2-18-2014

Are climate model simulations useful for forecasting precipitation trends? Hindcast and synthetic-data experiments

Nir Y. Krakauer
CUNY City College

Balázs M. Fekete
CUNY City College

Recommended Citation

Are climate model simulations useful for forecasting precipitation trends? Hindcast and synthetic-data experiments

Nir Y Krakauer and Balázs M Fekete

Department of Civil Engineering and NOAA CREST, The City College of New York, New York NY 10031, USA

E-mail: nkrakauer@ccny.cuny.edu

Abstract

Water scientists and managers currently face the question of whether trends in climate variables that affect water supplies and hazards can be anticipated. We investigate to what extent climate model simulations may provide accurate forecasts of future hydrologic nonstationarity in the form of changes in precipitation amount. We compare gridded station observations (GPCC Full Data Product, 1901–2010) and climate model outputs (CMIP5 Historical and RCP8.5 simulations, 1901–2100) in real and synthetic-data hindcast experiments. The hindcast experiments show that imputing precipitation trends based on the climate model mean reduced the root mean square error of precipitation trend estimates for 1961–2010 by 9% compared to making the assumption (implied by hydrologic stationarity) of no trend in precipitation. Given the accelerating pace of climate change, the benefits of incorporating climate model assessments of precipitation trends in water resource planning are projected to increase for future decades. The distribution of climate models' simulated precipitation trends shows substantial spatially coherent biases, suggesting that there may be room for further improvement in how climate models are parametrized and used for precipitation estimation. Linear extrapolation of observed trends in long precipitation records may also be useful, particularly for lead times shorter than about 25 years. Overall, our findings suggest that simulations by current global climate models, combined with the continued maintenance of *in situ* hydrologic observations, can provide useful information on future changes in the hydrologic cycle.

Keywords: hydrologic prediction, nonstationarity, climate change, precipitation, climate model, trend estimation, extrapolation

1. Introduction

In the present nonstationary regime associated with climate change, predicting trends in hydrologic variables such as precipitation would be of great value for water resources planning, with applications ranging from municipal decision making to energy modeling to ecological management to disaster preparedness [1]. Climate model simulations are widely used to generate scenarios of future precipitation change for such applications [2, 3]. However, the ability of global climate models (GCMs) to accurately represent the impacts of climate forcing on precipitation is unclear: GCMs have large known biases in representing present-day precipitation distributions, and different GCMs differ even as to the sign of expected precipitation change in many regions under particular climate forcing scenarios [4, 5]. Given continuing uncertainties as to the ability of GCMs to model precipitation change, some hydrologists and water managers have recommended maintaining the stationarity assumption for water resources planning, while building in robustness and resiliency whenever possible as precautionary measures, pending more definitive cues from observations and improvements in scientific understanding [6–8].

Here, we employ two ways to empirically assess the quality of GCM simulations of trends in precipitation: (1) synthetic-data experiments and (2) hindcasts with observational precipitation data. Synthetic-data experiments evaluate the ability of GCMs to predict precipitation trends simulated, for example by another GCM, under given forcing. Observation hindcasts evaluate the accuracy of GCMs in simulating past precipitation trends. Synthetic data have the advantage of completeness and ability to be fully characterized (being the output of a numerical model), and synthetic-data experiments can be carried out for any desired boundary conditions and climate forcing. Observations are limited in availability and have incompletely characterized errors, but have the important advantage of coming from the real earth system with which planners are faced. Combining the two approaches allows the more ambitious comparisons that can be carried out with synthetic data to be anchored by findings on how well synthetic data compare in predictability to actual observations.

We present an assessment of precipitation trend forecasting methods for (a) 1960–2010 hindcasts, using either observational or synthetic (climate model) data; (b) 2011–2100 forecasts, using synthetic data. Our goal is to determine whether and under what circumstances available GCM simulations enable better forecasts of precipitation changes compared to extrapolations based on historic data. Unlike many previous comparisons of modeled and observed precipitation trends [9, 10], we consider trends at close to the model grid scale, rather than aggregating to larger regions (latitude bands and continents) whose trends tend to be less useful for hydrologic decision makers.

2. Methods

2.1. Observational data

As an observation-based estimate of precipitation, we used the Global Precipitation Climatology Centre (GPCC) Full Data Product, Version 6, at 2.5° spatial resolution [11]. This monthly GPCC product is available for the years 1901–2010 and covers global land areas excluding Antarctica. The gridded precipitation estimates are based on a larger number of gauges than any other available product that covers a comparably long time span, which reduces bias for comparing to climate models [12]. This product has been previously compared to GCM precipitation trends [13].

2.2. Model simulations

We obtained monthly precipitation fields from GCM simulations undertaken for the Coupled Model Intercomparison Project Phase 5 (CMIP5) [14] as archived by the Earth System Grid Federation. We selected all available GCM simulations with complete monthly precipitation fields for the Intergovernmental Panel on Climate Change Historical (1901–2005) and RCP8.5 (2006–2100) runs [15]. The RCP8.5 forcing scenario was chosen because it extrapolates recent emissions trends [16]. This yielded 25 GCMs with complete model global precipitation fields for 1901–2100. A total of 56 ensemble members were available for these runs, although most GCMs

(16/25) had only one ensemble member available. Where multiple ensemble members were provided for the same GCM, we selected only one (the first by alphanumeric order of name). GCM precipitation fields were converted to the same 2.5° grid as in the GPCC product by amount-conserving bilinear interpolation [17].

2.3. Hindcast experiments: observational data

We assumed that the yearly precipitation time series $P(t)$ can be represented as

$$P(t) = \bar{P}(t) + \epsilon(t), \tag{1}$$

where $\bar{P}(t)$ is a smooth trend component, while $\epsilon(t)$ is a zero-mean high-frequency component with little year-to-year persistence. Based on available data up to a time t_1 and GCM simulations, our objective was to hindcast the precipitation change $\bar{P}(t_2) - \bar{P}(t_1)$ for various times $t_2 > t_1$. The observed change magnitude was determined using spline smoothing of the entire precipitation time series to estimate $\bar{P}(t)$.

We compared three extrapolation methods and one GCM-based method for this hindcasting:

(i) Hindcast zero change in precipitation, so that $\bar{P}(t_2) = \bar{P}(t_1)$ (stationarity or persistence forecast, *prs*).

(ii) Use linear regression for the period from 1901 to t_1 to estimate the rate of change in precipitation, and assume that the same linear trend continues to t_2 (linear extrapolation forecast, *elr*).

(iii) Fit a smoothing cubic spline (using Vapnik's method [18] for choosing the smoothing parameter) for the period from 1901 to t_1 to estimate the rate of change in precipitation at t_1, and linearly extrapolate the spline curve to t_2 (spline extrapolation forecast, *esp*). This method approaches the linear regression result if the data period is short or the trend is not significantly nonlinear. A similar form of spline extrapolation was previously used for estimating changes in cold extremes [19].

(iv) Fit a smoothing cubic spline to precipitation from each GCM run using the entire run from 1901 to 2010, and use the multimodel mean $\bar{P}(t_2) - \bar{P}(t_1)$ as the forecast (multimodel mean forecast, *mmm*). Using the multimodel mean is consistent with previous findings that multimodel averages typically performed better than individual GCMs in comparisons of observations with climate model ensembles [20–23].

Hindcasts were begun each year from 1960 to 2009 for subsequent years in the range 1961–2010 (1–50 years ahead). The main metric for hindcast quality was the root mean square error (RMSE) in the hindcast trend, averaged across land grid cells and than across hindcast lead times. Averages across grid cells were weighted by cell area.

2.4. Hindcast and forecast experiments: synthetic data

The hindcasts that began each year from 1960 to 2009 were also carried out using one of the GCM precipitation fields as synthetic data instead of the observations. The hindcasts were

carried out in exactly the same way as those with observational data, and repeated with each of the 25 GCMs in turn serving as synthetic data and the other 24 used to construct the multimodel mean forecast. Synthetic-data results are reported as averages across these 25 realizations.

Forecasts were also begun each year from 2010 to 2099 for subsequent years in the range 2011–2100 (1–90 years ahead). For these, the model precipitation fields since 1901 were used to estimate trends.

3. Results

3.1. Mean square error for hindcast experiments

For the 1961–2010 hindcasts, the multimodel mean *mmm* showed the lowest mean square error with both the GPCC observations and with synthetic data. Of the extrapolation methods, the linear and spline extrapolations *elr* and *esp* both had worse average performance than simply assuming no trend (*prs*) (first two columns of table 1). For the period up to 2010, the smoothing criterion for spline extrapolation resulted in modeled and observed precipitation trends that were very close to linear, which is why *elr* and *esp* give almost the same results. *mmm* outperformed *prs* on average at all lead times from 1 to 50 years, and for both methods, the hindcast RMSE increased linearly with lag time (figures 1(a) and (b)). The performance advantage of *mmm* over *prs* was comparable in the observations and in the synthetic data, although larger in the synthetic-data experiment (16%) than in the GPCC observations experiment (9%) (table 1). Despite their worse average performance, the linear and spline extrapolations *elr* and *esp* both performed better than *prs* at short lead times (<15 yr) (figures 1(a) and (b)).

For the 2011–2100 forecasts with synthetic data, *mmm* again outperformed *prs*, this time by a larger amount (31%). *elr* and *esp* also outperformed *prs* (table 1). *esp* even slightly outperformed *mmm* at shorter lead times (<25 yr). For each forecast method, RMSE increased with forecast lead time, and the increase was somewhat faster than for the earlier period, reflecting larger between-GCM differences in precipitation trends over the 21st century (figure 1(c)).

We conducted several alternative analyses to assess the sensitivity of our hindcast results to model and observation data selection. If we averaged all available CMIP5 ensemble members for each GCM instead of only one ensemble member per GCM, the model hindcast RMSE increased slightly (by 1%). If we excluded data from before 1950 (when fewer precipitation measurements are available) for estimating precipitation trends over the 1961–2010 hindcast period, RMSE for all methods increased by over 70%, reflecting the importance of a long baseline period for accurate trend estimation, but the relative performance advantage of *mmm* over *prs* was similar to our base case (10% as compared to 9%).

3.2. Trends in observations versus models

The performance of the multimodel mean *mmm* in the observation hindcasts reflects the degree to which the GCM precipitation trends agree with the GPCC observations. Figure 2(a) shows the spatial distribution of normalized precipitation
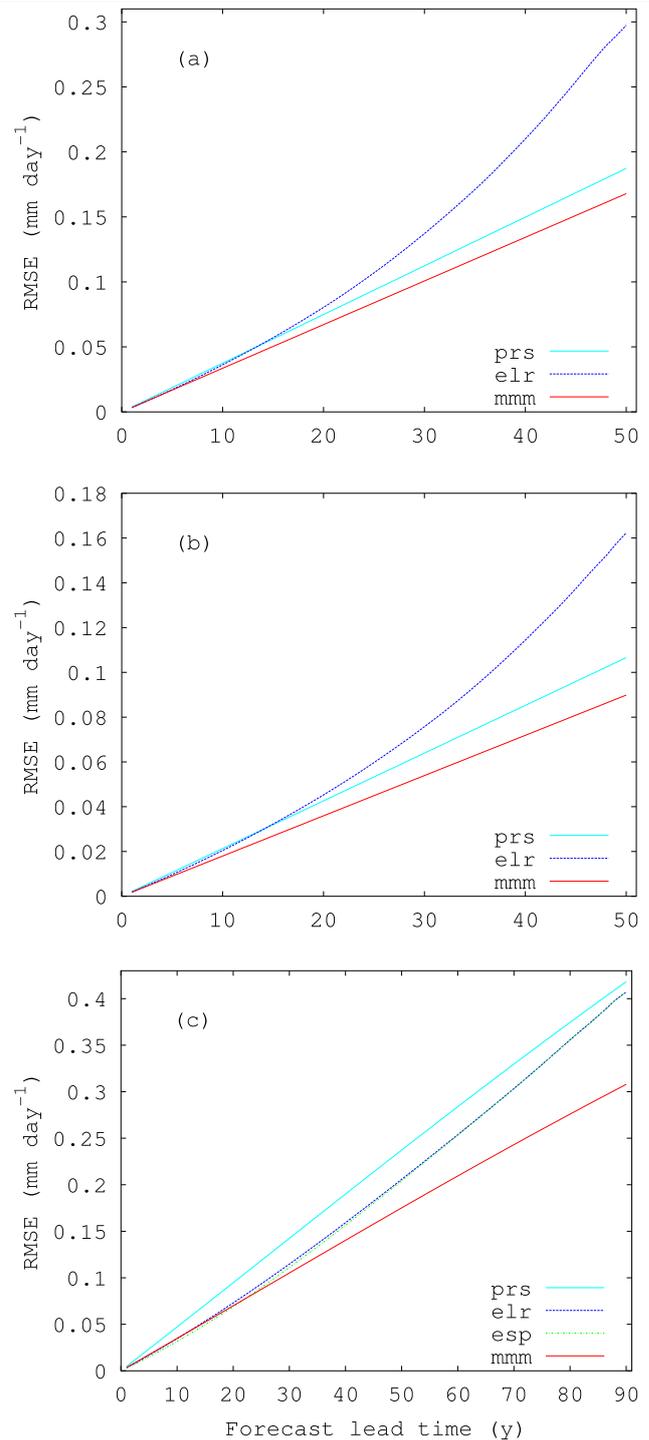
Figure 1. (a) Root mean square error (RMSE) of precipitation trend estimates for 1961–2010 compared to observational data (from GPCC), as a function of lead time. The estimates were made using either only extrapolation from observations (*prs*, *elr*, *esp* methods) or the mean of climate models (*mmm*). (b) Same, but compared to synthetic precipitation data from climate models for 1961–2010. (c) Same, but compared to synthetic precipitation data for the period 2011–2100. In panels a and b, *esp* is not plotted because it is effectively identical to *elr*.

trends over the observation period ($\bar{P}(2010) - \bar{P}(1901)$ divided by the mean precipitation from GPCC). Large increases in precipitation of 20–40% (about 1–2 standard deviations) can be seen, for example in northern Eurasia, northern North
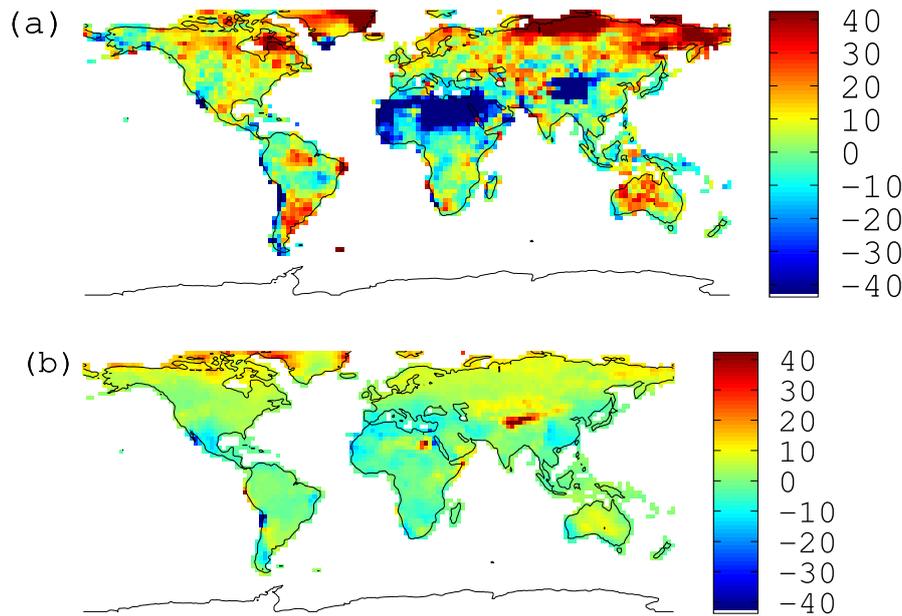
Figure 2. (a) Normalized precipitation change (in % of the mean amount) between 1901 and 2010 according to GPCC. (b) Same, but for the CMIP5 climate model mean. The color scale is the same for both panels.

Table 1. Root mean square error (mm day^{-1}) for precipitation trend estimates using different methods.

	Hindcast: observations	Hindcast: synthetic data	Forecast: synthetic data
prs	0.0788	0.0448	0.1766
elr	0.0991	0.0546	0.1551
esp	0.0991	0.0546	0.1536
mmm	0.0706	0.0378	0.1303

America (except Alaska), and Argentina. Similarly substantial decreases in precipitation can be seen, for example, in southwest Asia, Africa north of the Equator, the Pacific coast of South America, and southwestern Australia. Figure 2(b) shows that the CMIP5 climate model runs also show precipitation increases in the north, although of considerably smaller magnitude than observed (on the order of 10%). In other parts of the world, the CMIP5 average shows discrepancies in the sign of precipitation trends compared to GPCC: for example, eastern China's precipitation decreases in the model runs but not in GPCC, and the model runs do not capture the drying of northern Africa seen in GPCC.

Figure 3 shows the relative performance of *mmm* and *prs* in the observation hindcasts for each grid cell, expressed as

$$\log_e\left(\frac{\text{RMSE}_{prs}}{\text{RMSE}_{mmm}}\right).$$

In general, *mmm* does better than the *prs* forecast of no change (orange and red areas in figure 3) where the modeled precipitation changes are of the same sign as observations, for example in most boreal regions (except Alaska). *mmm* does worse than *prs* (green and blue in figure 3) where the modeled precipitation changes disagree with observations, as in much of China and Africa. Overall, *mmm* outperforms *prs* over 63% of the land area. The multimodel mean trend also outperforms all individual CMIP5 models (not shown), as measured by the

correlation between modeled and GPCC precipitation trends and by the root mean square error compared to the GPCC precipitation trend, supporting the use of multimodel averages for precipitation trend prediction.

4. Discussion

Our results show that global climate models, as represented by CMIP5 submissions, have some skill in representing precipitation trends over recent decades, even when evaluated at a spatial scale close to the original model grid scale (2.5°). Using individual climate model outputs as synthetic data, as compared to using observations, results in estimates of this skill which are inflated, but of the right order of magnitude (e.g. reduction in RMSE of 16% versus 9%). This lends confidence to the 21st century forecast evaluations using synthetic data which show substantial skill for the model mean compared to the *prs* assumption of stationarity. Thus, it appears likely that applying precipitation change factors based on the CMIP5 ensemble mean to water resources planning would lead to better decisions than making no use of these models' precipitation estimates.

The 9% reduction in RMSE seen for land precipitation trends is similar to the magnitude of the correlation found between observed and modeled 20th century streamflow trends [24]. Interpreting streamflow trends is more complicated than for precipitation trends, however, since streamflow is affected
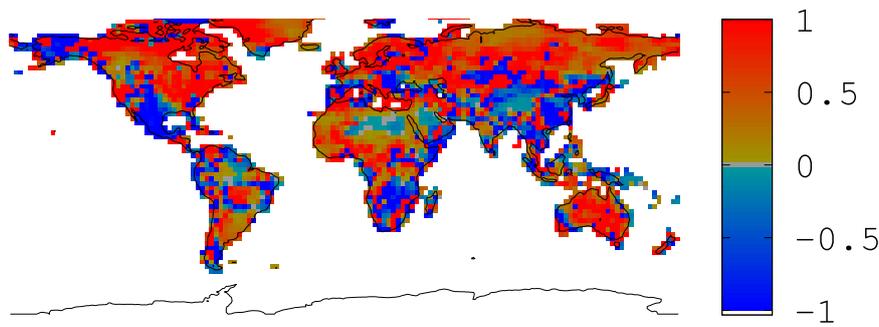
Figure 3. Log of the ratio of persistence (*prs*) to model mean (*mmm*) root mean square error for precipitation trend estimates over 1961–2010. Where values are positive, the model mean is sufficiently skillful that *mmm* outperforms *prs*. Refer to the text for detailed explanation of the forecast methods *prs* and *mmm*.

not only by precipitation but also by temperature, atmospheric CO_2 (regulating plant stomatal conductance and hence transpiration rate), and anthropogenic land use change and water regulation and diversion [25–28]. The ability of modeled climate trends to directly inform projections of streamflow, reservoir storage, soil moisture, and other land-surface hydrologic variables thus remains an important practical question. Other climate variables such as precipitation intensity are also important for water resource planning and natural hazard preparedness, and the ability of GCMs to contribute to estimates of future trends in these quantities could be explored using the methods used here for mean precipitation.

Here we only studied GCM abilities to predict precipitation *trends* ($\bar{P}(t_2) - \bar{P}(t_1)$ in the notation used above). Absolute precipitation *amounts* ($\bar{P}(t_1)$) are of course also vital to water resource management, and here long, accurate observation time series are clearly needed. Thus, maintaining and expanding networks for precipitation and other hydrologic observations is critical for informed hydrologic decision making, as well as for evaluating climate model outputs over the coming decades [29, 30]. With long, high-quality observation time series, extrapolating the observed trend may in fact outperform GCM projections for shorter lead times (e.g. less than about 25 years), as seen with the 21st century synthetic-data forecast experiments here (compare *elr* to *mmm* in figure 1(c)). Checking against local observations and hydrological experience is also recommended to confirm that historical precipitation trends estimated using gridded global datasets such as GPCC are valid for the particular area of interest.

One caveat is that our analyses only considered one radiative forcing trajectory over the 21st century (RCP8.5). Uncertainty about emission pathways will tend to increase the RMSE of the multimodel mean as an estimate of precipitation trends, particularly at long lead times (greater than several decades) when the choice of future emission pathway has an appreciable impact on projected radiative forcing [31].

The skill of climate models in hindcasting historic precipitation changes derives from their ability to simulate the observed changes, which is clearly imperfect. Even when simulated changes are of the right sign, they are rarely of the right magnitude—for example, precipitation increases over most of the Arctic are substantially underestimated. The causes for these discrepancies are being studied. For example, discrepancies between modeled and observed precipitation trends over

Europe seem to be associated with model underestimation of warming in North Atlantic sea surface temperature and trends in a Mediterranean–Scandinavia pressure difference [32, 33]. In fact, poor simulation of tropical ocean warming may be an important reason for discrepancies between modeled and observed precipitation trends over North America and north Africa as well as Europe [34]. Over tropical land areas, variability in the Southern Oscillation may be an important cause of observation-model discrepancies in precipitation trends [35]. The response of Northern Hemisphere land precipitation over the 20th century to anthropogenic aerosols appears to be underestimated in most GCMs, providing a complementary possible explanation of observation-model discrepancies [36–38].

Hopefully, some problems with GCM simulations of precipitation will be ameliorated with improvements in the representation of relevant atmosphere, ocean, and land-surface microphysical and dynamical processes. Pending such model improvements, it may also be worth investigating whether empirical corrections of the multimodel mean for biases using the historic record can be used to generate improved forecast products, particularly for the nearer term (lead times up to several decades). Such data-adaptive use of climate models may be explored, for example, using machine learning methods [39].

5. Conclusions

Our hindcast and forecast experiments suggest that climate models have some skill in simulating precipitation change over the next few decades. Given the large regional changes (in many places well in excess of 1 standard deviation) in precipitation seen over the last century, the assumption of hydrologic stationarity is a suboptimal one for planning even given the limitations of current climate models. Precipitation forecasts could be further improved by addressing model deficiencies or by empirical correction of the simulated spatial patterns of precipitation change.

Acknowledgments

References

[1] Milly P C D, Betancourt J, Falkenmark M, Hirsch R M, Kundzewicz Z W, Lettenmaier D P and Stouffer R J 2008 Stationarity is dead: whither water management? *Science* **319** 573–4

[2] Towler E, Rajagopalan B, Gilleland E, Summers R S, Yates D and Katz R W 2010 Modeling hydrologic and water quality extremes in a changing climate: a statistical approach based on extreme value theory *Water Resources Res.* **46** W11504

[3] Jakob D 2013 Nonstationary in extremes and engineering design *Extremes in a Changing Climate (Water Science and Technology Library* vol 65) (Dordrecht: Springer) pp 363–417

[4] Phillips T J and Gleckler P J 2006 Evaluation of continental precipitation in 20th century climate simulations: the utility of multimodel statistics *Water Resources Res.* **42** W03202

[5] Trenberth K 2011 Changes in precipitation with climate change *Clim. Res.* **47** 123–38

[6] Anagnostopoulos G G, Koutsoyiannis D, Christofides A, Efstratiadis A and Mamassi N 2010 A comparison of local and aggregated climate model outputs with observed data *Hydrol. Sci. J.* **55** 1094–110

[7] Stakhiv E Z 2011 Pragmatic approaches for water management under climate change uncertainty *J. Am. Water Resour. Assoc.* **47** 1183–96

[8] Fekete B M and Stakhiv E 2013 Water management preparation strategies for adaptation to changing climate *Climate Change and Inland Waters: Impacts and Mitigation for Ecosystems and Societies* ed C R Goldman, M Kumagai and R D Robarts (New York: Wiley) chapter 23

[9] Zhang X, Zwiers F W, Hegerl G C, Lambert F H, Gillett N P, Solomon S, Stott P A and Nozawa T 2007 Detection of human influence on twentieth-century precipitation trends *Nature* **448** 461–6

[10] Min S-K, Zhang X, Zwiers F W and Hegerl G C 2011 Human contribution to more-intense precipitation extremes *Nature* **470** 378–81

[11] Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U and Ziese M 2013 A description of the global land-surface precipitation data products of the global precipitation climatology centre with sample applications including centennial (trend) analysis from 1901–present *Earth Syst. Sci. Data* **5** 71–99

[12] Wan H, Zhang X, Zwiers F W and Shiogama H 2013 Effect of data coverage on the estimation of mean and variability of precipitation at global and regional scales *J. Geophys. Res.* **118** 534–46

[13] van Oldenborgh G J, Reyes F J D, Drijfhout S S and Hawkins E 2013 Reliability of regional climate model trends *Environ. Res. Lett.* **8** 014055

[14] Taylor K E, Stouffer R J and Meehl G A 2012 An overview of CMIP5 and the experiment design *Bull. Am. Meteorol. Soc.* **93** 485–98

[15] van Vuuren D P *et al* 2011 The representative concentration pathways: an overview *Clim. Change* **109** 5–31

[16] Peters G P, Andrew R M, Boden T, Canadell J G, Ciais P, Le Quere C, Marland G, Raupach M R and Wilson C 2013 The challenge to keep global warming below 2 °C *Nature Clim. Change* **3** 4–6

[17] Hahmann A N and Dickinson R E 2001 A fine-mesh land approach for general circulation models and its impact on regional climate *J. Clim.* **14** 1634–46

[18] Cherkassky V and Mulier F 2007 *Learning From Data: Concepts, Theory, and Methods* (New York: Wiley)

[19] Krakauer N Y 2012 Estimating climate trends: application to US plant hardiness zones *Adv. Meteorol.* **2012** 404876

[20] Lambert S J and Boer G J 2001 CMIP1 evaluation and intercomparison of coupled climate models *Clim. Dyn.* **17** 83–106

[21] Gleckler P J, Taylor K E and Doutriaux C 2008 Performance metrics for climate models *J. Geophys. Res.* **113** 27

[22] Pierce D W, Barnett T P, Santer B D and Gleckler P J 2009 Selecting global climate models for regional climate change studies *Proc. Natl Acad. Sci.* **106** 8441–6

[23] Sanderson B M and Knutti R 2012 On the interpretation of constrained climate model ensembles *Geophys. Res. Lett.* **39** L16708

[24] Milly P C D, Dunne K A and Vecchia A V 2005 Global pattern of trends in streamflow and water availability in a changing climate *Nature* **438** 347–50

[25] Gedney N, Cox P M, Betts R A, Boucher O, Huntingford C and Stott P A 2006 Detection of a direct carbon dioxide effect in continental river runoff records *Nature* **439** 835–8

[26] Piao S, Friedlingstein P, Ciais P, de Noblet-Ducoudre N, Labat D and Zaehle S 2007 Changes in climate and land use have a larger direct impact than rising CO_2 on global river runoff trends *Proc. Natl Acad. Sci. USA* **104** 15242–7

[27] Krakauer N Y and Fung I 2008 Mapping and attribution of change in streamflow in the coterminous US *Hydrol. Earth Syst. Sci.* **12** 1111–20

[28] Dai A 2011 Characteristics and trends in various forms of the palmer drought severity index (PDSI) during 1900–2008 *J. Geophys. Res.* **116** D12115

[29] Fekete B M, Vörösmarty C J, Roads J O and Willmott C J 2004 Uncertainties in precipitation and their impacts on runoff estimates *J. Clim.* **17** 294–304

[30] Fekete B M, Looser U, Pietroniro A and Robarts R D 2012 Rationale for monitoring discharge on the ground *J. Hydrometeorol.* **13** 1977–86

[31] Stott P A and Kettleborough J A 2002 Origins and estimates of uncertainty in predictions of twenty-first century temperature rise *Nature* **416** 723–6

[32] Bhend J and von Storch H 2008 Consistency of observed winter precipitation trends in northern Europe with regional climate change projections *Clim. Dyn.* **31** 17–28

[33] van Haren R, van Oldenborgh G J, Lenderink G, Collins M and Hazeleger W 2013 SST and circulation trend biases cause an underestimation of European precipitation trends *Clim. Dyn.* **40** 1–20

[34] Shin S-I and Sardeshmukh P D 2011 Critical influence of the pattern of Tropical Ocean warming on remote climate trends *Clim. Dyn.* **36** 1577–91

[35] Liu C and Allan R P 2013 Observed and simulated precipitation responses in wet and dry regions 1850–2100 *Environ. Res. Lett.* **8** 034002

[36] Lambert F H, Stine A R, Krakauer N Y and Chiang J C H 2008 How much will precipitation increase with global warming? *Eos* **89** 193–4

[37] Wu P, Christidis N and Stott P 2013 Anthropogenic impact on Earth's hydrological cycle *Nature Clim. Change* **3** 807–10

[38] Wilcox L J, Highwood E J and Dunstone N J 2013 The influence of anthropogenic aerosol on multi-decadal variations of historical global climate *Environ. Res. Lett.* **8** 024033

[39] Monteleoni C, Schmidt G A, Saroha S and Asplund E 2011 Tracking climate models *Statist. Anal. Data Min.* **4** 372–92