Supplemental Material for:

Influence of internal variability on population exposure to hydroclimatic changes

Justin S. Mankin\textsuperscript{a,b,c*}, Daniel Viviroli\textsuperscript{d}, Mesfin M. Mekonnen\textsuperscript{e}, Arjen Y. Hoekstra\textsuperscript{e}, Radley M. Horton\textsuperscript{f,b}, Jason E. Smerdon\textsuperscript{a}, and Noah S. Diffenbaugh\textsuperscript{g,h}

\textsuperscript{a} Lamont-Doherty Earth Observatory, Columbia University, New York, USA
\textsuperscript{b} NASA Goddard Institute for Space Studies, New York, USA
\textsuperscript{c} Emmett Interdisciplinary Program in Environment & Resources, Stanford University, California, USA
\textsuperscript{d} Department of Geography, University of Zurich, Zurich, Switzerland
\textsuperscript{e} Dept. of Water Engineering & Management, University of Twente, Enschede, The Netherlands
\textsuperscript{f} Center for Climate Systems Research, Columbia University, New York, USA
\textsuperscript{g} Department of Earth System Science, Stanford University, California, USA
\textsuperscript{h} Woods Institute for the Environment, Stanford University, California, USA

* Corresponding author: Justin S. Mankin
Ocean and Climate Physics
Lamont-Doherty Earth Observatory
Columbia University
61 Route 9W, P.O. Box 1000
Palisades, NY 10964, USA
Email: jsmankin@ldeo.columbia.edu
1 Supplemental Methods

1.1 Basin-scale analysis

We modify the STN-30p hydrological basin demarcations following the methods of ref. 1, aggregating smaller basins, particularly along coastlines, into larger, physically-meaningful basins using the coastal basins provided by ref. 2. We also mask small basins for which the GLDAS baseline data and the CESM LENS are too coarse, giving 676 distinct catchments. This method is also consistent with ref. 3.

The GLDAS-2 baseline data provides a closure of the water budget at the basin and annual scales (Fig. S1d) consistent with theory and provides a good basis on which to estimate historical deficits.

1.2 Projections of net deficits

We assess the CESM LENS monthly and annual basin-scale trends in P and ET against the GLDAS-2 baseline for 1948-2010 (Fig. S1a). We highlight the basins bounded by the range of the ensemble in dark grey, and those not bounded by the ensemble with the absolute magnitude of their bias (mm/day) (Fig. S1b,c). The CESM LENS performs well at characterizing basin-scale hydroclimatic trends over the 20th century, denoted by the small difference in baseline trends, giving us confidence to use CESM basin-scale trend projections to estimate future hydroclimate. Each LENS ensemble member represents a combination of forced change and unforced variability. We estimate monthly-scale trends in P and ET from the CESM LENS from 2011-2080, expressing them as relative changes (% change per 50-years, thus change at ~2060). The differences in trends among the ensemble arise from persistent atmospheric noise.
Following ref. 3, we correct for any minor systematic biases in the LENS baseline simulation (Fig. S1): we first multiply relative monthly trends (%) for each ensemble member by the GLDAS-2 baseline monthly mean value for P and ET (mm). These absolute monthly mean changes are then added to the GLDAS-2 baseline monthly means, providing estimates of future P and ET for each month in each basin for each ensemble member. We then calculate the net deficits (as described in the manuscript Methods section), and difference the future net deficits from the observed net deficits. To highlight the importance of climate variability in driving the range of uncertainty we show the spatial pattern of future change in natural hydroclimatic deficits (not considering any scenario of human water demand). We show the full range of the 30-member ensemble, highlighting the ensemble minimum, mean and maximum changes across the ensemble for each basin (Fig. S3) as well as the ensemble minimum and maximum combination of P and ET, to include their statistical significances (Fig. S5).

1.3 Statistical tests

We test the global difference in natural and net deficits using three statistical tests (one parametric and two nonparametric): (1) a two-tailed Welch’s t-test of equal means (used when the data samples have unequal variances); (2) a two-tailed Mann-Whitney-Wilcoxon test of equal distribution centers (under the assumption that data are not Gaussian); and (3) a bootstrapped two-tailed Kolmogorov-Smirnov test of equal probability distributions. We reject the null hypotheses of equal values using \( \alpha = 0.05 \). Results from these tests are presented in Table 1. We bold those quantities that are significant in two of the three tests.
We perform several different statistical tests at the basin-scale, depending on the quantity presented. All linear time trends have autocorrelation-corrected two-tailed Student’s test rejecting p>0.05, following ref. 4 (Fig. S5). Significance for ensemble-mean trends is presented as the ensemble signal-to-noise (Fig. 2a,b), where significant change is an ensemble-mean response greater than ensemble variability. To test for significant ensemble-mean change in basin net deficits, we employ a one-tailed Student’s t-test, rejecting the null hypothesis if the ensemble-mean change (relative to the GLDAS-2, presented in Fig. 1) is insignificant at the 5% level (Fig. 2e, S3). The basin-scale risks presented in Figs. 2e and S3 serve as a final robustness test. We present the percent of the ensemble agreeing on the sign of change. We present the population exposed to varied levels of risk in net deficit magnitudes and durations in Tables S1 and S2.

2 Supplemental Results and Discussion

2.1 The importance of the LENS characterization of uncertainty

The CESM LENS ensemble allows analyses that seek to identify the range of uncertainty from internal climate variability. This is in contrast to multi-model projections that allow estimates of response uncertainty (namely, how model physics influence model response). In multi-model ensembles, such as the CMIP5 there are not enough simulations with each individual model to effectively sample each model’s representation of internal variability. This can be seen in Fig. S5: the CESM can have statistically-significant trends of opposite signs at basin-scales arising from internal variability in the same forcing pathway and thus a confounding picture of the response from forcing.
Impacts analyses have largely focused on “best-estimate predictions” to aid adaptation decision-making. The requirement of robust decision-making under uncertainty\(^8\) nevertheless requires the identification of adaptation strategies that produce benefits under the broadest range of climate outcomes\(^9\), such as those from internal variability. “Uncertainty means more than one outcome is consistent with expectations”\(^{10}\), and expectations are shaped by what we identify to be the full range of possible outcomes from internal variability (not those outcomes arising from differing parameterization choices in multi-model ensembles). This suggests that poorly-characterized estimates of internal variability from multi-model ensembles could bias expectations of global warming.

Adaptations to hydroclimate changes will likely require regional and local responses encompassing institutional, social, and economic changes\(^4\), and therefore require significant time horizons for effective planning. Because adaptations are best made proactively to offset damages, however, their implementation should necessary occur when internal variability has the potential to mask or reverse long-term trends. This can create the appearance of maladaptation or be truly maladaptive and thus increase vulnerability to climate extremes. Furthermore, such adaptations will need benchmarks for evaluation as part of a strategy of robust decision-making under uncertainty. Many of these benchmarks are a function of the climate system itself, such as when impacts manifest, which are subject to irreducible uncertainty from internal variability.

2.2 **Net deficits**
Our quantification of future net deficits is a simple and transparent method to estimate future change in water availability with large ensembles of climate models and scenarios of population growth. To avoid over-speculation of future pathways in water consumption, we assume that per capita water consumption at the basin-scale remains stationary; that is, population projections are the only driver of water consumption changes. While this increase in population may come with improved water-use efficiency, increasing population combined with changes in diet and water use preferences in the developing world could increase per capita consumption that would in turn increase global-scale net deficits within a given climate pathway\textsuperscript{11}. Second, because our measure of net deficits is based on long-run sub-annual climatological relationships between water supply (P) and water demand (ET-H), it highlights climatological changes to sub-annual variability, and thus deficits. Because variability on shorter or longer timescales can also play an important role in local water stress, the long-run mean projections that we show can potentially encompass large shifts in short-term hydroclimate variability and therefore management responses.

2.3 Projected trends in hydroclimate

The spatial pattern of the CESM LENS ensemble-mean annual basin-scale precipitation response (Fig. 2a) is similar to the CMIP5 ensemble response in the mid-21\textsuperscript{st}-century of the RCP8.5 pathway (Figure 3, Box CC-RC from ref. \textsuperscript{7}; Figure AI.SM8.5.9 from ref. \textsuperscript{12}). The CESM LENS shows consistent 50-year decreases in precipitation along southwestern Australia (~8-12%), South Africa (12-18%), western Patagonia (12-14%), northeastern Brazil and the Amazon basin (2-18%), western Sahara
(2-10%) and the Mediterranean (2-4%), as well as robust mid- and high-latitude increases over Eurasia (Fig. 2a,c). The magnitude and robustness of the decreases are heavily influenced by the LENS expression of internal variability (Fig. 2c): in many of these regions responses are not robust across the ensemble (S/N<1), denoted by the stippling in those basins (Fig. 2a).

Annual evapotranspiration projections show a similar spatial pattern to both the CESM change in precipitation (Fig. 2a) and the late-21st-century CMIP5 ensemble change in evaporation under the RCP8.5 (Figure 12.25 from ref. 6). For example, there are marked, but not robust (S/N<1), warming-induced ET increases over eastern and southeastern Europe by 2060. These eastern and southeastern European increases in ET are collocated with small or negative precipitation trends, which are consistent with increases in the extent of the Hadley Circulation6,13,14.

We present a scatterplot of the maximum and minimum single-realization combination of P and ET change, as well as the ensemble-mean response, and their associated natural deficits (Fig. S5). Multi-decadal hydroclimatic variability creates the potential for statistically significant trends that deviate far from the ensemble mean response. Increases in annual-scale ET that exceed increases in annual-scale P drive increases in net deficits (Fig. S5a), as do decreases in annual-scale P or ET (Fig. S5a. c). The annual-scale changes in P and ET generally indicate a basin’s drying or wetting (i.e., its change in net deficits that are estimated based on monthly changes), as seen by the colors preferentially falling on one side of the line of no change (1:1 line, Fig. S5b), consistent with expectation15. The consistent relationship between annual-scale P and ET trends and sub-annually-calculated changes in net deficits, allows us to present the
annual-scale trends as a proxy for the direction of net deficit change (as presented in Fig. 2 and in the main).
References


3 Supplemental figure captions

Fig. S1 Data validation. a, Comparison of distributions of historical (1948-2010) monthly precipitation (P, blue) and evapotranspiration (ET, red) trends in the GLDAS data (solid lines) and the CESM LENS (dotted lines). b-c, Historical annual trend difference maps for precipitation [b] and evapotranspiration [c]. In [b] and [c] for the basins with GLDAS trends not bounded by the LENS trends, we show the magnitude of the difference between the GLDAS trend and the maximum or minimum LENS trend. Dark grey basins have GLDAS trends that fall within the distribution of LENS trends. d, The annual-scale, basin-scale water balance (P-ET-Q≈0) in the GLDAS data for 1948-2010.

Fig. S2 Global deficits in the present and future climates. a, Empirical cumulative distribution functions for the magnitude of natural deficits and net deficits (mm) in the present and future climates. We show the ensemble mean ECDF for SSP5, though the other SSP distributions stack atop the P-ET-H future distributions. b, As in [a], but for the duration of deficits (months).

Fig. S3 As in Fig. 2e, the risk of net deficit increases for all three SSP scenario in magnitude ([a]-[c], left column) and duration ([d]-[f], right column). The magnitudes are calculated as a percent of annual precipitation. We apply a threshold for inclusion for each run in each basin: the absolute magnitude of net deficit change must be greater than present-day water storage infrastructure in the GRanD database. Hatched basins have insignificant ensemble mean changes (p>0.05) based on a one-sample Student’s t-test.
**Fig. S4** Full range of projected changes in magnitude [a-d] and duration [d-f] of natural deficits (P-ET<0), not including human demand. These figures reflect forced climate and climate variability only.  

**a,e,** Ensemble composite of the basin-maximum change in magnitude (mm) and duration (months) of net deficit projected for each basin as a difference from the 1948-2010 baseline.  

**b,f,** As in [a,d], but for the basin-mean change in net deficit. Stippling in [b,f] indicates insignificant mean change (p>0.05) based on a single-sample Student’s t-test.  

**c,g,** As in [a,d], but for the basin-minimum change.  

**d,h,** The difference between [a,c] and [e,g], showing the full range of possible change within the LENS. Note that [a,e] and [e,g] are composites from the ensemble and not a single realization.

**Fig. S5** Basin scale internal variability in P and ET trends. We show annual mean trends in precipitation and evapotranspiration estimated from 2011-2080 from the CESM LENS, for:  

**a,** each basin’s maximum single-run combination of 50-year precipitation and evapotranspiration trends,  

**b,** each basin’s ensemble mean combination, and  

**c,** each basin’s single-run minimum combination. Colors are the projected change in the magnitudes of natural deficit for that P and ET combination (ensemble member) and basin (% of annual precipitation). Large open circles in [a] and [c] indicate basins where both the P and ET trends are statistically significant (p<0.05), whereas the large open circles in [b] are where the ensemble mean exceeds one standard deviation of the ensemble variability. Small circles are for trends that are not significant. We annotate [b]
with the classification of Greve et al. 2014, highlighting that general relationship of values falling on either side of the 1:1 line (no change), show expected changes in integrated deficits and thus are physically meaningful measures of wetting and drying. Natural (and net) deficits are calculated off of monthly trends (rather than the annual trends presented here), highlighting that annual-scale changes in P and ET can inform the direction of change in deficits. Finally, statistically-significant trends far from the ensemble mean is possible, denoted by the many large circles in [a] and [c].

Table captions

**Table S1** Probability (as percent of simulations, with corresponding population, and basins exposed) of any increase in net deficit magnitude greater than present-day water storage infrastructure in the GRanD database.

**Table S2** Probability (as percent of simulations, with corresponding population percentages, and basins exposed) of any increase in net deficit durations, but whose magnitude is greater than present-day water storage infrastructure in the GRanD database.
<table>
<thead>
<tr>
<th>PROBABILITY</th>
<th>SSP2 population (# of basins)</th>
<th>SSP3 population (# of basins)</th>
<th>SSP5 population (# of basins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;3%</td>
<td>87 (266)</td>
<td>87 (268)</td>
<td>87 (266)</td>
</tr>
<tr>
<td>45-55%</td>
<td>12 (45)</td>
<td>22 (43)</td>
<td>7 (38)</td>
</tr>
<tr>
<td>&gt;66%</td>
<td>21 (78)</td>
<td>20 (82)</td>
<td>22 (81)</td>
</tr>
<tr>
<td>&lt;66%</td>
<td>79 (598)</td>
<td>80 (594)</td>
<td>78 (595)</td>
</tr>
<tr>
<td>&gt;75%</td>
<td>17 (56)</td>
<td>17 (55)</td>
<td>17 (54)</td>
</tr>
<tr>
<td>&lt;25%</td>
<td>32 (462)</td>
<td>33 (462)</td>
<td>44 (467)</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>31 (125)</td>
<td>35 (131)</td>
<td>31 (122)</td>
</tr>
<tr>
<td>&lt;50%</td>
<td>66 (535)</td>
<td>61 (529)</td>
<td>68 (541)</td>
</tr>
<tr>
<td>&gt;90%</td>
<td>10 (16)</td>
<td>10 (18)</td>
<td>7 (18)</td>
</tr>
<tr>
<td>100%</td>
<td>&lt;1 (2)</td>
<td>5 (3)</td>
<td>1 (3)</td>
</tr>
</tbody>
</table>

*Table S1* Probability (as percent of simulations, with corresponding population, and basins exposed) of any increase in net deficit magnitude greater than present-day water storage infrastructure in the GRanD database.
<table>
<thead>
<tr>
<th>PROBABILITY</th>
<th>SSP2 % of global population (# of basins)</th>
<th>SSP3 % of global population (# of basins)</th>
<th>SSP5 % of global population (# of basins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;3%</td>
<td>83 (265)</td>
<td>84 (267)</td>
<td>84 (265)</td>
</tr>
<tr>
<td>45-55%</td>
<td>10 (41)</td>
<td>13 (39)</td>
<td>9 (33)</td>
</tr>
<tr>
<td>&gt;66%</td>
<td>20 (70)</td>
<td>21 (71)</td>
<td>19 (33)</td>
</tr>
<tr>
<td>&lt;66%</td>
<td>79 (606)</td>
<td>78 (605)</td>
<td>79 (610)</td>
</tr>
<tr>
<td>&gt;75%</td>
<td>13 (44)</td>
<td>17 (45)</td>
<td>16 (47)</td>
</tr>
<tr>
<td>&lt;25%</td>
<td>45 (468)</td>
<td>32 (466)</td>
<td>44 (469)</td>
</tr>
<tr>
<td>&gt;50%</td>
<td>29 (118)</td>
<td>31 (121)</td>
<td>28 (112)</td>
</tr>
<tr>
<td>&lt;50%</td>
<td>68 (550)</td>
<td>62 (543)</td>
<td>69 (554)</td>
</tr>
<tr>
<td>&gt;90%</td>
<td>9 (13)</td>
<td>8 (12)</td>
<td>5 (13)</td>
</tr>
<tr>
<td>100%</td>
<td>&lt;1 (1)</td>
<td>5 (2)</td>
<td>1 (2)</td>
</tr>
</tbody>
</table>

*Table S2* Probability (as percent of simulations, with corresponding population percentages, and basins exposed) of any increase in net deficit durations, but whose magnitude is greater than present-day water storage infrastructure in the GRanD database.
CESM v GLDAS
PRECIPITATION

ET GLDAS
P GLDAS
P LENS
ET LENS

P - ET - Q
GLDAS WATER BALANCE

PDF (%)

mm/day

mm/day

mm/day
Change in natural deficit magnitude

Maximum trend
Mean trend
Minimum trend

ET [% 50 yrs⁻¹]
Precipitation [% 50 yrs⁻¹]

-120 -80 0 80 120
-120 -80 0 80 120
-120 -80 0 80 120

a  
Maximum trend

b  
Mean trend

Dry
Wet

Significant
Not significant

Change in natural deficit magnitude