doi:10.1038/nature09762

Supplementary Methods

a. Estimating fraction of attributable England and Wales flood risk in autumn 2000 using a simple thermodynamic framework

A popular simple thermodynamic argument assumes precipitation extremes are constrained to change with the water vapour capacity of the atmosphere that can be determined, under conditions of constant relative humidity, using change in mean surface temperature alone according to the Clausius-Clapeyron relation[38]. This argument is typically invoked in the aftermath of floods as an explanation for possible increases in such severe wet events under an anthropogenically warming climate.

While this is an oversimplified treatment[8,9] not fully accounting for the complex hydrometeorology typically associated with UK flooding[4,10,15], it may nevertheless provide a physically plausible first guess of increases in mid-latitude precipitation extremes under warming[39,40,41]. Indeed, a recently updated analysis of observed atmospheric column water vapour for past decades finds increasing trends over the UK and western Europe, and a significant autumnal increase more generally over Europe and the Northern Hemisphere[42]; and this appears in agreement with a similar analysis finding increases in observed atmospheric humidity under warming for these regions that are within expected moistening rates for near-constant relative humidity[43].This latter analysis in particular appears broadly consistent with observations of Clausius-Clapeyron scale increases in surface specific humidity (the principle source for the free-troposphere) under warming over past decades, again with near-constant relative humidity – including for an European region incorporating the UK[44,45]. Since these surface specific humidity increases have been attributed to mainly anthropogenic drivers[46], this lends support to a thermodynamic mechanism for increasing UK precipitation, and hence flooding, under anthropogenic warming.

Here we use this thermodynamic argument to deduce the reduction in observed[3] England and Wales total daily precipitation extremes for an autumn 2000 climate, had estimated[21,22] twentieth-century surface warming attributable to anthropogenic greenhouse gas emissions not occurred. Then regarding this reduction in precipitation extremes as a direct measure of reduction in flooding, we calculate the fraction of attributable risk (FAR)[13,16] of these extremes, and compare it to the FAR of autumn 2000 flooding explicitly modelled in terms of severe daily river runoff using our more rigorous multi-step probabilistic event attribution (PEA) framework of the main text.

One limitation of this simple thermodynamic framework is that only the single observed state of precipitation in autumn 2000 is available for sampling extremes in that climate

(cf. the PEA framework where many possible states are generated). To try to overcome this we assume observed precipitation from autumns 1990-2000 is representative of precipitation under the conditions that obtained in autumn 2000, and hence representative of the climate for that time. Alternative nearby autumns could have been chosen, but we specifically choose these years because the attributable twentieth-century surface warming, according to which the precipitation is to be reduced, is for the 1990s relative to the 1900s (see Methods). The blue curve in Supplementary Fig. 1 shows the distribution of observed precipitation extremes from autumns 1990-2000.

Applying the Clausius-Clapeyron relation (which at mid-latitudes typically dictates a change in daily precipitation extremes of approximately 6.5% per degree temperature change[39]), we scale down the distribution of observed precipitation according to the mean England and Wales portion of the attributable twentieth-century surface warming. Since the relation only applies to precipitation extremes[38] we only scale down precipitation above the 90th percentile. The resulting distributions of thermodynamically reduced precipitation extremes, had the warming not occurred, are shown by non blue-coloured curves in Supplementary Fig. 1.

Given this reduction in precipitation extremes, we now estimate the FAR of a severe precipitation event in the observed climate and so directly, in this simple framework, of flooding. However, because the precipitation distributions exhibit noise due to limited sampling of extremes in the relatively short 1990-2000 observation period, because the Clausius-Clapeyron relation can reasonably apply to a range of extreme percentiles, and because 'severe' is generally not well defined, we consider a range of percentiles above the 90th percentile of the observed distribution as defining severe precipitation event thresholds, and calculate the FAR each time. Similarly to the main text, we estimate uncertainty in these FAR calculations using a Monte Carlo bootstrap sampling procedure on precipitation in pairs of observed and thermodynamically reduced distributions.

Supplementary Fig. 2 shows the variation of FAR with percentile of the observed precipitation distribution. Clearly there is noise due to limited sampling, but the best estimate (median) aggregate FAR is roughly stable at 0.25 (corresponding to a 33% increase in flood risk) – which is within range of the aggregate FAR of autumn 2000 flooding estimated using our PEA framework (10-90% confidence interval of the black histogram in Fig. 4 is 0.2-0.9 (1 s.f), as indicated by the horizontal bars in Supplementary Fig. 2). This consistency of increased flood risk between the simple thermodynamic framework and more rigorous PEA framework supports the physical plausibility of our model-based PEA result.

However an exact comparison of attributable autumn 2000 flood risk between frameworks is difficult because the PEA framework's estimate uses a severe event threshold corresponding to approximately a 99.9th percentile event (horizontal line in

Fig. 3 intersects A2000 curve at approximately an one-in-nine autumn occurrence frequency, as indicated by the solid vertical line in Supplementary Fig. 2), whereas limited sampling in the simple framework prevents reliable estimates for approximately 99th percentile events and above. In particular, the simple framework's best estimate aggregate FAR of roughly 0.25 for lower percentile events falls appreciably short of the PEA framework's best estimate FAR of 0.6 (median of the black histogram in Fig. 4 to 1 s.f., as indicated by the square in Supplementary Fig. 2), and it is difficult to assess whether there would be better agreement were simple framework estimates closer to the 99.9th percentile possible, and thus whether a thermodynamic change in precipitation extremes is a reasonable determinant of change in flood risk there or if other factors such as dynamics and hydrology become more important. This is an inherent limitation in the simple framework of having only a few observed realisations of precipitation extremes available from which to estimate attributable flood risk – and demonstrates why the PEA framework, with many realisations all of which account for thermodynamics, dynamics, and hydrology, can better quantify this risk.

b. Robustness of modelled response to greenhouse gas emissions: thermodynamic versus dynamic contributions

The analysis of Section a. showed that the response of England and Wales total daily autumn river runoff to twentieth-century anthropogenic greenhouse emissions in an explicit-modelling framework is approximately consistent with a simple thermodynamic response. However, the response in the explicit-modelling framework was generally higher, suggesting that dynamics were tending to reinforce the simple thermodynamic response. This could happen on many scales: from feedback in local moisture convergence to changes in continental-scale circulation regimes. But simulation of the statistics of circulation regimes with climate models, even at a seasonal-forecast resolution such as used here, is strongly model-dependent[47]. Here we assess, therefore, to what extent our key results in the main text might have arisen from increasing greenhouse gas emissions inducing a regime-shift in our climate model – which would make those results acutely model-dependent.

Supplementary Fig. 3a shows the runoff response in going from the A2000N to the A2000 climate, via the change in occurrence frequency curves (identical to those in Fig. 3, but with the curves from all four A2000N climate estimates amalgamated onto a single panel for simplicity). For each climate model simulation contributing to these curves, the prevalence of the weather regime relevant to UK autumn 2000 weather may be represented by the projection of the simulated mean autumn 500-hPa geopotential height onto the observed synoptic structure in Fig. 1a that is closely related to the 'Scandinavia' circulation pattern[10]. So to estimate the impact of any changes in the strength of the Scandinavia pattern on our results, we attempt to first eliminate this impact from the runoff response by extracting and comparing only sub-ensembles of A2000N and A2000 simulations in Supplementary Fig. 3a having similar autumn-mean projections onto the Scandinavia pattern. Thereby, we isolate changes due only to smaller scale dynamics and thermodynamics. We define 'similar' here as being within the same pentiles of the distribution of projections of A2000 simulations.

Supplementary Figs. 3b-f show the runoff response in going from the A2000N to the A2000 climate, with simulations now grouped in sub-ensembles according to the pentile projection ranges (from weakest to strongest projection range respectively). We see that runoff in all A2000 and A2000N sub-ensembles generally increases with projection strength, as would be expected for wetter UK conditions under a stronger Scandinavia pattern[10]. Moreover, we see that for A2000 and A2000N sub-ensembles corresponding to the different near-fixed strengths of this weather regime, runoff response is generally similar to that for the full ensemble (Supplementary Fig. 3a) in which changes in this weather regime were free to also influence that response. This suggests the flood risk response in the full ensemble is dominated by small-scale dynamics and thermodynamics, not by a shift in the occupation-frequency of the key atmospheric mode of variability for UK autumnal precipitation.

This gives some assurance that our results regarding FAR would be robust to use of a different (for example, higher resolution) atmospheric climate model: because while thermodynamic responses are generally consistent across models of varying resolution, at least at mid-latitudes[39,40,41], the response of weather regimes to external driving is not even consistent in sign. Hence, although another model might simulate a different response of the Scandinavia pattern to anthropogenic greenhouse gas emissions, our results do not appear to depend on it. Indeed, to reverse our key results, such a hypothetical model would have to simulate a substantial weakening of the Scandinavia pattern in response the emissions, which is the opposite of what is seen in our ensemble.

It remains important that we used a climate model capable of simulating the synoptic conditions observed in autumn 2000, including the magnitude of the Scandinavia pattern, but our results do not depend on the simulated response of that pattern to anthropogenic greenhouse gas emissions. While we would strongly welcome efforts by other modelling groups to test our results using other atmospheric models, the above sensitivity studies provide some assurance that our key results are likely to be robust across models.

c. Subtracting twentieth-century SST warming attributable to greenhouse gas emissions

To construct A2000N SSTs we subtracted, from the observed[26] A2000 SSTs, estimated twentieth-century warming attributable to greenhouse gas emissions. We did this for all weeks in the April 2000 to March 2001 simulation period, but illustrate the procedure here using autumn (September to November) mean values.

Supplementary Fig. 4 shows the autumn mean for the observed A2000 SSTs, as well as sea ice. These gridded weekly-mean observations are supplied with SSTs in the marginal ice zone generated from sea ice coverage using a quadratic relationship and set to -1.8^{o}C where coverage in a grid box is at least 90%[26]. We then spatio-temporally interpolate these observations to suit the HadAM3-N144 model.

The attributable twentieth-century greenhouse gas warming in these SSTs cannot be found directly from observations, because observations also contain the signal of both

natural (e.g. solar and volcanic) and other anthropogenic (e.g. sulphate aerosol) drivers, and internal variability. Instead we derived this warming from estimates in prior studies[21,22] that used established 'optimal fingerprinting' analysis[1,2]. This uses multiple linear regression to compare observed surface temperature change, $\boldsymbol{Y}$, against that simulated, $\boldsymbol{X}$, by a climate model under m external drivers (greenhouse gas, sulphate aerosol, solar, volcanic, etc.). Thus $\boldsymbol{Y}$ is considered a linear superposition of separately simulated spatio-temporal responses, $\boldsymbol{X}_i$, to m external drivers, giving

$$\boldsymbol{Y} = \sum_{i=1}^{m} \beta_i (\boldsymbol{X}_i - \boldsymbol{\nu}_i) + \boldsymbol{\nu}_0 \quad (1)$$

where the regression co-efficient β_i is a scaling factor on the amplitude of each response, $\boldsymbol{\nu}_i$ represents noise in each response due to its estimation from a finite sample of simulations, and $\boldsymbol{\nu}_0$ represents noise in the observations due to natural internal variability of the climate system. Accordingly, the regression is performed in the ('optimal') direction of maximum signal-to-noise of a response ('fingerprint') rather than the direction of maximum signal[50]. It returns an uncertainty distribution on β, and if the hypothesis $\beta = 0$ can be confidently rejected then the response is detected and the observed change can be attributed in some part to the corresponding driver.

Supplementary Fig. 5 shows uncertainty distributions on β for the twentieth-century surface temperature response to greenhouse gas driving, estimated in the prior studies using four coupled atmosphere-ocean climate models (HadCM3, GFDLR30, NCARPCM1, and MIROC3.2). All four distributions demonstrate a response is detected in observed change, with differences reflecting the relative strength of response (for example, the NCARPCM1 β is overall highest because the response is weakest in that model).

To what extent responses from other differently formulated coupled models are also detectable it is to date undetermined; even though more advanced optimal fingerprinting methods exist[51] to account for such inter-model variance. Likewise it is difficult to assess how this would affect the robustness of our main result (Fig. 4), without repeating our whole study. Here we assumed the four available detected responses, scaled by their respective β distribution, effectively span uncertainty in the true twentieth-century surface warming attributable to greenhouse gas emissions.

The β distributions were estimated for annual responses, but given the timescale of England and Wales autumn 2000 flooding we required seasonal resolution at the very least. In the absence of clear evidence of a trend in the magnitude of the seasonal cycle, we further assumed these scaling factors apply equally to seasonal response patterns. These seasonal patterns were constructed for each coupled model by subtracting surface temperature averaged over a lengthy pre-industrial control simulation having constant external drivers (assumed equivalent to 1900-09 decadal-mean climatology), from the 1990-99 decadal-mean surface temperature averaged over the small ensembles of simulations used in the prior optimal fingerprinting analyses having historical greenhouse gas driving up to the end of that decade.

Supplementary Fig. 6 shows our constructed response patterns for autumn. They are generally similar across coupled models due to domination of the land-sea contrast, with differences mainly reflecting different formulations of sea ice. We scale these patterns by deciles of the respective β distributions (5^{th}, 15^{th},…, 95^{th} percentiles; solid vertical lines in Supplementary Fig. 5) yielding 10 equiprobable estimates of attributable warming, per coupled model, that are subtracted from the A2000 SSTs.

Ideally one would repeat this procedure to construct A2000N sea ice. However, optimal fingerprinting analysis is limited in its ability to detect changes in variables having relatively small spatial coverage and sharp gradients, and both coupled model simulations and early twentieth-century observations of sea ice are unreliable. Instead we used a simple empirical SST-sea ice relationship, determined from observations, to construct sea ice consistent with our constructed A2000N SSTs and ensure it was consistent with the characteristics of sea ice coverage (see Methods).

Supplementary Notes

38. Allen, M.R. & Ingram, W.J. Constraints on future changes in climate and the hydrologic cycle. *Nature* **419**, 224-232 (2002).

39. Pall, P., Allen, M.R. & Stone, D.A. Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO_2 warming. *Clim. Dyn.* **28**, 351-363 (2007).

40. Emori, S. & Brown, S.J. Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophys. Res. Lett.* **32**, L17706 (2005).

41. Sugiyama, M., Shiogama, H. & Emori, S. Precipitation extreme changes exceeding moisture content increases in MIROC and IPCC climate models. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 571-575 (2010).

42. Durre, I., Williams, C.N., Yin, X & Vose, R.S. Radiosonde-based trends in precipitable water over the Northern Hemisphere: An update. *J. Geophys. Res.* **114**, D05112 (2009).

43. McCarthy, M.P., Thorne, P.W. & Titchner, H.A. An analysis of tropospheric humidity trends from radiosondes. *J. Clim.* **22**, 5820-5838 (2009).

44. Willett, K.M., Jones, P.D., Gillett, N.P. & Thorne, P.W. Recent changes in surface humidity: Development of the HadCRUH dataset. *J. Clim.* **21**, 5364-5383 (2008).

45. Willett, K.M., Jones, P.D., Thorne, P.W. & Gillett, N.P. A comparison of large scale changes in surface humidity over land in observations and CMIP3 general circulation models. *Environ. Res. Lett.* **5**, 025210 (2010).

46. Willett, K.M., Gillett, N.P., Jones, P.D. & Thorne, P.W. Attribution of observed surface humidity changes to human influence. *Nature.* **449**, 710-713 (2007).

47. Randall, D.A. *et al*. in *Climate change 2007: The physical science basis.* (eds Solomon, S. *et al*.) Ch.8, 589-662 (Cambridge University Press, 2007).

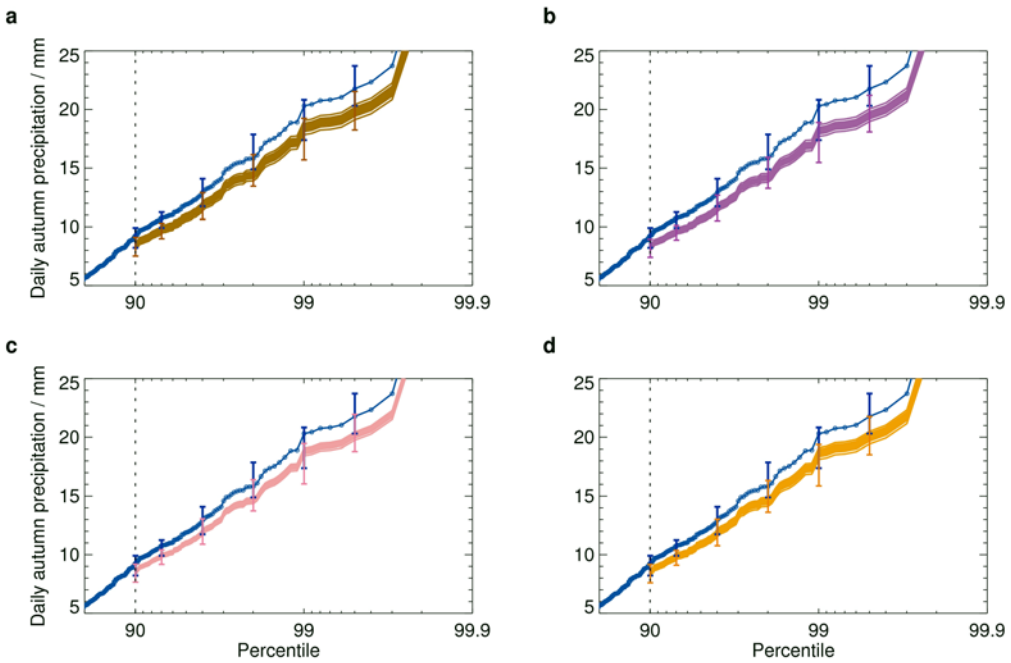
48. Goddard, D.M. Interpolation techniques and grid transformation used in the Unified Model. UK Met Office Unified Model Documentation Paper S1, 66 pp. (1998).

49. Taylor, K.E., Williamson, D. & Zwiers, F. The sea surface temperature and sea-ice concentration boundary conditions for AMIP II simulations. PCMDI Report No. 60, 24 pp. (2000).

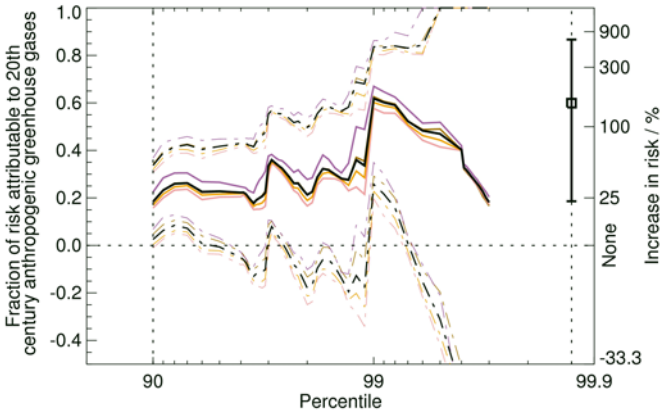
50. Allen, M.R. & Stott, P.A. Estimating signal amplitudes in optimal fingerprinting, part I: theory. *Clim. Dyn.* **21**, 477-491 (2003).

51. Huntingford, C., Stott, P.A., Allen, M.R. & Lambert, F.H. Incorporating model uncertainty into attribution of observed temperature change. *Geophys. Res. Lett.* **33**, L05710 (2006).

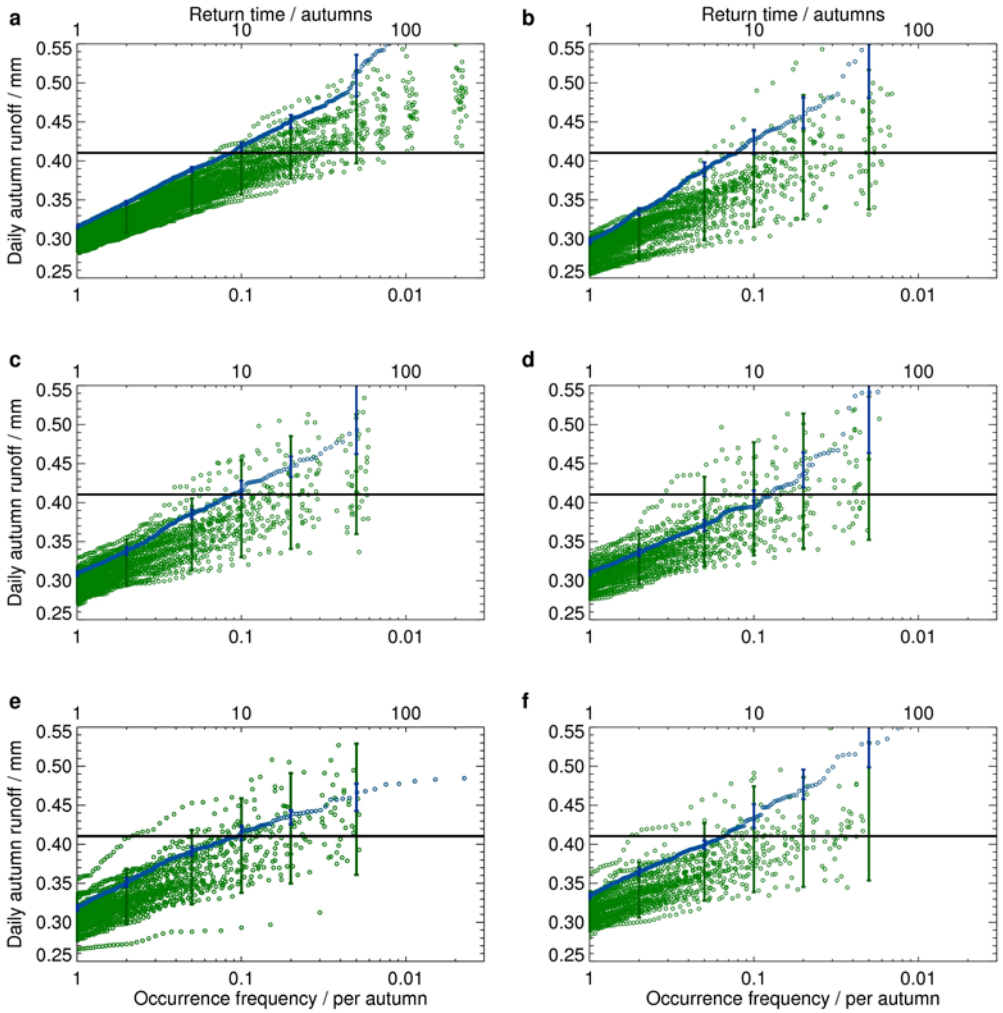
Supplementary Figures and Legends



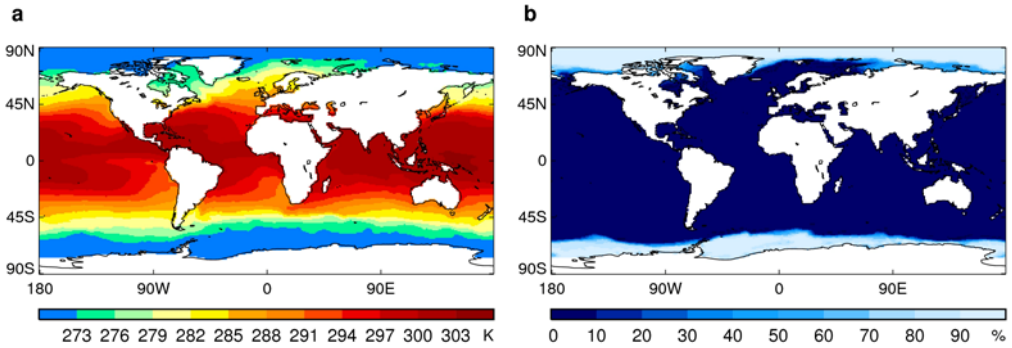
Supplementary Figure 1. Thermodynamic change in daily precipitation extremes for England and Wales autumns. a-d, Each panel identically shows the distribution of observed daily precipitation extremes from autumns 1990-2000 (blue circles, with interpolated curve). Non-blue coloured curves are different for each panel and show the distribution of thermodynamically reduced daily precipitation extremes for this same period, which is deduced by scaling down the 90th percentile and above (delineated by the dotted vertical line) of the observed distribution according to the Clausius-Clapeyron relation and pattern-estimates of attributable twentieth-century surface warming from, HadCM3 (**a**; brown), GFDLR30 (**b**; purple), NCARPCM1 (**c**; pink), and MIROC3.2 (**d**; orange). Ten such non-blue curves per panel result from the ten amplitude-scalings per pattern-estimate (see Methods). Bars represent 5-95% confidence intervals estimated using a Monte Carlo bootstrap sampling procedure similar to that in Fig. 3.



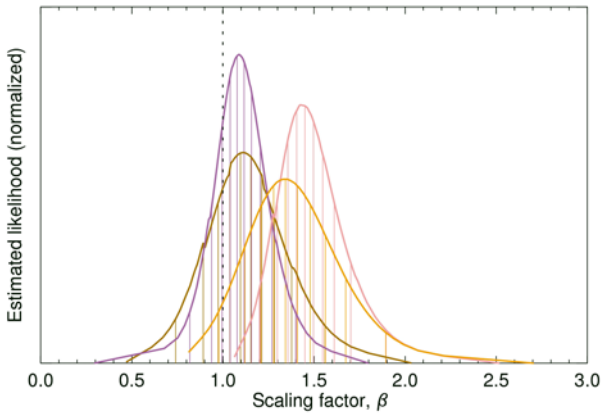
Supplementary Figure 2. A range of changes in risk of daily precipitation extremes for England and Wales autumns. Shown is the range of fraction of attributable risk (FAR) of daily precipitation extremes, estimated using the distributions of observed and thermodynamically reduced autumns 1990-2000 precipitation of Supplementary Fig. 1. These FAR estimates are made at the 90th percentile of the observed distribution onwards. Colours are as in Fig. 4, with black again representing the aggregate FAR. Dashed curves mark the 10-90% confidence intervals estimated using a Monte Carlo bootstrap sampling procedure similar to that for Fig. 4. Solid curves mark the 'best estimate' (median) FAR. Right hand axis shows the equivalent increase in risk. Also marked is the aggregate FAR of autumn 2000 flooding estimated in terms of daily runoff using the PEA framework of the main text; with horizontal bars indicating the 10-90% confidence interval and square indicating the best estimate (median), all located on the solid vertical line that denotes the flooding was a 99.87th percentile event in the A2000 climate.



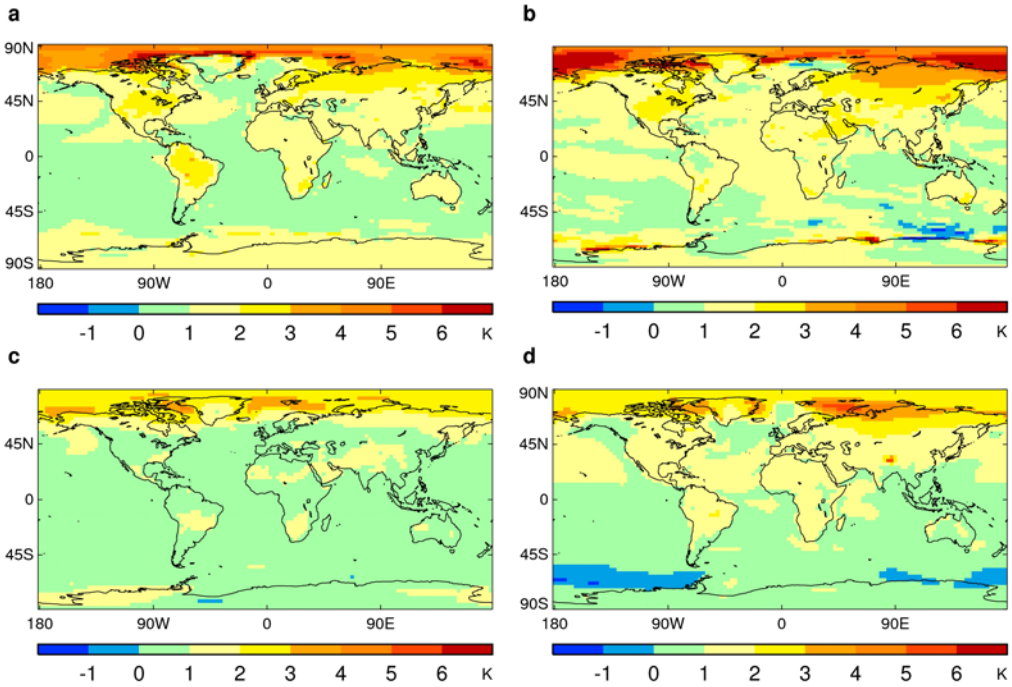
Supplementary Figure 3. As for Fig. 3 but, **a**, amalgamating the curves (now green) from all four A2000N climate estimates; then grouping all simulations according to the strength of projection (dot product) of their mean autumn 500-hPa geopotential height onto the regression pattern in Fig. 1a over the region bounded 90E, 90N, 90W, 30N, where grouping ranges are defined by the 0-20% (**b**), 20-40% (**c**), 40-60% (**d**), 60-80% (**e**), and 80-100% (**f**) pentiles of the distribution of projections of A2000 simulations.



Supplementary Figure 4. Mean autumn A2000 SSTs and sea ice. Observed[27] SSTs are interpolated from the supplied 1° gridded weekly-mean values to 6-day 'weekly'-mean values suited to the HadAM3-N144 grid and 360-day year, using methods[48,49] attempting to preserve those observed means. The autumn 2000 mean of these re-gridded 6-day weekly-mean values is shown in **a**; likewise for sea ice coverage in **b**.



Supplementary Figure 5. Scaling factors for modelled twentieth-century surface temperature response to greenhouse gas driving. Curves show the uncertainty distribution on the amplitude-scaling factor, β, for the twentieth-century surface temperature response to greenhouse gas driving; estimated[21,22] using optimal fingerprinting analysis of HadCM3 (brown), GFDLR30 (purple), NCARPCM1 (pink), and MIROC3.2 (orange) simulations. Solid vertical lines mark deciles of a distribution, ranging from the 5th to the 95th percentile. β=0 implies a response is not detected in observed change; β=1 implies the response is identical to that change.



Supplementary Figure 6. Modelled twentieth-century surface temperature response to greenhouse gas driving. Maps show 1990s decadal-mean autumn surface temperature anomalies, constructed from the average of a small ensemble of historical greenhouse gas driven simulations relative to the average of a lengthy pre-industrial control simulation, using HadCM3 (**a**), GFDLR30 (**b**), NCARPCM1 (**c**), and MIROC3.2 (**d**).

Supplementary Tables

Catchment	RMSE	Correlation
Anglian	0.83	0.82
Dee	0.73	0.85
Northumbria	1.55	0.79
North West	2.22	0.73
Severn	0.38	0.91
Solway Tweed	0.84	0.89
South East	0.79	0.80
South West	0.37	0.86
Thames	0.25	0.89
Trent	0.48	0.74
Wales	1.95	0.83

Supplementary Table 1. Evaluation of the coupled hydrologic-hydraulic scheme. Shown are the root-mean-squared error (RMSE) in mm and coefficient of correlation for daily river runoff from the coupled hydrologic-hydraulic scheme against observations[36], for 11 England and Wales catchments. Evaluation is for 1986-95. Area-weighted mean RMSE and correlation across catchments is 0.61 and 0.85 respectively (see Methods).

		Correlation (all quantiles)	Correlation ($\geq 75^{th}$ quantile)	Correlation ($\geq 90^{th}$ quantile)
Against scheme	**Annual**	0.82	0.76	0.77
	Autumnal	0.86	0.76	0.74
Against obs.	**Annual**	0.76	0.71	0.59
	Autumnal	0.82	0.72	0.64

Supplementary Table 2. Evaluation of the precipitation-runoff model. Shown is the coefficient of correlation for England and Wales total daily river runoff from the precipitation-runoff model, against both the coupled hydrologic-hydraulic scheme and observations[36]. This is shown annually and autumnally both for all runoff quantiles, and for all quantiles at or above the 75^{th} and 90^{th} quantiles. Evaluation is for 1986-95 (see Methods).