Supplementary Information

Estimating the impacts of climate change: reconciling disconnects between physical climate and statistical models

Supplemental Methods

Data

For Figure 1 we use data from the Global Historical Climatology Network database (Menne et al., 2012). We download daily data and aggregate to monthly intervals, requiring that at least 90% of days for that month are available. For the map in Fig S2a, we use stations with either average or maximum daily temperature available, because many stations only report one or the other. We limit correlation calculations to a maximum of 150 randomly sampled station pairs for computational efficiency and presentation.

For surface air temperature and monthly precipitation, we used the monthly University of Delaware (UDEL) datasets, which have a 0.5 degree resolution. We take the annual mean for T and the annual sum for P. For our uncertainty estimate in gridded T and P, we used data from the NOAA/CIRES/DOE Twentieth Century Reanalysis (V3). Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Earth System Research Laboratory Physical Sciences Laboratory. These data were provided by the NOAA PSL website. Reanalysis data was interpolated to half degree resolution to match the other datasets. To quantify uncertainty, we use the model ensemble spread since model standard deviations are not provided directly. We approximate the standard deviation by dividing the spread by 4, assuming that the spread is approximately equal to $\pm 2\sigma$. Annual data are not available, so we combine monthly spread into an annual spread by taking the square root of the mean squared monthly standard deviation. We then divide by the annual grid cell standard deviation of the mean over time to obtain the uncertainty metric. We use data from 1980 to 2015 as this is a common time-frame for empirical climate impact studies. This metric is then applied to the UDEL data my multiplying the metric by UDEL grid cell standard deviation. The UDEL standard deviation and the spread metric are both shown in Figure S3. Climate data for projections to 2050 were taken from the r1i1p1f1 member of CESM2-WACCM SSP245 simulation. We take the mean from 2045 to 2055 as the average 2050 condition to smooth internal variability.

We use the purchasing power parity adjusted gross domestic product (GDP PPP) data from Kummu et al. (Kummu et al., 2018), which is allocated to 5 arc minute grid cells from administrative 1 and country data, depending on data availability. For our pixel-level analysis we aggregate to 0.5 degrees by summing grid 6x6 grid cells to the UDEL grid. As is done for other variables in the same dataset, we then normalize by population using the HYDE 3.2 dataset (Klein Goldewijk et al., 2017), which was aggregated to the pixel level in the same way as the GDP dataset. All data (including T and P) were aggregated to the administrative 1 and country levels using a rasterized map created from shapefiles from the Global Administrative Areas dataset (GADM 3.6) using the Python package geocube. When spatially aggregating we applied population weighting.

Idealized Model Framework

In our idealized model, we generate an outcome variable Y_{it} that has a known quadratic relationship with T and P at the grid cell level. We then add different types of measurement errors to T and P, and then estimate the fit parameters and predicted outcomes with the simulated imperfect simulated measurements (see Methods).We implement several slightly different versions of this regression model but in all cases we create the outcome variable as an exact function of T and P and then add error to one or both T and P, estimate the coefficients, and compare them to the known relationships:

$$y_{it} = \beta_T T_{it} + \beta_{T^2} T_{it}^2 + \beta_P P_{it} + \beta_{P^2} P_{it}^2 + \epsilon_{it} \tag{8}$$

$$\hat{y}_{it} = \hat{\beta}_T (T_{it} + \delta_{T,i,t}) + \hat{\beta}_{T^2} (T_{it} + \delta_{T,i,t})^2 + \hat{\beta}_P (P_{it} + \delta_{P,i,t}) + \hat{\beta}_{P_{it}}^2 (P_{it} + \delta_{P,i,t})^2 \tag{9}$$

δ is applied as a function of the standard deviation of the annual mean for T and annual sum for P. This means that if, for example, a standard deviation of error applied to a grid cell, for each year in the dataset we draw randomly from a standard normal distribution and multiply the result by the grid-cell level standard deviation. The linear case in Figure 4 does not include the quadratic terms in Eqn. 8 or Eqn. 9 but is otherwise the same. In some cases we select T and P from a random normal distribution and vary $cor(T, P)$. In the cases where we use the UDEL dataset as the underlying T and P, the correlational structure of that dataset is retained.

GDP regressions

The GDP regressions for analysis prior to the mean state section are quite similar to the idealized regression framework. Regressions in the mean state section are described in the main text. To estimate the impact of T and P on GDP, we use the equation:

$$log(GDPpc_{it}) = \beta_T(T_{it}) + \beta_{T^2}(T_{it})^2 + \beta_P(P_{it}) + \beta^2_{P_{it}}(P_{it})^2 + \mu_i + \nu_t + \epsilon_{it} \tag{10}$$

where GDPpc is GDP per capita and μ_i and ν_t are spatial and time fixed effects. This closely follows the design of Damania et al. (2020). To project future GDP damages, we then use the estimated coefficients to calculate:

$$log(GDPpc_{it}) = \hat{\beta}_T T_{it} + \hat{\beta}_{T^2} T^2_{it} + \hat{\beta}_P P_{it} + \hat{\beta}_{P^2} P^2_{it} \tag{11}$$

$$f_i = exp(logGDPpc_{i,fut}) - exp(logGDPpc_{i,hist}) \tag{12}$$

where f_i is the fractional damage to per capita GDP at each location, $T_{i,hist}$ and $P_{i,hist}$ are the temporal averages of T and P from 1990 to 2015 (the estimation period in the full panel) and $T_{i,fut}$ and $P_{i,fut}$ are the temporal averages of T and P from 2045 to 2055. We use these averages to diminish the influence of interannual variability for average projections. Regressions are population weighted. For the aggregated regressions (i.e. Admin Level 1, country), T and P are population-weighted when aggregating. Future (post-2015) gridded population distributions for population-weighting are from Wang and Sun (2022).

Projections in the mean state section follow (Burke et al., 2015b) where GDPpc change is estimated and added to the baseline SSP2 GDP change. This is iterated forward. As in the first row of Figure 8, we demean the predicted values.

Iterative correction

We invert the idealized framework for the correction applied in Figure 6. To do this, we

1. estimate pixel-level influence of T and P on GDPpc using UDEL T and P
2. generate an outcome variable as an exact function of UDEL T and P
3. add random normal errors to T and P in accordance with the CIRES ensemble spread
4. estimate the coefficients using the now erroneous T and P
5. slightly adjust the coefficients used to generate the outcome variable in Step 2, and repeat until the estimated coefficients match those in Step 1

This yields a new set of corrected coefficients, assuming the distribution of errors we estimate from CIRES model output.

Table S1. Summary of regression results for GDP per capita growth (gdppcg) using four different models

Dependent Variable:	gdppcg	gdppcg	gdppcg	gdppcg
Spatial unit:	country	country	country	country
Model:	(1)	(2)	(3)	(4)
T	0.0153***	0.0069		0.0115**
	(0.0059)	(0.0048)		(0.0051)
T2	-0.0005***	-0.0097***		-0.0004***
	(0.0001)	(0.0031)		(0.0001)
P	5.677e-05	0.0001		-3.577e-05
	(0.0001)	(0.0002)		(0.0001)
P2	-1.097e-07	-1.167e-06		-3.986e-08
	(2.808e-07)	(1.079e-06)		(2.888e-07)
T:Tmean		0.0195***		
		(0.0066)		
T2:Tmean		-1.346e-05		
		(1.715e-05)		
P:Pmean		9.541e-07		
		(1.821e-06)		
P2:Pmean		1.78e-09		
		(3.89e-09)		
Tanom			0.0117**	
			(0.0050)	
Tanom2			-0.0073*	-0.0096***
			(0.0038)	(0.0032)
Panom			7.188e-05	
			(0.0001)	
Panom2			-2.568e-06	-7.076e-07
			(2.342e-06)	(7.408e-07)
Tanom:Tmean			-0.0008***	
			(0.0003)	
Tanom2:Tmean			-0.0002	
			(0.0003)	
Panom:Pmean			-2.444e-07	
			(6.218e-07)	
Panom2:Pmean			8.858e-09	
			(9.328e-09)	
Fixed-effects				
country	Yes	Yes	Yes	Yes
country*time	Yes	Yes	Yes	Yes
Fit statistics				
Observations	5184	5184	5184	5184
R^2	0.1153	0.1206	0.1208	0.1205

Standard errors in parentheses. Clustered by country.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

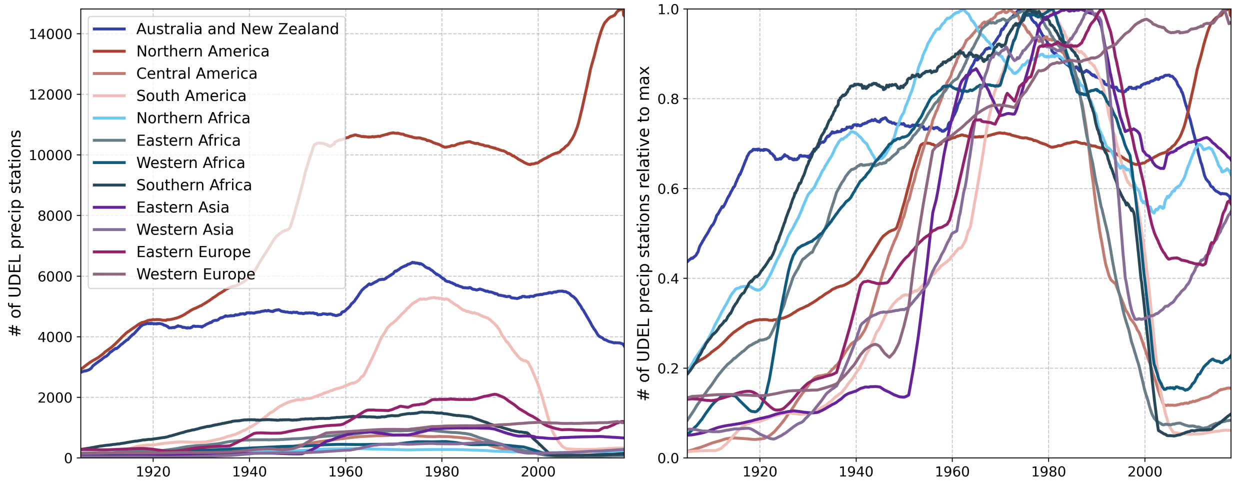

Figure S1: Time series of the number of stations with available data used for the University of Delaware (UDEL) gridded precipitation product in various regions of the world, with and without normalization to the maximum value of each region.

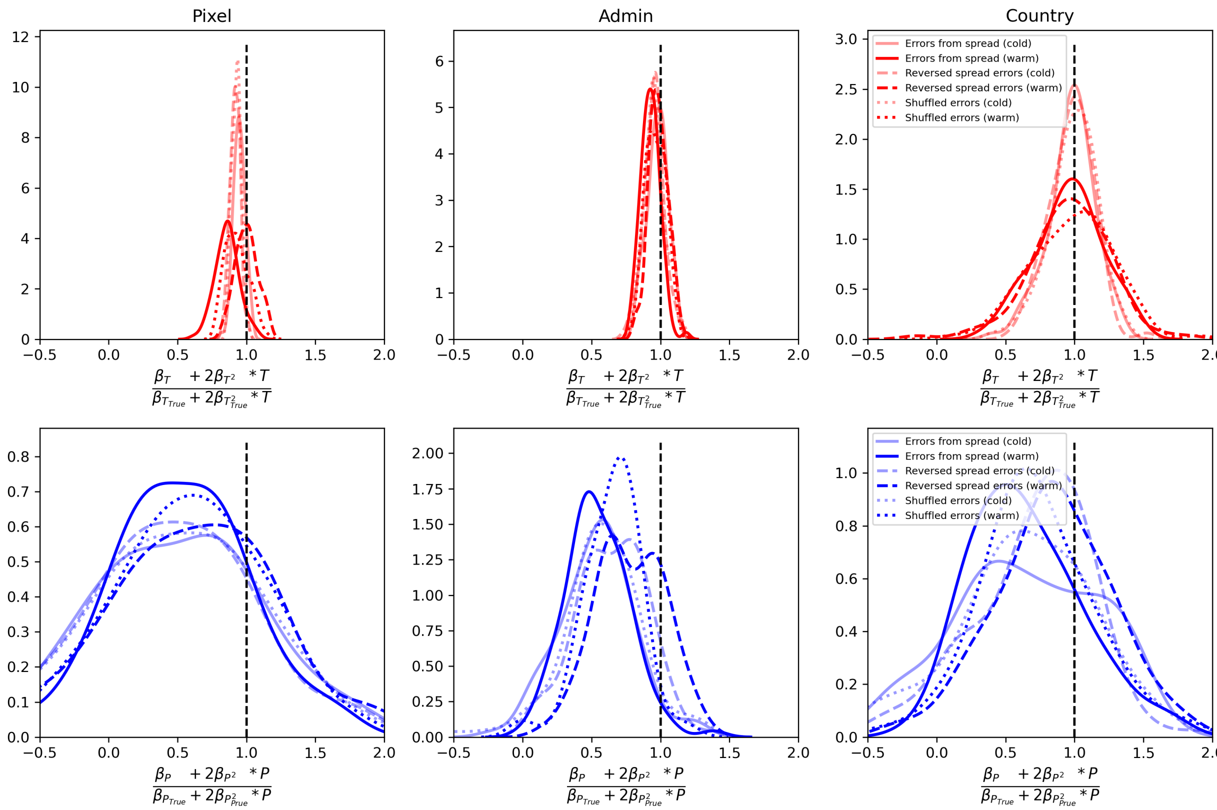

Figure S2: Bias in marginal impacts of T (top row, red) and P (bottom row, blue). Light colors are for cooler climates (normalized T=0.2) and dark colors are for warmer climates (normalized T=0.8). Each plot shows the ratio of estimated to true quadratic coefficients. The first column is the same as in Figure 2. The other two columns are the same, using administrative and country aggregation instead of pixel-level analysis. Errors from spread are defined using Figure 2. Swapped errors are the same except the location of high and low errors have been reversed. Shuffled errors have randomized locations of the same errors. The columns contain the same information for three levels of spatial aggregation. Random errors are added at the pixel level, prior to aggregation. Distributions of bias are narrowest for P at the administrative level, probably because there is a balance in number of observations and averaging to reduce measurement error.

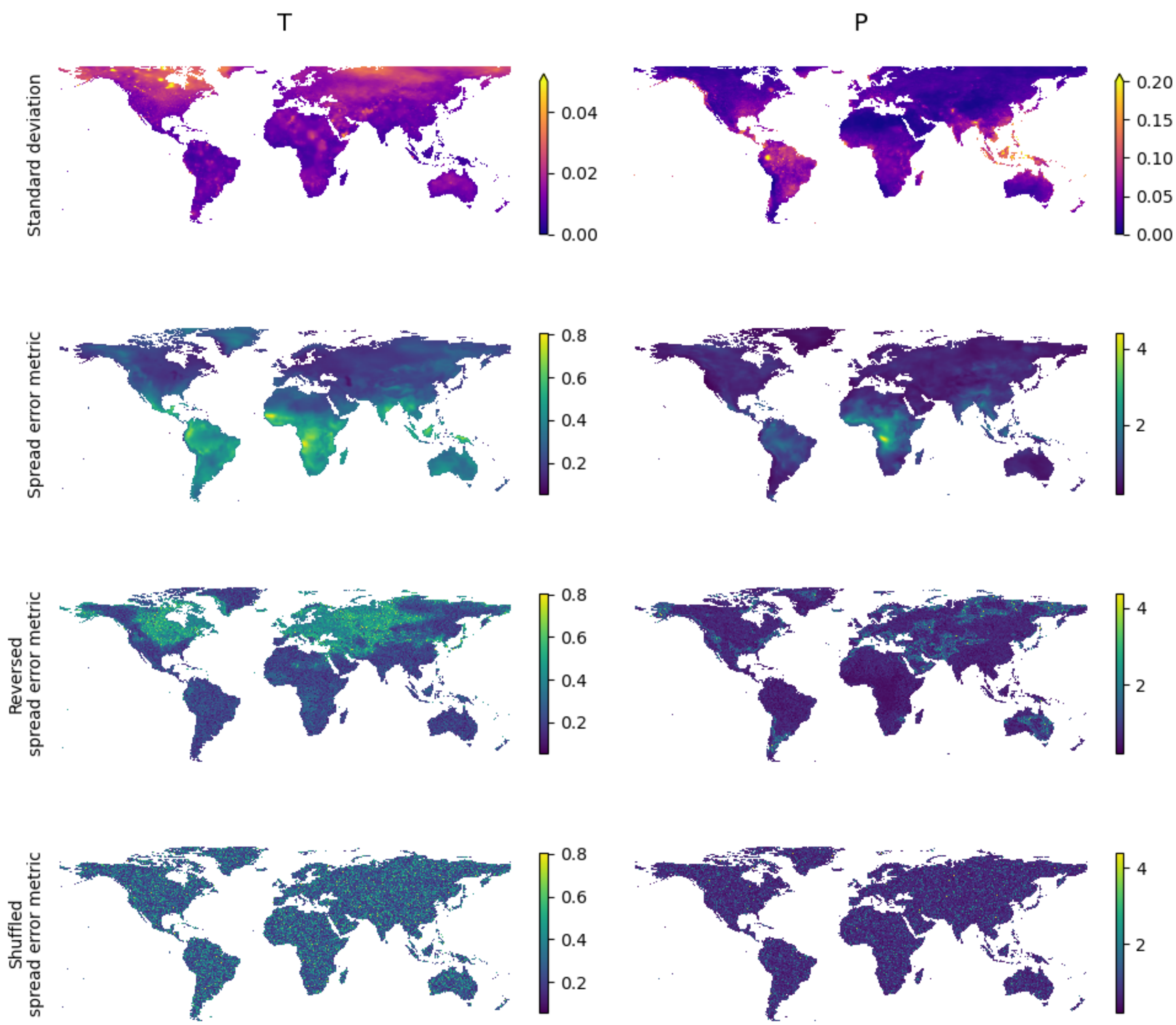

Figure S3: Maps of the standard deviations of T and P from the UDEL dataset and of spread metrics from the NOAA-CIRES-DOE Twentieth Century Reanalysis Project. The spread metric was calculated as the reanalysis spread (i.e. range), divided by four to get an approximate standard deviation of the uncertainty, divided by the model interannual standard deviation. This represents the uncertainty relative to local interannual variability. The reversed spread metric swaps the location of the highest and lowest uncertainties and the shuffled spread metric randomizes the spatial distribution globally. These metrics are multiplied by the standard UDEL grid cell interannual standard deviation (top row) to add regional uncertainties to the idealized model simulation. The Spread error metric (second row) contains the same information as Figure 2

.

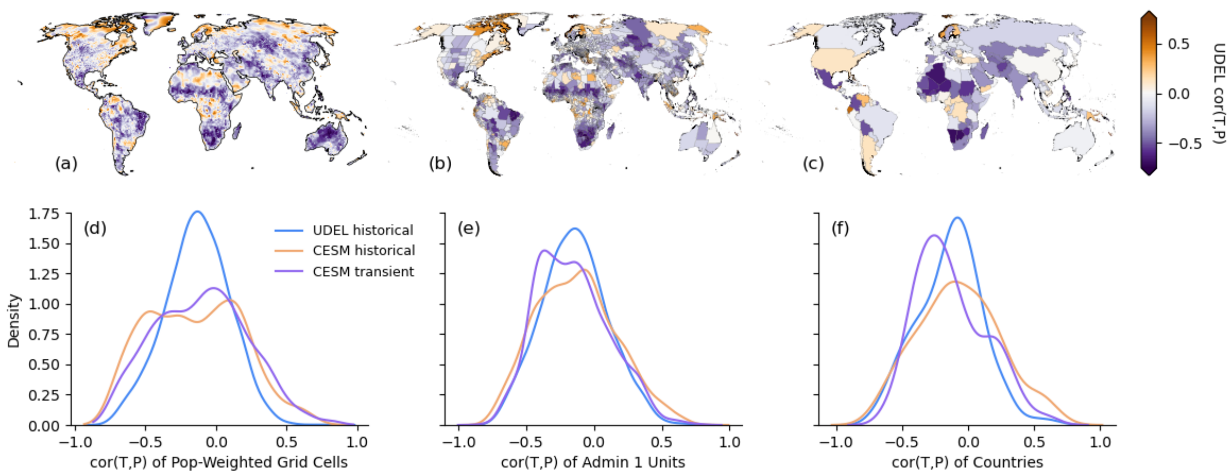

Figure S4: The correlation between annual average temperature and total annual precipitation at different spatial scales of aggregation. (a-c) show maps of UDEL historical $cor(T, P)$ from 1960 to 1990 using population-weighting for aggregation to administrative 1 and country levels. (d-f) show histograms of the same maps as well as the equivalents for a CESM2 historical simulation and a transient warming scenario (SSP2) from 2040 to 2070.

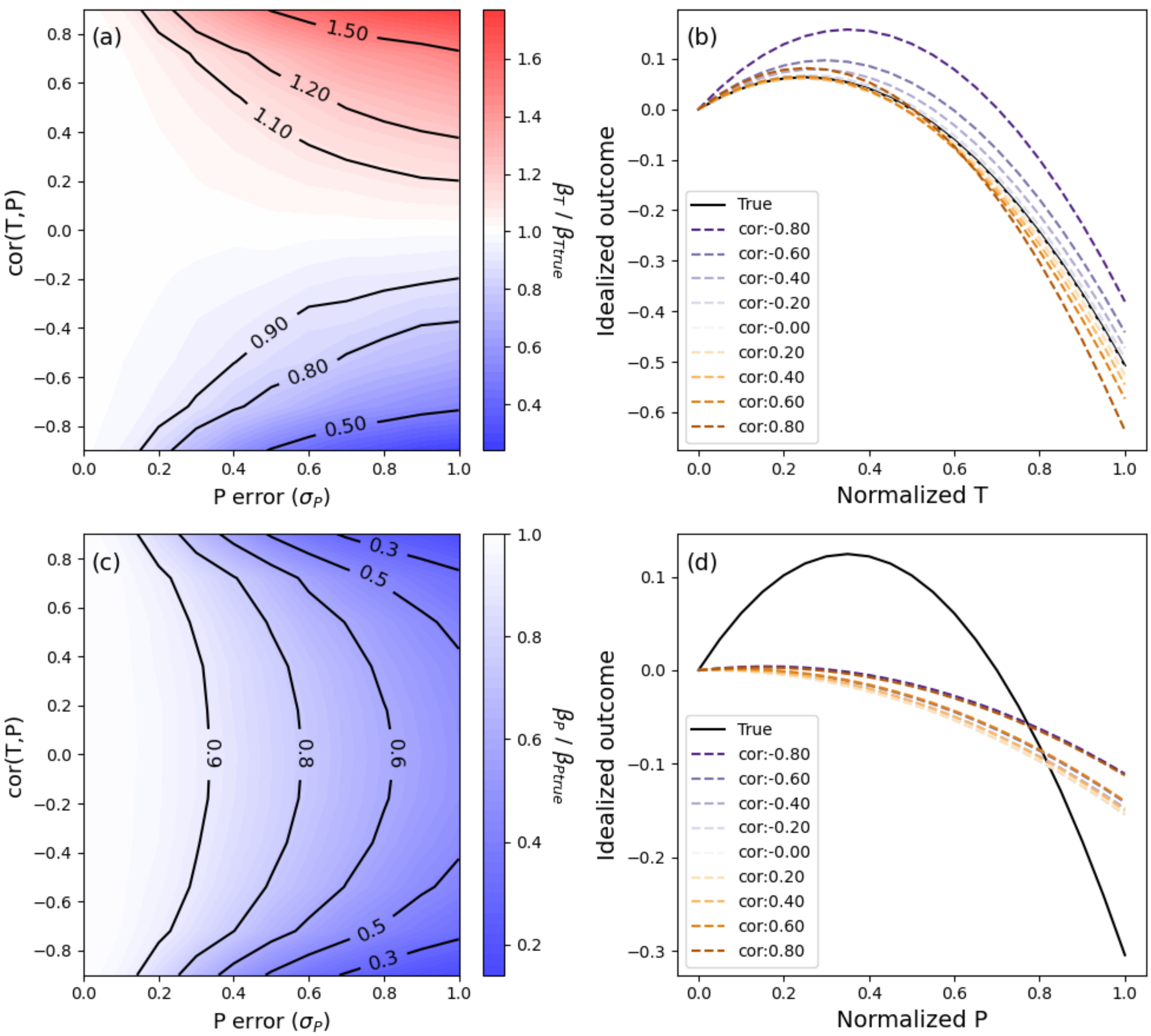

Figure S5: Bias in regression results resulting from idealized measurement error and $cor(T, P)$. A linear formulation of impacts is shown on the left and a quadratic formulation on the right. The top row shows temperature bias, the bottom row shows precipitation bias. The left column is the same as in Figure 4. As in Figure 4, T and P contribute approximately equally to impacts and random normal error is only added to P.

.

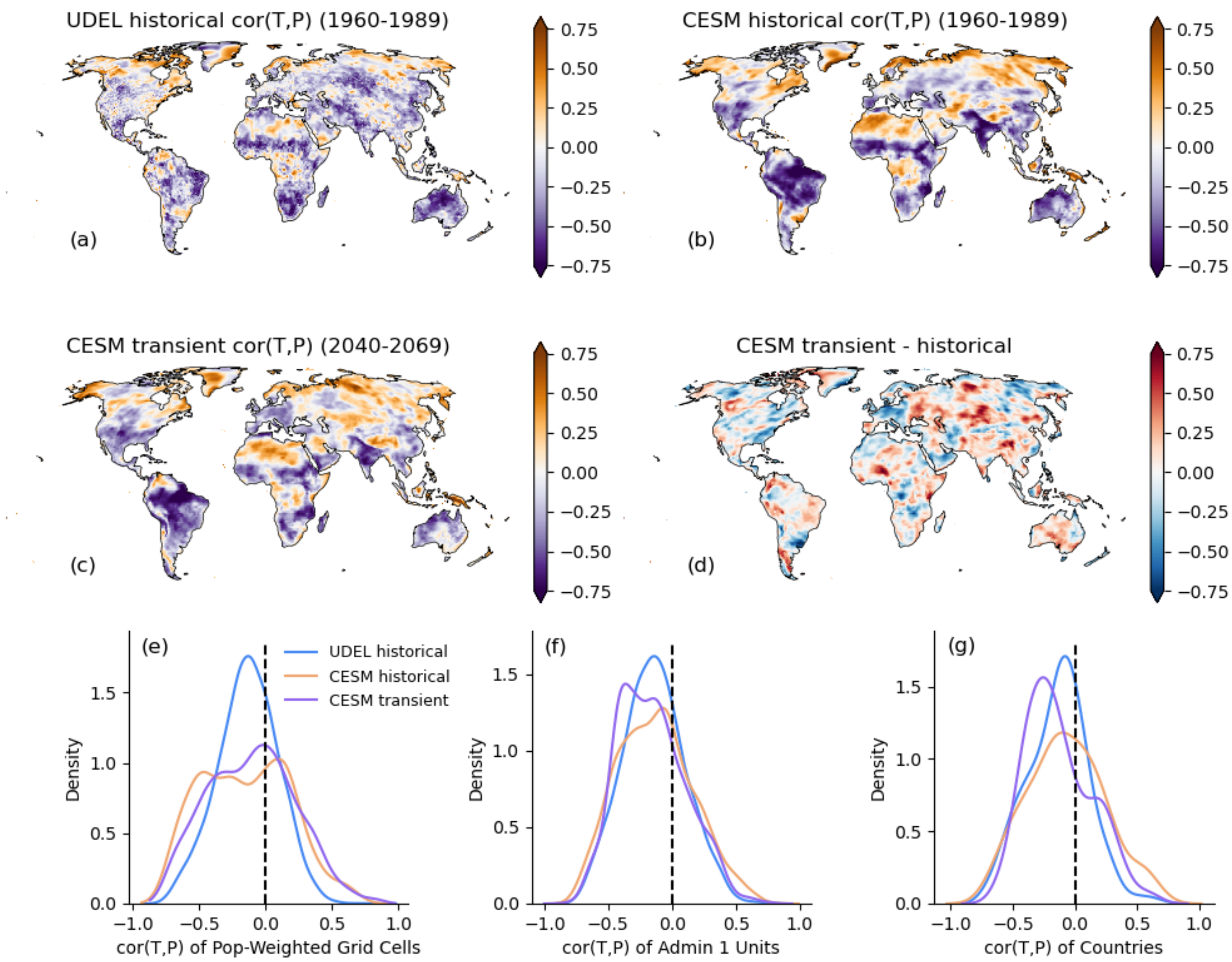

Figure S6: Annual correlation between T and P. **(a)** Historical cor(T,P) in the University of Delaware gridded dataset. **(b)** Historical cor(T,P) from the CESM2-WACCM historical simulation. **(c)** Future cor(T,P) as predicted by the CESM2-WACCM middle-of-the-road climate change simulation (SSP245). **(d)** the difference between panels c and b. **(e-g)** Distribution of cor(T,P) in panels (a-c) at different levels of aggregation: **(e)** grid cell, **(f)** administrative 1 units (e.g. US states), and **(g)** countries).

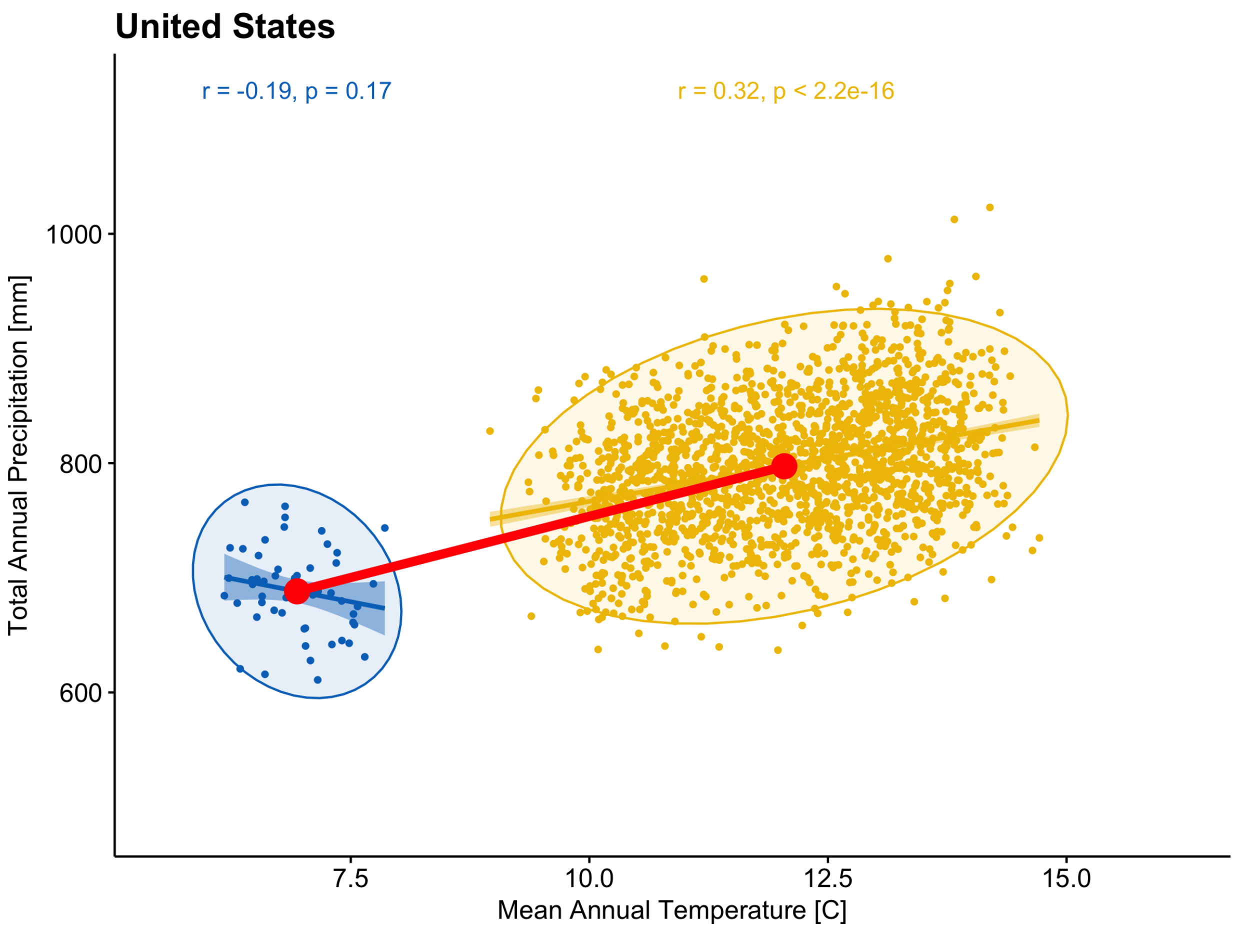

Figure S7: Relationship between T and P in historical CESM simulations and in CESM-LENS2, aggregated to the United States. This figure demonstrates that the relationship between T and P in historical, transient, and future climate, may change significantly in some places, and even change sign.

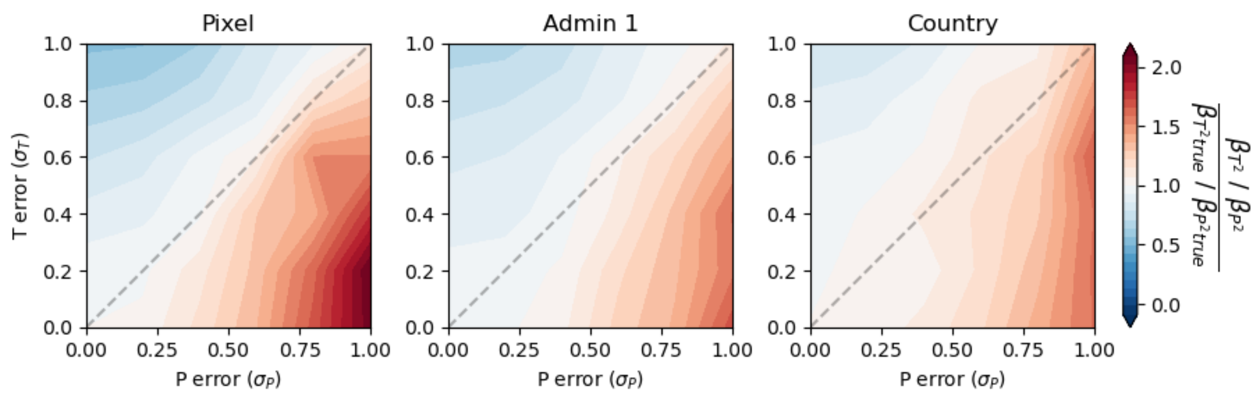

Figure S8: Relative bias in the strength of estimated temperature versus precipitation impacts on an outcome Y in an idealized modeling framework. Contours show the estimated ratio of the quadratic components of T and P, under uniform added error in T and P, to the true ratio of those components (specified ex-ante). Red shading indicates that T is overvalued relative to P. Error is added as a function of unit-level interannual standard deviation of T (vertical axis) and P (horizontal axis). Regressions are population-weighted at the pixel level, while T and P are aggregated using population weighting for the administrative 1 and country-level regressions.

Here we describe some of the statistical theory explored numerically in this study to demonstrate how statistical problems can arise.

Measurement error

The linear model with coefficient β and error term ϵ is:

$$y_{it} = \alpha + \beta x_{it} + \epsilon_{it} \tag{13}$$

and the best-fit coefficient estimated with ordinary least squares regression is:

$$\hat{\beta} = \frac{cov(x_{it}, y_{it})}{var(x_{it})} \tag{14}$$

If we then allow the independent variable x_{it} to have measurement error:

$$\tilde{x}_{it} = x_{it} + u_{it} \tag{15}$$

and use Eq.13, the estimated coefficient becomes:

$$\hat{\beta} = \frac{cov(x_{it} + u_{it}, \beta x_{it} + \epsilon_{it})}{var(x_{it} + u_{it})} = \beta \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} \tag{16}$$

Meaning that the magnitude of the estimated coefficient $\hat{\beta}$ is biased towards zero as the measurement error increases.

If we add fixed effects* the estimating equation becomes:

$$y_{it} = \beta x_{it} + \mu_i + \epsilon_{it}. \tag{17}$$

Inclusion of these unit-specific intercepts is equivalent to demeaning both the right- and left-sides of the estimating equation:

$$y_{it} - \overline{y_{it}} = \beta(x_{it} - \overline{x_{it}}) + \epsilon_{it} \tag{18}$$

or

$$\delta y_{it} = \beta \delta x_{it} + \epsilon_{it} \tag{19}$$

Here for demonstration purposes we follow Pischke (Pischke, 2007) and eliminate the fixed effect by differencing between two time periods.

$$y_{it} - y_{i,t-1} = \beta(x_{it} - x_{i,t-1}) + (\epsilon_{it} - \epsilon_{i,t-1}) \tag{20}$$

*For demonstration here we consider only location-specific time-invariant intercepts meant to capture cross-sectional differences, and omit time-specific global intercepts meant to capture temporal features common to the entire system, although the logic is similar.

Inclusion of measurement error in x_{it} in this case exacerbates the attenuation bias:

$$\hat{\beta} = \beta \frac{\sigma^2_{\delta x}}{\sigma^2_{\delta x} + \sigma^2_u} \tag{21}$$

where

$$\sigma^2_{\delta x} = var(x_{it} - x_{i,t-1}) = var(x_{it}) - 2cov(x_{it}, x_{i,t-1}) + var(x_{i,t-1}) = var(x_{it}) - 2cov(x_{it}, x_{i,t-1}) \tag{22}$$

$$\sigma^2_{\Delta u} = var(u_{it} - u_{i,t-1}) = var(u_{it}) - 2cov(u_{it}, u_{i,t-1}) + var(u_{i,t-1}) = var(u_{it}) - 2cov(u_{it}, u_{i,t-1}) \tag{23}$$

The last term of both x and u are stationary (presumed to be the case if fixed effects were included). The expression for the estimated coefficient then becomes:

$$\hat{\beta} = \beta \frac{\sigma^2_x(1-\rho)}{\sigma^2_x(1-\rho) + \sigma^2_u(1-r)} \tag{24}$$

where ρ and r are the first order autocorrelation coefficients for x and u, respectively. Attenuation bias thus becomes worst in the case when x has high correlation over time ($\rho \to 1$), but u is truly stochastic ($r = 0$). This is the case for many climate impacts questions, where human outcomes (e.g., GDP, crop yields) are serially correlated, but measurement error in x is likely not highly autocorrelated.

Correlated independent variables and measurement error

A key finding in this paper is that measurement error can, in a sense, propagate between correlated variables to bias estimated coefficients: for example, measurement error in precipitation (P) when $cor(T, P) \neq 0$, leads to bias in the estimated coefficients for both P (as demonstrated above), but also T. Here we again follow Pischke (Pischke, 2007) to show mathematically why this occurs:

We begin with a standard linear model for an impact y that is a function of T and P (and eliminating subscripts for clarity):

$$y = \beta_T T + \beta_P P + \epsilon \tag{25}$$

If we imperfectly observe $\tilde{P} = P + u$ then the formula for $\hat{\beta}_P$ is

$$\hat{\beta}_P = \frac{var(T)cov(y, \tilde{P}) - cov(T, \tilde{P})cov(y, T)}{var(\tilde{P})var(T) - cov(T, \tilde{P})^2} \tag{26}$$

which in the limit becomes

$$\text{plim}\hat{\beta}_P = \frac{\sigma^2_T(\beta_P\sigma^2_P + \beta_T\sigma_{TP}) - \sigma_{\tilde{P}T}(\beta_T\sigma^2_T + \beta_P\sigma_{TP})}{\sigma^2_T(\sigma^2_P + \sigma^2_u) - (\sigma_{\tilde{P}T})^2} \tag{27}$$

IF P is only correlated with T can be simplified because $\sigma_{TP} = \sigma_{T\tilde{P}}$ and

$$\text{plim}\hat{\beta}_P = \frac{\beta_P(\sigma_T^2\sigma_P^2 - (\sigma_{TP})^2)}{\sigma_T^2(\sigma_P^2 + \sigma_u^2) - (\sigma_{TP})^2} = \beta_P\lambda' \tag{28}$$

Since $\sigma_T^2\sigma_P^2 > (\sigma_{TP})^2$, $\hat{\beta}_P$ is biased towards zero. λ', the attenuation factor, can be rewritten as

$$\lambda' = \frac{\lambda - R^2_{T\tilde{P}}}{1 - R^2_{T\tilde{P}}} \tag{29}$$

where $\lambda = \sigma_P^2/(\sigma_P^2 + \sigma_u^2)$. This shows that as the correlation between T and $\tilde{P}$ goes up, the attenuation bias increases. If we assume that $\sigma_T^2 = \sigma_P^2$ (which is generally not the case, but simplifies the math), then

$$\text{plim}\hat{\beta}_T - \beta_T = \rho_{TP}(1 - \lambda')\beta_P = -\rho_{TP}(\text{plim}\hat{\beta}_P - \beta_P) \tag{30}$$

where ρ_{TP} is the correlation coefficient between T and P. This shows that the correlation between T and P determines how much the bias in β_P influences the bias in β_T.

Non-stationary mean state

Finally, we note that fixed effects are often included in climate impacts models to account for average differences between observational units, like countries. The intent here is intuitive, in that each country has its own systems and and history and economic structure. Countries (or states, or districts) have unique structures to their economies, governments, etc. that can be hard to capture through data, and have different mean levels of a range of human systems outcome variables (e.g., GDP, agricultural productivity, infant mortality rate, etc.). As discussed above, the inclusion of unit-specific intercepts demeans both sides of the equation and turns the analysis into one that is focused on within-unit variation: that is, how does GDP within a country vary compared to average when that country is hotter or cooler than average?

This poses three interconnected problems from a methodological standpoint. The first is straightforward, and discussed in the main manuscript: the mean climate state is nonstationary due to anthropogenic warming. Therefore, while estimates from the recent past may be (relatively) unbiased, climate models tell us that the assumption of stationary mean climate is incorrect moving into the future. We address this above by allowing the mean state to evolve on its own, and thus separating out variability-based responses to climate from longer run responses. While there are many ways to do this, the motivation and tradeoffs are similar – it may be necessary to sacrifice some strict causal identification in empirical models to credibly link impact estimates to future projections.

Second, the existence of spatial gaps in modern measurement networks (Figure 1) means that many of our impact coefficient estimates may be biased, with that bias exacerbated by the use of fixed effects models, as demonstrated above. And finally, since temperature and precipitation have been negatively correlated over land (Figure S6), the use of fixed effects models to estimate climate impacts poses an additional problem. As discussed, fixed effects jointly demean both the outcome and the right-hand side of an estimating equation. Therefore, fixed effects subtract off the mean climate state – which includes temperature and precipitation, but also the correlation between temperature and precipitation. This relationship is not expected to remain stationary, but projections vary heterogeneously around the world (Figure S6, Figure S7).