Hydrologic Research and Development Program. We are particularly grateful to the USGS Hydrologists and Hydrologic Technicians who worked tirelessly to collect, quality assure, and store these flood data. We also thank Jery Stedinger and Chris Milly for their very helpful review comments.

REFERENCES

Bates, B.C., Kundzewicz, Z.W., Wu, S. and Palutikof, J.P., 2008. Climate change and water. Geneva: IPCC Secretariat, Technical paper of the Intergovernmental Panel on Climate Change. Available at: http://www.ipcc.ch/publications_and_data/publications_and_data_technical_papers.shtml.

Breusch, T.S. and Pagan, A.R., 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47, 1287–1294.

Cohn, T.A. and Lins, H.F., 2005. Nature's style: naturally trendy. *Geophysical Research Letters*, 32, L23402, doi:10.1029/2005GL024476.

Dettinger, M.D., 2011. Climate change, atmospheric rivers, and floods in California—a multimodel analysis of storm frequency and magnitude changes. *Journal of the American Water Resources Association*, 47, 514–523, doi:10.1111/j.1752-1688.2011.00546.x.

Dettinger, M.D. and Cayan, D.R., 1995. Large-scale atmospheric forcing of recent trends toward early snowmelt runoff in California. *Journal of Climate*, 8, 606–623.

Douglas, E.M., Vogel, R.M. and Kroll, C.N., 2000. Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology*, 240, 90–105.

Durbin, J. and Watson, G.S., 1950. Testing for serial correlation in least squares regression I. *Biometrika*, 37, 409–428.

Gutowski, W.J., *et al*., 2008, Causes of observed changes in extremes and projections of future changes. *In:* T.R. Karl, *et al*., eds. *Weather and climate extremes in a changing climate—regions of focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*, Washington, DC: US Climate Change Science Program, Synthesis and Assessment Product 3.3, 81–116.

Hannaford, J. and Marsh, T.J., 2008. High-flow and flood trends in a network of undisturbed catchments in the UK. *International Journal of Climatology*, 28 (10), 1325–1338, doi:10.1002/joc.1643.

Hodgkins, G.A. and Dudley, R.W., 2006. Changes in the timing of winter–spring streamflows in eastern North America,

1913–2002. *Geophysical Research Letters*, 33, L06402, doi:10.1029/2005GL025593.

IPCC (Intergovernmental Panel on Climate Change), 2007. *Climate change 2007: The physical science basis*, S. Solomon, *et al*., eds. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press.

Kendall, M.G., 1938. A new measure of rank correlation. *Biometrika*, 30, 81–93.

Knox, J.C., 2000. Sensitivity of modern and Holocene floods to climate change. *Quaternary Science Review*, 19, 439–457.

Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon and hydrological statistics. *Hydrological Sciences Journal*, 48 (1), 3–24.

Kundzewicz, Z.W., 2004. Searching for change in hydrologic data. *Hydrological Sciences Journal*, 49 (1), 3–6.

Kundzewicz, Z.W., *et al*., 2005. Trend detection in river flow series: 1. Annual maximum flow. *Hydrological Sciences Journal*, 50 (5), 797–810.

Kundzewicz, Z.W. and Robson, A.J., 2004. Change detection in river flow records—review of methodology. *Hydrological Sciences Journal*, 49 (1), 7–19.

Lindström, G. and Bergström, S., 2004. Runoff trends in Sweden 1807–2002. *Hydrological Sciences Journal*, 49 (1), 69–83.

Lins, H.F., and Slack, J.R., 1999. Streamflow trends in the United States. G*eophysical Research Letters*, 26 (2), 227–230, doi:10.1029/1998GL900291.

McCabe, G. and Wolock, D.M., 2002. A step increase in streamflow in the conterminous United States. *Geophysical Research Letters*, 29 (24), 2185, doi:10.1029/2002GL015999.

Milly, P.C.D., *et al*., 2008. Stationarity is dead: whither water management?, *Science*, 319, 573–574.

Milly, P.C.D., Wetherald, R.T., Dunne, K.A. and Delworth, T.L., 2002. Increasing risk of great floods in a changing climate. *Nature*, 415, 514–517.

Mudelsee, M., Börngen, M., Tetzlaff, G and Grünewald, U., 2003. No upward trends in the occurrence of extreme floods in central Europe. *Nature*, 421, 166–169.

National Research Council, 2007. *Colorado River Basin water management: Evaluating and adjusting to hydroclimatic variability*. Washington, DC: National Academies Press.

National Research Council, 2009. *Informing decisions in a changing climate*. Washington, DC: National Academies Press.

Politis, D.N. and Romano, J.P., 1994. The stationary bootstrap. *Journal of the American Statistical Association*, 89 (428), 1303–1313.

R Development Core Team, 2010. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0, http://www.R-project.org.

Rannie, W.F., 1998. A survey of hydroclimate, flooding, and runoff in the Red River Basin prior to 1870. Ottawa, ON: *Geological Survey of Canada Open-File Report 3705*.

Seaber, P.R., Kapinos, F.P. and Knapp, G.L., 1987. Hydrologic unit maps. Reston, VA: *US Geological Survey Water-Supply Paper 2294*, http://pubs.usgs.gov/wsp/wsp2294/.

Trenberth, K.E., 1999. Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climate Change*, 42, 327–339, doi:10.1023/A:1005488920935.

Vecchia, A.V., 2008. Climate simulation and flood risk analysis for 2008–40 for Devils Lake, North Dakota. Reston, VA: *US Geological Survey Scientific Investigations Report 2008-5011*.

Villarini, G., Serinaldi, F., Smith, J.A. and Krajewski, W.F., 2009. On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, 45, W08417, doi:10.1029/2008WR007645.

Villarini, G. and Smith, J.A., 2010. Flood peak distributions for the Eastern United States. *Water Resources Research*, 46, W06504, doi:10.1029/2009WR008395.

Wilson, D., Hisdal, H. and Lawrence, D., 2010. Has streamflow changed in the Nordic countries? Recent trends and comparisons to hydrological projections. *Journal of Hydrology*, 394, 334–346, doi:10.1016/j.hydrol.2010.09.010.

APPENDIX

Site and peak selection

Peak-streamflow data for US Geological Survey (USGS) streamgauges in the National Water Information System (NWIS) online database, http://nwis.waterdata.usgs.gov/nwis/peak, accessed 27 January 2010, were examined for the 48 coterminous States and the District of Columbia. Streamgauge records were considered for inclusion in the study based on four criteria:

(1) They must have had at least 85 years of annual peak streamflow data through the end of water year 2008 and there must have had at least five annual peak discharges in the years 1999 through 2008 (because we wanted the study to consider only those records that include recent data).

(2) The peak streamflows must not be significantly affected by regulation. This is typically indicated by USGS qualification codes 5 or 6 in the database, although we recognize that there is a degree of subjectivity to these criteria and some streamgauges were removed based on additional information, including consultation with US Geological Survey employees familiar with the watersheds associated with specific streamgauges and examination of the ratio of upstream storage to mean annual discharge. One flood (the 1964 peak for the Marias River near Shelby, Montana) was deleted from the record because it was a result of a dam failure, although the dam is not considered to have a substantial impact on floods in other years.

(3) The amount of urbanization of the watershed upstream of the streamgauge must be low. Only those watersheds with average population densities of fewer than 150 persons per square kilometre (in the 2000 Census; data obtained from Daren Carlisle, US Geological Survey, March 2009) were considered.

(4) Streamgauges with tidal influences were also eliminated based on consultation with U.S. Geological Survey employees familiar with the site.

Some of the selected long-term flood records contain "non-systematic peaks". These are peaks that took place prior to the time when continuous streamgauging began at these sites. They are typically estimated by indirect methods and they exist in the database typically because they are very large floods. Including them in this study would have introduced a bias towards an excessive number of large floods early in the record. Thus, non-systematic peaks were removed. The criterion used to do this was the following: peaks were included if they come from groups of three or more consecutive years of peak streamflow values. Individual peaks or two consecutive peaks followed by a gap in data collection were deleted because of concern that they were collected in response to a large flood event.

Some streamgaging sites have been discontinued in the past and then reactivated in response to renewed interest in the sites. Because of this, there are some missing years of data at some sites. From the first peak to the last peak, sites range from 0 to 25 years of missing peaks (mean 3.2 missing peaks, median 0 missing peaks). On average, the number of peaks used in the analysis was 96.7% of the total possible years between the beginning of the record and 2008 and in all cases was at least 85 peak streamflow values.

After removing peaks affected by dam failure and non-systematic peaks, any site with fewer than 85 peaks remaining was removed from the study. In addition, streamgauges were removed from the data set if they were among a set of streamgauges in the data set that were closely spaced and/or highly correlated with each other. If the same stream had multiple streamgauges within an 8-digit hydrologic unit (Seaber *et al*. 1987), the streamgauge with the shorter period of record was removed. In the case where the record lengths were equal, the site with the smallest drainage area was removed. For the remaining sites, Kendall's tau (Kendall 1938) was computed for all pairs of streamgauges. For streamgauge pairs with Kendall's tau greater than 0.7, the streamgauge with the shorter period of record was removed. Thus, the data set was designed to eliminate cases where records were highly correlated with each other, but the data set still contains substantial cross-correlation. The subsequent maps and analysis do reveal strong regional patterns.

The data set used for this study is provided in Table S1. The definitive source for the data is the publicly available USGS National Water Information System online database at http://nwis.waterdata.usgs.gov/nwis/peak, which also contains the nonsystematic peaks removed for this study and is updated as new annual peaks are determined.

For the 200 streamgauge records used in the study the median record length is 93 years and the longest is 127 years. The watersheds range in size from 41 to 179 000 km^2 (median size is 1940 km^2).

Regression diagnostics and residuals analysis

Arguably, the regression could have used the log of the global mean carbon dioxide (GMCO2) concentration as its explanatory variable (radiative forcing being proportional to ln(GMCO2)), but over the range of GMCO2 concentrations in the data set, the regression results using either GMCO2 or ln(GMCO2) would be virtually identical in terms of attained significance levels or estimated change in flood magnitude per unit change in GMCO2 (for this time period the linear correlation coefficient between annual GMCO2 and the log of annual GMCO2 is 0.999). The only transformation that the data suggests is needed is the log transformation of the dependent variable. In addition to visual examination of residuals plots, two formal statistical tests were performed on all 200 regressions to determine the overall appropriateness of the regression model. The first was the Breusch-Pagan test (Breusch and Pagan 1979) to determine if the variance of the residuals is constant. The results show that in most cases (182 out of 200) one would not reject the null hypothesis of constant variance ($\alpha = 0.05$). Of the remaining cases, eight show decreasing variance and 10 show increasing variance. Under the null hypothesis we would expect five decreasing and five increasing. As such, we considered the constant variance model to be appropriate for the data. Note that under the assumption of constant variance, the percentage change in flood magnitudes for a given change in GMCO2 concentration would be the same across all quantiles of the annual flood peak distribution (e.g. the 2-year flood, 50-year flood, or 100-year flood).

The second test was the Durbin-Watson test (Durbin and Watson 1950) to determine if there is serial correlation in the residuals. Significant ($\alpha = 0.05$) serial correlation of residuals was found at 20 sites. This result is one reason to discount

conclusions about the significance of individual β_1 values and to use the stationary bootstrap procedure, described in the body of the paper, to consider the significance of associations between flood peaks and GMCO2.

Global mean carbon dioxide concentration

Data were obtained 25 January 2010, from the National Aeronautics and Space Administration Goddard Institute for Space Studies, http://data.giss.nasa.gov/modelforce/ghgases/Fig1A.ext.txt. Monthly values were averaged over each water year (the 12-month period from 1 October for any given year to 30 September of the following year, designated by the calendar year in which it ends) to produce mean annual values that correspond with the water years for which there are peak streamflow values.

El Niño/Southern Oscillation

The NINO 3.4 index, sea-surface temperature anomalies were obtained 25 January 2010, from the Royal Netherlands Meteorological Institute, Climate Explorer, http://climexp.knmi.nl/data/inino5.dat.

Atlantic Multidecadal Oscillation

The Atlantic Multidecadal, unsmoothed, long data set was obtained from the NOAA Earth System Research Laboratory (ESRL) 25 January 25 2010, at http://www.esrl.noaa.gov/psd/data/correlation/amon.us.long.data.

Pacific Decadal Oscillation

The Pacific Decadal Oscillation data set was obtained 25 January 2010, from the Joint Institute for the Study of the Atmosphere and Ocean at http://jisao.washington.edu/pdo/PDO.latest.

Transformation of slope to units of percent change in flood magnitude per 10 parts per million increase in CO_2

The expression $100(e^{10\beta_1} - 1)$ is a useful way to transform the β_1 coefficient. This expression defines a slope in units of percent change in flood magnitude per 10 parts per million (ppm) increase in GMCO2. Additionally, for low values of β_1, multiplying by 1000 results in an approximation to this transformation. As a point of reference, GMCO2 concentrations are currently increasing at a rate of about 10 ppm every 5 years.

ANOVA test of simple atmospheric carbon dioxide linear regression model against a more complex linear regression model for peak streamflow

Analysis of variance (ANOVA) was used to test a simple global mean carbon dioxide (GMCO2) linear regression model for peak streamflow against a more complex linear regression model for peak streamflow in which GMCO2, Pacific Decadal Oscillation (PDO), Atlantic Multidecadal Oscillation (AMO), and El Niño/Southern Oscillation (ENSO) were the explanatory variables. The comparison of these two models is valid only if they are fitted to the same data set. The data set for GMCO2 has values for the earliest peak streamflow values used (beginning in 1882), however, the PDO data set began in 1901. Therefore, the results described here are for peaks in the period 1901–2008 only.

When using GMCO2 as an explanatory variable for peak streamflow at 200 sites from 1901–2008, GMCO2 was a statistically significant explanatory variable at 48 sites (significance level $\alpha = 0.05$).

This simple model was tested against the more complex model using an F-test within ANOVA (R Development Core Team 2010). At the 0.05 significance level, the more complex model was significant for 34 sites (results shown in Table S2, which is visible online).