PERSPECTIVE
https://doi.org/10.1038/s41558-019-0639-x

nature climate change

The potential to reduce uncertainty in regional runoff projections from climate models

Flavio Lehner[1,2,3*], Andrew W. Wood[2], Julie A. Vano[2,4], David M. Lawrence[1], Martyn P. Clark[5] and Justin S. Mankin[6,7,8]

Increasingly, climate change impact assessments rely directly on climate models. Assessments of future water security depend in part on how the land model components in climate models partition precipitation into evapotranspiration and runoff, and on the sensitivity of this partitioning to climate. Runoff sensitivities are not well constrained, with CMIP5 models displaying a large spread for the present day, which projects onto change under warming, creating uncertainty. Here we show that constraining CMIP5 model runoff sensitivities with observed estimates could reduce uncertainty in runoff projection over the western United States by up to 50%. We urge caution in the direct use of climate model runoff for applications and encourage model development to use regional-scale hydrological sensitivity metrics to improve projections for water security assessments.

Despite a robust theoretical understanding of changes in the global hydrological cycle, it has been challenging to reduce the uncertainties in regional-scale hydrological projections[1]. Much of this uncertainty arises from scientific and observational gaps in describing the climate system and representing it in climate and Earth system models (collectively called ESMs here). Whereas long-term terrestrial warming is an expected outcome from increasing greenhouse gas concentrations (Fig. 1a), precipitation changes are more uncertain at regional scales (Fig. 1b) which are critical for determining future water security. For example, in the Columbia River basin of the Pacific Northwest United States, 90% of CMIP5 models project an increase in precipitation in response to warming. For a regional warming of approximately 2 °C, the multi-model mean change (the signal) is +3.5%, but the standard deviation across individual model projections (the noise) is 3%, yielding a low signal-to-noise ratio (S/N, indicated by stippling in Fig. 1b). In parts of the Upper Colorado River basin, CMIP5 models do not even agree on the sign of projected precipitation changes (indicated by hatching in Fig. 1b).

Although future changes in precipitation are important, long-term changes in surface water availability (principally runoff; here runoff is precipitation minus evapotranspiration, ET) are arguably more relevant for future water security assessments. Warming in the western United States has been linked to an increased probability of a decline in runoff due both to increased ET and to earlier snowmelt-driven runoff[2–6]. However, increased ET and increased precipitation might balance each other, especially for low to moderate future warming scenarios[7,8]. As a result, uncertainty in projected runoff is even larger than uncertainty in precipitation in many regions (compare hatching in Fig. 1b and c).

Increased uncertainty in runoff relative to precipitation arises partly from uncertainties in representing the land surface partitioning of precipitation, runoff and ET, a dynamic often characterized by a metric called runoff efficiency—the fraction of precipitation that becomes runoff (Fig. 1d). Runoff efficiency generally shows more widespread projection uncertainty than precipitation or runoff (Fig. 1d). Exceptions are regions such as parts of the southwestern United States and other semi-arid regions where a combination of warming, snowpack declines and future vegetation greening combine to robustly partition precipitation away from runoff towards ET[9]. Even so, many of the regions with robust declines in runoff efficiency have average changes that fall within one standard deviation of the ensemble spread (Fig. 1d). In current streamflow forecasting systems, runoff efficiencies are often assumed to be constant, but several recent studies have identified changes in runoff efficiency as a risk to water resources[10,11]. Narrowing uncertainty in the change in runoff and runoff efficiency is thus crucial.

Uncertainty in hydroclimate projections has persisted over several generations of coupled climate models[12], yet ESMs have progressed in resolution and process-level detail. Since the early 2000s, ESM land modelling schemes have evolved to include detailed hydrological modelling approaches, incorporating more physically motivated parameterizations of processes for snow, vegetation and terrain influences, and towards greater capacity for representing subgrid heterogeneity in hydrologic processes[13–16]. The increasing realism of ESMs coupled with ever finer model resolution has led to the direct use of ESM hydrological fields as a basis for studying water security and climate change impacts[17–20], which is likely to continue with CMIP6.

Yet ESMs have substantial regional biases in quantities such as temperature, precipitation, ET and runoff. These biases stem from errors in simulated atmospheric circulation and microphysics, coarse spatial resolution affecting orographic processes, and local hydrological and ecological processes. To make useful statements in the face of these biases, most ESM-based climate impact studies resort to using relative changes rather than absolute changes projected by ESMs[18,21,22], or—historically more common—to downscaling and bias-correcting ESM climate outputs for indirect use in uncoupled hydrological model simulations[23–25].

Critically, when using projected absolute or relative changes from an ESM, it is assumed—often without explicit testing—that

[1]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. [2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. [3]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland. [4]Aspen Global Change Institute, Basalt, CO, USA. [5]Coldwater Laboratory, University of Saskatchewan, Canmore, Alberta, Canada. [6]Department of Geography, Dartmouth College, Hanover, NH, USA. [7]Department of Earth Sciences, Dartmouth College, Hanover, NH, USA. [8]Division of Ocean and Climate Physics, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA. *e-mail: flehner@ucar.edu

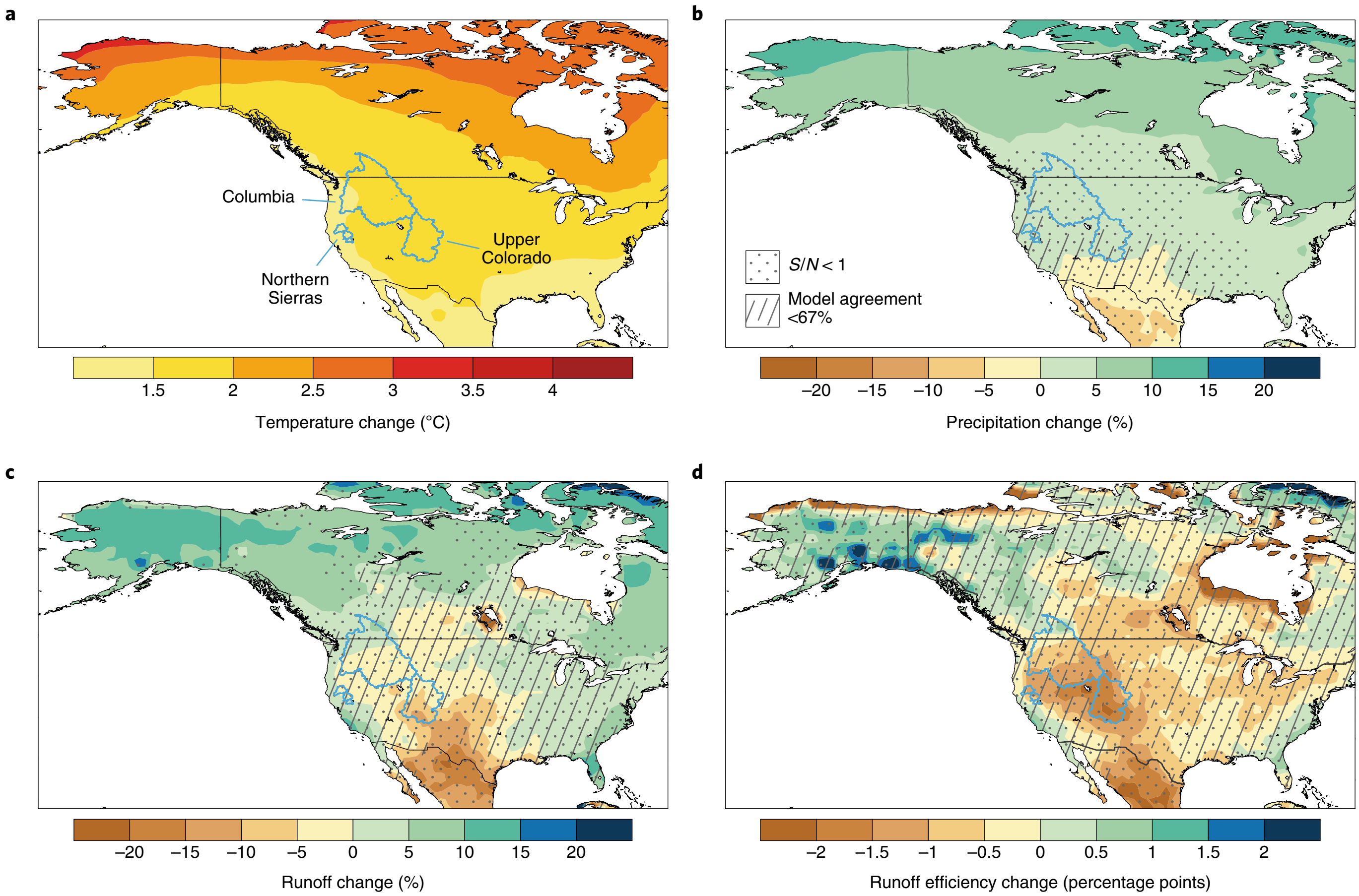

Fig. 1 | Projected changes in temperature, precipitation, runoff and runoff efficiency. **a**, Temperature change; **b**, precipitation change; **c**, runoff change; **d**, runoff efficiency change. The maps show the differences of water-year mean in 2000–2058 relative to 1950–2008 (corresponding to approximately 2 °C warming across the study basins, outlined in blue) from the CMIP5 model projections under a high-emissions scenario (RCP 8.5). Hatching indicates where <67% of the models agree on the sign of change. Stippling indicates where the multi-model mean change signal is less than one standard deviation of the inter-model spread of the 2000–2058 mean (signal-to-noise S/N <1).

a model's sensitivity to climate change is trustworthy despite biases in the mean states and fluxes[26]. In the context of runoff projections, there are at least two key sensitivities for which this assumption is typically made: the precipitation sensitivity of runoff (sometimes termed 'runoff elasticity'; for example, ref. [27]), defined as the per cent change in runoff induced by a 1% change in precipitation ($\Delta Q/\Delta P$); and the temperature sensitivity of runoff, defined as the per cent change in runoff for a 1 °C change in temperature ($\Delta Q/\Delta T$).

Previous studies have quantified regional runoff sensitivities in CMIP models, demonstrating wide inter-model spread in sensitivities but consistency in model behaviour across emissions scenarios and model generations[19,28]. In uncoupled configurations, land models have received ample scrutiny and tuning in an operational context, such as the National Land Data Assimilation System[29,30]. However, calibration of an uncoupled land model to present-day observations does not always guarantee more reliable results in coupled simulations (owing to the strong influence of coupling on simulated climate[31]) or in out-of-sample cases (for example, future climate). Realistic behaviour of physically based fully coupled models would thus offer advantages[32]. To the best of our knowledge, no systematic effort has been made to assess the credibility of the regional runoff sensitivity in coupled ESM simulations. Runoff sensitivities in ESMs are rarely explicitly tuned, offering an opportunity to develop an observational constraint independent of calibration efforts. Such a constraint should be grounded in hydrological theory and would ideally operate similarly on multiple time scales (for example, interannual and decadal), thus adhering to established rules of observational or emergent constraints that rely on the fluctuation–dissipation theorem[33].

In this Perspective, we aim to explore regional runoff sensitivity biases in ESMs, highlight their importance for runoff projections and illustrate the potential for reducing uncertainty in future runoff projections. We advocate for caution when using ESMs for regional runoff projections and for incorporating sensitivity metrics in land surface model evaluation and development.

Our study focuses on three river basins in the western United States, motivated by their contrasting hydroclimate regimes and the availability of high-quality observations (see Methods). The Upper Colorado River and the Northern Sierras, for example, straddle a latitude at which future precipitation changes are highly uncertain, yet the two basins have different vulnerability of the snowpack to warming, owing to their different elevations[34]. In contrast, the Columbia River covers a higher latitude in the northwestern United States, at which there is a stronger climate model consensus that precipitation will increase with warming. Despite the regional focus, we argue that the framework for understanding model biases in runoff sensitivity presented hereafter has implications for and applicability to global modelling efforts.

Model mean and sensitivity biases

Biases of factors of 2–3 are common among state-of-the-art ESMs for water year (October–September) absolute precipitation and runoff in the three focal basins over the historical period of 1950–2008 (Fig. 2a,c,e; see Methods for data and model description). The covariances in relative precipitation and runoff look similar to observations (Fig. 2b,d,f), which might instil confidence in the models' representation of runoff sensitivities. A closer look, however, shows that even in relative space, the models display a wide range of relationships between precipitation and runoff (as measured by the regression coefficient; insets in Fig. 2), some clearly different from the observed relationship. Despite the focus on precipitation and runoff here, it is important to keep in mind that in basins like the Colorado, where ET is approximately 85% of precipitation, small model biases in ET can have large consequences for biases in runoff. Since ET is difficult to observe, it has become common practice to use temperature as a proxy for ET when developing multivariate models for runoff sensitivity.

We characterize runoff sensitivities using multiple linear regression of the form $Q = a\Delta P + b\Delta T + c(\Delta P\Delta T) + e$, where a and b are estimates of the precipitation and temperature sensitivity, respectively, ΔP and ΔT are precipitation and temperature anomalies, $(\Delta P\Delta T)$ is an interaction term and e is the residual (see Methods). We do not present this formulation as an authoritative definition of runoff sensitivities, and we acknowledge that other formulations are valid and informative[35], including theoretical frameworks that provide an energetics perspective and incorporate interannual storage[36], results from sensitivity experiments with offline land surface model simulations[37,38] and factual–counterfactual simulations with atmosphere–land-only models[39]. Because our goal is to demonstrate the bias and spread of sensitivities in coupled ESMs, the choice of sensitivity metric does not alter the study's findings as long as the definition is physically reasonable and applied consistently across models and observations. There are other known uncertainties in estimating runoff sensitivities, such as the regression itself relating to the length of record for which the sensitivity is calculated, the observational datasets used, and internal climate variability. Generally, all of these sources of uncertainty are important, and it is worth noting that formulating a unified and robust method to estimate runoff sensitivities is an active area of research[36,39,40]. Key methodological uncertainties are tested and discussed in Supplementary Section 1.

The three basins show distinct runoff sensitivities in observations (Fig. 3), with the Colorado showing a stronger temperature sensitivity (−11.1%/°C [5–95% confidence interval of regression: −18.1, −4.2]) than the Columbia (−3.1%/°C [−7.5, 1.3]) and the Northern Sierras (−4.3%/°C [−8.4, −0.2]). The precipitation sensitivity is more similar across basins, ranging from 1.1% per 1% change in precipitation [0.9, 1.3] in the Columbia to 1.2%/ % [1.0, 1.5] in the Colorado and 1.3%/% [1.3, 1.4] in the Northern Sierras.

The CMIP5 models exhibit a wide range of runoff sensitivities with regard to both temperature and precipitation (Fig. 3). Although observational estimates of runoff sensitivities contain uncertainty themselves, there are a number of ESMs that display sensitivities that are significantly higher or lower (and outside the uncertainty range) than estimated from observations, leading to the conclusion that systematic biases in runoff sensitivities are indeed present in the CMIP5 model archive.

A key reason to pay attention to a model's runoff sensitivities to temperature and precipitation is the potential for biases to offset one another. An illustrative example is model 9 in the Colorado, which has a positive temperature sensitivity (meaning that runoff increases with temperature, which is counter to theory and observations[37,41]), but also an overestimated precipitation sensitivity (Fig. 3a). Nonetheless, this model produces a runoff change under future climate change that lies close to the middle of the CMIP5 distribution (not shown). Thus, although this model appears to be representative of the CMIP5 multi-model mean—and therefore close to what is often considered our best estimate of the future—it is so for unrealistic reasons.

The case for caring about runoff sensitivity

Together with the local climate, runoff sensitivities are the key parameter that determines surface water availability, which in turn gives rise to the diversity of ecosystems found across river basins today. A central goal of ESM development is to increase accuracy in projections of future changes in regional temperature and precipitation; it is thus critical to develop confidence in the translation of projected climate into projected hydrology. It is instructive to obtain a first-order estimate of the extent to which runoff sensitivity biases may influence ESM projections of changes in hydrology, and consequently, future surface water availability. We tackle this question by comparing the projected CMIP5 runoff with projections constrained by observed runoff sensitivities.

We first investigate whether model-simulated changes in runoff can be explained solely by combining changes in precipitation and temperature with each model's precipitation and temperature sensitivities[42,43]. The premise here is that sensitivities derived from historical interannual variability are similar to the (a priori unknown) sensitivities acting on longer timescales relevant to climate change projections. To this end, each model's 'simulated' change in runoff (ΔQ_{m}), for a future 40-year period that is 2 °C warmer than historical (1950–2008; see Supplementary Table 1), is plotted against a 'predicted' change in runoff constructed as $\widehat{\Delta Q_{\mathrm{m}}} = a_{\mathrm{m}}\Delta P_{\mathrm{m}} + b_{\mathrm{m}}\Delta T$, where a_{m} and b_{m} are each model's historical runoff sensitivities for precipitation and temperature, respectively, ΔP_{m} is each model's future precipitation change in %, and ΔT the temperature change (2 °C; Fig. 4).

In all three regions, the simulated and predicted runoff changes, ΔQ_{m} and $\widehat{\Delta Q_{\mathrm{m}}}$, align well with explained variances of 0.69, 0.77 and 0.9 (all $p < 0.05$) across the 19 CMIP5 models (Fig. 4a,c,e). Thus, historical runoff sensitivities are important first-order drivers of future runoff changes, given a particular future temperature and precipitation. Evidence is accumulating that in certain regions, the vegetation response to both warming and elevated CO_2 can be an important driver of changes in the surface water availability, through its control on the partitioning of precipitation into runoff and ET[44–50]. To account for such a potential effect, we conduct an additional evaluation that expands the regression model for $\widehat{\Delta Q_{\mathrm{m}}}$ to include a term for global CO_2 concentration. We find that the prediction skill improves in two out of the three basins (Fig. 4a,c,e), consistent with the increased fidelity of offline ET calculations with inclusion of CO_2 effects[47,50].

When split into contributions from temperature changes and precipitation changes, it becomes evident that in the Colorado, changes in temperature and precipitation explain almost equal but opposing amounts of the future changes in runoff, as well as similar fractions of variance across models (Fig. 4a; the latter being consistent with CMIP3 simulations in ref. [42]). Being less temperature-sensitive than the Colorado, both the Columbia and Northern Sierras show that precipitation changes are more important for explaining the model spread in simulated runoff changes (Fig. 4c and e).

Future changes in runoff can thus be estimated from model-simulated temperature and precipitation changes combined with observed runoff sensitivities a_{o} and b_{o}, such that $\widehat{\Delta Q_{\mathrm{o}}} = a_{\mathrm{o}}\Delta P_{\mathrm{m}} + b_{\mathrm{o}}\Delta T$ ('observationally constrained'; Fig. 4b). When compared to the simulated changes ΔQ_{m}, they provide an estimate of the potential uncertainty reduction and improved accuracy that might be achievable if model runoff sensitivity biases were alleviated. We do not include an observational constraint on the CO_2 sensitivity here, owing to limited experimental observations of its magnitude[51]. Whereas the simulated and predicted multi-model mean runoff changes in the Colorado are similar (about −5%), the observationally constrained runoff change is −17.2%, owing to higher temperature sensitivity and lower precipitation sensitivity in observations compared with the models (Fig. 3a). Note that both the simulated and predicted runoff changes have wide probability density

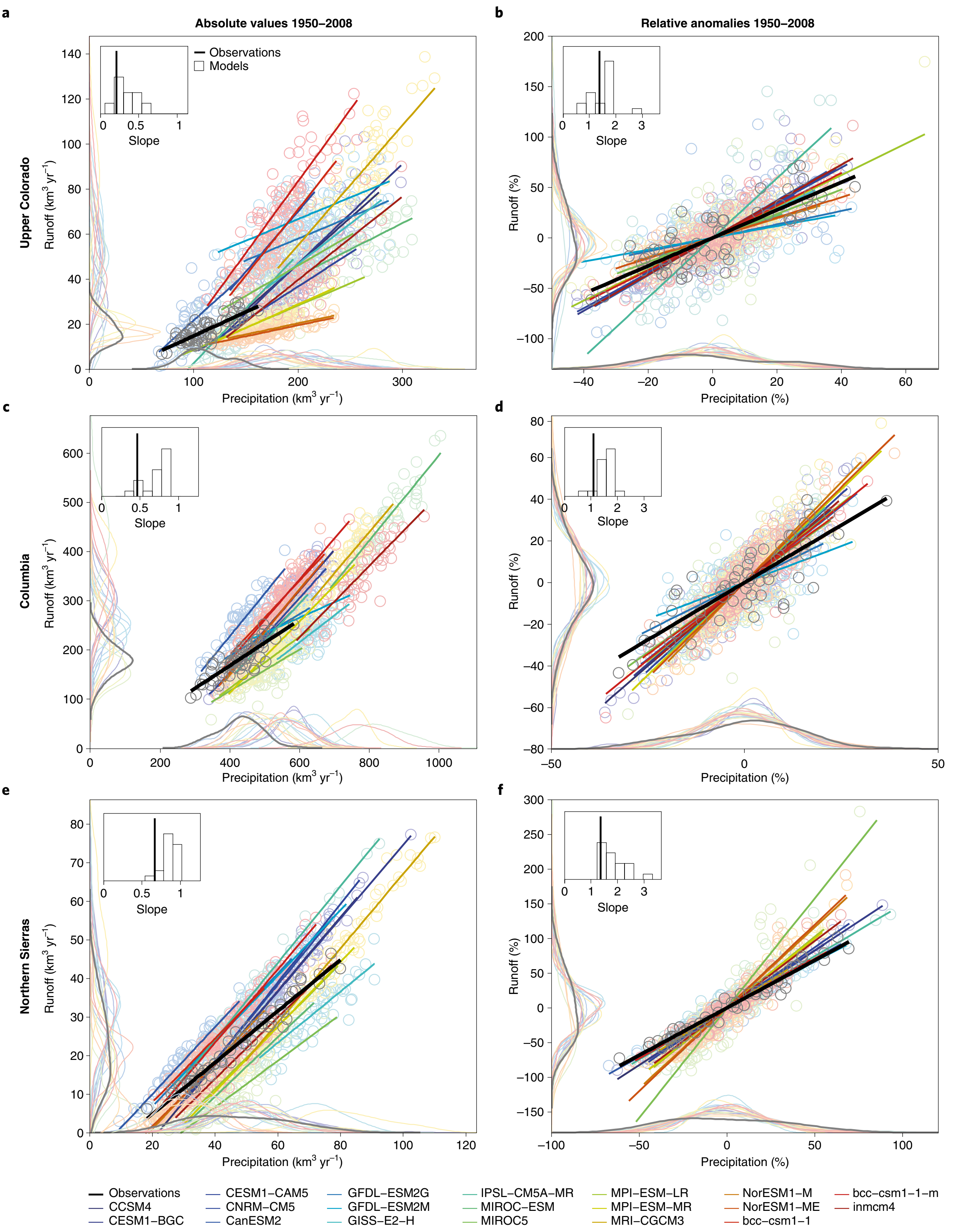

Fig. 2 | Absolute and relative bias in precipitation and runoff in Earth system models. a, Water-year runoff as a function of water-year precipitation for the Upper Colorado from observations (grey/black) and CMIP5 models (colours). **b**, Same as **a**, but values relative to their respective 1950–2008 mean (%). Circles denote individual water years, and solid lines are linear fits to the data. Probability density functions of the water-year data are given on the panel sides; their magnitude is arbitrary, but they are relative to each other. The inset shows a histogram of the linear fit slopes from models; the black line gives the observed value. **c**, **d**, Same as **a**, **b**, but for the Columbia. **e**, **f**, Same as **a**, **b**, but for the Northern Sierras.

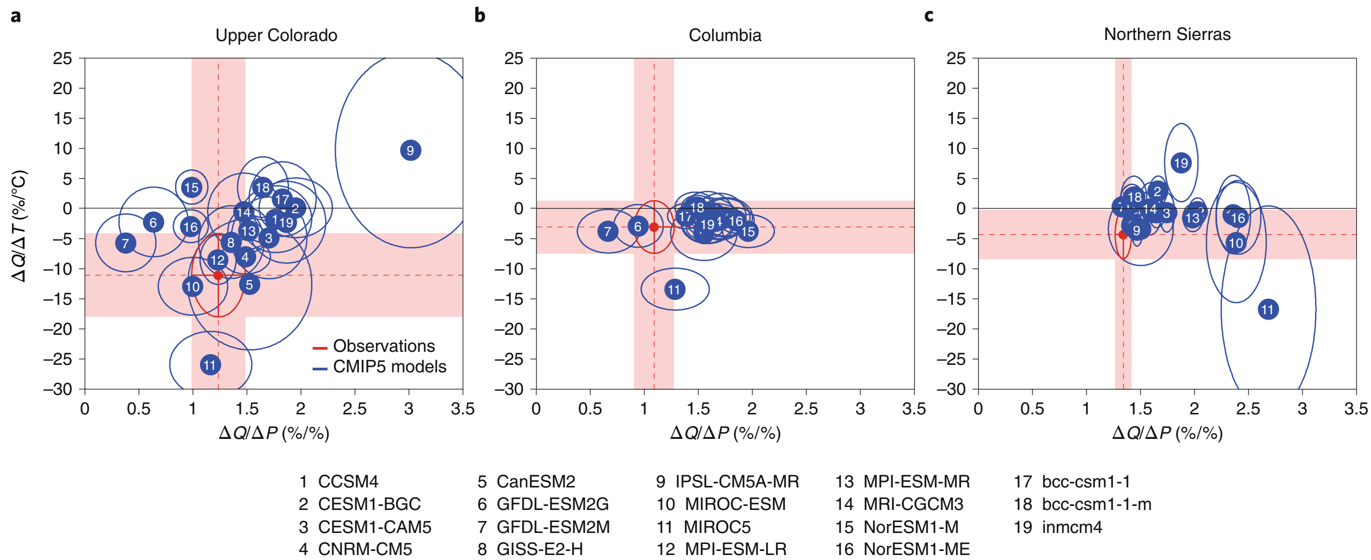

Fig. 3 | Runoff sensitivities. a–c, Precipitation sensitivity ($\Delta Q/\Delta P$) and temperature sensitivity ($\Delta Q/\Delta T$) for 1950–2008 from observations and CMIP5 models for (**a**) the Upper Colorado, (**b**) the Columbia and (**c**) the Northern Sierras. The ellipses for models and pink bars for observations correspond to the 5–95% confidence interval of the sensitivities.

functions (PDFs) that encompass the observationally constrained estimates. However, relative to the full spread (39.6 percentage points), the observationally constrained approach reduces the uncertainty in future runoff projections to 17 percentage points, which is a 57% reduction (Fig. 4b). Including the uncertainty in the observational runoff sensitivities themselves (Methods) still results in a reduction of the ΔQ uncertainty by 30% to 27.8 percentage points. The reduced uncertainty also strengthens the declining signal for Colorado runoff.

In the Columbia, temperature- and precipitation-related changes cancel out such that the simulated and predicted multi-model mean runoff changes are similar (1.4% and −0.1%), while the observationally constrained runoff change is −2.4%. In the Northern Sierras, simulated and predicted runoff changes are −0.6% and −0.4%, while the observationally constrained runoff change is −5.4%. The spread in future runoff projections is reduced by 45% (22%, when observational uncertainty is taken into account) and 53% (47%) in the two basins, respectively, when observed sensitivities (with observed uncertainties) are applied.

Caveats and limitations

The estimates above assume that historical sensitivities would not change substantially with other environmental changes that are expected to accompany climate change, an assumption that appears to hold for these basins and warming levels, but has been suggested to be more generally true[28,42,52] (see also Supplementary Fig. 2 for a global map). However, there are good reasons to expect sensitivities to change with climate. For example, the possible role of vegetation response to CO_2 increases in precipitation partitioning has been mentioned before. For the basins here, sensitivity experiments with models in which the plant physiological response to CO_2 is isolated (with little to no concomitant warming) reveal that under increased CO_2, more precipitation gets partitioned towards runoff because of higher water-use efficiency in plants, although regional differences and model uncertainty are large (Supplementary Section 2 and Supplementary Fig. 3; see also ref. [46]). For the Columbia, these particular model simulations suggest that plant physiological effects will partition precipitation towards runoff, thus indeed leading to notable changes in runoff sensitivities, whereas for the Colorado they suggest no such changes. Scenario-driven simulations (for example, under RCP8.5, thus including increased CO_2 and warming), on the other hand, consistently indicate that plants over the western United States will consume a greater fraction of precipitation in the future than in the present, thereby partitioning precipitation away from runoff and leading to absolute runoff declines[48]. The key point is that model runoff sensitivities are mediated to a non-negligible degree by active areas of model development such as the structural assumptions made about vegetation processes and plant responses to anthropogenic forcing.

As model development continues, it is not immediately clear whether increased complexity in the treatment of land ecosystems will decrease or increase model runoff sensitivities, or affect their stationarity. Having a baseline from which to explore these changes will be valuable. Still, the sensitivity biases explored here are typically larger than the uncertainty in future changes of the sensitivities themselves, attesting to the usefulness of an observational constraint on present-day sensitivities, despite our imperfect stationarity assumption.

Another caveat to linear prediction of ΔQ based on historical sensitivities and future precipitation and temperature changes is the lack of temporal and spatial discrimination. Runoff sensitivities vary by season and sub-basin[43,53], and small-scale spatial patterns of precipitation and temperature changes may alter the annual and basin-wide estimate of sensitivities and runoff changes[54]. Thus extrapolation into different climates is not recommended without robust process-level understanding and physically based modelling[35,55]. Changes to land cover, whether human-induced or natural, also have the potential to change runoff sensitivities[56,57]. Finally, a notable concern is that the simulated climate may be highly unrealistic (for example, precipitation overestimated by a factor of 2–3), necessitating the assumption that sensitivities under such conditions are informative about true sensitivities as well. Given the dependence of runoff efficiency on surface climate[27,58], this assumption should be scrutinized.

Moving forward

A model's runoff sensitivity emerges as a property of the coupled system, since it is not tuned in any CMIP5 ESM development. This raises the question: "Why do models have such different runoff sensitivities?" Strikingly, the answer seems to be unknown. Although it is unlikely that a single culprit can be identified that

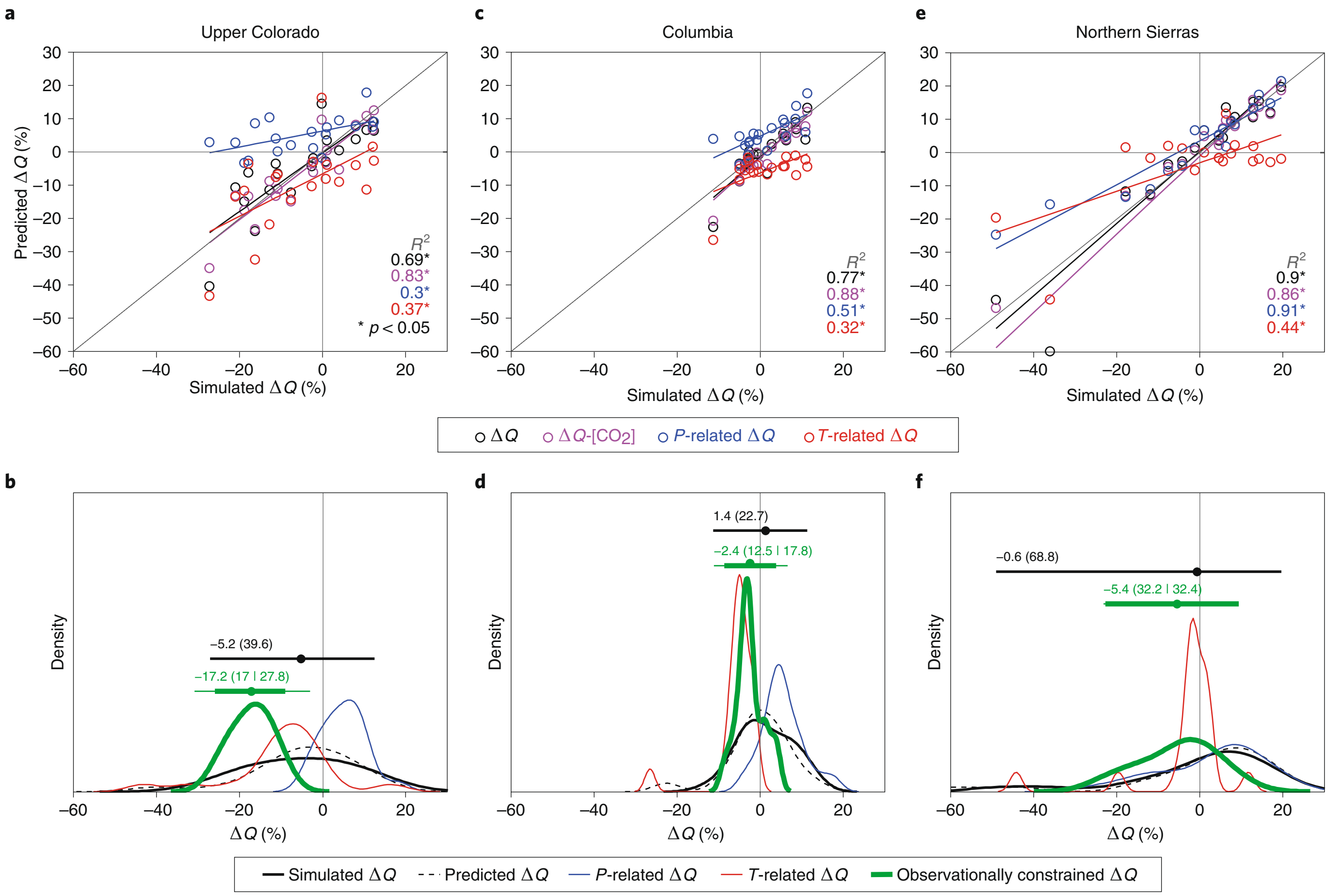

Fig. 4 | Runoff projections and their relationship to runoff sensitivities. a, Model-simulated runoff change (ΔQ) as a function of predicted ΔQ for a future 40-year period that is 2 °C warmer than the reference period, 1950–2008, in the Upper Colorado basin. Predicted ΔQ is based on the statistical model $\widehat{Q_m} = a_m \Delta P_m + b_m \Delta T$, where a_m and b_m are each model's historical runoff sensitivities for precipitation and temperature, respectively, ΔP_m is each model's future precipitation change (%) and ΔT the temperature change of 2 °C. Predicted ΔQ-[CO_2] is based on the same model but with an added term $c_m \Delta CO_{2,m}$, using the time series of historical and RCP 8.5 CO_2 concentration as a predictor. Each circle represents one climate model. Precipitation (P)-related ΔQ assumes no temperature change, and temperature (T)-related ΔQ assumes no precipitation change, when using the statistical model. Explained variances (R^2) are given and are all significant ($p < 0.05$). **b**, Kernel density functions of the simulated and predicted ΔQ, also including an observationally constrained estimate of ΔQ when using the statistical model, but substituting the observed runoff sensitivities a_o and b_o for each model's sensitivities a_m and b_m. The mean and (in brackets) full range of the simulated and observationally constrained ΔQ are given above the black and green bars; the larger green range includes uncertainty in the observational constraint itself (see Methods). **c**, **d**, Same as **a**, **b**, but for the Columbia basin. **e**, **f**, Same as **a**, **b**, but for the Northern Sierras.

causes runoff sensitivity biases across CMIP5 models, some common relationships can help to lead the way to diagnosing bias and eventually to improving models. Across the CMIP5 models, there is a weak correlation between temperature bias and runoff sensitivity to temperature ($r = -0.31$, -0.44^* and -0.55^* for the Colorado, Columbia and Northern Sierras; $^*p < 0.05$)—that is, warmer models have a stronger temperature sensitivity—suggesting that biases in surface radiation balance could be a common cause of biases in temperature and temperature sensitivity[59]. Also, wetter models tend to have a weaker precipitation sensitivity, although this relationship is weak across models and not present in all basins. Recent research[49] identified model biases in canopy light use, interception loss and root water uptake processes as key drivers of bias in surface water partitioning, with implications for the future runoff changes that are associated directly with CO_2. Another recent study found regional relationships between model biases in precipitation (P) and projected changes in P − ET, enabling an observational constraint on the absolute future change in P − ET via observed precipitation[60], similar to another study tackling the same question by means of a constraint directly from observed runoff[61]. Unsurprisingly, these results indicate that biases in model components other than the land surface model might also contribute to biases in runoff sensitivities.

Although not addressed in this study, model interdependency[62] is another potential issue typically not accounted for in most CMIP5 studies which exploit the multi-model ensemble of opportunity[63]; seven out of 19 ESMs used here have fundamentally similar land model components (CLM; see Supplementary Table 1).

Biases in hydrological sensitivities are an important source of uncertainty in future projections of surface water availability when taken directly from ESMs, as shown here by the nearly linear relationship between each model's sensitivities and its future runoff. The observational constraint discussed here, although conceptual in nature, follows guiding principles from the emergent constraint literature[33]. In the basins of the western United States, projected runoff uncertainty could in principle be reduced by up to about 50% if model runoff sensitivity biases were eliminated, which suggests consequent improvements in our ability to project runoff changes. Information on sensitivity biases and their ramifications for runoff projections are typically not included in ESM-based water security or climate change impact assessments, yet would provide valuable perspective for decision makers reliant on climate change information[64].

These results argue strongly for including sensitivity metrics in coupled land model benchmarking and intercomparison projects—for example, the International Land Model Benchmarking (ILAMB) project[65,66]—to support improving the fidelity of ESM land model components and potentially ESM performance more comprehensively[67]. Other examples of such metrics include the 'efficiency space' of joint evaporation and runoff behaviour[68,69], the snow albedo feedback[70,71] or land–atmosphere moisture feedbacks[72]. Together with new benchmarking methods based on information theory[73], such metrics allow one to track biases and improvements in land surface models over time. In all of these efforts, uncertainties in the definition of a particular sensitivity metric arising from different methods, length of record, or datasets need to be considered.

Conclusions

We have highlighted present-day biases in runoff sensitivities in ESMs and their implications for the runoff component of future water security assessments, which are increasingly based directly on ESM output. At present, caution is warranted when using ESMs directly for regional hydroclimate impact and water security assessments. Coordinated model development efforts between the climate and hydrology communities are strongly encouraged, in the expectation that this would increase the accuracy of runoff and its sensitivity in coupled models, thereby increasing their utility for projecting regional surface water availability directly and making observational constraints as presented here obsolete. Focusing on the bias in sensitivity as well as bias in the mean state is critical, and this perspective has illustrated several evaluation metrics to support these efforts that are ready for deployment.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41558-019-0639-x.

Received: 23 January 2019; Accepted: 23 October 2019;
Published online: 26 November 2019

References

1. Collins, M. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. E. et al.) 1029–1136 (IPCC, Cambridge Univ. Press, 2013).
2. Milly, P. C. D., Dunne, K. A. & Vecchia, A. V. Global pattern of trends in streamflow and water availability in a changing climate. *Nature* **438**, 347–350 (2005).
3. Wood, A. W., Lettenmaier, D. P. & Palmer, R. N. Assessing climate change implications for water resources planning. *Clim. Change* **37**, 203–228 (1997).
4. Christensen, N. S., Wood, A. W., Voisin, N., Lettenmaier, D. P. & Palmer, R. N. The effects of climate change on the hydrology and water resources of the Colorado River basin. *Clim. Change* **62**, 337–363 (2004).
5. Barnett, T. P. et al. Human-induced changes in the hydrology of the Western United States. *Science* **319**, 1080–1083 (2008).
6. Mankin, J. S. & Diffenbaugh, N. S. Influence of temperature and precipitation variability on near-term snow trends. *Clim. Dyn.* **45**, 1099–1116 (2015).
7. Nijssen, B., O'Donnell, G. M., Hamlet, A. F. & Lettenmaier, D. P. Hydrologic sensitivity of global rivers to climate change. *Clim. Change* **50**, 143–175 (2001).
8. Lehner, F. et al. Projected drought risk in 1.5 °C and 2 °C warmer climates. *Geophys. Res. Lett.* **44**, 7419–7428 (2017).
9. Mankin, J. S., Smerdon, J. E., Cook, B. I., Williams, A. P. & Seager, R. The curious case of projected twenty-first-century drying but greening in the American West. *J. Clim.* **30**, 8689–8710 (2017).
10. Lehner, F. et al. Mitigating the impacts of climate nonstationarity on seasonal streamflow predictability in the U.S. Southwest. *Geophys. Res. Lett.* **44**, 12208–12217 (2017).
11. Woodhouse, C. A. & Pederson, G. T. Investigating runoff efficiency in Upper Colorado River streamflow over past centuries. *Water Resour. Res.* **54**, 286–300 (2018).
12. Knutti, R. & Sedláček, J. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Change* **3**, 1–5 (2012).
13. Pitman, A. J. The evolution of, and revolution in, land surface schemes designed for climate models. *Int. J. Climatol.* **23**, 479–510 (2003).
14. Lawrence, D. M. et al. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **3**, M03001 (2011).
15. Clark, M. P. et al. Improving the representation of hydrologic processes in Earth system models. *Water Resour. Res.* **51**, 5929–5956 (2015).
16. Clark, M. P. et al. Characterizing uncertainty of the hydrologic impacts of climate change. *Curr. Clim. Change Rep.* **2**, 55–64 (2016).
17. Seager, R. et al. Model projections of an imminent transition to a more arid climate in southwestern North America. *Science* **316**, 1181–1184 (2007).
18. Seager, R. et al. Projections of declining surface-water availability for the southwestern United States. *Nat. Clim. Change* **3**, 482–486 (2013).
Example of an impact assessment based on results from Earth system models expressed in relative space; partly motivated this Perspective.
19. Zhang, X., Tang, Q., Zhang, X. & Lettenmaier, D. P. Runoff sensitivity to global mean temperature change in the CMIP5 models. *Geophys. Res. Lett.* **41**, 5492–5498 (2014).
20. van der Wiel, K. et al. 100-year Lower Mississippi floods in a global climate model: characteristics and future changes. *J. Hydrometeorol.* **19**, 1547–1563 (2018).
21. Mankin, J. S., Viviroli, D., Mekonnen, M., Hoekstra, A. Y. & Horton, R. M. Influence of internal variability on population exposure to hydroclimatic changes. *Environ. Res. Lett.* **12**, 044007 (2017).
22. Kam, J., Knutson, T. R. & Milly, P. C. D. Climate model assessment of changes in winter–spring streamflow timing over North America. *J. Clim.* **31**, 5581–5593 (2018).
23. Schewe, J. et al. Multimodel assessment of water scarcity under climate change. *Proc. Natl Acad. Sci. USA* **111**, 3245–3250 (2014).
24. Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Change* **62**, 189–216 (2004).
25. Pierce, D. W., Cayan, D. R. & Thrasher, B. L. Statistical downscaling using localized constructed analogs (LOCA). *J. Hydrometeorol.* **15**, 2558–2585 (2014).
26. Hall, A. Projecting regional change. *Science* **346**, 1461–1462 (2014).
27. Schaake, J. C. in *Climate Change and US Water Resources* (ed. Waggoner, P. E.) 177–206 (Wiley, 1990).
Classic illustration of the relationship between climate and runoff sensitivity.
28. Tang, Q. & Lettenmaier, D. P. 21st century runoff sensitivities of major global river basins. *Geophys. Res. Lett.* **39**, 1–5 (2012).
29. Mitchell, K. E. et al. The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.* **109**, D07S90 (2004).
30. Xia, Y. et al. Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *J. Geophys. Res. Atmos.* **117**, https://doi.org/10.1029/2011JD016051 (2012).
31. Laguë, M. M., Bonan, G. B. & Swann, A. L. S. Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and un-coupled land–atmosphere system. *J. Clim.* https://doi.org/10.1175/JCLI-D-18-0812.1 (2019).
32. Clark, M. P. et al. The evolution of process-based hydrologic models: Historical challenges and the collective quest for physical realism. *Hydrol. Earth Syst. Sci.* **21**, 3427–3440 (2017).
33. Hall, A., Cox, P., Huntingford, C. & Klein, S. Progressing emergent constraints on future climate change. *Nat. Clim. Change* **9**, 269–278 (2019).
Formulated a framework to assess the robustness of emergent constraints; partly guided the assessment of the observational constraints in this Perspective.
34. Musselman, K. N. et al. Projected increases and shifts in rain-on-snow flood risk over western North America. *Nat. Clim. Change* **8**, 808–812 (2018).
35. Andreassian, V., Coron, L., Lerat, J. & Le Moine, N. Climate elasticity of streamflow revisited — an elasticity index based on long-term hydrometeorological records. *Hydrol. Earth Syst. Sci.* **20**, 4503–4524 (2016).
Detailed investigation of the challenge to constrain runoff sensitivities from observations.
36. Milly, P. C. D., Kam, J. & Dunne, K. A. On the sensitivity of annual streamflow to air temperature. *Water Resour. Res.* https://doi.org/10.1002/2017WR021970 (2018).
Energetics perspective on runoff sensitivities, illustrates limitations of regression-based approaches.
37. Vano, J. A., Das, T. & Lettenmaier, D. P. Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature. *J. Hydrometeorol.* **13**, 932–949 (2012).
38. Vano, J. A. et al. Understanding uncertainties in future Colorado River streamflow. *Bull. Am. Meteorol. Soc.* **95**, 59–78 (2014).

39. Hoerling, M. et al. Causes for the century-long decline in Colorado River flow. *J. Clim.* https://doi.org/10.1175/jcli-d-19-0207.1 (2019).
40. Barsugli, J. J., Hoerling, M. P. & Livneh, B. Is the recent drought on the Colorado River the new normal? *EOS* **100**, https://doi.org/10.1029/2019EO117173 (2019).
41. Woodhouse, C. A., Pederson, G. T., Morino, K., McAfee, S. A. & McCabe, G. J. Increasing influence of air temperature on upper Colorado River streamflow. *Geophys. Res. Lett.* **43**, 2174–2181 (2016).
42. Vano, J. A. & Lettenmaier, D. P. A sensitivity-based approach to evaluating future changes in Colorado River discharge. *Clim. Change* **122**, 621–634 (2014).
43. Vano, J. A., Nijssen, B. & Lettenmaier, D. P. Seasonal hydrologic responses to climate change in the Pacific Northwest. *Water Resour. Res.* **51**, 1959–1976 (2015).
44. Betts, R. A. et al. Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature* **448**, 1037–1041 (2007).
45. Roderick, M. L., Greve, P. & Farquhar, G. D. On the assessment of aridity with changes in atmospheric CO_2. *Water Resour. Res.* https://doi.org/10.1002/2015WR017031 (2015).
46. Swann, A. A. L. S., Hoffman, F. M., Koven, C. D. & Randerson, J. T. Plant responses to increasing CO_2 reduce estimates of climate impacts on drought severity. *Proc. Natl Acad. Sci. USA* **113**, 10019–10024 (2016).
47. Milly, P. C. D. & Dunne, K. A. Potential evapotranspiration and continental drying. *Nat. Clim. Change* **6**, 946–949 (2016).
48. Mankin, J. S. et al. Blue water tradeoffs with vegetation in a CO_2-enriched climate. *Geophys. Res. Lett.* https://doi.org/10.1002/2018GL077051 (2018).
49. Lian, X. et al. Partitioning global land evapotranspiration using CMIP5 models constrained by observations. *Nat. Clim. Change* **8**, 640–646 (2018).
50. Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R. & Donohue, R. J. Hydrologic implications of vegetation response to elevated CO_2 in climate projections. *Nat. Clim. Change* **9**, 44–48 (2019).
51. De Kauwe, M. G. et al. Forest water use and water use efficiency at elevated CO_2: a model-data intercomparison at two contrasting temperate forest FACE sites. *Glob. Change Biol.* **19**, 1759–1779 (2013).
52. Roderick, M. L. & Farquhar, G. D. A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties. *Water Resour. Res.* **47**, 1–11 (2011).
53. Das, T., Pierce, D. W., Cayan, D. R., Vano, J. A. & Lettenmaier, D. P. The importance of warm season warming to western U.S. streamflow changes. *Geophys. Res. Lett*. **38**, https://doi.org/10.1029/2011GL049660 (2011).
54. Xiao, M., Udall, B. & Lettenmaier, D. P. On the causes of declining Colorado River streamflows. *Water Resour. Res.* **2**, 6739–6756 (2018).
55. Hoerling, M., Lettenmaier, D., Cayan, D. & Udall, B. Reconciling projections of Colorado River streamflow. *Southwest Hydrol*. 20–22 (May/June 2009).
56. Inbar, M., Tamir, M. & Wittenberg, L. Runoff and erosion processes after a forest fire in Mount Carmel, a Mediterranean area. *Geomorphology*. https://doi.org/10.1016/S0169-555X(97)00098-6 (1998).
57. Edburg, S. L. et al. Cascading impacts of bark beetle-caused tree mortality on coupled biogeophysical and biogeochemical processes. *Front. Ecol. Environ.* **10**, 416–424 (2012).
58. Lehner, F., Wahl, E. R., Wood, A. W., Blatchford, D. B. & Llewellyn, D. Assessing recent declines in Upper Rio Grande runoff efficiency from a paleoclimate perspective. *Geophys. Res. Lett.* **44**, 4124–4133 (2017).
59. Best, M. J. et al. The plumbing of land surface models: benchmarking model performance. *J. Hydrometeorol.* **16**, 1425–1442 (2015).
60. Padrón, R. S., Gudmundsson, L. & Seneviratne, S. I. Observational constraints reduce likelihood of extreme changes in multidecadal land water availability. *Geophys. Res. Lett.* https://doi.org/10.1029/2018GL080521 (2019). **Example of an observational constraint on model projections of precipitation minus evapotranspiration**.
61. Yang, H. et al. Regional patterns of future runoff changes from Earth system models constrained by observation. *Geophys. Res. Lett.* **44**, 5540–5549 (2017).
62. Knutti, R., Masson, D. & Gettelman, A. Climate model genealogy: generation CMIP5 and how we got there. *Geophys. Res. Lett.* **40**, 1194–1199 (2013).
63. Knutti, R., Baumberger, C. & Hirsch Hadorn, G. in *Computer Simulation Validation* (eds. Beisbart, C. & Saam, N.) 835–855 (Springer, 2019).
64. Vano, J. A. et al. DOs and DON'Ts for using climate change information for water resource planning and management: guidelines for study design. *Clim. Serv.* **12**, 1–13 (2018).
65. Hoffman, F. M. et al. *International Land Model Benchmarking (ILAMB) 2016 Workshop Report*. Technical Report DOE/SC-0186 (2016).
66. Collier, N. et al. The International Land Model Benchmarking (ILAMB) system: design, theory, and implementation. *J. Adv. Model. Earth Syst.* **10**, 2731–2754 (2018).
67. Eyring, V. et al. Taking climate model evaluation to the next level. *Nat. Clim. Change* **9**, 102–110 (2019).
68. Koster, R. D. & P. Mahanama, S. P. Land surface controls on hydroclimatic means and variability. *J. Hydrometeorol.* **13**, 1604–1620 (2012).
69. Koster, R. 'Efficiency space': a framework for evaluating joint evaporation and runoff behavior. *Bull. Am. Meteorol. Soc.* **96**, 393–396 (2015).
70. Hall, A. & Qu, X. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.* **33**, 1–4 (2006).
71. Thackeray, C. W., Qu, X. & Hall, A. Why do models produce spread in snow albedo feedback? *Geophys. Res. Lett.* **45**, 6223–6231 (2018).
72. Levine, P. A., Randerson, J. T., Swenson, S. C. & Lawrence, D. M. Evaluating the strength of the land–atmosphere moisture feedback in Earth system models using satellite observations. *Hydrol. Earth Syst. Sci.* **20**, 4837–4856 (2016).
73. Nearing, G. S., Ruddell, B. L., Clark, M. P., Nijssen, B. & Peters-Lidard, C. Benchmarking and process diagnostics of land models. *J. Hydrometeorol.* **19**, 1835–1852 (2018).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Observations. For observed streamflow (hereafter termed runoff, consistent with the variable used from ESMs), monthly records from 1950 to 2008 are used for the gauges at Lee's Ferry, Arizona, an outlet control point for the Upper Colorado River basin (source: Bureau of Reclamation; go.nature.com/2pcnglb), at The Dalles, Oregon, the outlet control point for the drainage area of the Columbia River (source: Bonneville Power Administration; go.nature.com/2osfnYh), and a four-gauge index in the Northern Sierras (source: California Department of Water Resources; https://cdec.water.ca.gov/), summing the American River at Folsom, the Feather River at Oroville, the Sacramento River at Bend Bridge and the Yuba River near Smartville. The flows are naturalized; that is, major flow alterations due to human diversions and water management are removed empirically from the actually gauged flow. This yields an estimate of the natural flow, making it more interpretable with respect to climate information such as precipitation and temperature, and more directly comparable to ESM output (CMIP5-generation models typically do not represent human diversions or water management). We considered using the Global Runoff Data Centre (GRDC) database but found the lack of flow naturalization in those data to compromise our analysis; however, for analysis of less impaired basins, the GRDC data might be feasible. We use gridded monthly mean precipitation and temperature data from the Livneh dataset[74] as primary observational datasets owing to length of record and coverage of the Columbia basin, with other datasets being tested in Supplementary Section 1. The Livneh forcing dataset is created through the interpolation of ground-based measurements (see refs. [74,75] for more details). Observational datasets are spatially aggregated over the watershed upstream of the respective gauge location (that is, spatially averaged and multiplied by the drainage area). The observational uncertainty is non-zero, but typically the smallest of a list of uncertainty sources (see Supplementary Section 1).

Earth system models. We use the following monthly mean output variables from all ESMs in the CMIP5 archive that provide the experiments 'piControl' (preindustrial control simulation with perpetual 1850 external forcing), '1pctCO2' (branched from piControl, but with an increase in CO_2 concentration of 1% per year—that is, a doubling after 70 years), 'historical' (typically 1850 to 2005, with best estimates of all forcings as per CMIP5 protocol) and 'rcp85' (from 2006 onwards, with a strong increase in greenhouse gas concentrations as per CMIP5 protocol): total runoff ('mrro'), precipitation ('pr'), evapotranspiration (converted from latent heat flux 'hfls'; ET), and surface air temperature ('tas'), conservatively regridded to a common 1° × 1° horizontal resolution. Streamflow for the different basins is calculated as runoff aggregated over the same watersheds as used for observations after mapping the watershed shapefile to 1° × 1° (performing calculations on each ESM's native grid did not alter the results notably), similar to ref. [18]. From all available ESMs, we select the ones for which we were able to balance $P-\mathrm{ET}$ with runoff within 3% in the long-term mean of piControl over the watersheds considered. It is unclear why some models appear to be unbalanced in terms of runoff $= P - \mathrm{ET}$, but we speculate that either additional runoff terms from these models are missing in the CMIP5 calculation of 'mrro' or substantial drift exists in storage terms such as soil moisture or lakes. Application of our criterion leaves 19 CMIP5 models (listed in Supplementary Table 1).

Thus, for the purpose of this study, the term runoff is interchangeable with $P-\mathrm{ET}$. This assumes no significant role for interannual basin storage, an assumption supported by the fact that dynamically routed streamflow and basin-aggregated runoff are typically not significantly different in terms of their long-term mean or their autocorrelation; they also exert almost perfect correlation on interannual timescales (Supplementary Fig. 4). Further, for the basins in this study, simulated interannual variability in land water storage is typically at least an order of magnitude smaller than interannual variability in runoff (Supplementary Table 2), although lag-1 autocorrelation in observed streamflow time series has been argued to be indicative of storage influence[36,76]. Overall, storage variability is unlikely to affect our conclusions significantly, an assessment supported by observations and hydrologic modelling[77], but a comprehensive exploration of this issue is beyond the scope of this paper.

We use the Community Earth System Model Large Ensemble (CESM LE[78]), a single-model 40-member ensemble from 1920 to 2100 under historical and RCP 8.5 forcing, to illustrate the potential uncertainty in a given ESM's sensitivity estimate due to internal climate variability, absent any model differences. A 1,800-year long 'piControl' simulation with the same model is also used.

Finally, we investigate seven ESMs from the Coupled Carbon Cycle Climate Model Intercomparison Project (C4MIP[79]) for which a particular set of three sensitivity experiments is available: (1) plant physiology sees constant preindustrial CO_2 concentrations, as in piControl, but the radiation scheme sees a CO_2 increase of 1% per year, as in 1pctCO2 (CMIP5 technical name 'esmFdbk1'; here called 'CO2rad'), (2) plant physiology sees a CO_2 increase of 1% per year, but the radiation scheme sees preindustrial CO_2 concentrations ('esmFixClim1'; here called 'CO2phys'), (3) plant physiology and radiation scheme both see a CO_2 increase of 1%/year ('1pctCO2'). These simulations are used to assess to what extent plant physiological responses under elevated CO_2 contribute to changes in surface water partitioning in the study basins. The simulations are the same as in ref. [46].

Runoff sensitivity definitions. Runoff sensitivities are difficult to calculate from observations directly[40], being subject to similar challenges as observation-based calculations of transient climate sensitivity: short observational record, system-internal variability, conflation of driving factors, potential nonlinearities and measurement uncertainty (see, for example, ref. [35] for a discussion). Some studies focus on precipitation sensitivity alone or use bivariate regression models to describe joint precipitation and temperature sensitivities[35,36,80–83], others use hydrological or land surface model sensitivity experiments, in which inputs of precipitation and temperature are systematically perturbed to assess that specific model's sensitivity[37].

We characterize runoff sensitivities using multiple linear regression of the form $Q = a\Delta P + b\Delta T + c(\Delta P\Delta T) + e$, where a and b are estimates of the precipitation and temperature sensitivity, respectively, ΔP and ΔT are precipitation and temperature anomalies ($\Delta P = \frac{P-\bar{P}}{\bar{P}}$; $\Delta T = T - \bar{T}$; overbars denote long-term mean), $(\Delta P\Delta T)$ is an interaction term and e is the residual. This approach considers the possibility of interdependency between precipitation and temperature (see, for example, refs. [84,85] for similar concepts), although c was not found to be significant here. The sensitivities are calculated on water year basis for the period 1950–2008 (water year = Oct–Sep; total volumes for precipitation, and mean values for temperature) for observations and the concatenated 'historical' and 'rcp85' simulations. Calculations based on water years aim to minimize the influence of storage carry-over effects; in the western United States in particular, the water year is closely tied to the annual cycle of snow accumulation and melt.

Several methodological uncertainties in estimating runoff sensitivities are explored in Supplementary Section 1. A key uncertainty is length of record, which is generally described by the confidence interval on the regression coefficients in above equation. We propagate this uncertainty to the observational constraint in Fig. 4 by populating a PDF of ΔQ using plausible combinations of runoff sensitivities from observations (in addition to the single value obtained from the observed regression coefficients a_o and b_o; see green PDF in Fig. 4). Specifically, 1,000 combinations of a_o and b_o are randomly chosen from a joint PDF of a_o and b_o, which is based on the 5–95% confidence intervals on a_o and b_o. These 1,000 combinations are used to estimate 1,000 $\widehat{\Delta Q_m}$, representing an observationally constrained ΔQ that includes uncertainty on the observational constraint itself (see wider green range in Fig. 4).

Data availability

All data used in this study are publicly available. The CMIP5 simulations are available through PCMDI, the CESM simulations are available on earthsystemgrid.org, and the observational data are available through the respective institutions. Post-processed data can be obtained from the corresponding author.

Code availability

Code to produce the figures is available from the corresponding author.

References

74. Livneh, B. et al. A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013. *Sci. Data* **2**, 150042 (2015).
75. Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P. & Nijssen, B. A Long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States. *J. Clim.* **15**, 3237–3251 (2002).
76. Milly, P. C. D. & Dunne, K. A. Macroscale water fluxes 2. Water and energy supply control of their interannual variability. *Water Resour. Res.* **38**, 24-1–24-9 (2002).
77. Rosenberg, E. A., Clark, E. A., Steinemann, A. C. & Lettenmaier, D. P. On the contribution of groundwater storage to interannual streamflow anomalies in the Colorado River basin. *Hydrol. Earth Syst. Sci.* **17**, 1475–1491 (2013).
78. Kay, J. E. et al. The Community Earth System Model (CESM) large ensemble project: a community resource for studying climate change in the presence of internal climate variability. *Bull. Am. Meteorol. Soc.* **96**, 1333–1349 (2015).
79. Friedlingstein, P. et al. Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *J. Clim.* **27**, 511–526 (2014).
80. Vogel, R. M., Wilson, I. & Daly, C. Regional regression models of annual streamflow for the United States. *J. Irrig. Drain. Eng.* **125**, 148–157 (1999).
81. Risbey, J. S. & Entekhabi, D. Observed Sacramento basin streamflow response to precipitation and temperature changes and its relevance to climate impact studies. *J. Hydrol.* **184**, 209–223 (1996).
82. Fu, G., Charles, S. P. & Chiew, F. H. S. A two-parameter climate elasticity of streamflow index to assess climate change effects on annual streamflow. *Water Resour. Res.* **43**, 1–12 (2007).
83. Sankarasubramanian, A., Vogel, R. M. & Limbrunner, J. F. Climate elasticity of stream flow in the United States. *Water Resour. Res.* **37**, 1771–1781 (2001).
84. Nowak, K., Hoerling, M., Rajagopalan, B. & Zagona, E. Colorado River basin hydroclimatic variability. *J. Clim.* **25**, 4389–4403 (2012).

85. Harding, B. L., Wood, A. W. & Prairie, J. R. The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River basin. *Hydrol. Earth Syst. Sci.* **16**, 3989–4007 (2012).

Acknowledgements

We thank A. Pendergrass, S. Swenson, E. Wahl, C. Milly, L. Gudmundsson, S. Seneviratne, M. Hoerling, J. Barsugli and N. Addor for discussions, and A. Swann for discussion and for providing the C4MIP simulations. This work benefited from discussions at a 2018 workshop on Colorado River climate sensitivity held at NOAA in Boulder, USA. We acknowledge the efforts of all those who contributed to producing the simulations and observational datasets. The National Center for Atmospheric Research is sponsored by the US National Science Foundation (NSF). F.L. is supported by NSF AGS-0856145, Amendment 87, by the Bureau of Reclamation under Cooperative Agreement R16AC00039, and the Regional and Global Model Analysis (RGMA) component of the Earth and Environmental System Modeling Program of the US Department of Energy's Office of Biological & Environmental Research (BER) via NSF IA 1947282. A.W. is supported by the Bureau of Reclamation (CA R16AC00039), by the US Army Corps of Engineers (CSA 1254557). A.W. and J.A.W. are supported by the NASA Advanced Information Systems Technology program (award ID 80NSSC17K0541). D.M.L is partially supported by NSF INSPIRE grant (NSF-EAR-1528298) and by the RUBISCO Scientific Focus Area (SFA), which is sponsored by the Regional and Global Climate Modeling (RGCM) Program in the Climate and Environmental Sciences Division of the Office of Biological and Environmental Research in the US Department of Energy Office of Science. J.A.V. is supported by grant 80NSSC17K0541 from the NASA AIST program.

Author contributions

F.L. and A.W.W. conceived the study. F.L. conducted all analyses, constructed the figures and led the writing. All authors contributed to the interpretation of the results and the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41558-019-0639-x.

Correspondence should be addressed to F.L.

Peer review information *Nature Climate Change* thanks Qiuhong Tang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.