

Downscaling and bias-correction contribute considerable
uncertainty to local climate projections in CMIP6

SUPPORTING INFORMATION

David C. Lafferty^{a*} & Ryan L. Sriver^a

^aDepartment of Atmospheric Sciences, University of Illinois Urbana-Champaign

*Corresponding author email: davidcl2@illinois.edu

Contents

1	Supplementary Notes	1
1.1	Data sources	1
1.2	Observational dataset comparisons	1
1.3	Uncertainty partitioning for alternative definitions of extremes	2
1.4	Methodological checks	2
1.5	Absolute uncertainty comparison	4
2	Supplementary Tables	6
3	Supplementary Figures	9

List of Figures

1	<i>Timeseries:</i> Annual average temperature (cities)	10
2	<i>Timeseries:</i> Annual total precipitation (cities)	11
3	<i>Timeseries:</i> Annual maximum of daily maximum temperature (cities)	12
4	<i>Timeseries:</i> Annual maximum 1-day precipitation (cities)	13
5	<i>Spatial corr.:</i> Annual average temperature	14
6	<i>Spatial corr.:</i> Annual total precipitation	15
7	<i>Spatial corr.:</i> Annual maximum of daily maximum temperature	16
8	<i>Spatial corr.:</i> Annual maximum 1-day precipitation	17
9	<i>Spatial corr.:</i> Annual vs. extreme temperature	18
10	<i>Spatial corr.:</i> Annual vs. extreme precipitation	19
11	<i>Obs. comparison:</i> Observational disagreement in annual temperature averages (cities) .	20
12	<i>Obs. comparison:</i> Observational disagreement in annual total precipitation (cities) . . .	21
13	<i>Obs. comparison:</i> Observational disagreement in annual daily temperature maxima (cities)	22
14	<i>Obs. comparison:</i> Observational disagreement in annual daily precipitation maxima (cities)	23
15	<i>Obs. comparison:</i> Observational disagreement in annual averages (global)	24
16	<i>Obs. comparison:</i> Observational disagreement in 1-day maxima (global)	25
17	<i>Obs. comparison:</i> Observational disagreement in dry days (global)	26
18	<i>Obs. comparison:</i> Observational disagreement in 99th tasmax percentile	27
19	<i>Obs. comparison:</i> Observational disagreement in 99th precipitation percentile	28
20	<i>Obs. comparison:</i> Observational disagreement in 99th tas percentile	29
21	<i>Obs. comparison:</i> Observational disagreement in 99th tasmin percentile	30
22	<i>Obs. comparison:</i> Coastal effects for projections of temperature extremes in Lagos . . .	31
23	<i>Obs. comparison:</i> Coastal effects for projections of temperature extremes in Mumbai . .	32
24	<i>Obs. comparison:</i> Coastal effects for projections of temperature extremes in Rio	33
25	<i>Obs. comparison:</i> Coastal effects for projections of precipitation extremes in Lagos . . .	34
26	<i>Obs. comparison:</i> Coastal effects for projections of precipitation extremes in Mumbai . .	34
27	<i>Obs. comparison:</i> Coastal effects for projections of precipitation extremes in Rio	35
28	<i>Alt. colormap:</i> Global variance decomposition of climate averages	36
29	<i>Alt. colormap:</i> Global variance decomposition of annual 1-day maxima	37
30	<i>Alt. colormap:</i> Global variance decomposition of threshold indices of climate extremes .	38

31	<i>Alt. extremes:</i> 1-day annual temperature maxima using daily tas, tasmin	39
32	<i>Alt. extremes:</i> 5-day temperature and precipitation maxima	40
33	<i>Alt. extremes:</i> Extremely hot days as tasmax exceeding the 99th percentile in GMFD .	41
34	<i>Alt. extremes:</i> Extremely hot days as tasmax exceeding the 99th percentile in ERA5 . .	42
35	<i>Alt. extremes:</i> Extremely hot days as tas exceeding the 99th percentile in GMFD	43
36	<i>Alt. extremes:</i> Extremely hot days as tas exceeding the 99th percentile in ERA5	44
37	<i>Alt. extremes:</i> Extremely hot days as tasmin exceeding the 99th percentile in GMFD . .	45
38	<i>Alt. extremes:</i> Extremely hot days as tasmin exceeding the 99th percentile in ERA5 . .	46
39	<i>Alt. extremes:</i> Extremely hot days as tasmax exceeding the 99.97th percentile in GMFD	47
40	<i>Alt. extremes:</i> Extremely hot days as tas exceeding the 99.97th percentile in GMFD . .	48
41	<i>Alt. extremes:</i> Extremely hot days as tasmin exceeding the 99.97th percentile in GMFD	49
42	<i>Alt. extremes:</i> Extremely wet days as pr exceeding the 99th percentile in GMFD	50
43	<i>Alt. extremes:</i> Extremely wet days as pr exceeding the 99th percentile in ERA5	51
44	<i>Alt. extremes:</i> Extremely wet days as pr exceeding the 99.97th percentile in GMFD . .	52
45	<i>Alt. extremes:</i> Extremely wet days as pr exceeding the 99.97th percentile in ERA5 . . .	53
46	<i>Alt. extremes:</i> Extremely hot & dry days as tasmax exceeding 99th percentile in GMFD	54
47	<i>Alt. extremes:</i> Extremely hot & dry days as tasmax exceeding the 99th percentile in ERA5	55
48	<i>Method checks:</i> Annual averages under BB13 scenario uncertainty	56
49	<i>Method checks:</i> 1-day maxima under BB13 scenario uncertainty	57
50	<i>Method checks:</i> 5-day maxima under BB13 scenario uncertainty	58
51	<i>Method checks:</i> Extremely hot days under BB13 scenario uncertainty	59
52	<i>Method checks:</i> Dry days under BB13 scenario uncertainty	60
53	<i>Method checks:</i> Extremely hot & dry days under BB13 scenario uncertainty	61
54	<i>Method checks:</i> Extremely wet days under BB13 scenario uncertainty	62
55	<i>Method checks:</i> Annual averages without output availability weighting	63
56	<i>Method checks:</i> 1-day maxima without output availability weighting	64
57	<i>Method checks:</i> 5-day maxima without output availability weighting	65
58	<i>Method checks:</i> Extremely hot days without output availability weighting	66
59	<i>Method checks:</i> Dry days without output availability weighting	67
60	<i>Method checks:</i> Extremely hot & dry days without output availability weighting	68
61	<i>Method checks:</i> Extremely wet days without output availability weighting	69
62	<i>Method checks:</i> Timeseries and decomposition for 1-day maxima (cities)	70

63	<i>Method checks:</i> Decomposition for threshold extremes indices (cities)	71
64	<i>Method checks:</i> ANOVA results for annual averages, 1-day maxima (cities)	72
65	<i>Method checks:</i> ANOVA results for threshold extremes indices (cities)	73
66	<i>Method checks:</i> Total uncertainty comparison for averages, 1-day maxima (cities)	74
67	<i>Method checks:</i> Total uncertainty comparison for threshold indices (cities)	75
68	<i>Method checks:</i> Total uncertainty comparison for all metrics (global)	76
69	<i>Absolute uncertainty:</i> Annual averages	77
70	<i>Absolute uncertainty:</i> 1-day maxima	78
71	<i>Absolute uncertainty:</i> 5-day maxima	79
72	<i>Absolute uncertainty:</i> Extremely hot days	80
73	<i>Absolute uncertainty:</i> Dry days	81
74	<i>Absolute uncertainty:</i> Extremely hot & dry days	82
75	<i>Absolute uncertainty:</i> Extremely wet days	83

1 Supplementary Notes

2 1.1 Data sources

3 As described in the main text, we leverage five ensembles of statistically downscaled model
4 outputs. Table 1 lists the parent GCMs available in each ensemble and denotes any missing
5 scenarios and/or variables. If available, for each GCM we gather outputs for four Shared
6 Socioeconomic Pathway (SSP) scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) and
7 three variables (tasmin, tasmax, pr). Table 2 provides some details on each of the downscaling
8 and bias-correction implementations.

9 1.2 Observational dataset comparisons

10 Differences in downscaled projections can stem in part from disagreements in the underlying
11 observational data used for training. The downscaled ensembles that we rely on in this study
12 were trained using three separate observational datasets (although some were trained on the
13 same dataset over different time periods; see Table 2). The observational datasets are: Global
14 Meteorological Forcing Dataset (GMFD) for Land Surface Modeling [1], ERA5 reanalysis from
15 the European Centre for Medium-Range Weather Forecasts [2], and W5E5v2.0 [3, 4].

16 Figs. 11–14 show timeseries of annual temperature and precipitation averages and annual
17 1-day maxima in New Delhi, Lagos, and Seattle for each observational dataset. Figs. 15 &
18 16 show global maps of a measure of the historical disagreement across datasets for annual
19 temperature and precipitation averages, annual 1-day maxima, and the annual number of dry
20 days. Figs. 18–21 show ERA5-GMFD differences in the historical (1980-2014) 99th percentiles
21 for temperature and precipitation.

22
23 **Coastal effects.** We found sizeable differences in the projections of temperature and precipita-
24 tion extremes in coastal locations. Here we provide some evidence in support of this being driven
25 by disagreements in the underlying observational data. Fig. 22 shows, on the left column, the
26 ERA5-GMFD ratio of the historical 99th percentiles of daily maximum, minimum, and average
27 temperature, in the grid points surrounding Lagos. This gives an indication of the magnitude
28 of disagreement between the two observational products for temperature extremes. The center
29 columns shows, for the Lagos gridpoint, all downscaled projections of annual daily temperature
30 maxima from a single parent GCM (CanESM5) under a single SSP scenario (SSP3-7.0). By

focusing on a single GCM and scenario, we highlight the differences that arise solely due to downscaling and bias-correction. The right column shows the full variance decomposition after accounting for all models and scenarios (i.e., our results from the main text). Fig. 22 clearly demonstrates that large disagreements in the observational quantiles drive large differences in the associated downscaled projections, which in turn leads to downscaling contributing more uncertainty relative to other sources. Note that for daily maximum temperature (top row), and to a lesser extent daily minimum temperature (middle row), the observational disagreement is particularly pronounced at coastal gridpoints. The datasets tend to agree farther inland. This may be related to biases in the representation of coastal temperature extremes in ERA5 [5].

We repeat this exercise for two other large coastal metros: Mumbai, India (Fig. 23) and Rio de Janeiro, Brazil (Fig. 24), as well as for precipitation extremes in each city (Figs. 25–27). These results generally confirm our hypothesis that large observational disagreements lead to considerable differences in downscaled projections. They also demonstrate that disagreements as stark as those in Lagos do not occur everywhere.

1.3 Uncertainty partitioning for alternative definitions of extremes

We provide global maps of the variance decomposition under alternative definitions of extremes. These include: 1-day temperature maxima based on daily minimum and daily average temperature (Fig. 31), 5-day temperature and precipitation maxima (Fig. 32), and alternative definitions of what constitutes an extremely hot day (Figs. 33–41) and extremely wet day (Figs. 42–45). We also report the results for a multivariate metric, extremely hot and dry days (Figs. 46 & 47).

1.4 Methodological checks

As discussed in the main text, we perform some methodological checks to test the robustness of our results. These include using the Brekke & Barsugli definition of scenario uncertainty [6] (Figs. 48–54) and not performing the weighting strategy that accounts for data availability (Figs. 55–61).

ANOVA. Our main results also neglect interaction effects among uncertainty sources. Here we implement two checks to estimate the magnitude of these effects. First, we apply a factorial analysis of variance (ANOVA) that facilitates an uncertainty partitioning that accounts for

interactions. Adopting the terminology of the main text, let $x(t, s, m, d)$ denote a given climate metric in some location at year t from scenario s , parent GCM m , and downscaling method d , and let $\hat{x}(t, s, m, d)$ denote the estimated forced response after applying the 4th order polynomial fit. We apply a linear ANOVA model to the forced response:

$$\begin{aligned}\hat{x}(t, s, m, d) = & \mu(t) + \alpha(t, s) + \beta(t, m) + \gamma(t, d) \\ & + \theta_{\alpha\beta}(t, s, m) + \theta_{\alpha\gamma}(t, s, d) + \theta_{\beta\gamma}(t, m, d) \\ & + \varepsilon(t, s, m, d).\end{aligned}\tag{1}$$

Here, $\mu(t)$ represents the entire ensemble mean, $\alpha(t, s)$ represents the average deviation of scenario s from the ensemble mean, $\beta(t, m)$ represents the average deviation of GCM m , and $\gamma(t, d)$ represents the average deviation of downscaling method d . The two-way interaction effects are given by $\theta_{\alpha\beta}(t, s, m)$, $\theta_{\alpha\gamma}(t, s, d)$, and $\theta_{\beta\gamma}(t, m, d)$ for scenario-GCM, scenario-downscaling, and model-downscaling interactions respectively. We neglect the three-way interaction. The error term is $\varepsilon(t, s, m, d)$ and assumed to be independent and identically distributed. The uncertainty partitioning proceeds by first calculating the total sum of squares $\widehat{SS}_{total}(t)$, which represents the total variation in the forced response $\hat{x}(t, s, m, d)$, and the sums of squares attributable to each variation term in Eq. 1: $SS_{\alpha}(t)$, $SS_{\beta}(t)$, $SS_{\gamma}(t)$, $SS_{\alpha\beta}(t)$, $SS_{\alpha\gamma}(t)$, $SS_{\beta\gamma}(t)$, and $SS_{\varepsilon}(t)$. We have an unbalanced design since not every combination of scenario, GCM, and downscaling method is available. To calculate the sums of squares we elect to use a Type II ANOVA and rely on the `car` package [7] in the R programming language. The sum of squares of the forced response can be expressed as:

$$\widehat{SS}_{total}(t) = SS_{\alpha}(t) + SS_{\beta}(t) + SS_{\gamma}(t) + SS_{\alpha\beta}(t) + SS_{\alpha\gamma}(t) + SS_{\beta\gamma}(t) + SS_{\varepsilon}(t).\tag{2}$$

However, so far we have neglected the influence of interannual variability. To account for interannual variability, we calculate the sum of squares $SS_{IV}(t)$ associated with the difference between the raw outputs and the forced response as

$$SS_{IV}(t) = \sum_{s,m,d} [x(t, s, m, d) - \hat{x}(t, s, m, d)]^2.\tag{3}$$

An alternative approach here would be to apply the ANOVA to the raw outputs and assume that the error term ε represents interannual variability. However, we found that this leads to

an extremely small contribution from interannual variability. Future work could explore this mismatch in more detail.

The total sum of squares in the raw outputs can then be expressed as

$$SS_{total}(t) = \widehat{SS}_{total}(t) + SS_{IV}(t) \\ \simeq SS_{\alpha}(t) + SS_{\beta}(t) + SS_{\gamma}(t) + SS_{\alpha\beta}(t) + SS_{\alpha\gamma}(t) + SS_{\beta\gamma}(t) + SS_{IV}(t). \quad (4)$$

In the second line we neglect the error term $SS_{\epsilon}(t)$ because it is typically very small. The relative fraction of variance attributed to each uncertainty source is obtained by dividing each term by $SS_{total}(t)$. These results are reported in Figs. 64 & 65, where we focus on our three example cities.

Total uncertainty comparison. The second test that we use to check for interaction effects is simpler but can be applied at the global scale. As described in the main text, we assume that total variance can be independently decomposed as the sum of each individual variance estimate (see Methods section in the main text). In general, this is not true due to the presence of interaction effects. Here we calculate the difference between our assumed total uncertainty (as the sum of individual variances) and the true total uncertainty (as the variance across all outputs). In Figs. 66 & 67 we show this ratio for our three example cities, where it can be seen that the city/metric combinations with the largest differences between assumed and true uncertainty also exhibit the largest interaction effects as measured by the ANOVA decomposition (Figs. 64 & 65). Fig. 68 shows the total uncertainty comparison for all metrics globally.

1.5 Absolute uncertainty comparison

As noted in the discussion surrounding Figure 5 in main text, most our main results show only the relative contribution of each uncertainty source, measured as a fraction of the total variance. This masks the fact that the contribution from each uncertainty source tends to grow over time in absolute terms. For example, scenario uncertainty is almost always small at the beginning of the 21st century since scenarios are most similar at these early time horizons—notable differences among scenarios do not emerge until later in the century. In Figs. 69–75 we show maps of the absolute uncertainty contributed by each source for all metrics. Similar to Figure 5 in the main text, we measure absolute uncertainty by calculating the standard deviation across each axis of

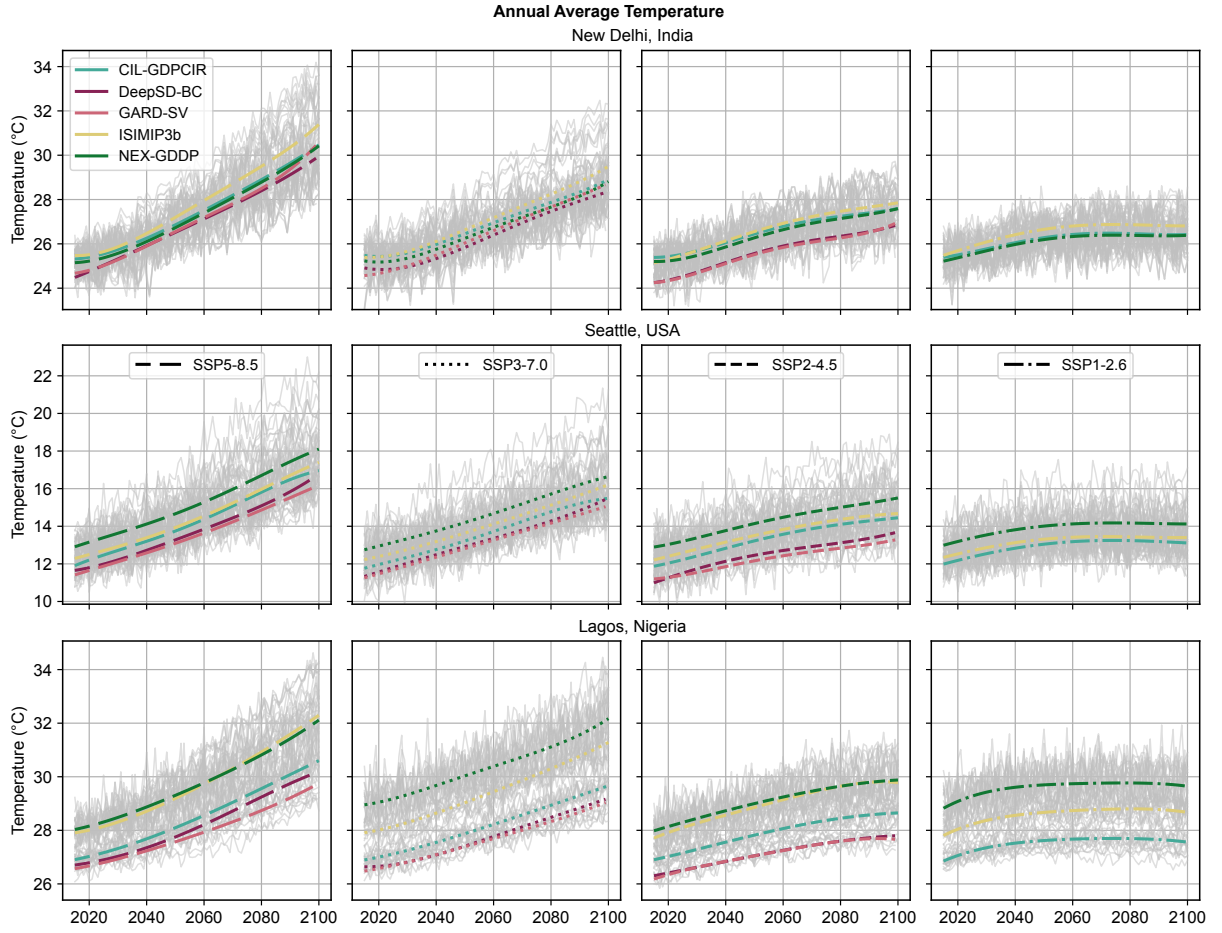
102 uncertainty, rather than the variance, in order to express the results in physically meaningful
103 units.

Supplementary Table 1: Downscaled outputs across GCMs, ensembles, and SSP scenarios. Check marks indicate complete data coverage (all SSPs and variables are available). Text indicates missing data (text describes missing SSP and/or variables). Blank boxes indicate GCM/downscaling combination does not exist.

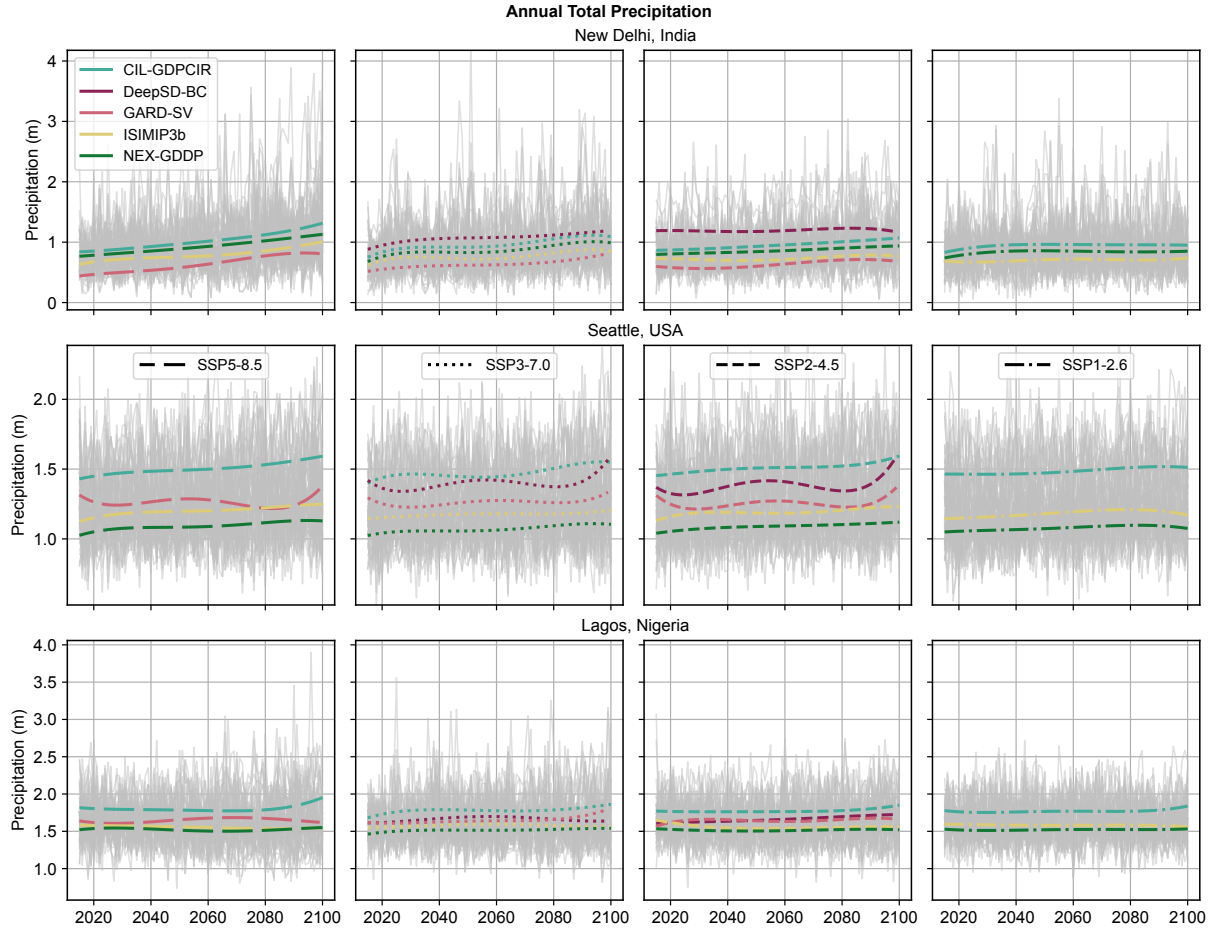
CMIP6 GCM	NEX-GDDP	CIL-GDPCIR	carbonplan		ISIMIP3b
			GARD-SV	DeepSD-BC	
ACCESS-ESM1-5	✓	SSP5-8.5			
BCC-CSM2-MR	✓	✓	SSP1-2.6 pr		
CanESM5	✓	✓	SSP1-2.6	SSP1-2.6	SSP2-4.5
CMCC-ESM2	✓	✓			
CNRM-CM6-1	✓				SSP2-4.5
CNRM-ESM2-1	✓				SSP2-4.5
EC-Earth3	✓	✓			SSP2-4.5
EC-Earth3-Veg-LR	✓	✓			
GFDL-ESM4	✓	✓			✓
HadGEM3-GC31-LL	SSP3-7.0	SSP3-7.0			
INM-CM4-8	✓	✓			
INM-CM5-0	✓	✓			
IPSL-CM6A-LR	✓				✓
MIROC-ES2L	✓	✓			
MIROC6	✓	✓	SSP1-2.6 pr		SSP2-4.5
MPI-ESM1-2-HR	✓		SSP1-2.6		✓
MPI-ESM1-2-LR	✓	✓			
MRI-ESM2-0	✓			SSP1-2.6 pr for SSP5-8.5	✓
NESM3	SSP3-7.0	SSP3-7.0			
NorESM2-LM	✓	✓			
NorESM2-MM	✓	✓			
UKESM1-0-LL	✓	✓			✓

Supplementary Table 2: Overview of different downscaling & bias-correction methodologies. Here we provide a highly condensed description of each algorithm; additional details can be found in the references provided.

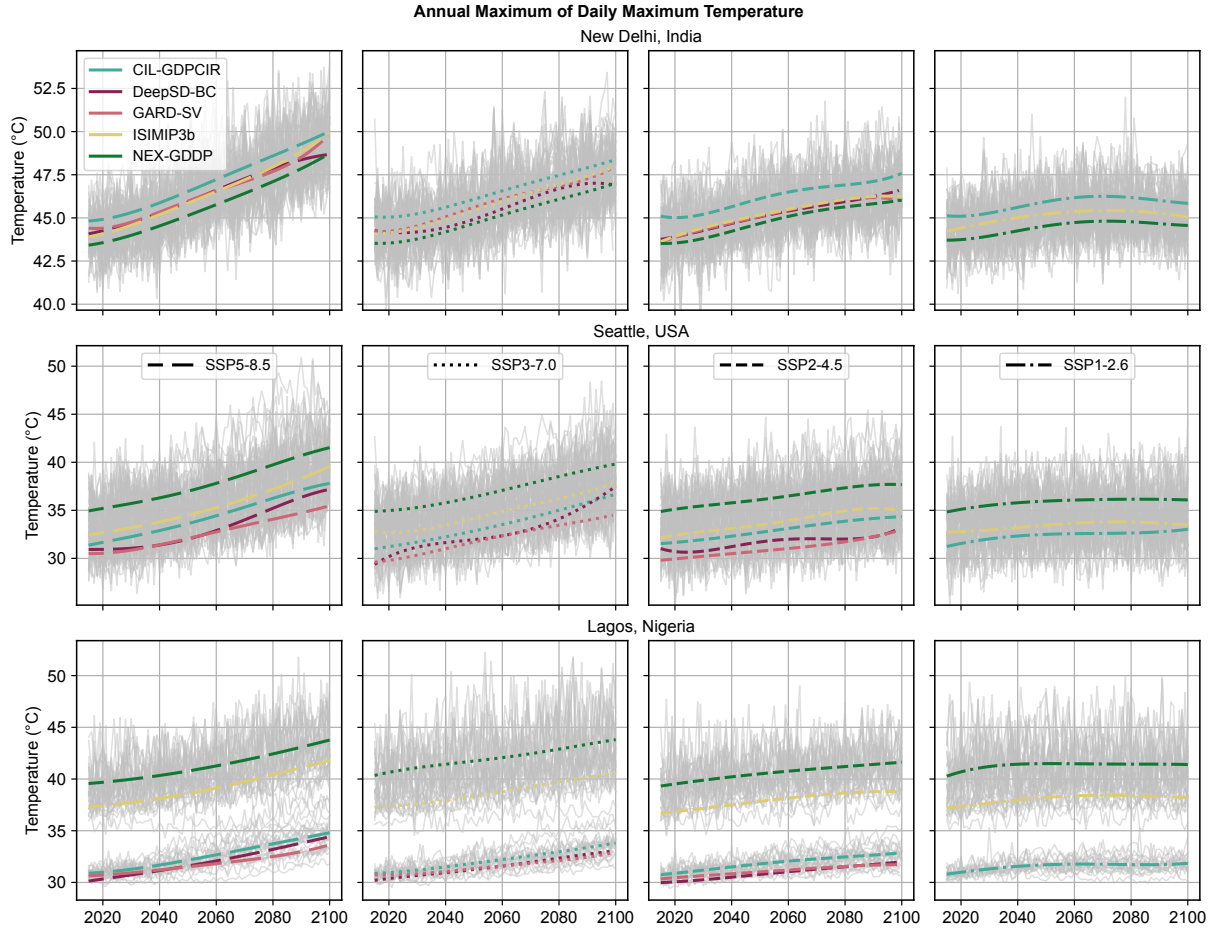
Ensemble	Total # GCMs	Bias-correction	Downscaling	Uni-/multi- variate	Accounts for spatial patterns?	Spatial resolution	Obs. dataset	Refs.
NEX-GDDP	36	Quantile mapping	Spatial disaggregation via bilinear interpolation	Univariate	No	1/4°	GMFD (1960-2014)	[8, 9]
CIL-GDPCIR	25	Quantile Delta Mapping	Quantile-Preserving Localized-Analog	Univariate	No	1/4°	ERA5 (1995-2014)	[10, 11]
ISIMIP3b	10	Non-parametric quantile mapping	Bilinear interpolation & quantile mapping	Univariate	No	1/2°	W5E5v2.0 (1979-2014)	[12, 13, 14]
carbonplan: GARD-SV	4	Quantile mapping	Generalized Analog Regression (pure regression variant)	Univariate	No	1/4°	ERA5 (1981-2010)	[15, 16, 17]
carbonplan: DeepSD-BC	2	Quantile mapping	Convolutional neural network	Multivariate	Yes	1/4°	ERA5 (1981-2010)	[15, 16, 18]



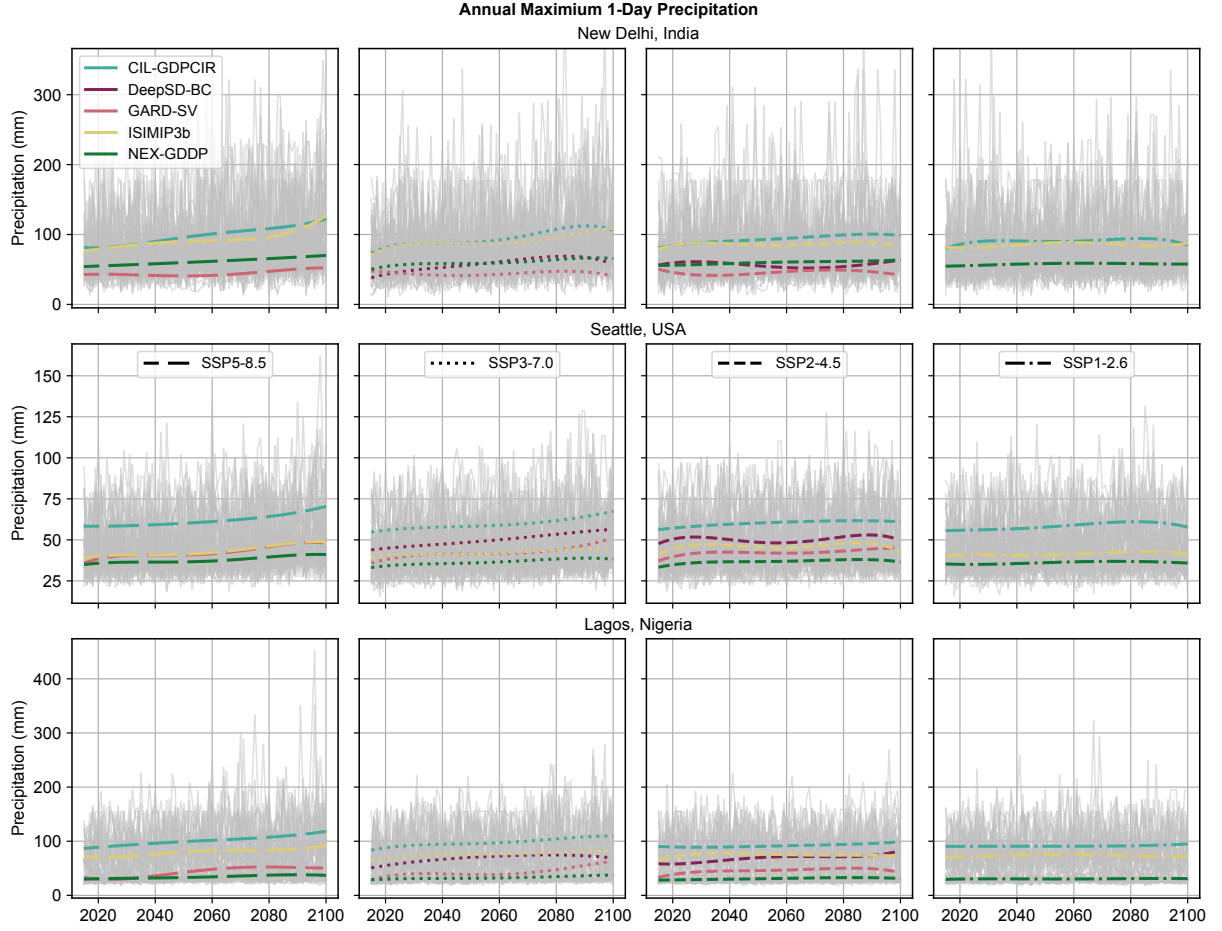
Supplementary Figure 1: **Timeseries:** Projections of annual average temperature observations in New Delhi, Seattle, Lagos. Similar to Figure 1 in the main text, gray lines show individual model outputs and colored lines show ensemble means. The columns demarcate different SSP scenarios.



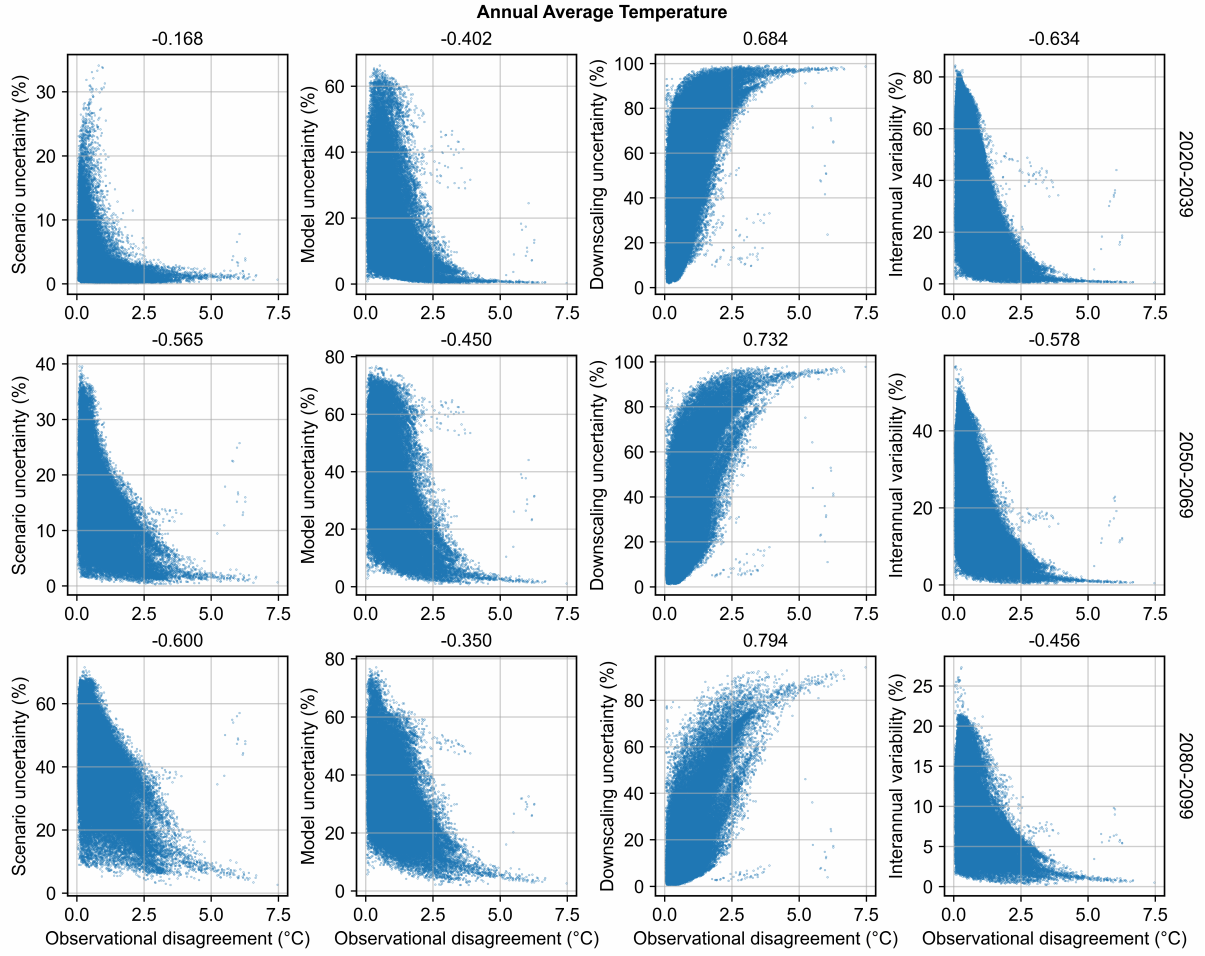
Supplementary Figure 2: **Timeseries:** Projections of annual total precipitation observations in New Delhi, Seattle, Lagos. Similar to Figure 1 in the main text, gray lines show individual model outputs and colored lines show ensemble means. The columns demarcate different SSP scenarios.



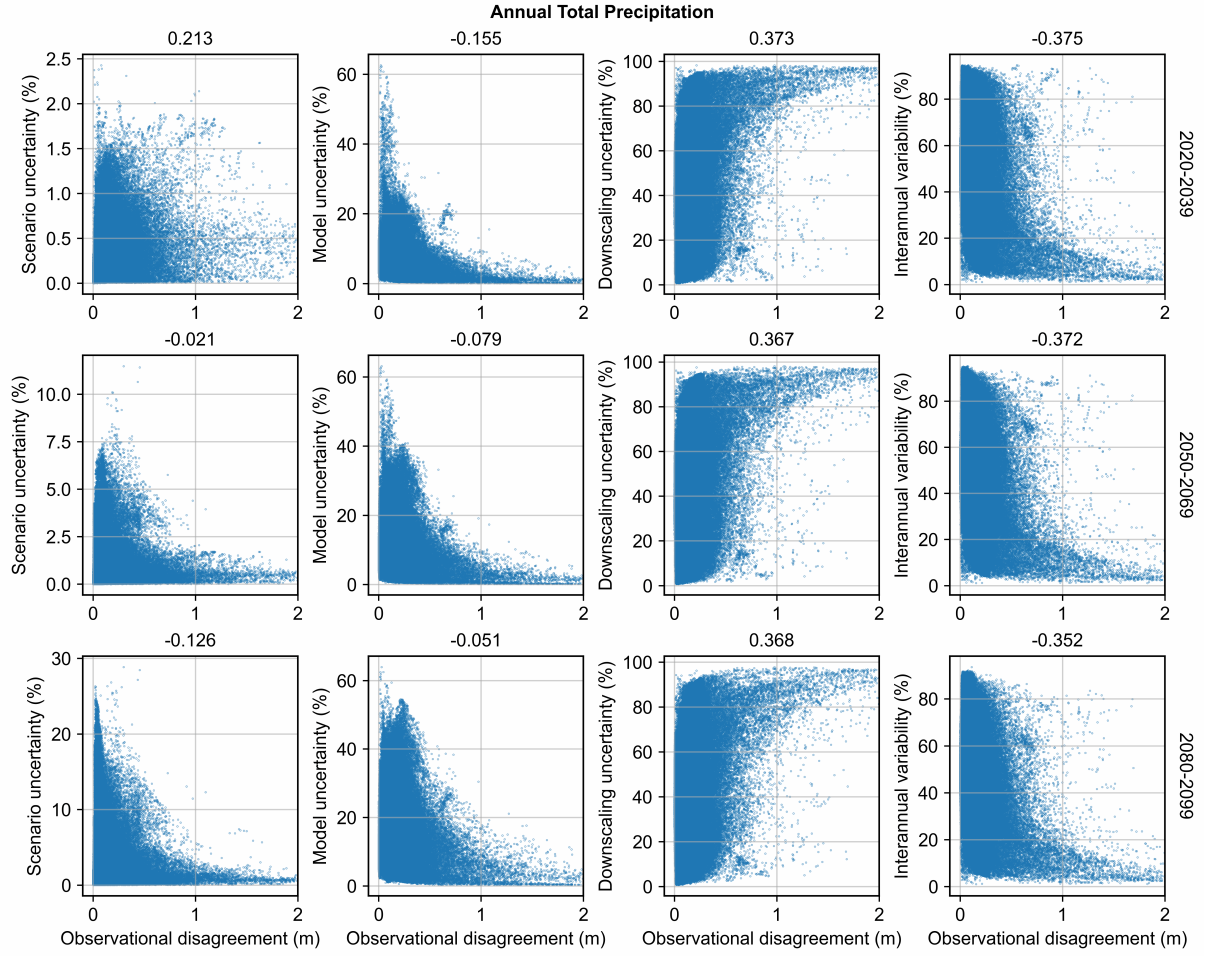
Supplementary Figure 3: **Timeseries:** Projections of annual maximum of daily maximum temperature observations in New Delhi, Seattle, Lagos. Similar to Figure 1 in the main text, gray lines show individual model outputs and colored lines show ensemble means. The columns demarcate different SSP scenarios.



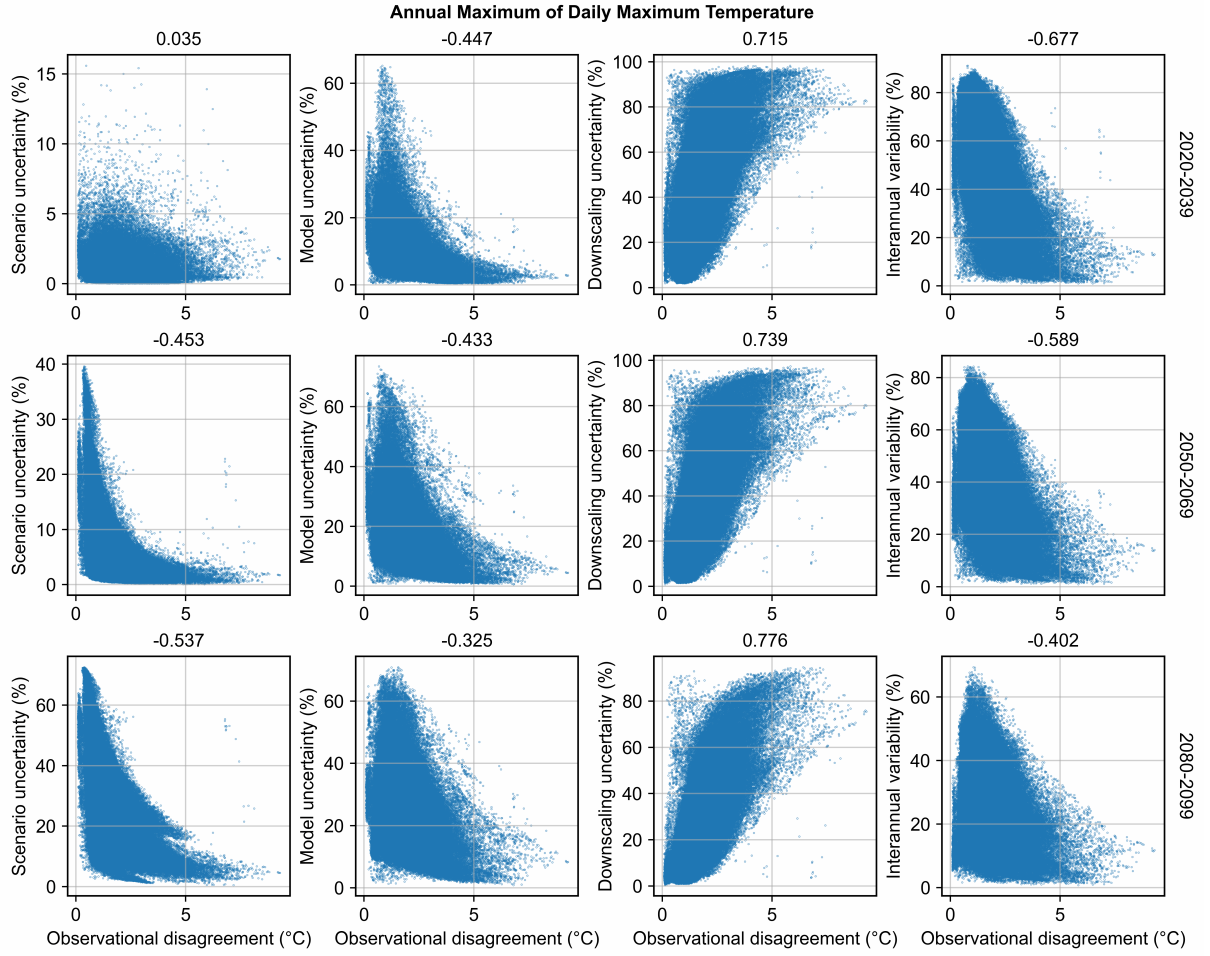
Supplementary Figure 4: **Timeseries:** Projections of annual 1-day precipitation maxima observations in New Delhi, Seattle, Lagos. Similar to Figure 1 in the main text, gray lines show individual model outputs and colored lines show ensemble means. The columns demarcate different SSP scenarios.



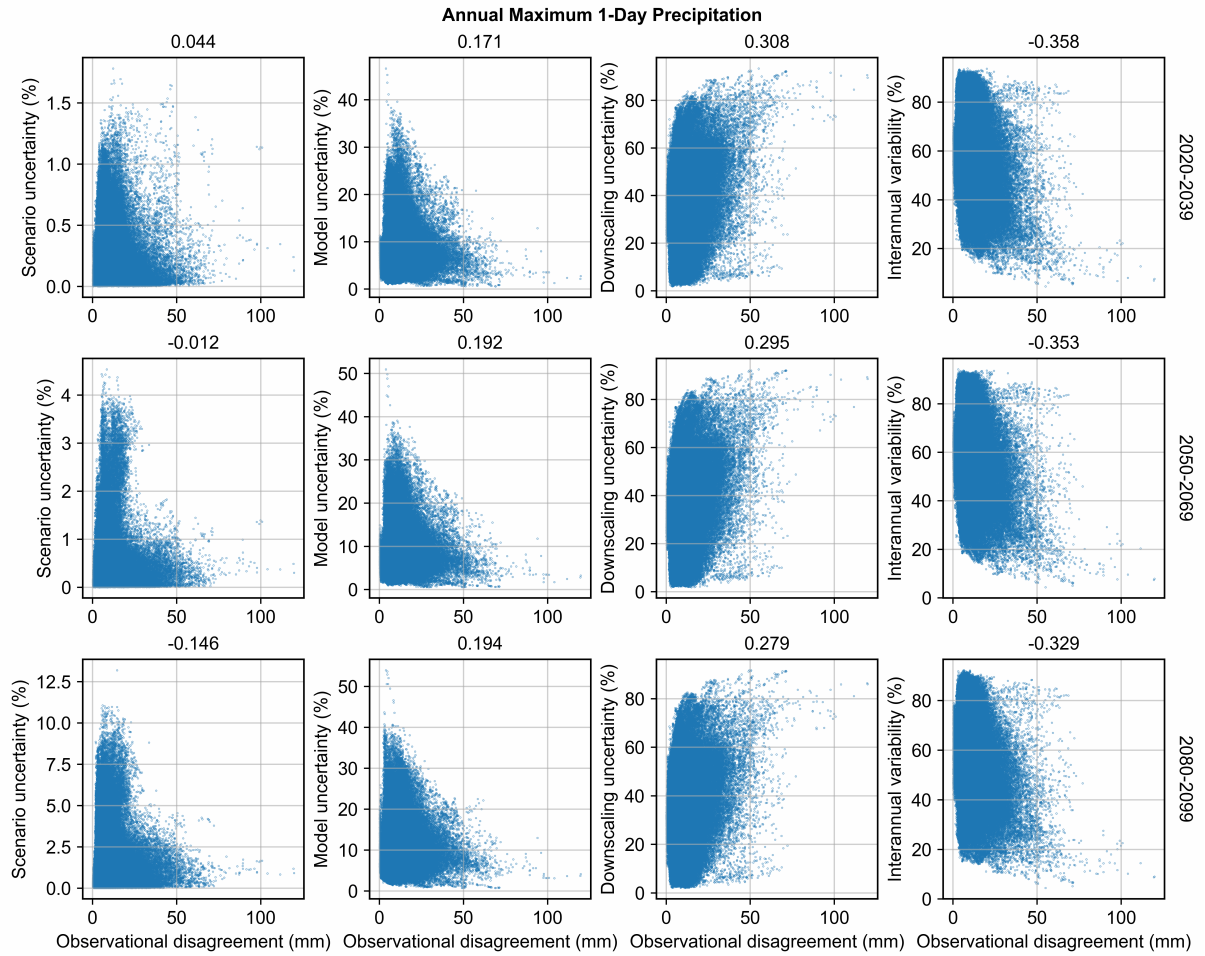
Supplementary Figure 5: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual average temperature and the associated observational uncertainty (measured by the standard deviation across observational datasets at that grid cell). The text above each subplot gives the Pearson correlation coefficient.



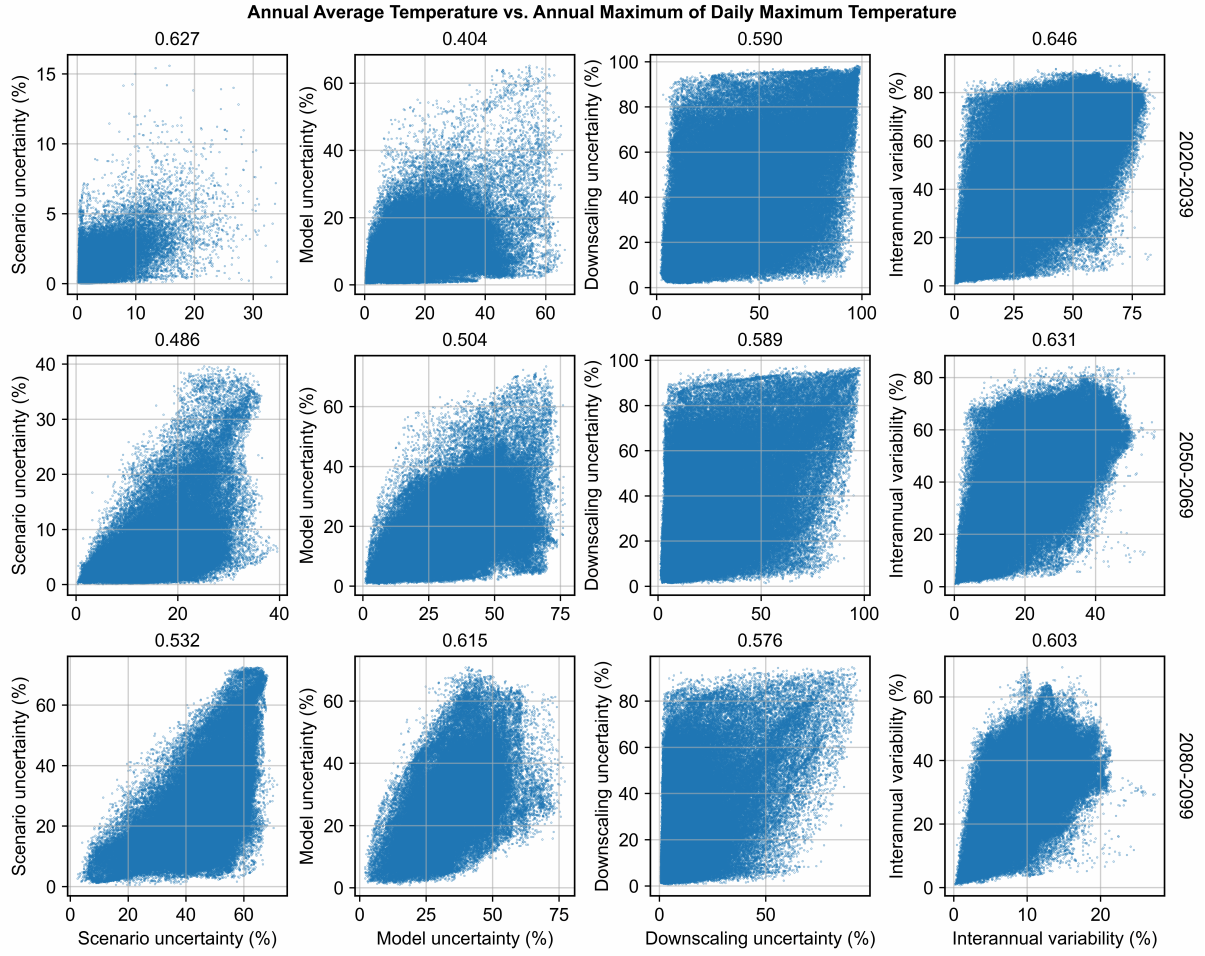
Supplementary Figure 6: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual total precipitation and the associated observational uncertainty (measured by the standard deviation across observational datasets at that grid cell). The text above each subplot gives the Pearson correlation coefficient.



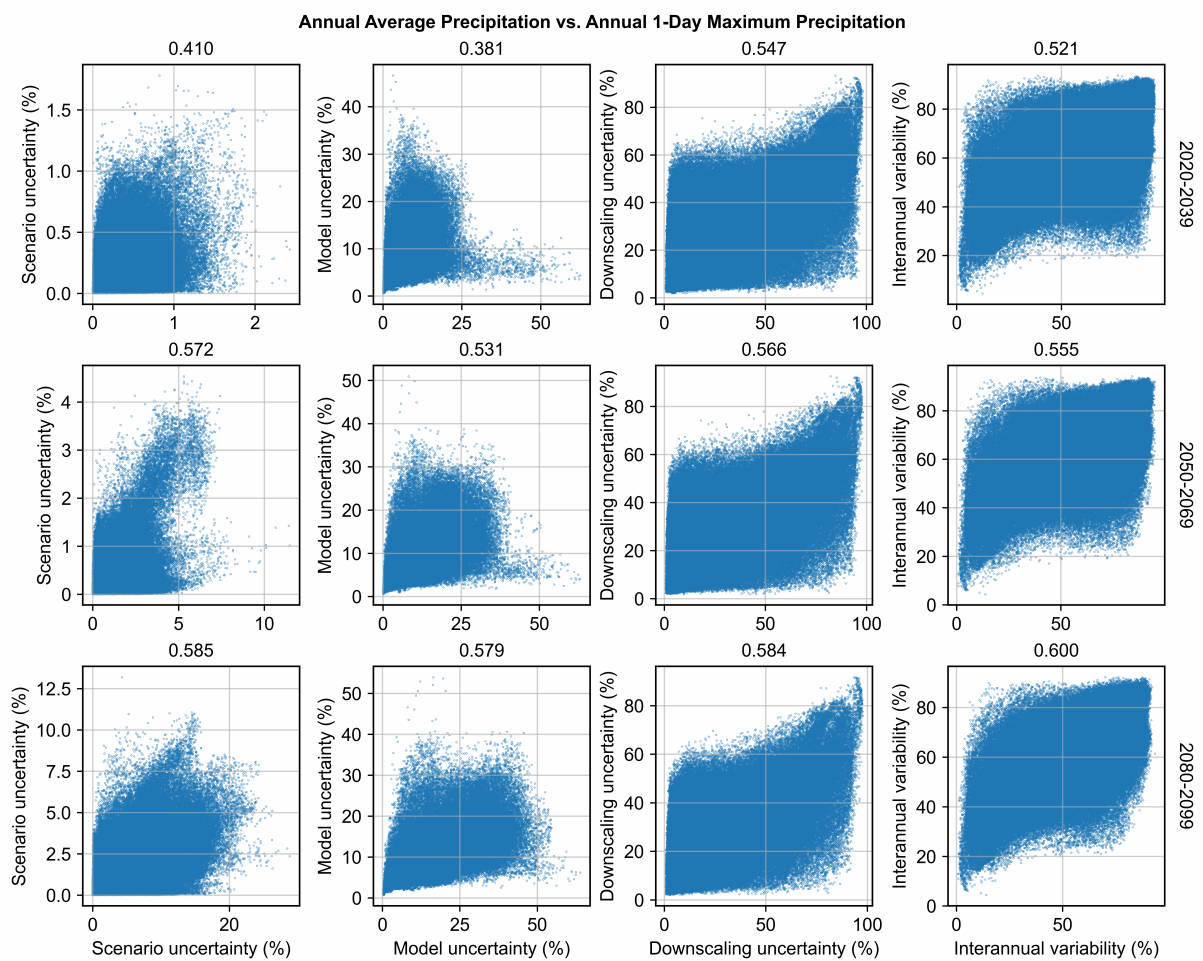
Supplementary Figure 7: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual maximum of daily maximum temperature and the associated observational uncertainty (measured by the standard deviation across observational datasets at that grid cell). The text above each subplot gives the Pearson correlation coefficient.



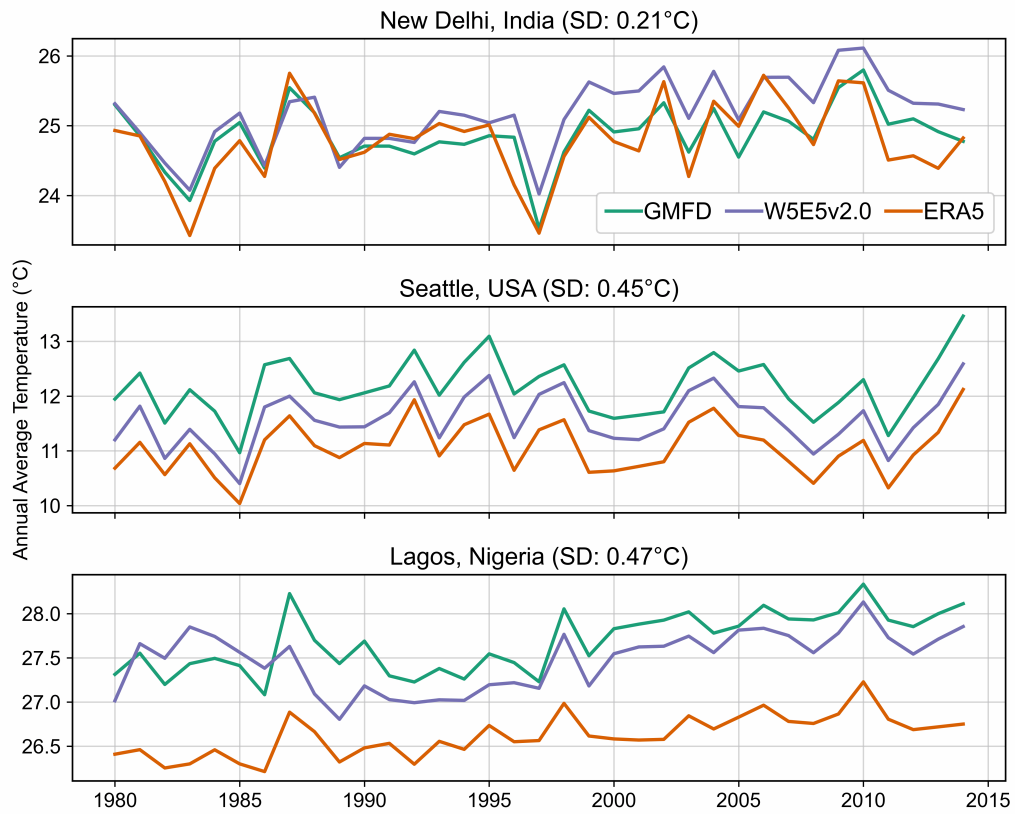
Supplementary Figure 8: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual 1-day precipitation maxima and the associated observational uncertainty (measured by the standard deviation across observational datasets at that grid cell). The text above each subplot gives the Pearson correlation coefficient.



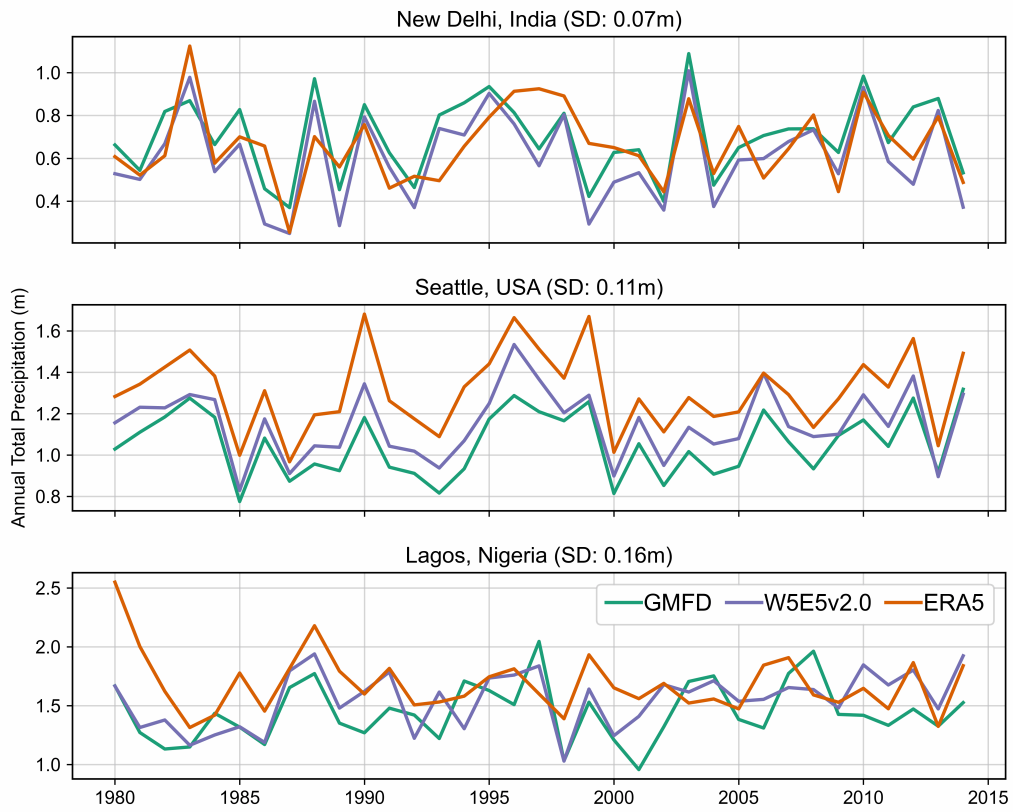
Supplementary Figure 9: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual average temperature and annual maximum of daily maximum temperature. The text above each subplot gives the Pearson correlation coefficient.



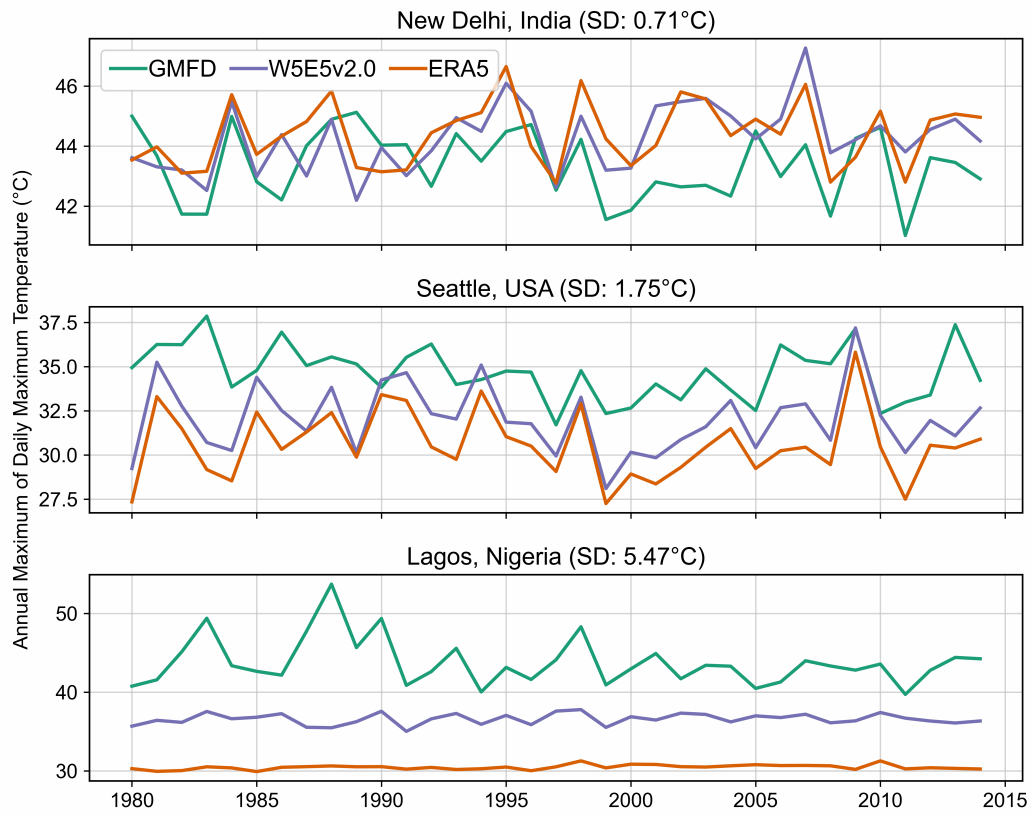
Supplementary Figure 10: **Spatial correlations:** Scatter plots showing the grid cell correlation between the uncertainty decomposition results for annual total precipitation and annual 1-day precipitation maxima. The text above each subplot gives the Pearson correlation coefficient.



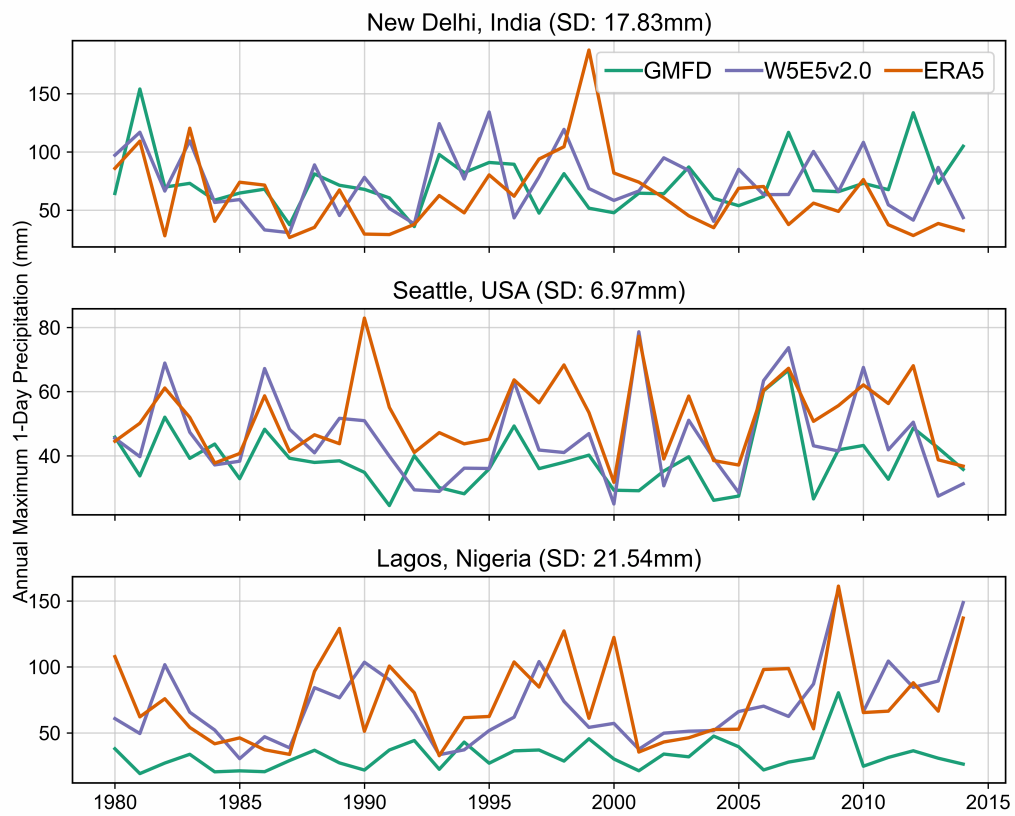
Supplementary Figure 11: **Observational dataset comparison:** Annual average temperature observations in New Delhi, Seattle, Lagos. GMFD observations are shown in green, W5E5v2.0 in purple, and ERA5 in orange. The standard deviation reported at the top of each subplot is calculated as the annual standard deviation across datasets, averaged over 1980-2014.



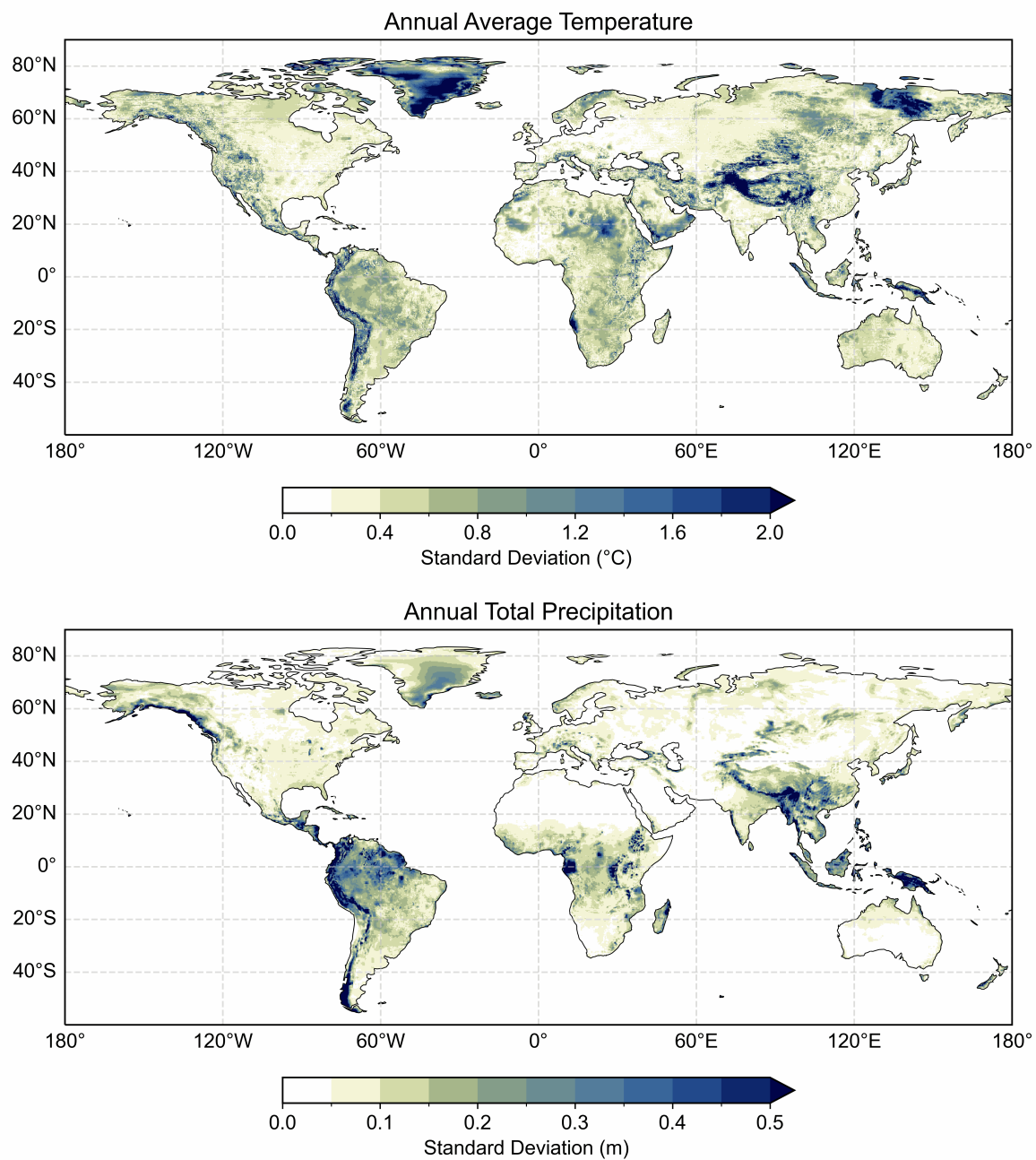
Supplementary Figure 12: **Observational dataset comparison:** Annual total precipitation observations in New Delhi, Seattle, Lagos. GMFD observations are shown in green, W5E5v2.0 in purple, and ERA5 in orange. The standard deviation reported at the top of each subplot is calculated as the annual standard deviation across datasets, averaged over 1980-2014.



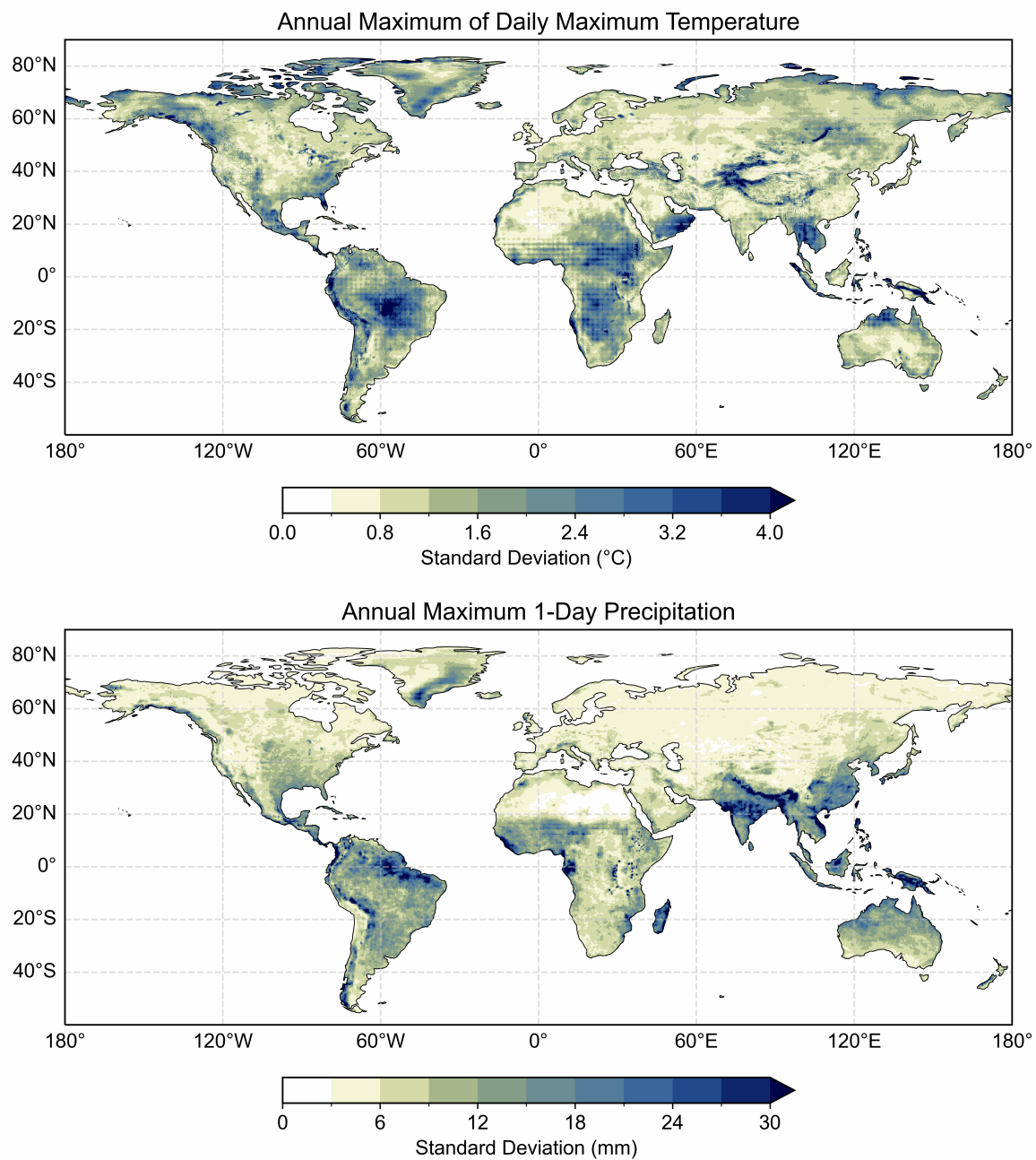
Supplementary Figure 13: **Observational dataset comparison:** Annual maximum of daily maximum temperature observations in New Delhi, Seattle, Lagos. GMFD observations are shown in green, W5E5v2.0 in purple, and ERA5 in orange. The standard deviation reported at the top of each subplot is calculated as the annual standard deviation across datasets, averaged over 1980-2014.



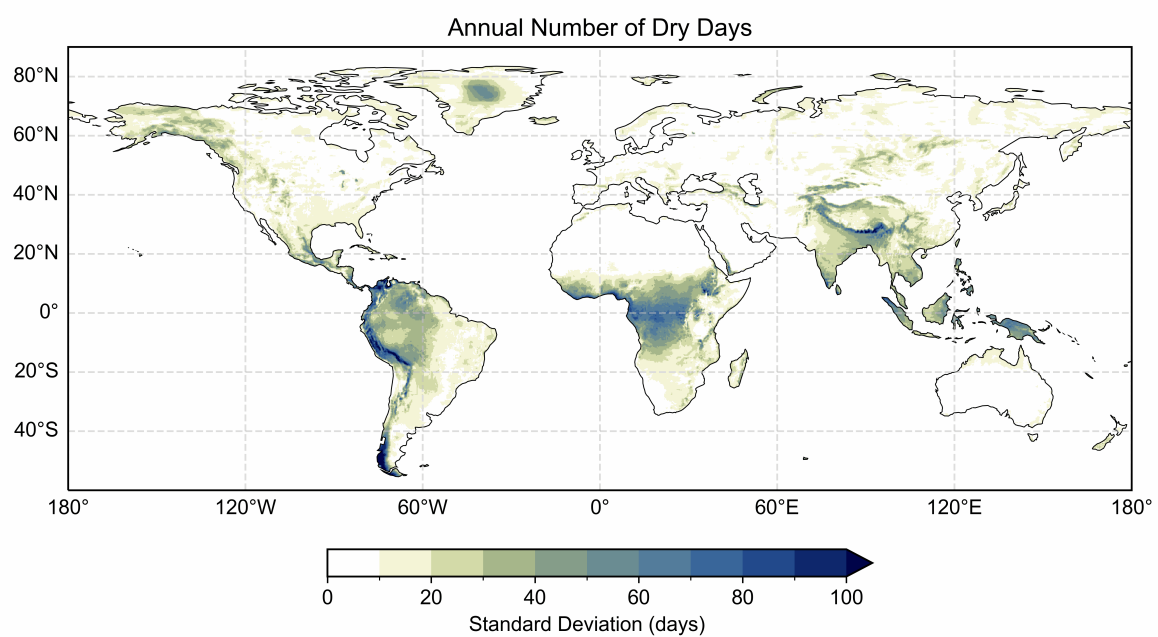
Supplementary Figure 14: **Observational dataset comparison:** Annual 1-day precipitation maxima observations in New Delhi, Seattle, Lagos. GMFD observations are shown in green, W5E5v2.0 in purple, and ERA5 in orange. The standard deviation reported at the top of each subplot is calculated as the annual standard deviation across datasets, averaged over 1980-2014.



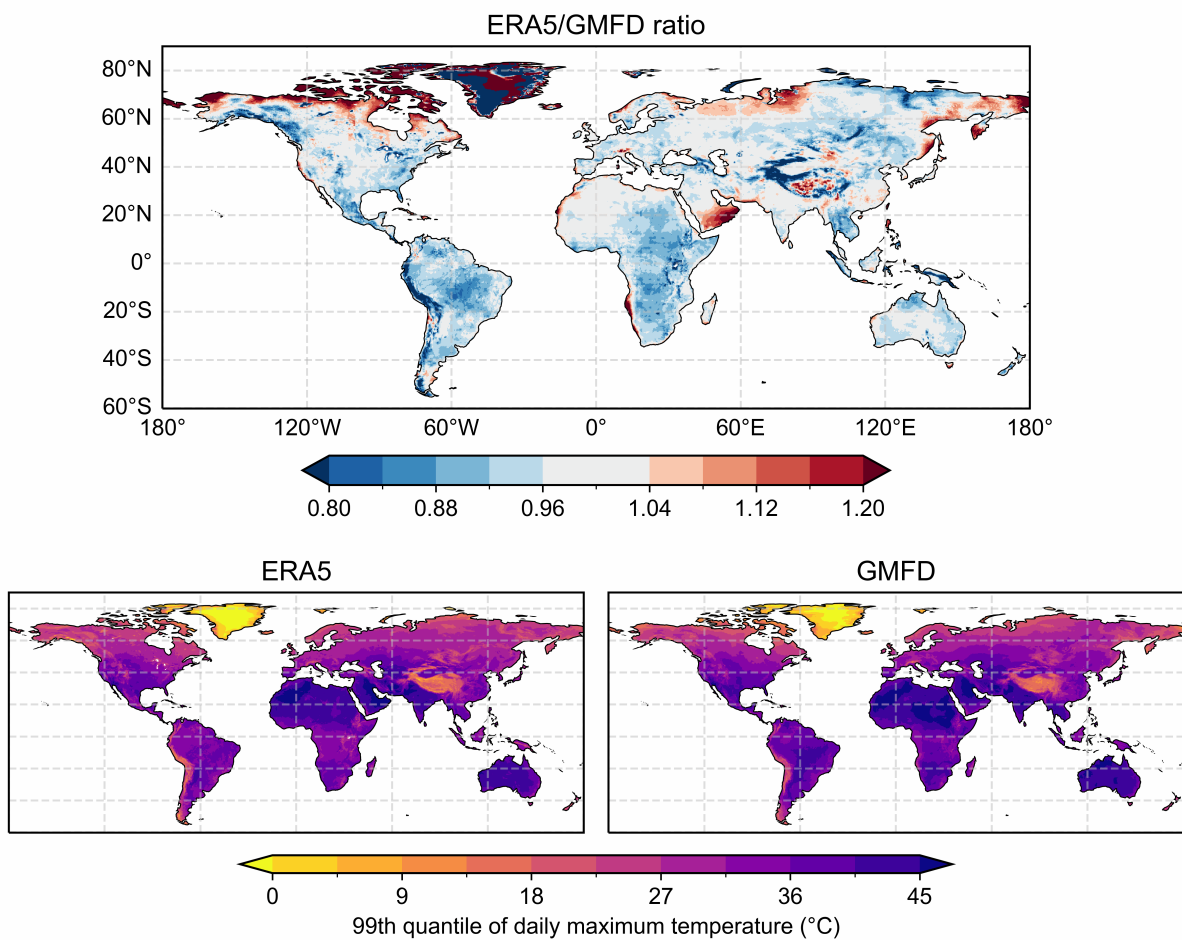
Supplementary Figure 15: **Observational dataset comparison:** Observational disagreement in annual averages. The standard deviation is calculated across the GMFD, W5E5v2.0, and ERA5 datasets then averaged over 1980-2014.



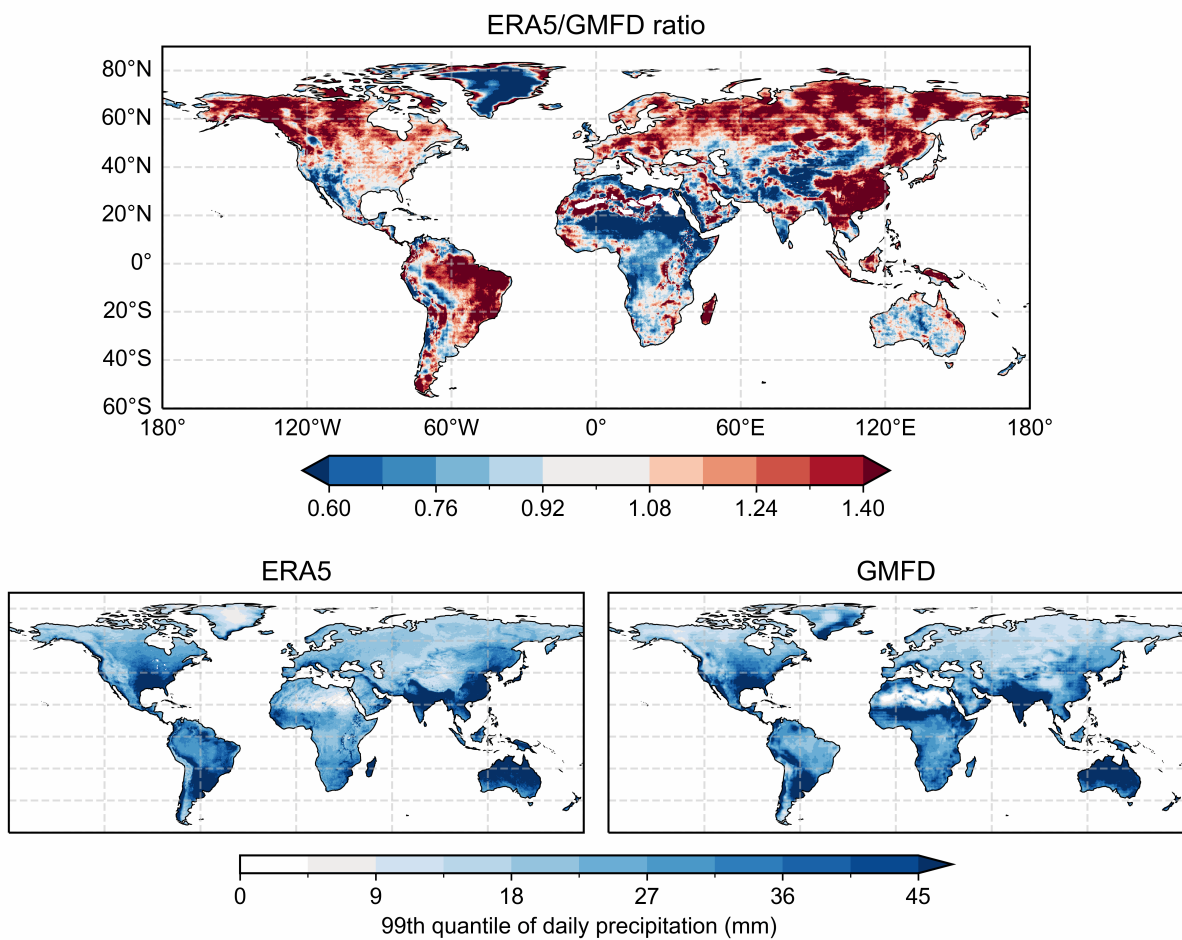
Supplementary Figure 16: **Observational dataset comparison:** Observational disagreement in 1-day maxima. The standard deviation is calculated across the GMFD, W5E5v2.0, and ERA5 datasets then averaged over 1980-2014.



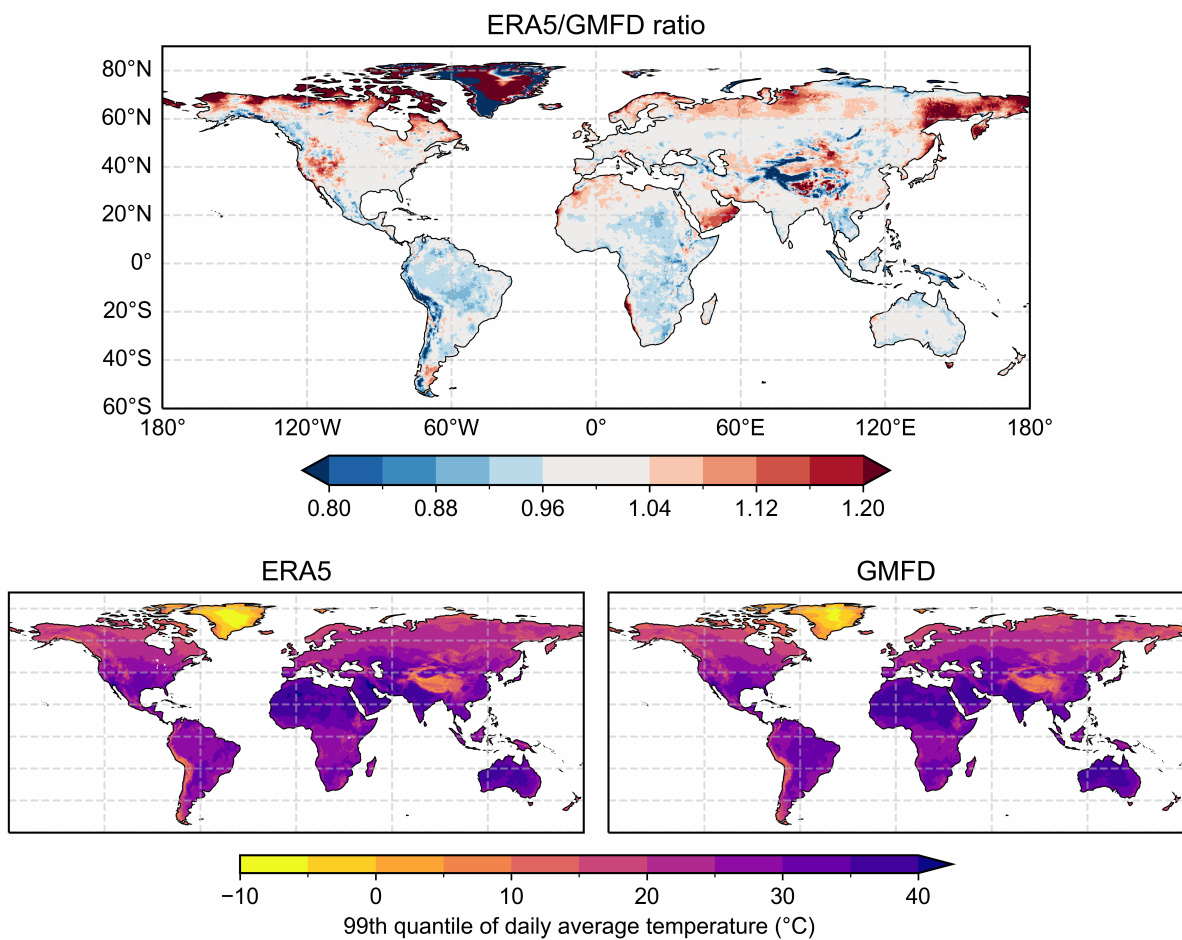
Supplementary Figure 17: **Observational dataset comparison:** Observational disagreement in dry days. The standard deviation is calculated across the GMFD, W5E5v2.0, and ERA5 datasets then averaged over 1980-2014.



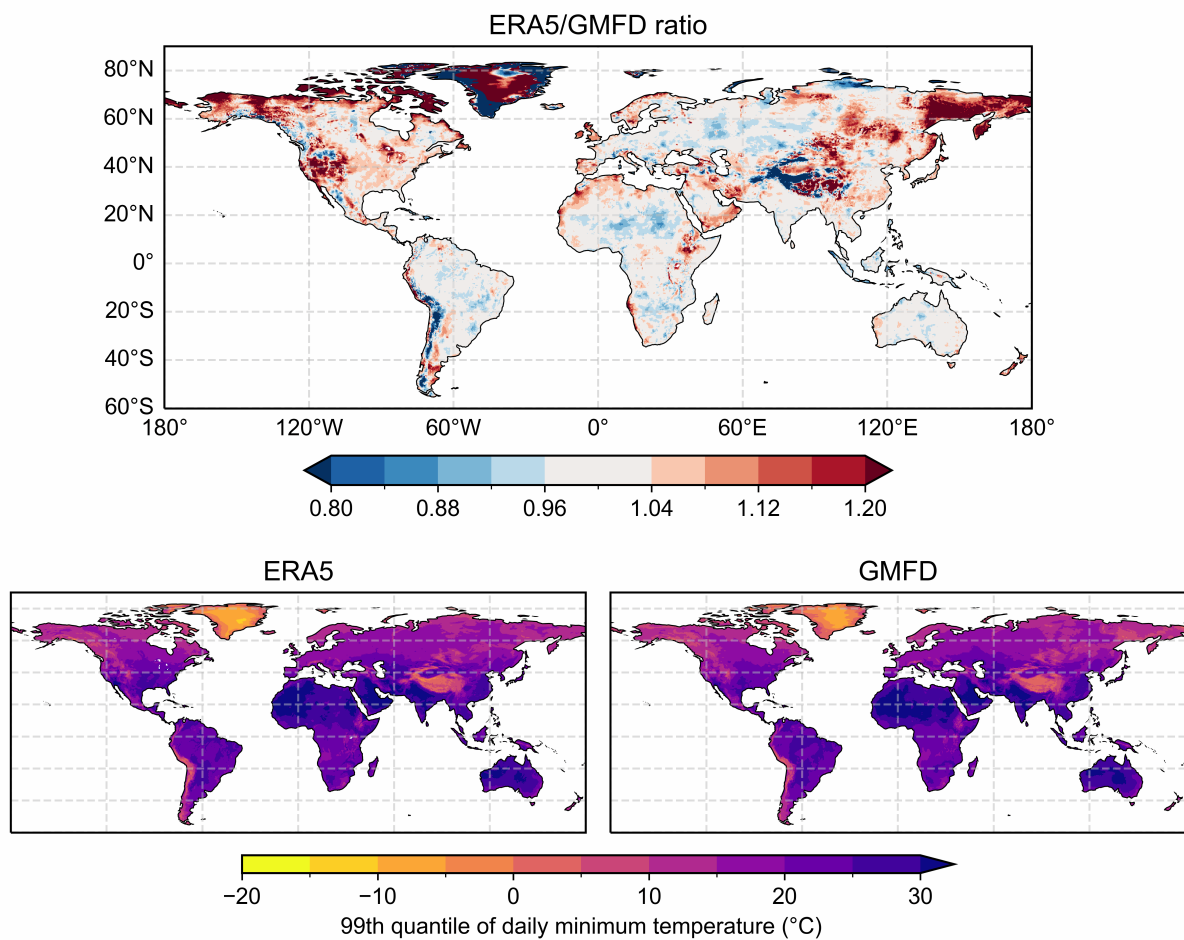
Supplementary Figure 18: **Observational dataset comparison:** Observational disagreement in 99th tasmax percentile. The top subplots shows the ERA5/GMFD ratio of the 99th daily maximum temperature percentile calculated over 1980-2014. The bottom subplots shows the percentile values for each dataset.



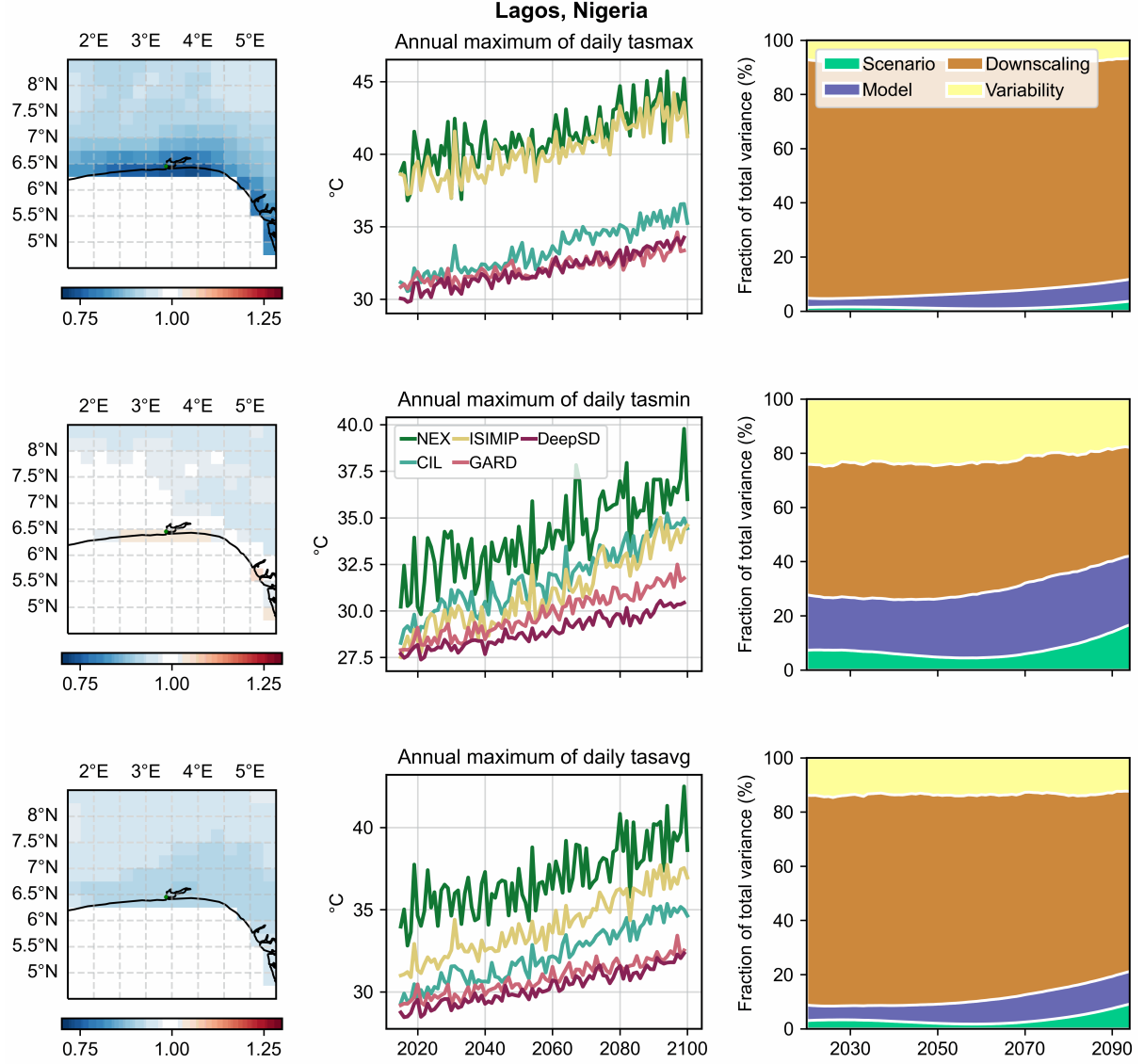
Supplementary Figure 19: **Observational dataset comparison:** Observational disagreement in 99th precipitation percentile. The top subplots shows the ERA5/GMFD ratio of the 99th daily precipitation percentile calculated over 1980-2014. The bottom subplots shows the percentile values for each dataset.



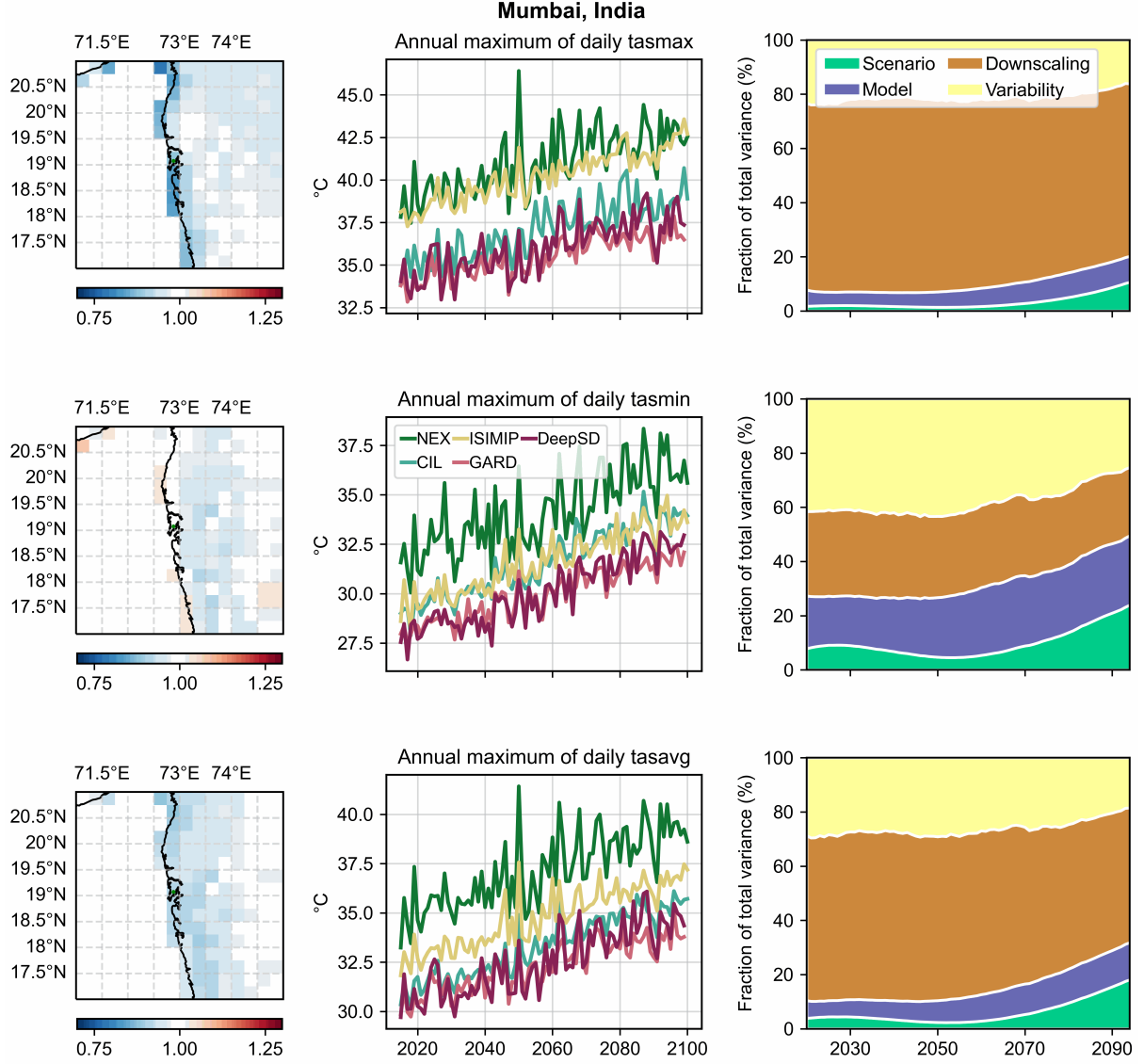
Supplementary Figure 20: **Observational dataset comparison:** Observational disagreement in 99th tas percentile. The top subplots shows the ERA5/GMFD ratio of the 99th daily average temperature percentile calculated over 1980-2014. The bottom subplots shows the percentile values for each dataset.



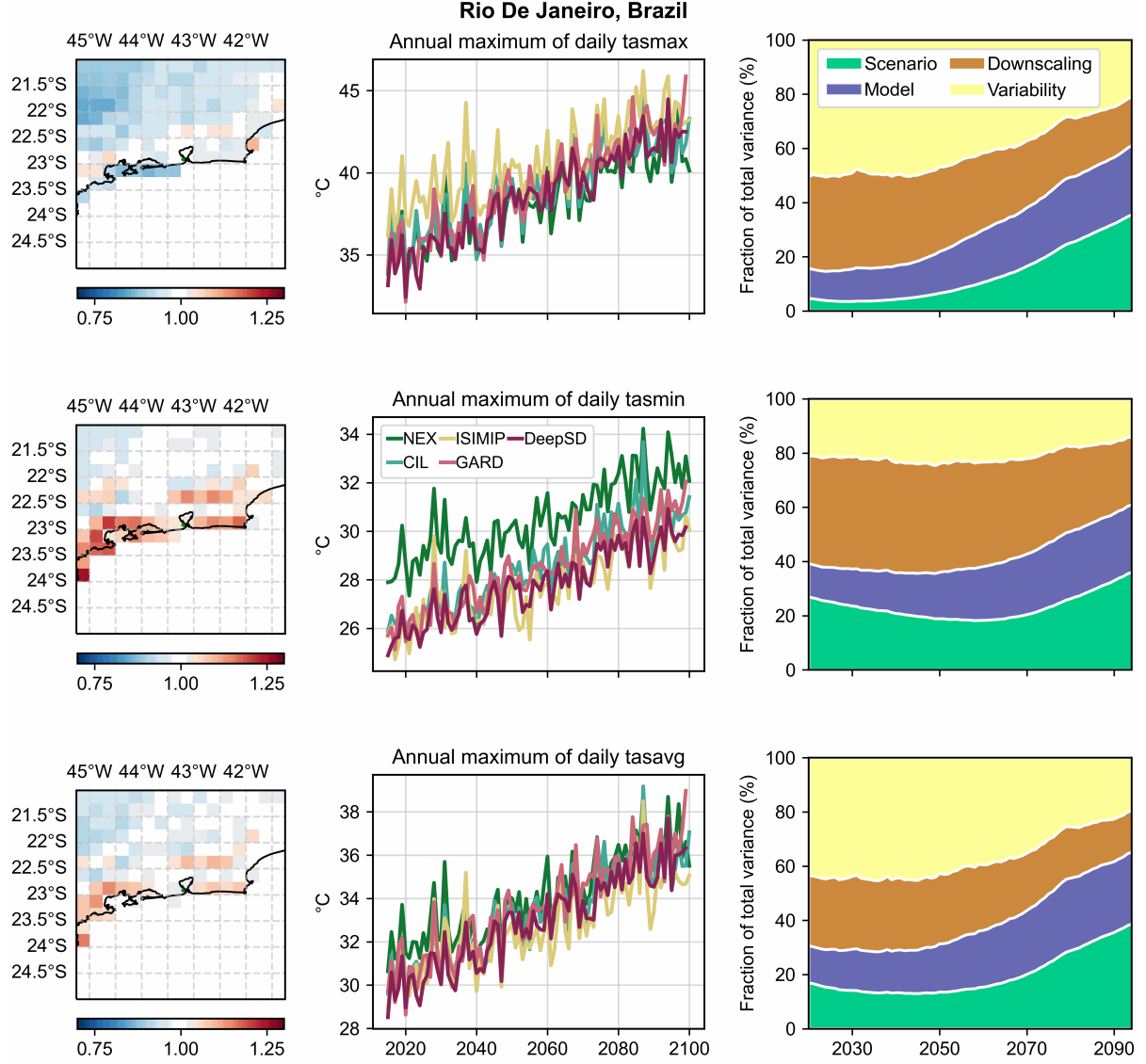
Supplementary Figure 21: **Observational dataset comparison:** Observational disagreement in 99th tasmin percentile. The top subplots shows the ERA5/GMFD ratio of the 99th daily minimum temperature percentile calculated over 1980-2014. The bottom subplots shows the percentile values for each dataset.



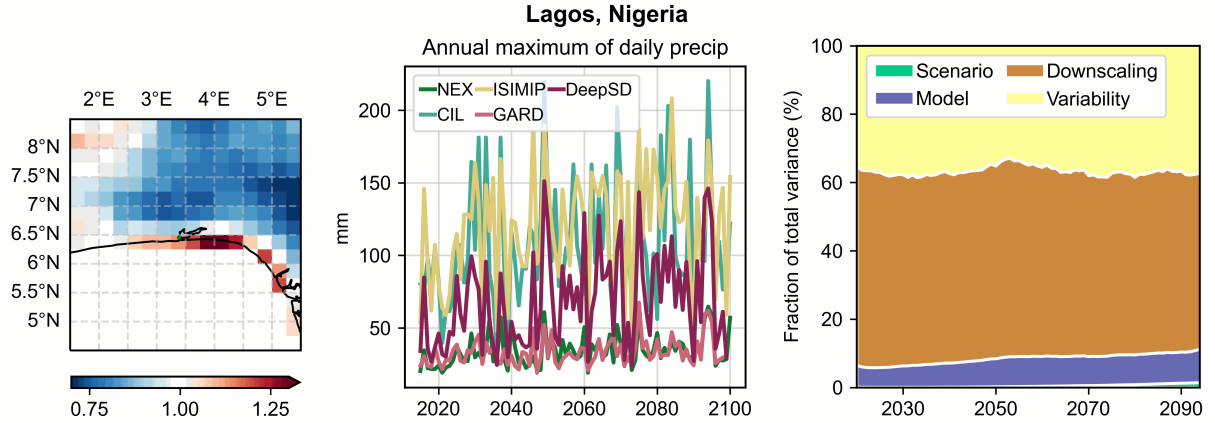
Supplementary Figure 22: **Observational dataset comparison:** Coastal effects for projections of temperature extremes in Lagos. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Lagos. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper). The top, middle, and bottom rows show the results for daily maximum, minimum, and average temperature, respectively.



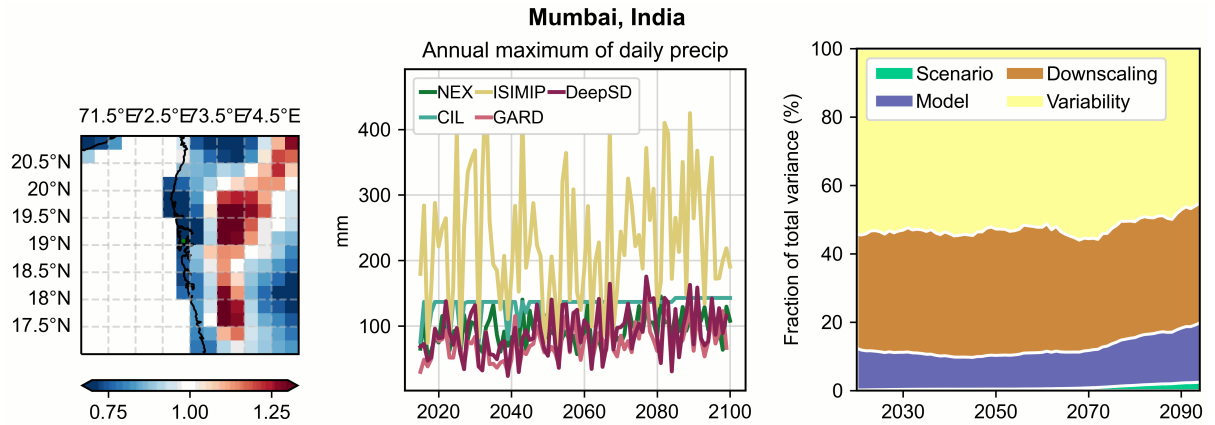
Supplementary Figure 23: **Observational dataset comparison:** Coastal effects for projections of temperature extremes in Mumbai. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Mumbai. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper). The top, middle, and bottom rows show the results for daily maximum, minimum, and average temperature, respectively.



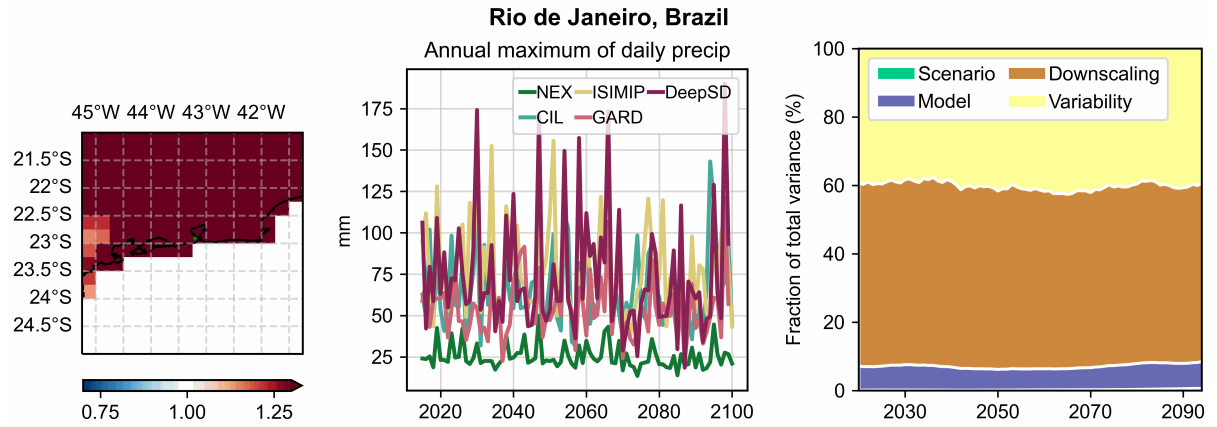
Supplementary Figure 24: **Observational dataset comparison:** Coastal effects for projections of temperature extremes in Rio. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Rio. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper). The top, middle, and bottom rows show the results for daily maximum, minimum, and average temperature, respectively.



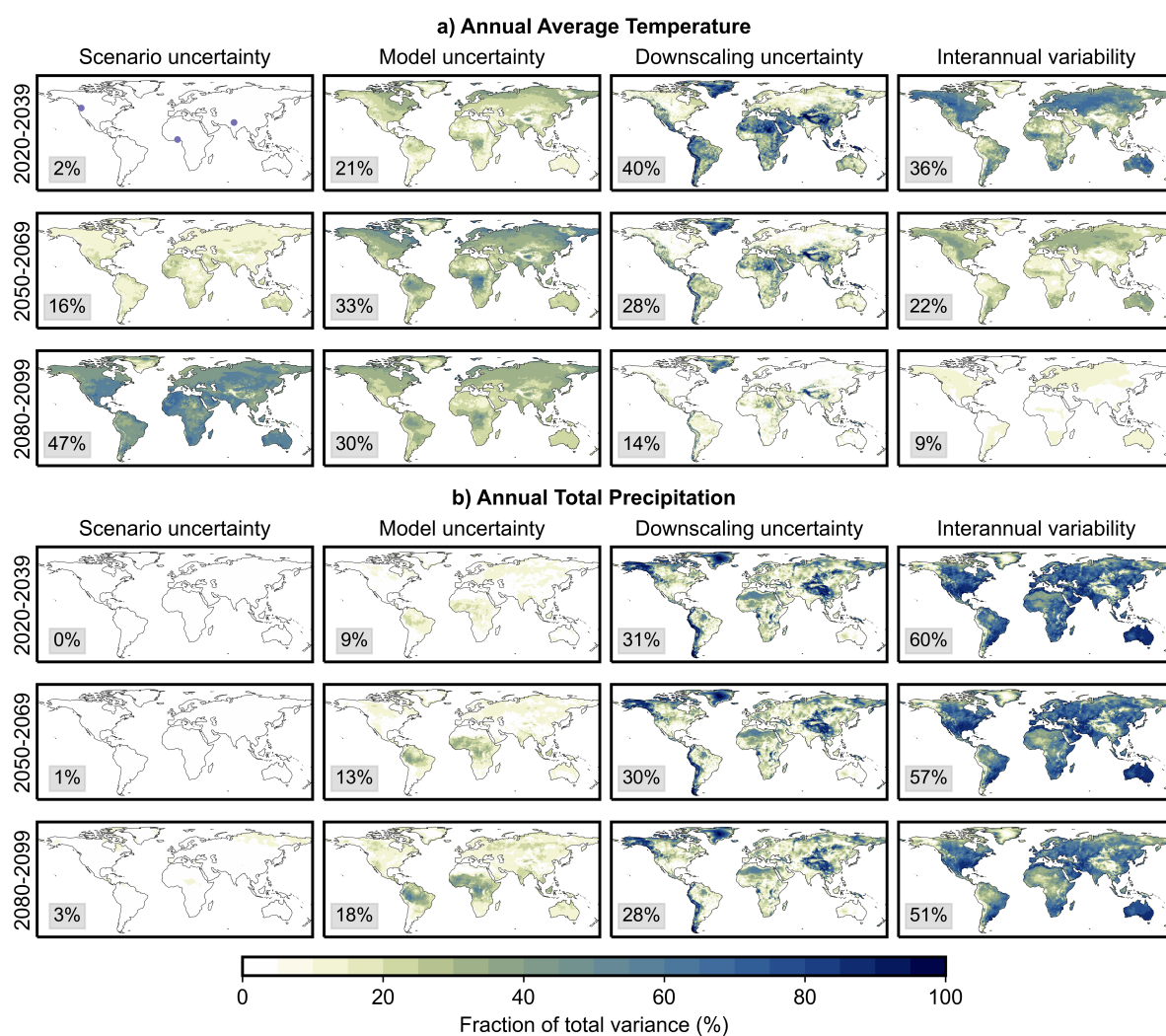
Supplementary Figure 25: **Observational dataset comparison:** Coastal effects for projections of precipitation extremes in Lagos. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Lagos. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper).



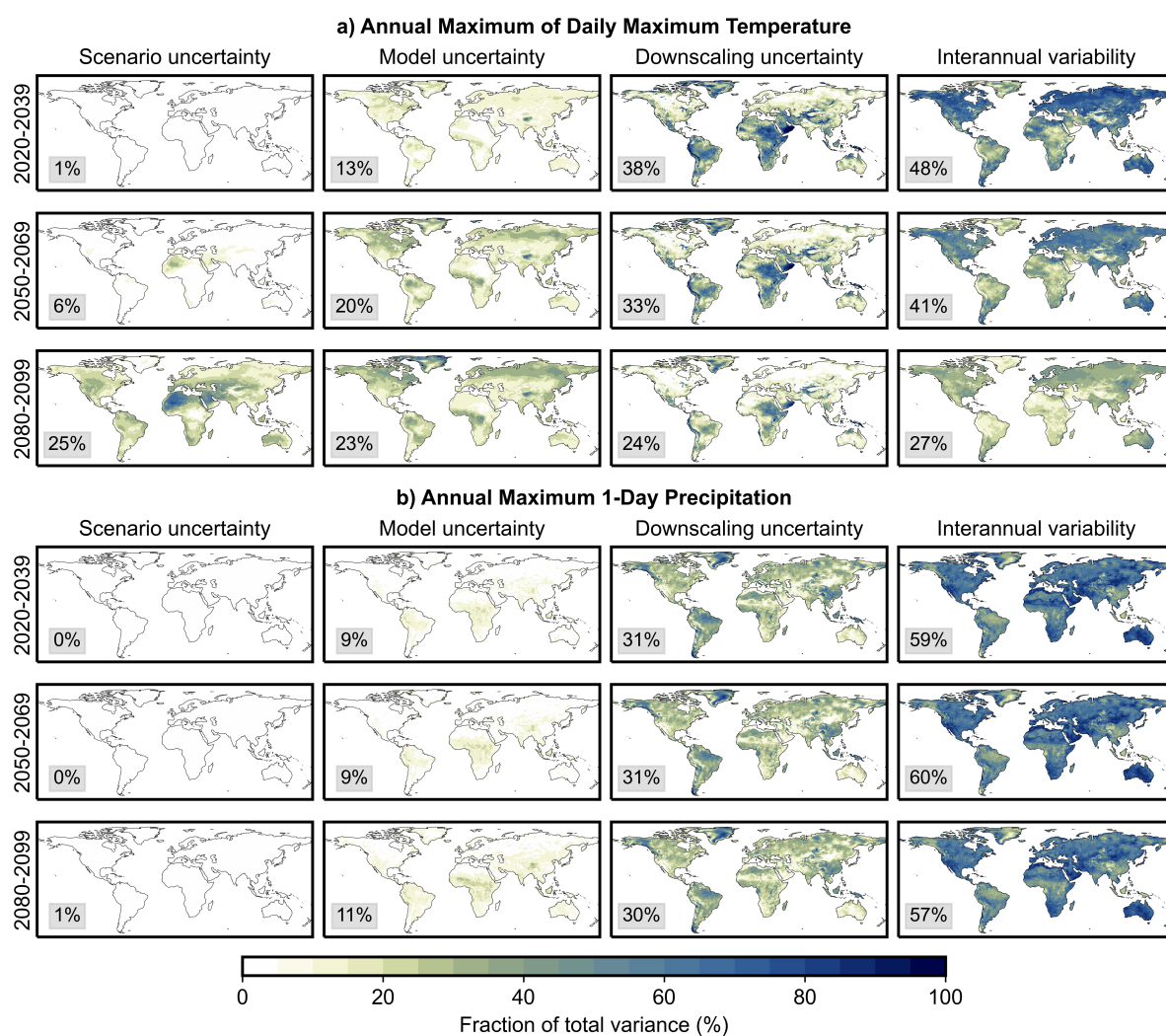
Supplementary Figure 26: **Observational dataset comparison:** Coastal effects for projections of precipitation extremes in Mumbai. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Mumbai. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper).



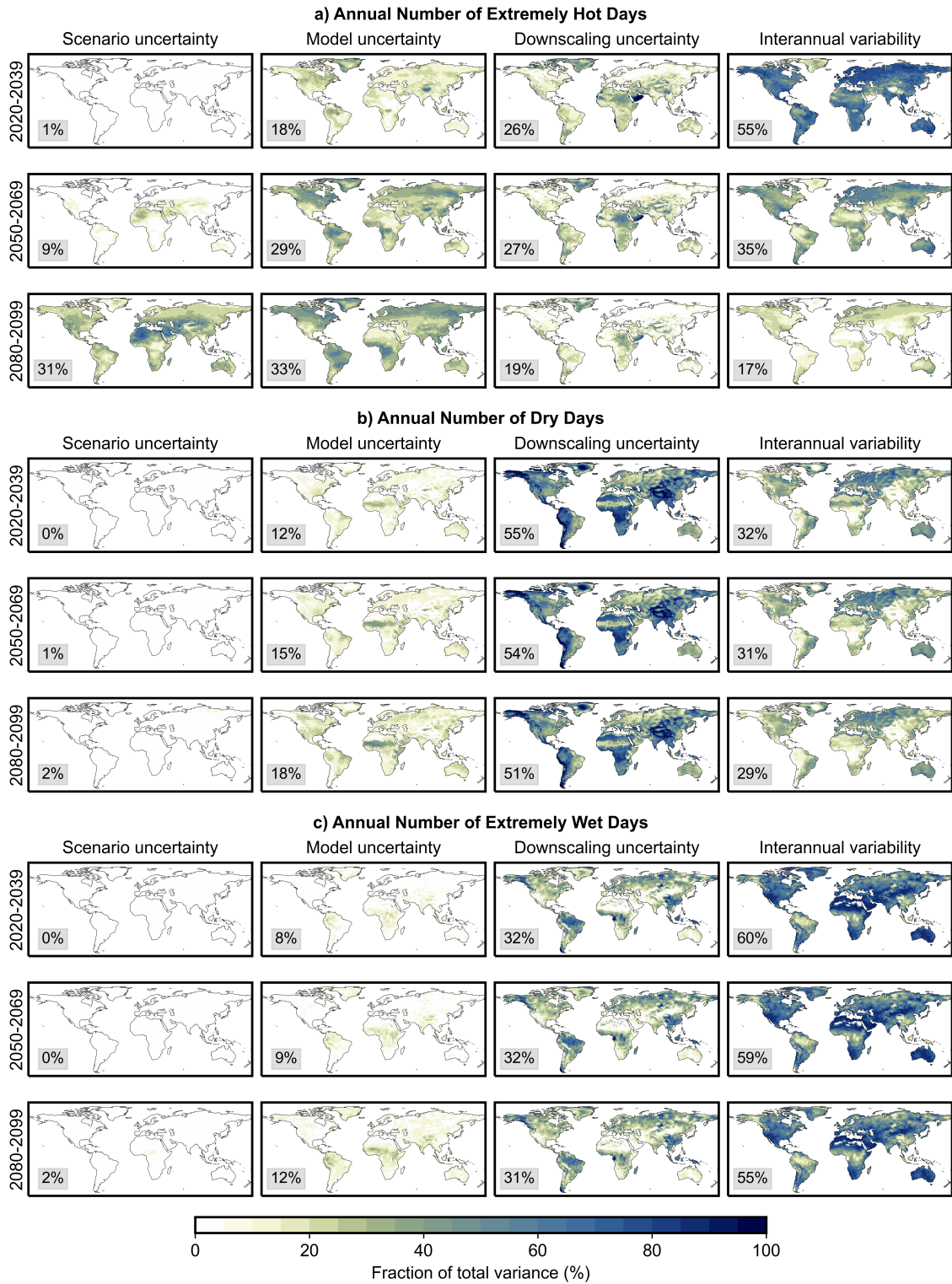
Supplementary Figure 27: **Observational dataset comparison:** Coastal effects for projections of precipitation extremes in Rio. Left column shows the ERA5/GMFD ratio of the 99th percentiles of each variable. The green dot on the map shows the location of Rio. Central column shows downscaled projections of annual maxima from CanESM5 under SSP3-7.0 for each ensemble. Recall that NEX-GDDP is downscaled and bias-corrected against GMFD and CIL-GDPCIR, GARD-SV, and DeepSD-BC are downscaled and bias-corrected against ERA5. Right column shows the variance decomposition when all models and scenarios are taken into account (i.e., the results from the main paper).



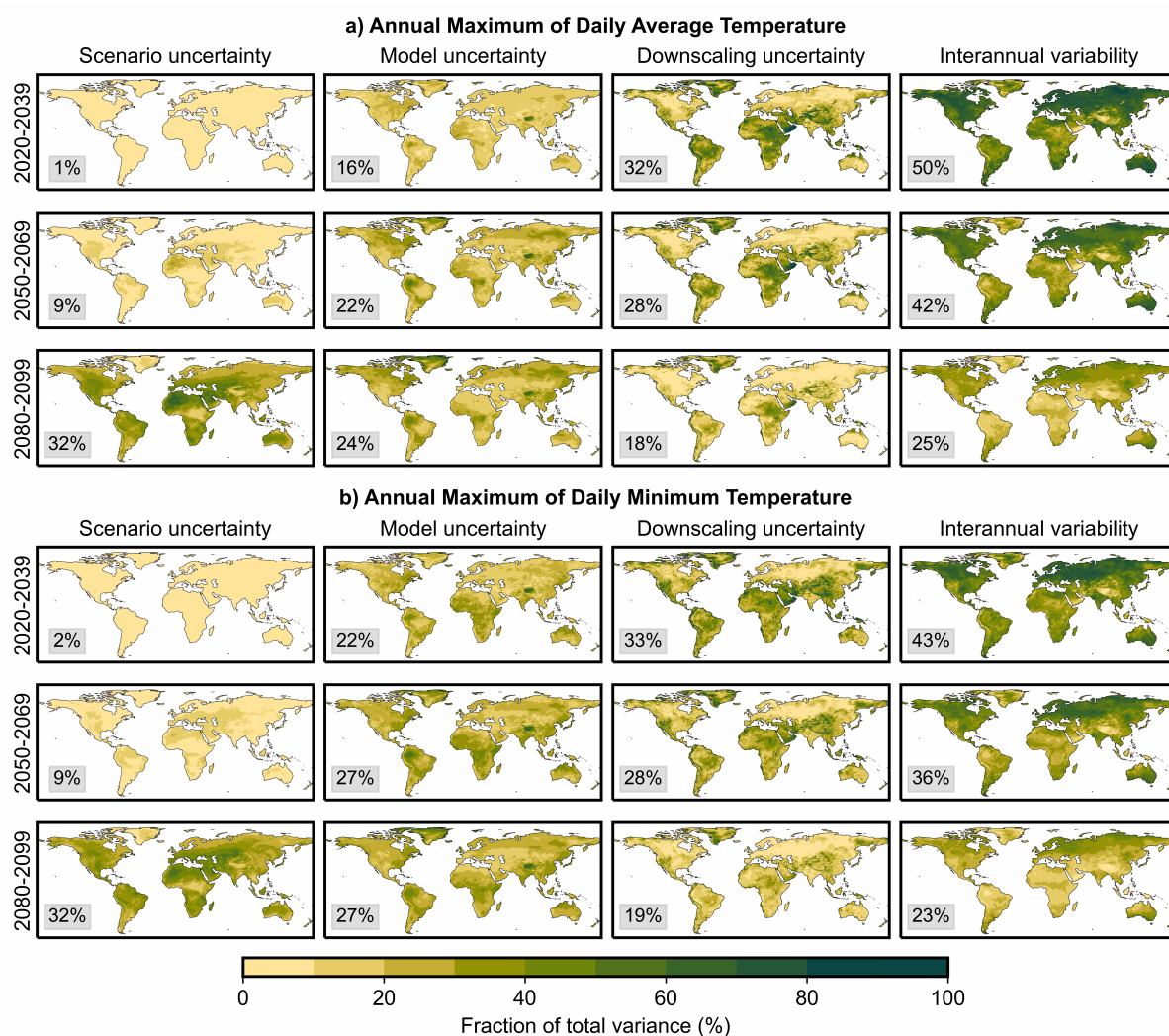
Supplementary Figure 28: **Alternative colormap:** As Figure 2 in the main text, but with an alternative and more granular colormap.



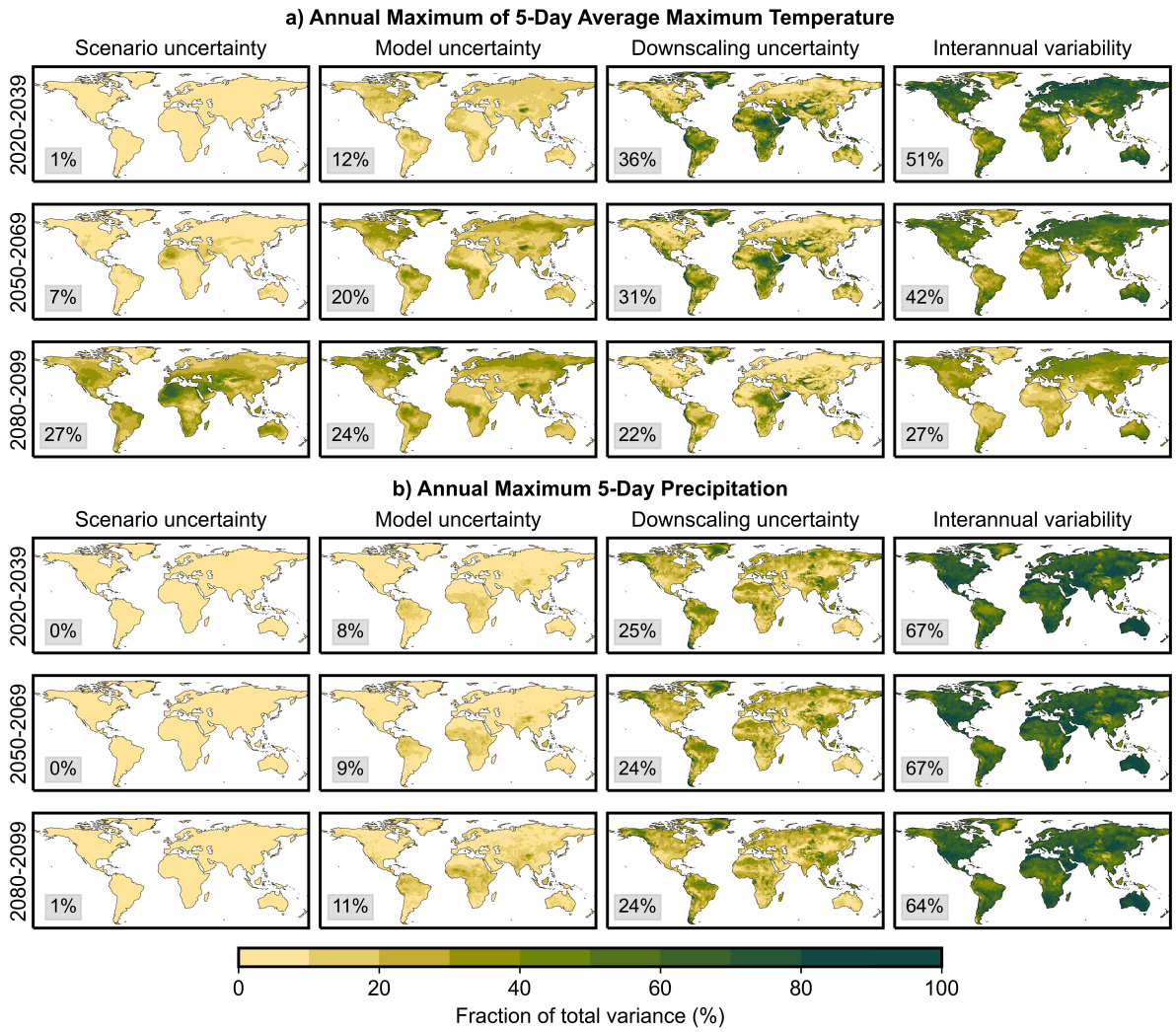
Supplementary Figure 29: **Alternative colormap:** As Figure 3 in the main text, but with an alternative and more granular colormap.



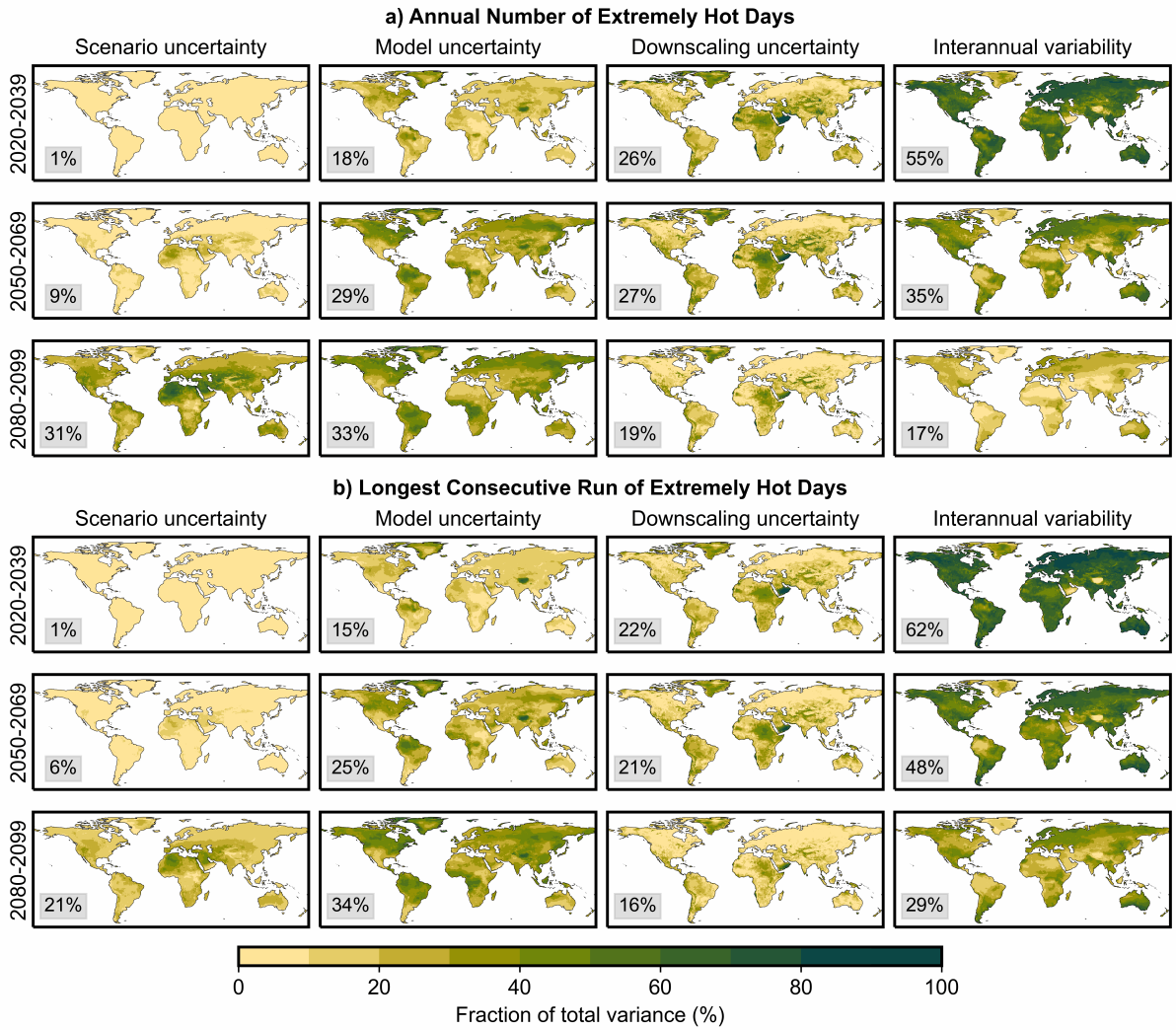
Supplementary Figure 30: **Alternative colormap:** As Figure 4 in the main text, but with an alternative and more granular colormap.



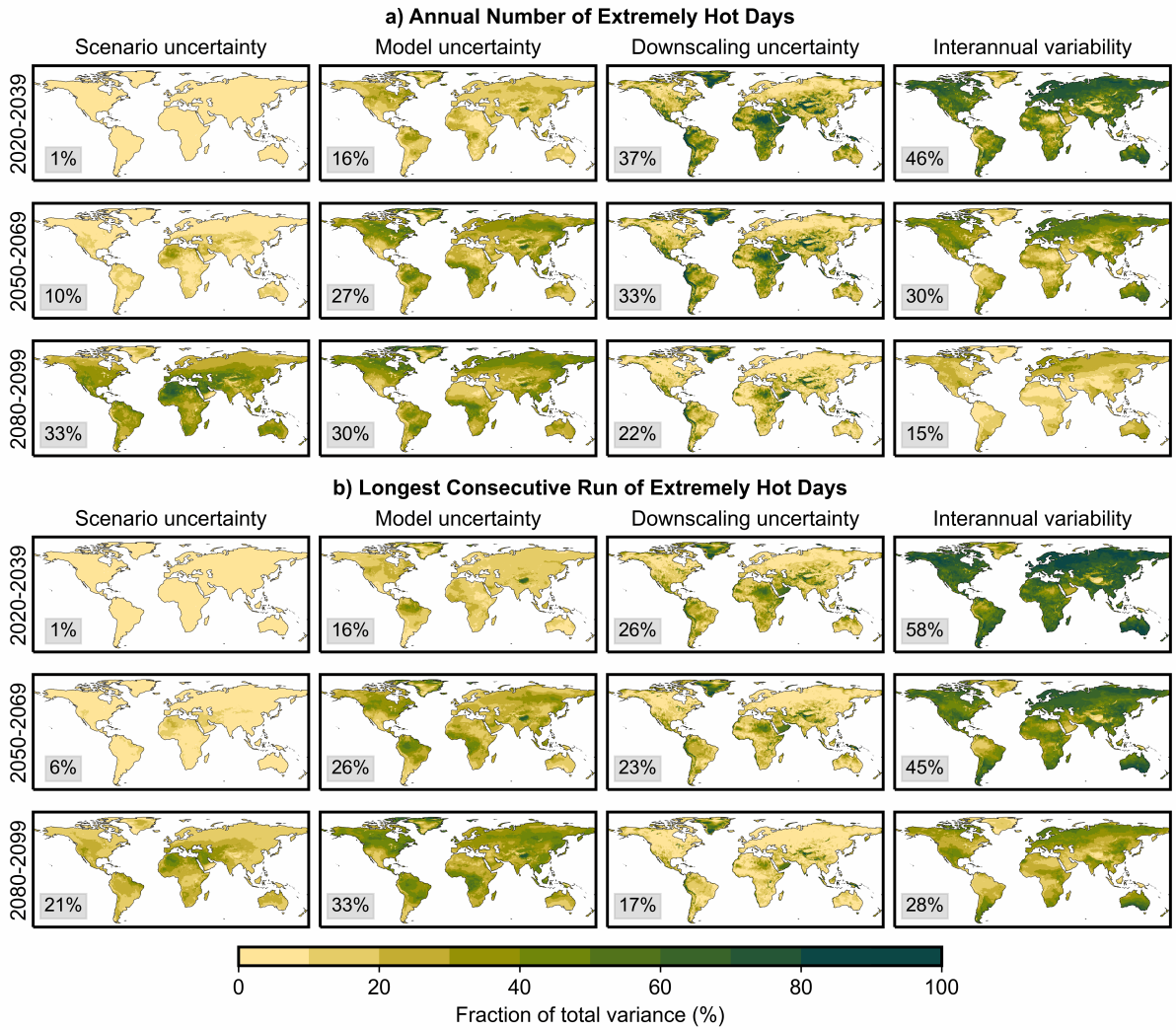
Supplementary Figure 31: **Alternative definition of extremes:** Variance decomposition for 1-day annual temperature maxima calculated using (a) daily average and (b) daily minimum temperatures.



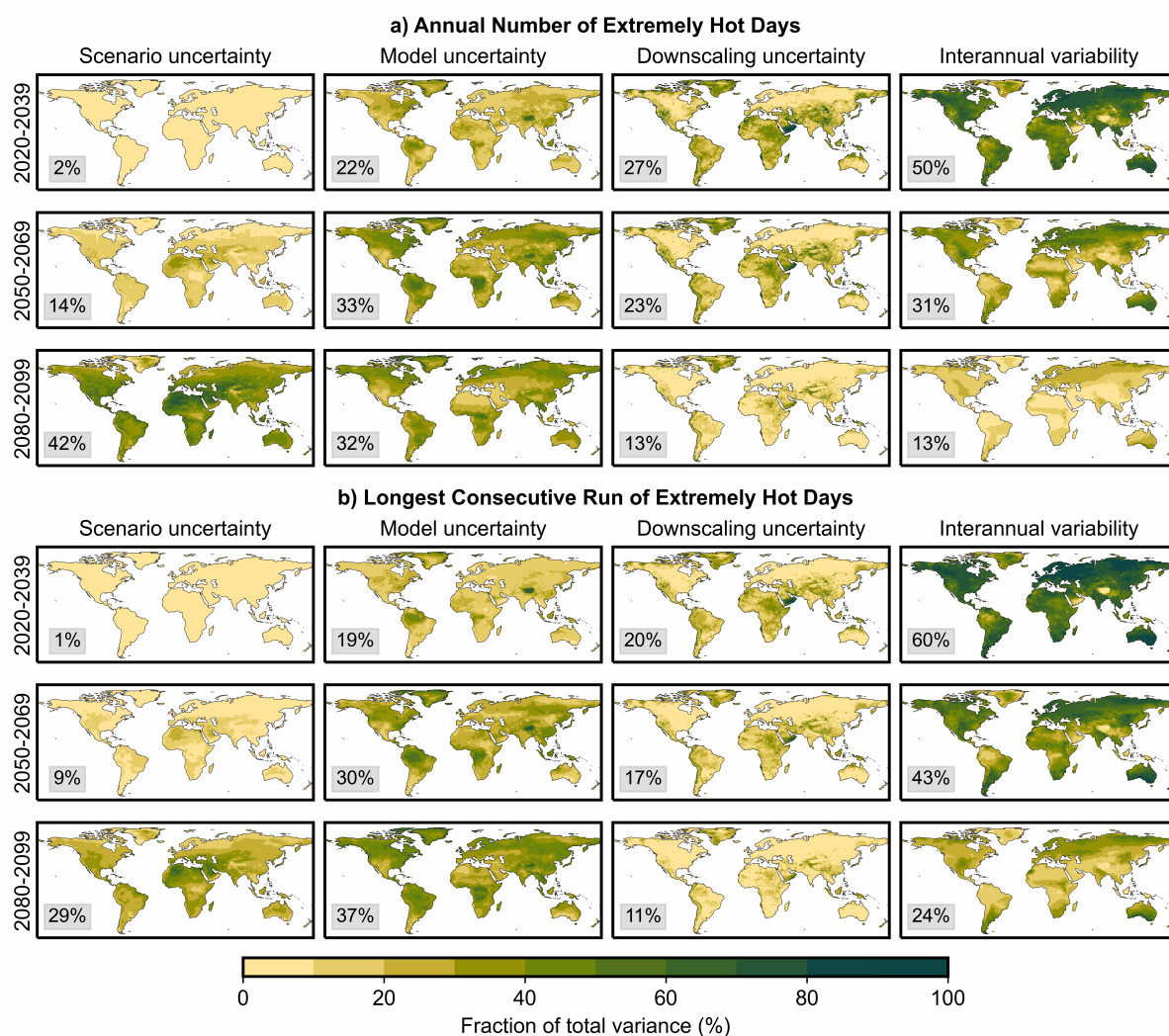
Supplementary Figure 32: **Alternative definition of extremes:** Variance decomposition for 5-day annual maxima of (a) daily maximum temperature and (b) daily precipitation.



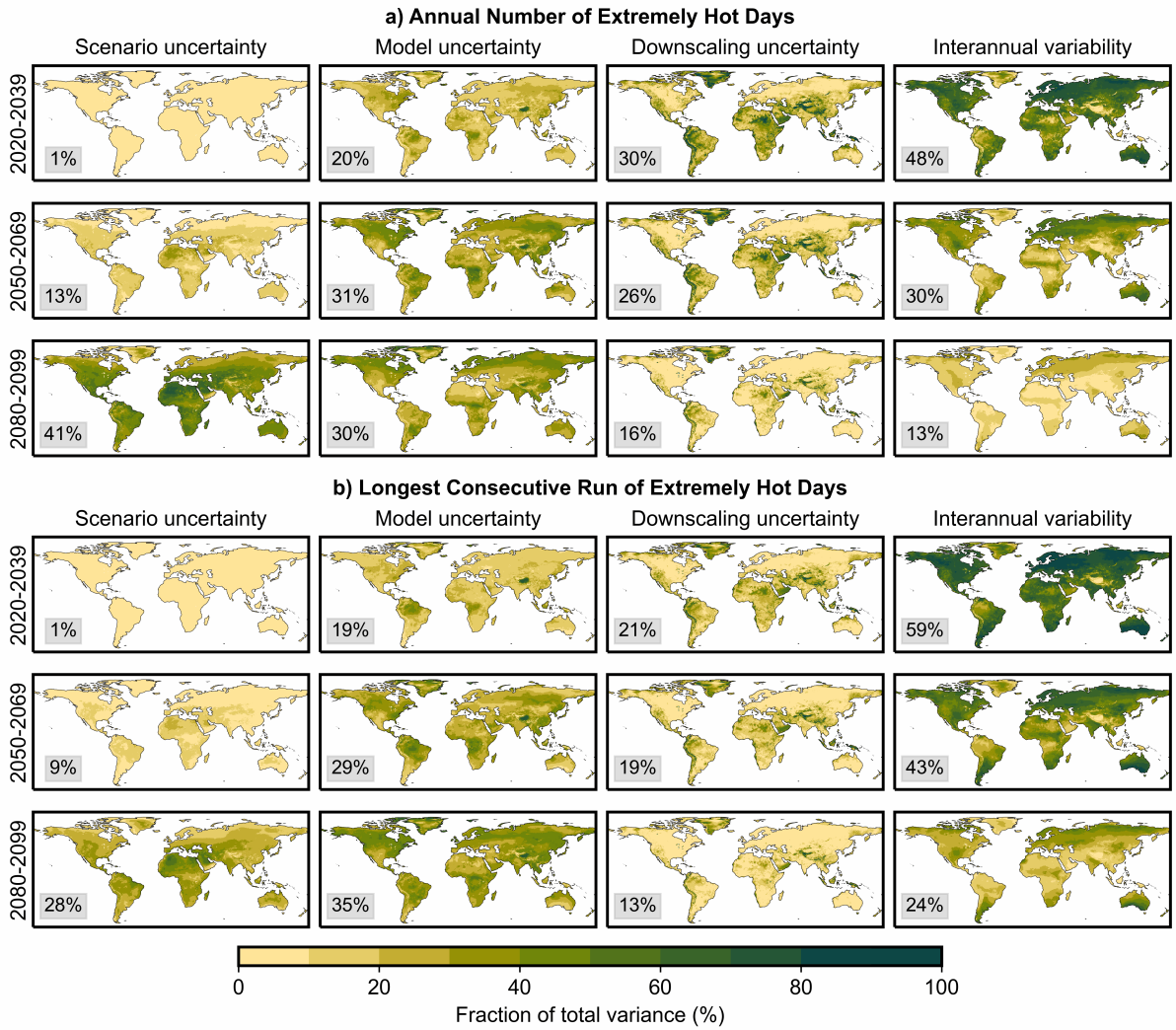
Supplementary Figure 33: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily maximum temperature exceeds the 99th percentile in GMFD.



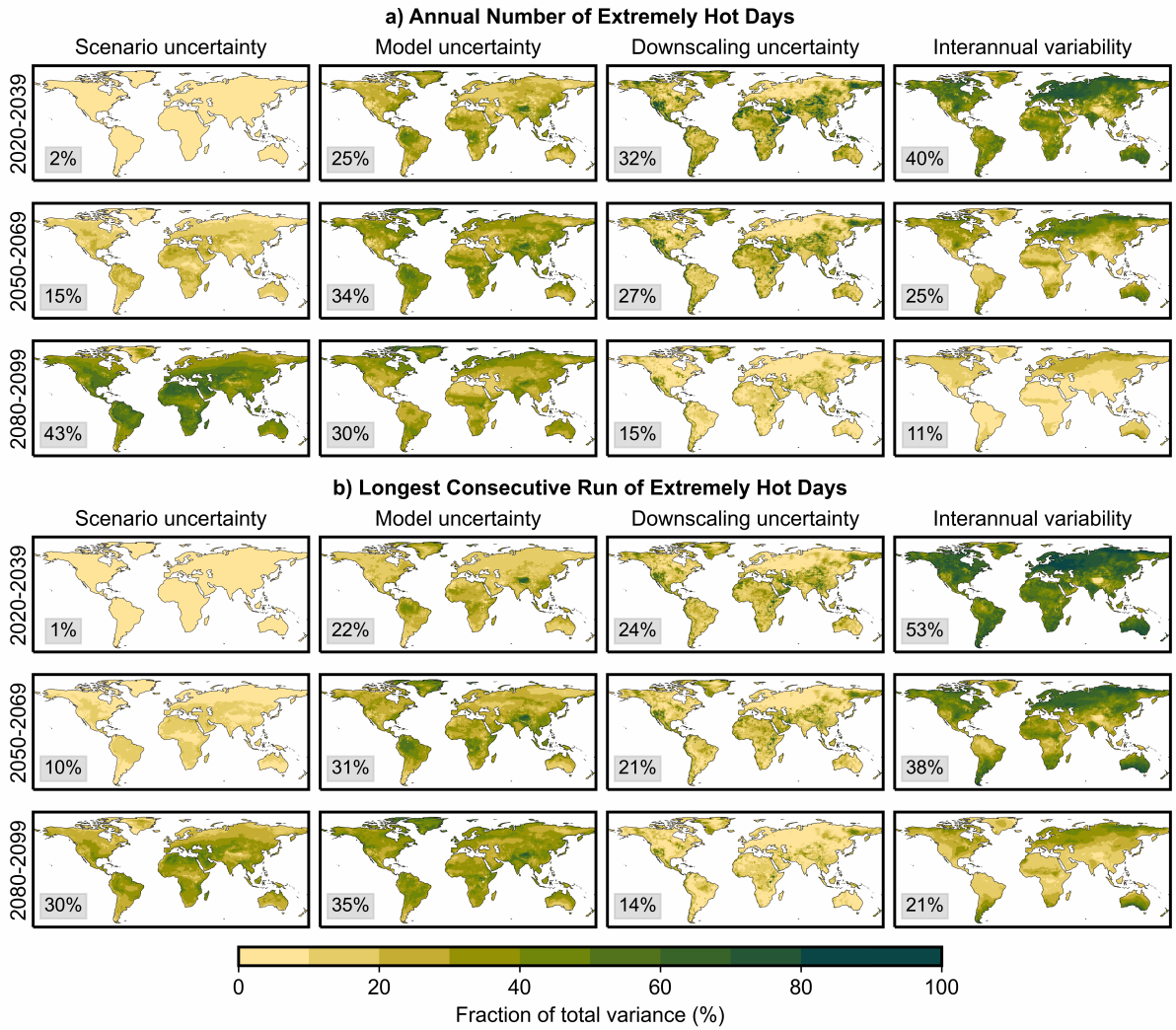
Supplementary Figure 34: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily maximum temperature exceeds the 99th percentile in ERA5.



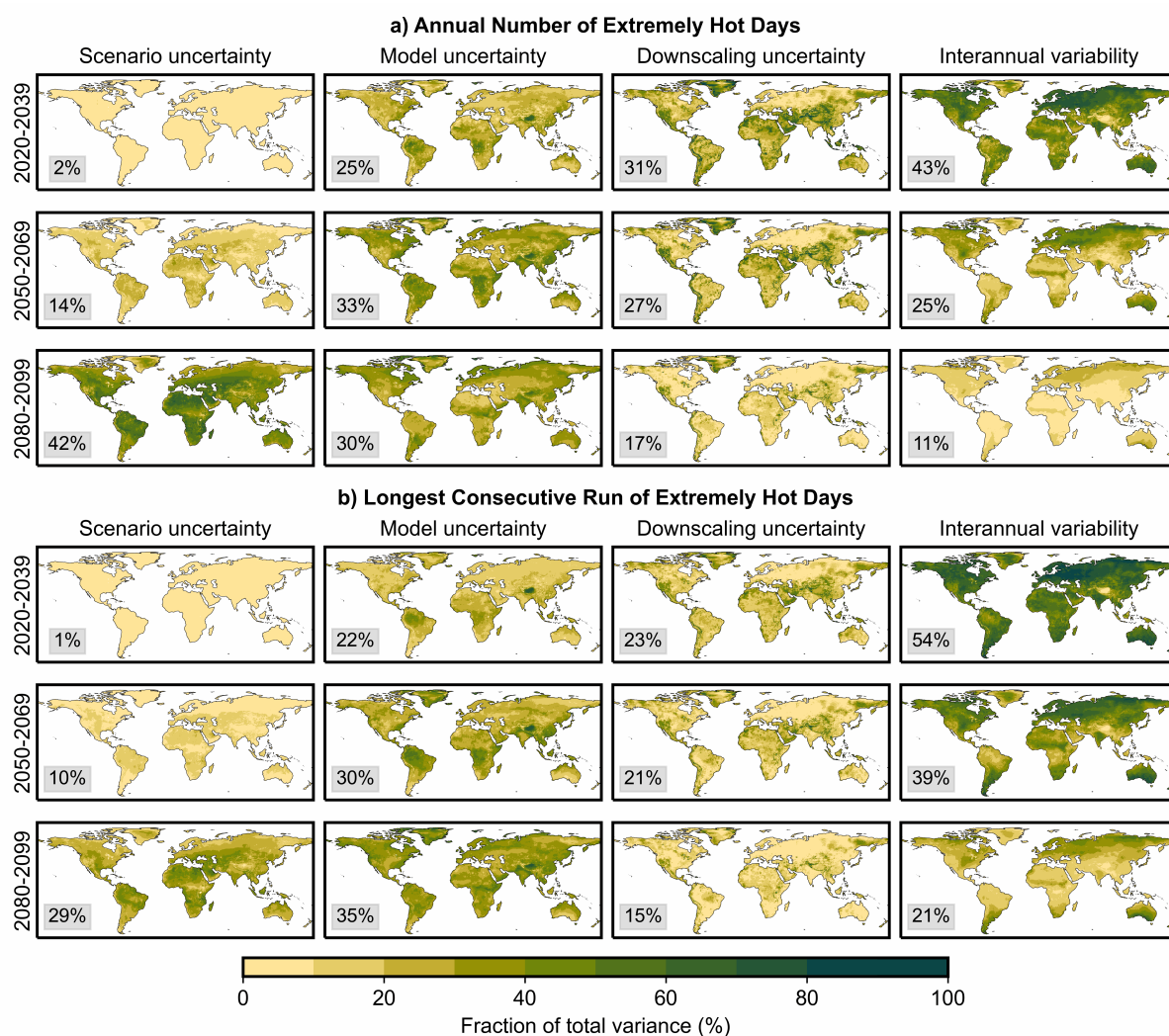
Supplementary Figure 35: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily average temperature exceeds the 99th percentile in GMFD.



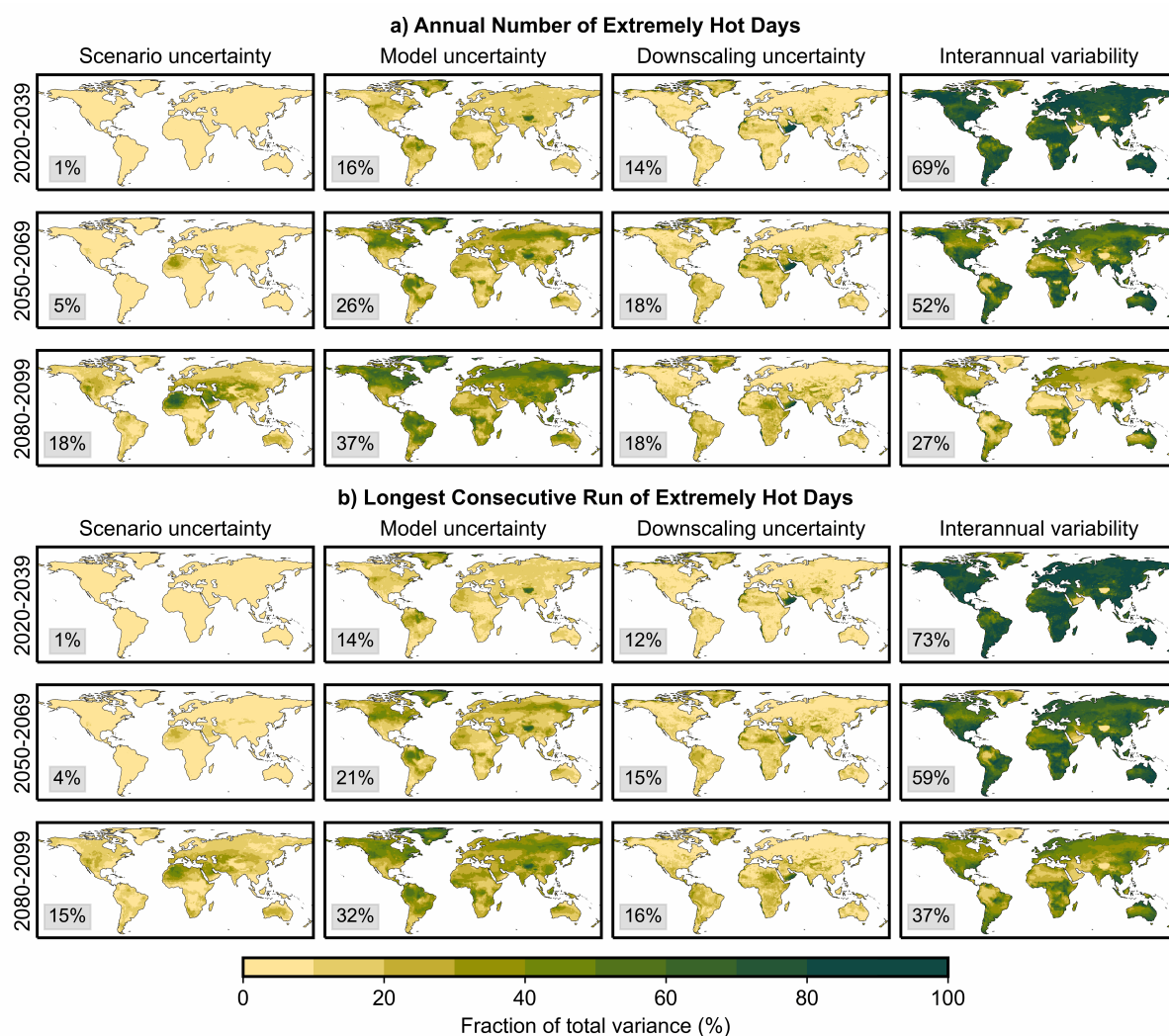
Supplementary Figure 36: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily average temperature exceeds the 99th percentile in ERA5.



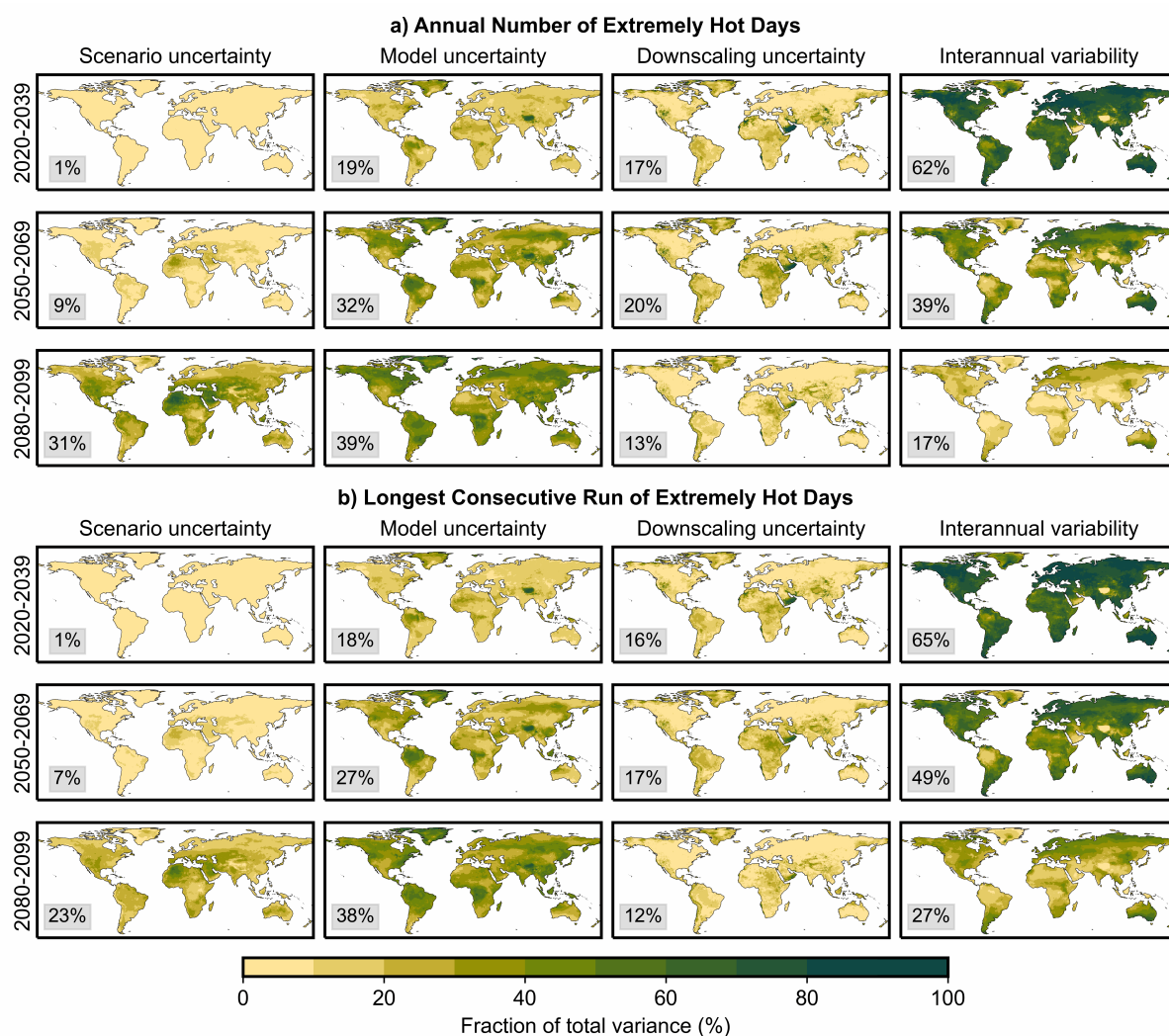
Supplementary Figure 37: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily minimum temperature exceeds the 99th percentile in GMFD.



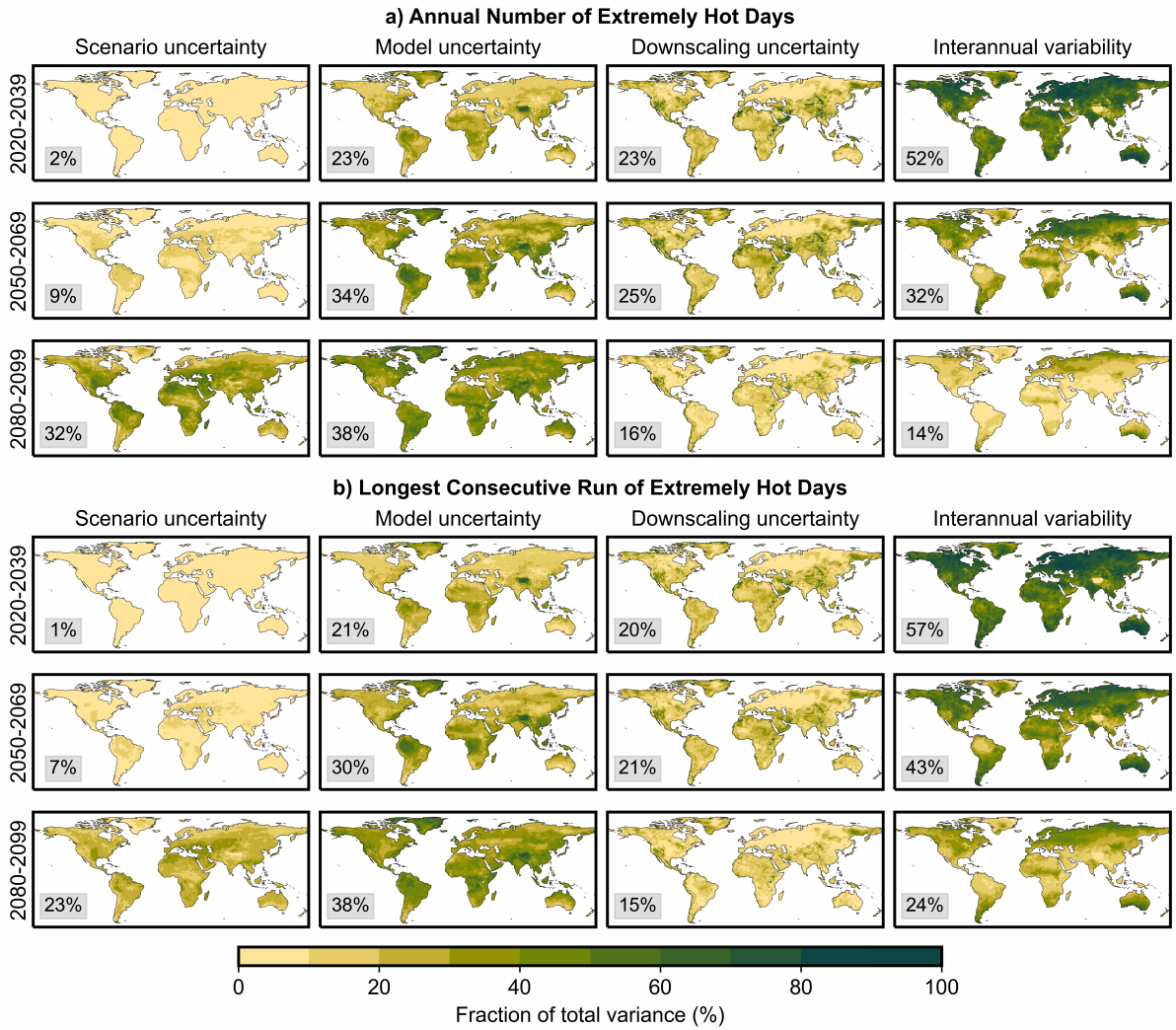
Supplementary Figure 38: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily minimum temperature exceeds the 99th percentile in ERA5.



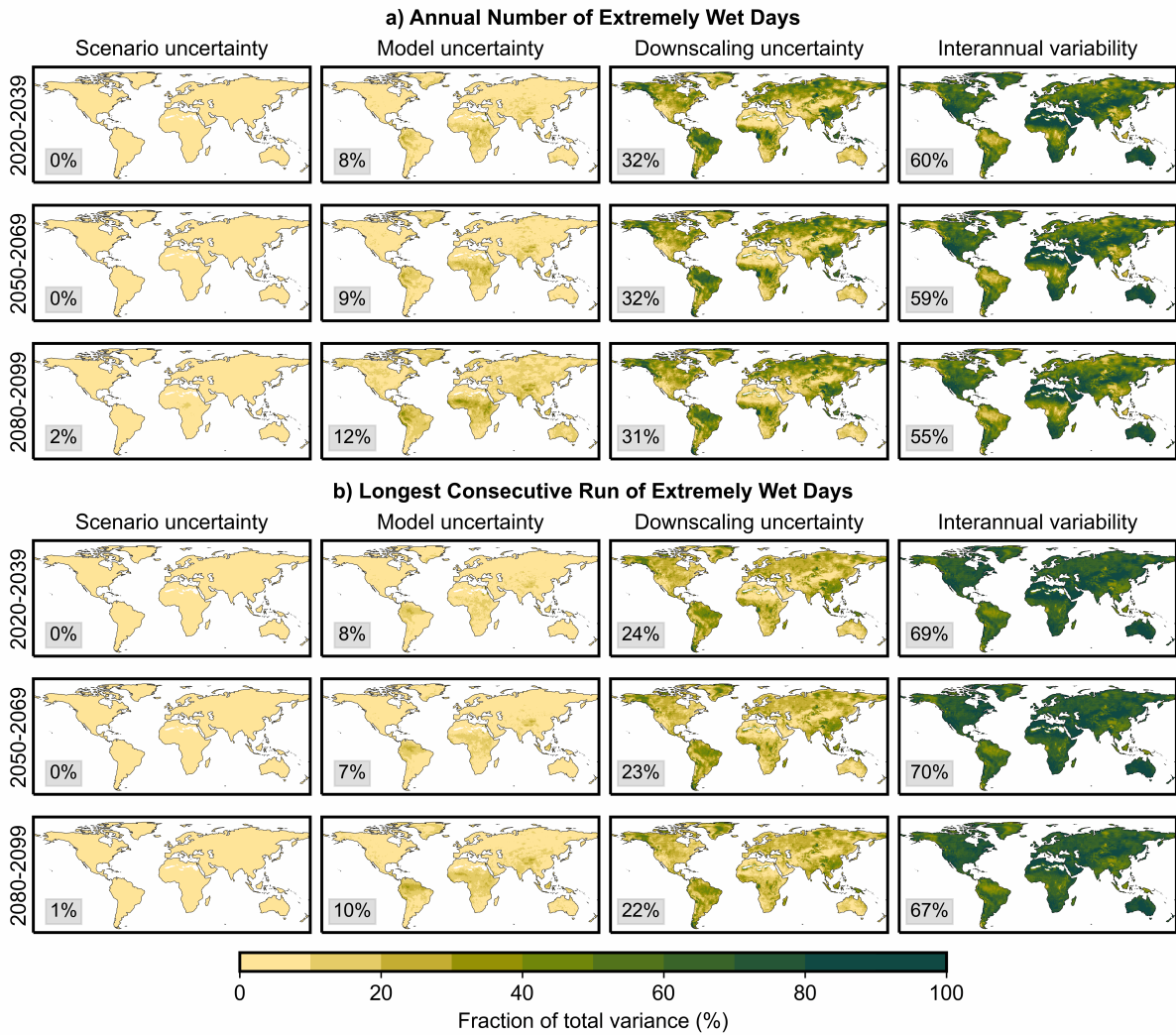
Supplementary Figure 39: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily maximum temperature exceeds the 99.97th percentile in GMFD.



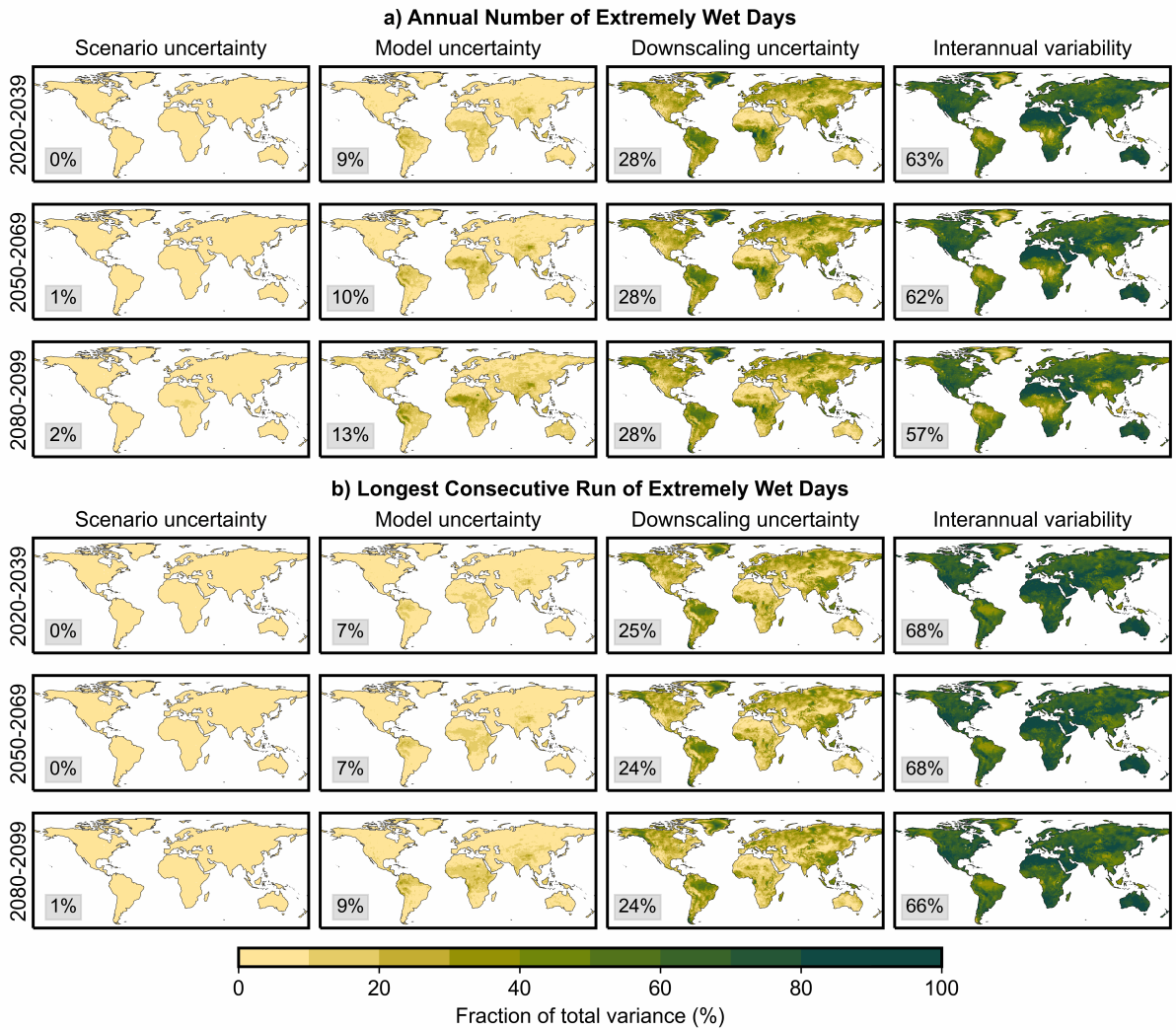
Supplementary Figure 40: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily average temperature exceeds the 99.97th percentile in GMFD.



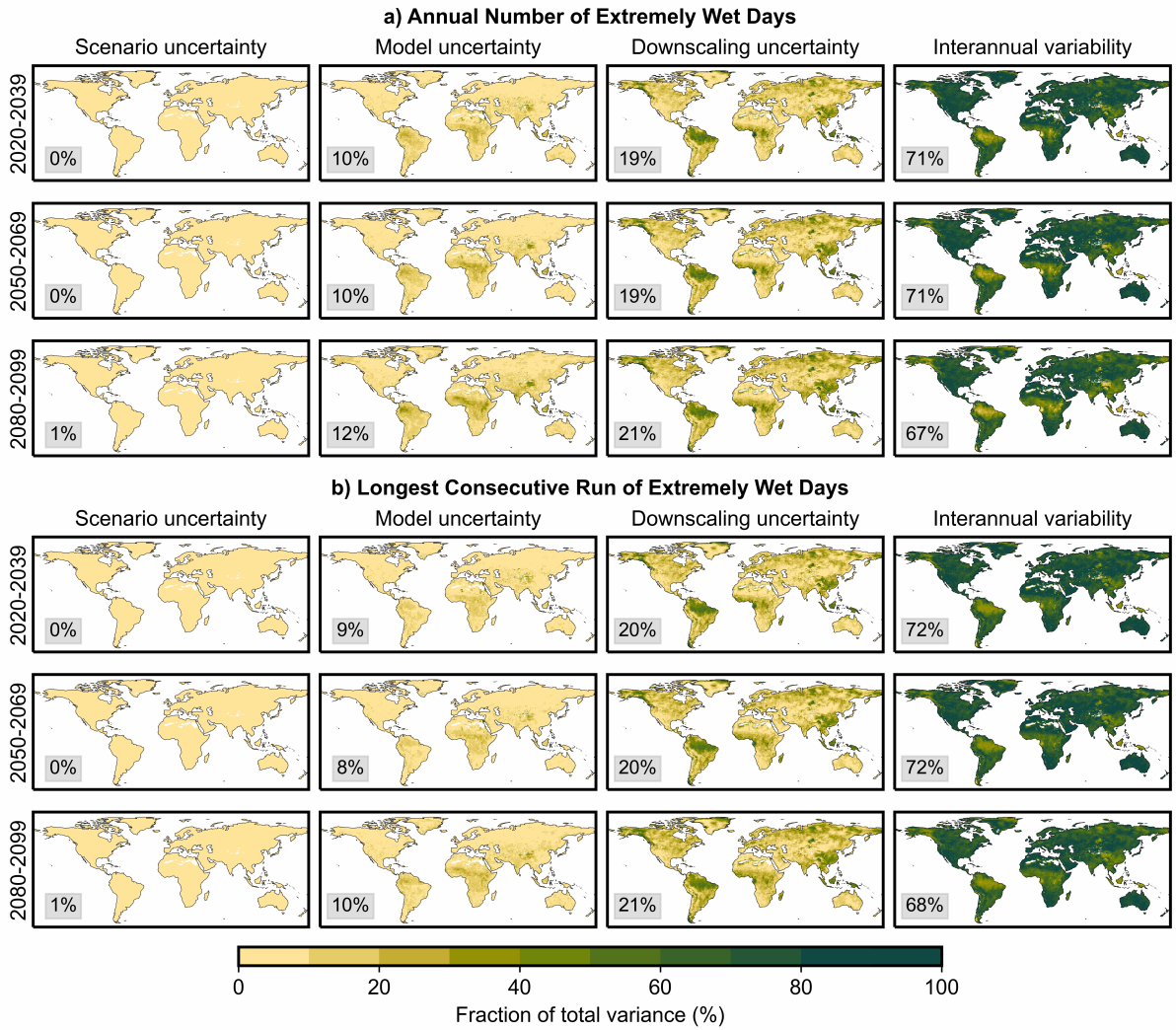
Supplementary Figure 41: **Alternative definition of extremes:** Variance decomposition for extremely hot days, defined as when daily minimum temperature exceeds the 99.97th percentile in GMFD.



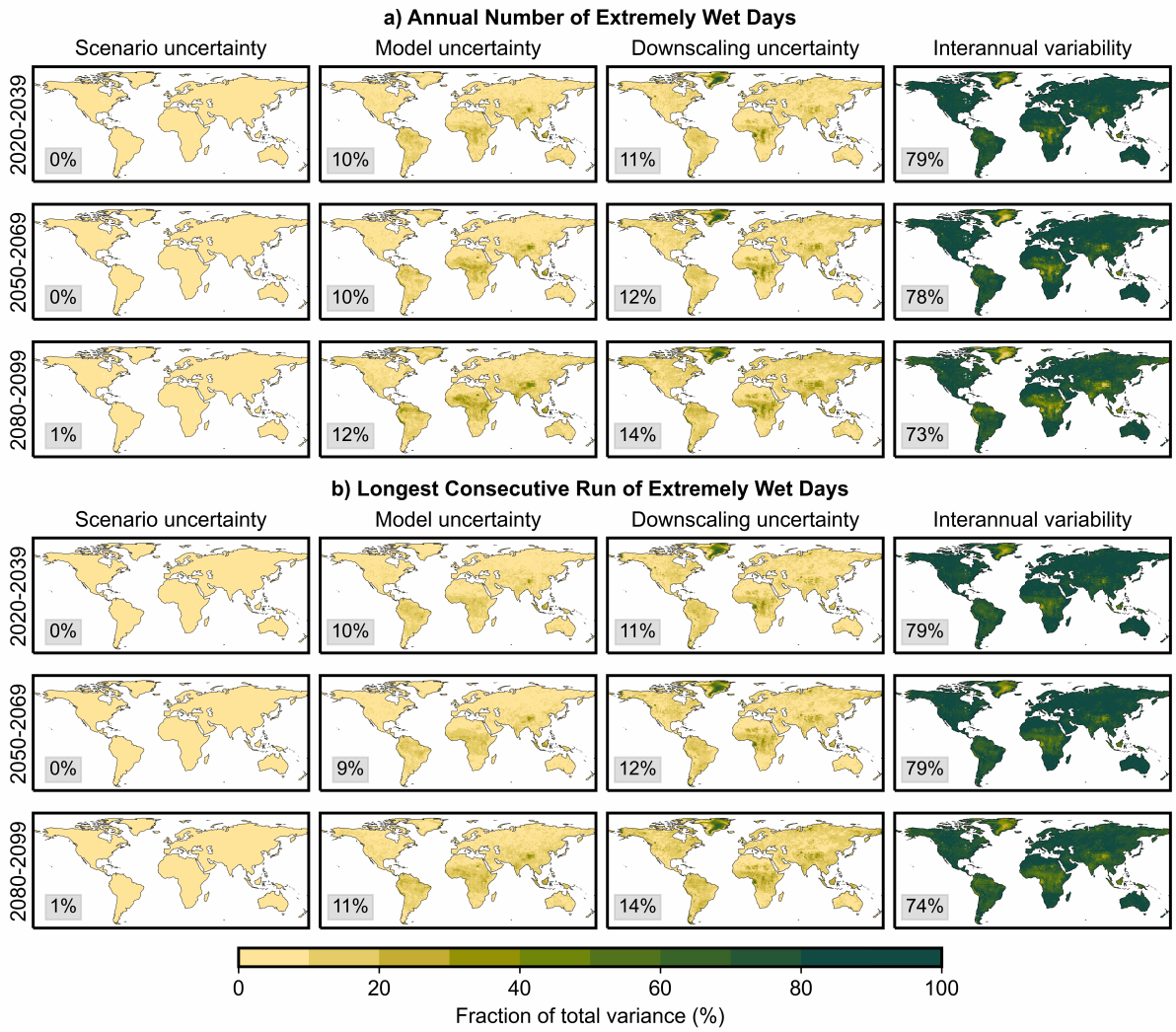
Supplementary Figure 42: **Alternative definition of extremes:** Variance decomposition for extremely wet days, defined as when daily precipitation exceeds the 99th percentile in GMFD.



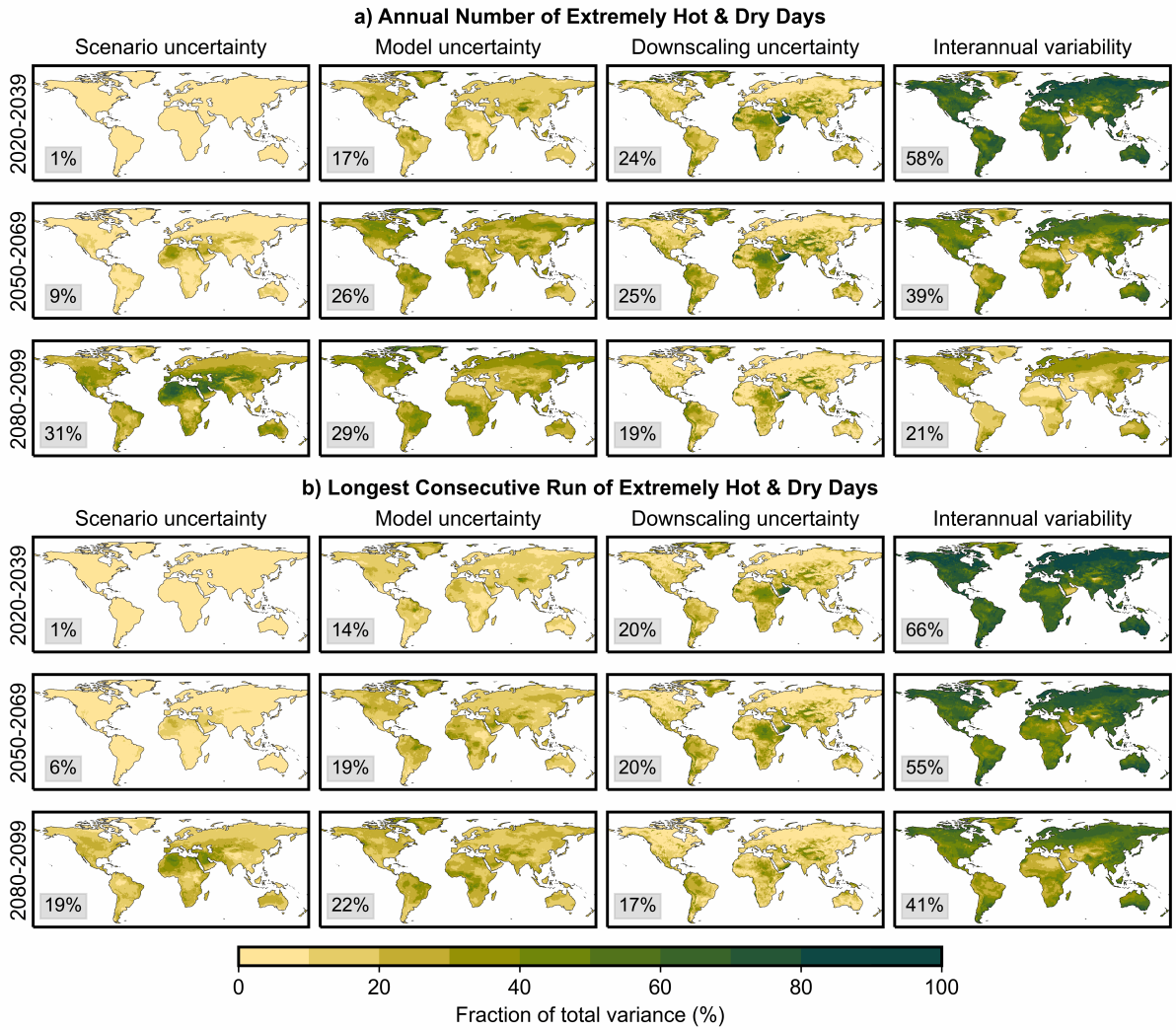
Supplementary Figure 43: **Alternative definition of extremes:** Variance decomposition for extremely wet days, defined as when daily precipitation exceeds the 99th percentile in ERA5.



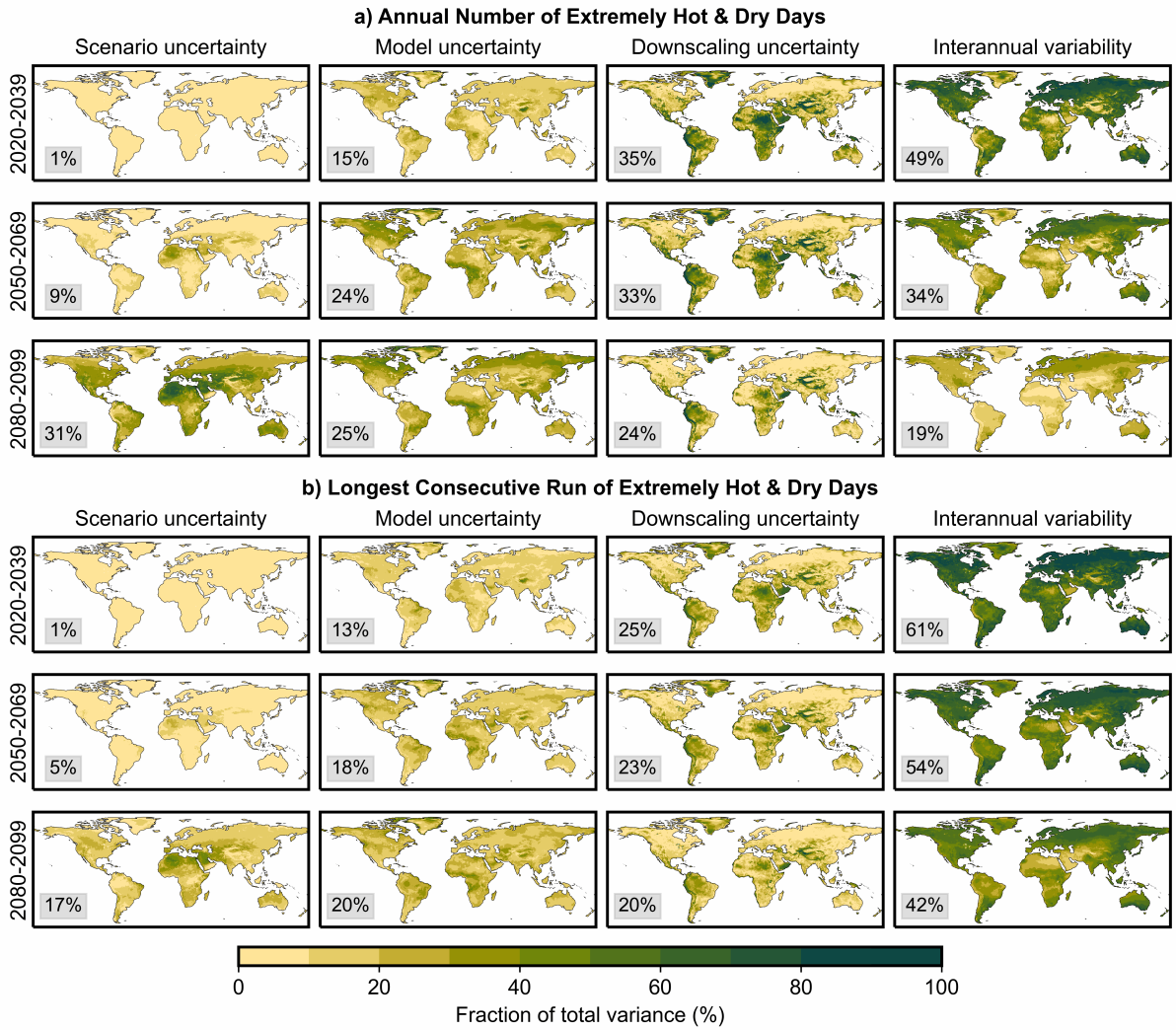
Supplementary Figure 44: **Alternative definition of extremes:** Variance decomposition for extremely wet days, defined as when daily precipitation exceeds the 99.97th percentile in GMFD.



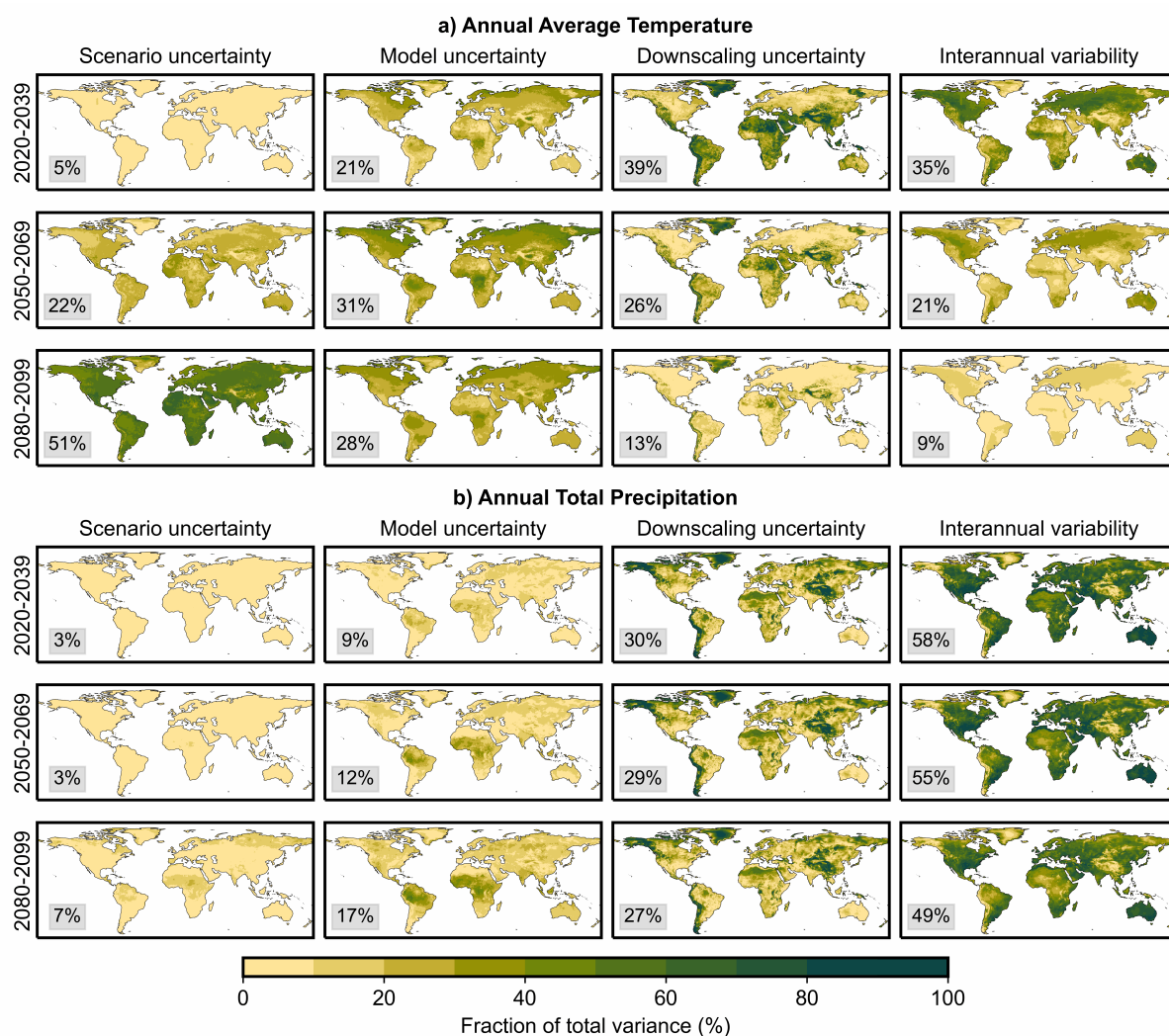
Supplementary Figure 45: **Alternative definition of extremes:** Variance decomposition for extremely wet days, defined as when daily precipitation exceeds the 99.97th percentile in ERA5.



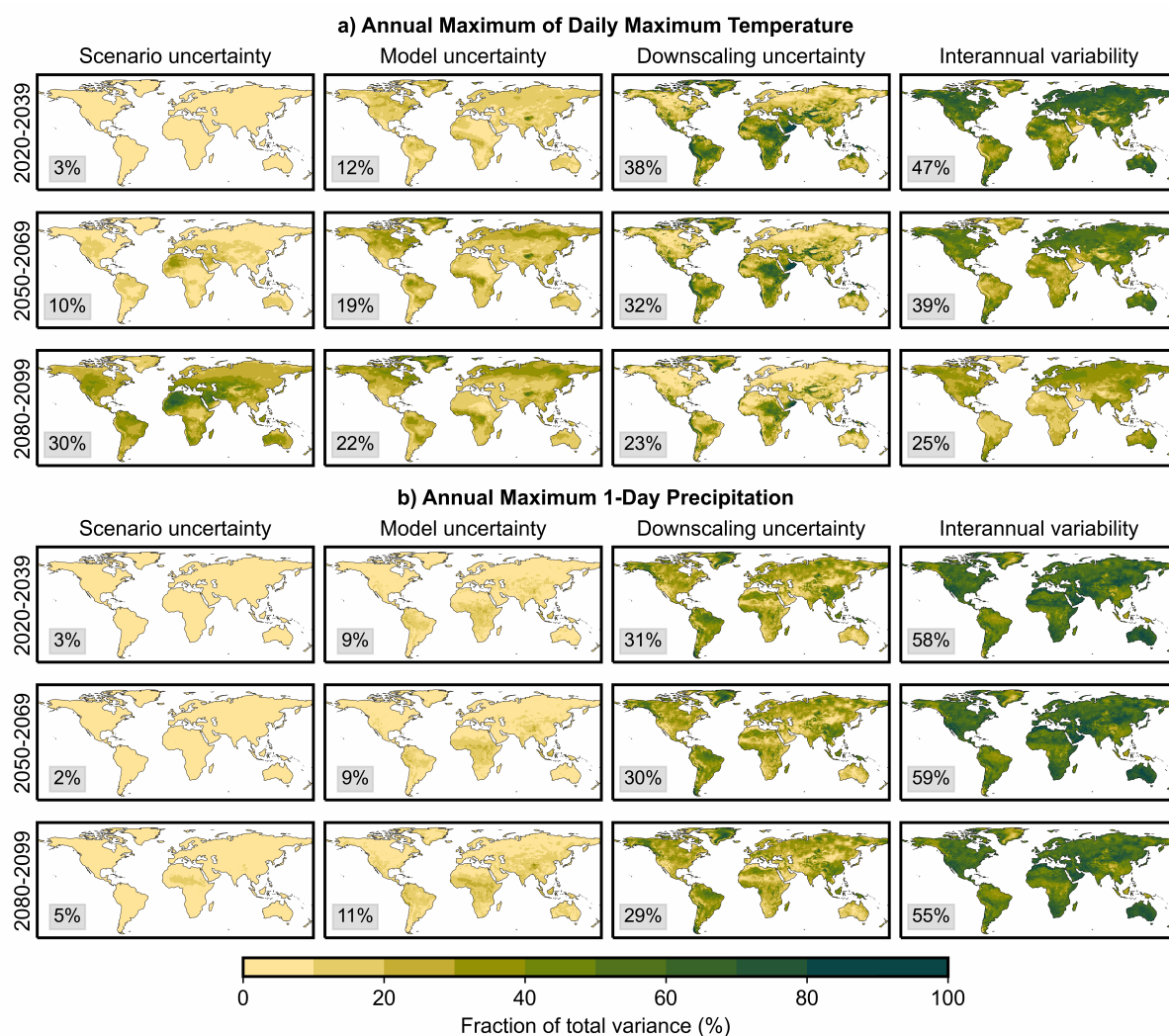
Supplementary Figure 46: **Alternative definition of extremes:** Variance decomposition for extremely hot & dry days, defined as when daily maximum temperature exceeds the 99th percentile in GMFD and daily precipitation is simultaneously less than 1mm.



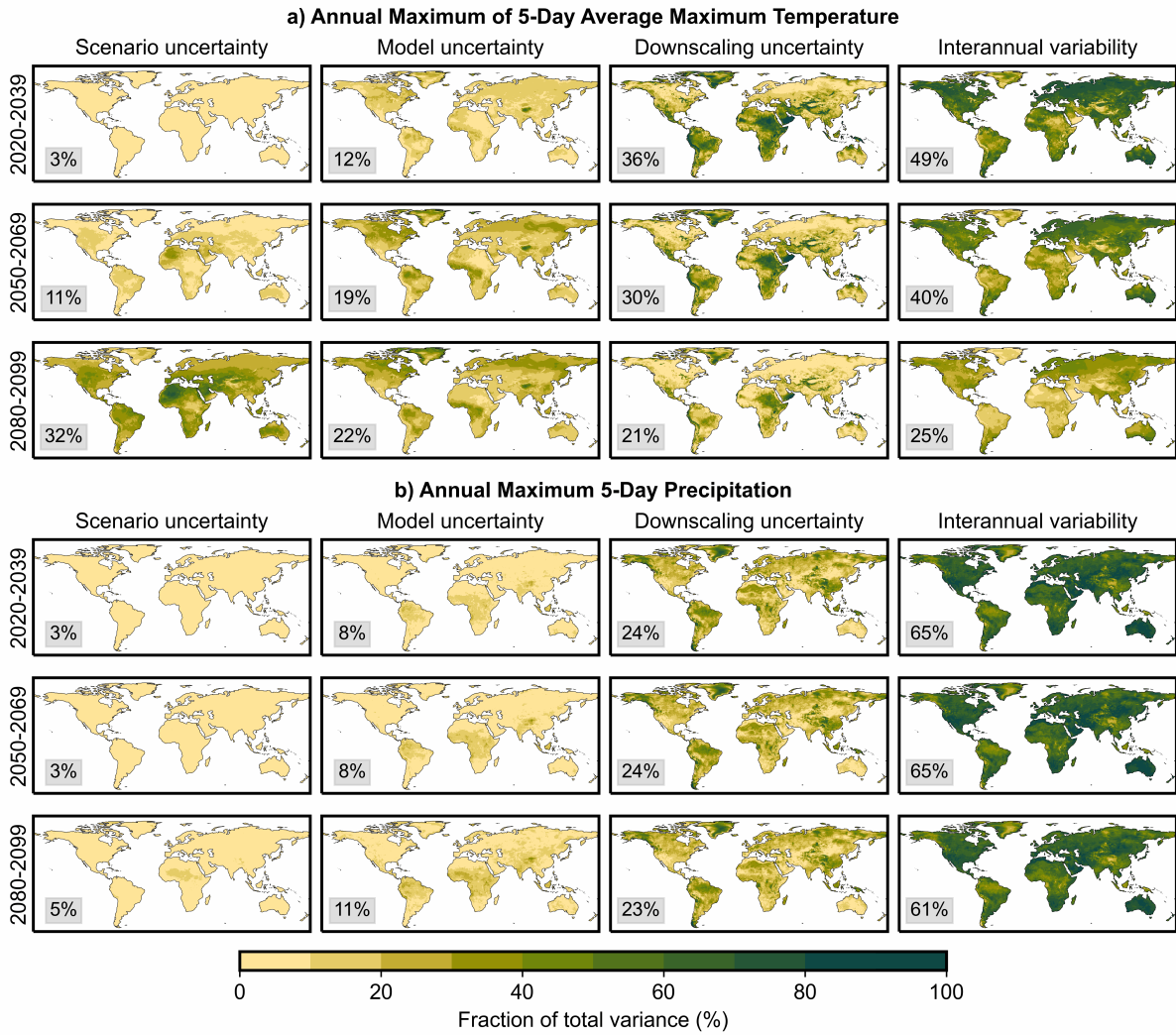
Supplementary Figure 47: **Alternative definition of extremes:** Variance decomposition for extremely hot & dry days, defined as when daily maximum temperature exceeds the 99th percentile in ERA5 and daily precipitation is simultaneously less than 1mm.



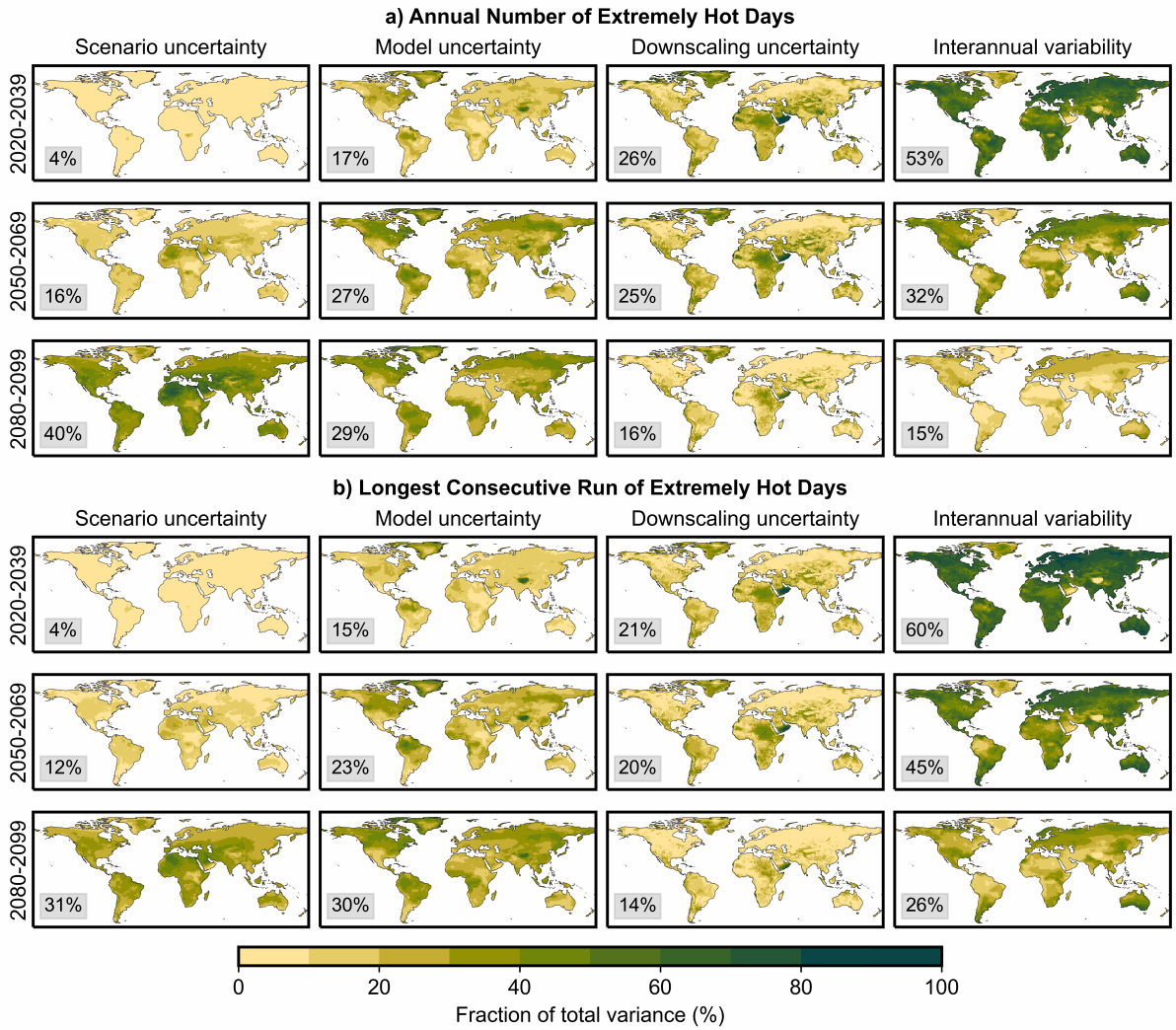
Supplementary Figure 48: **Methodological checks:** Variance decomposition for (a) annual average temperature and (b) annual total precipitation using the Brekke & Barsugli definition for scenario uncertainty.



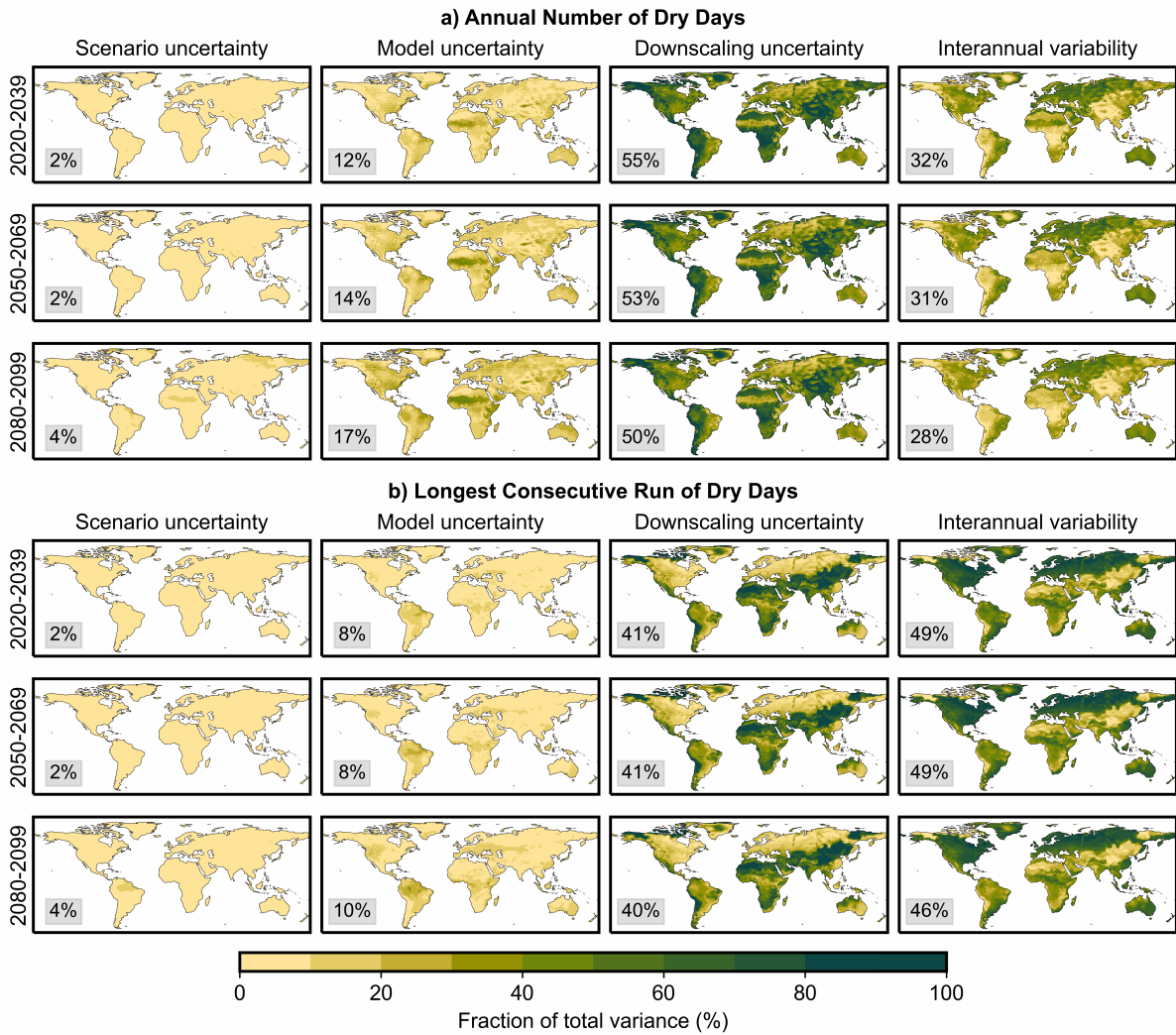
Supplementary Figure 49: **Methodological checks:** Variance decomposition for annual 1-day maxima of (a) daily maximum temperature and (b) daily precipitation using the Brekke & Barsugli definition for scenario uncertainty.



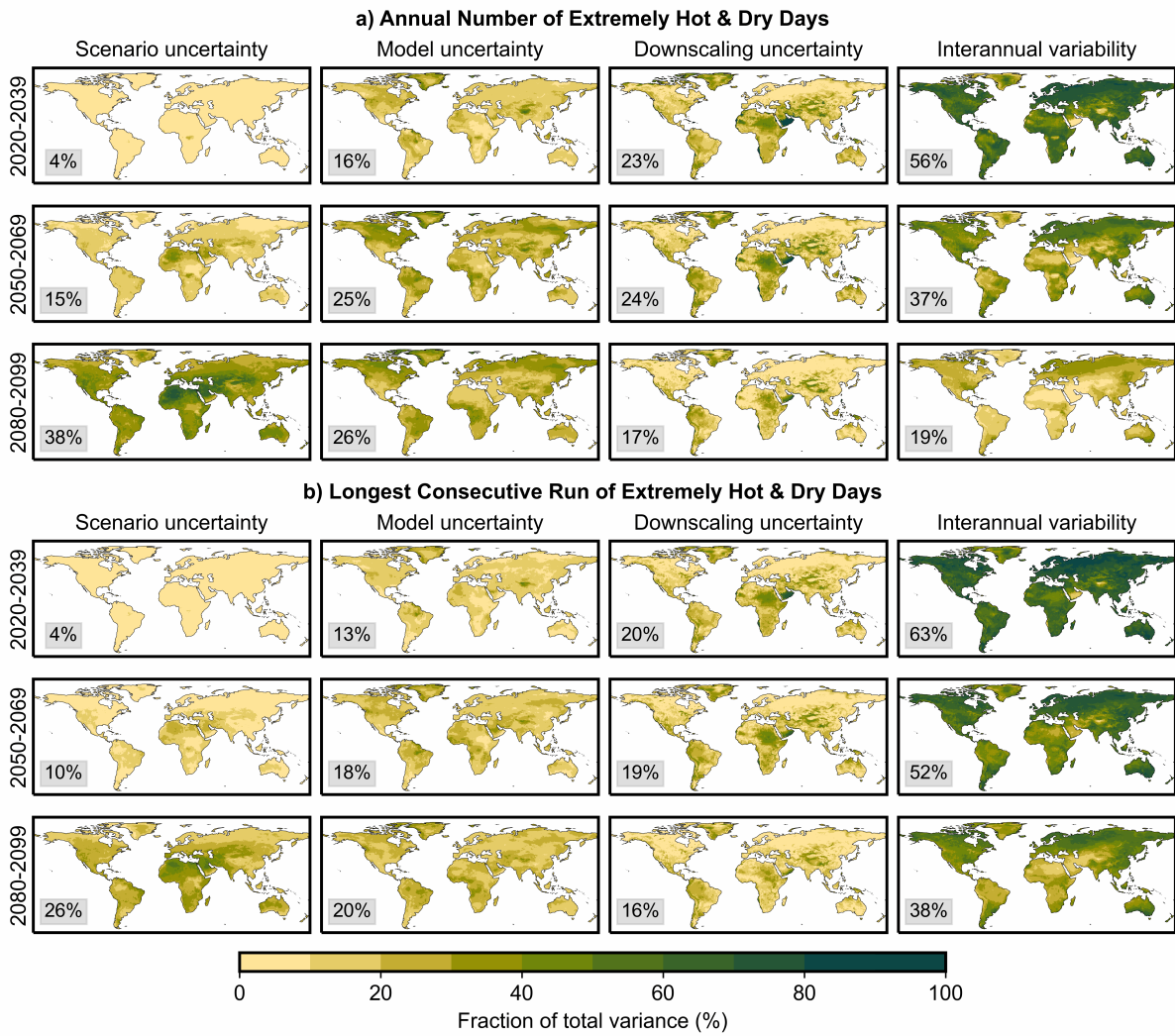
Supplementary Figure 50: **Methodological checks:** Variance decomposition for annual 5-day maxima of (a) daily maximum temperature and (b) daily precipitation using the Brekke & Barsugli definition for scenario uncertainty.



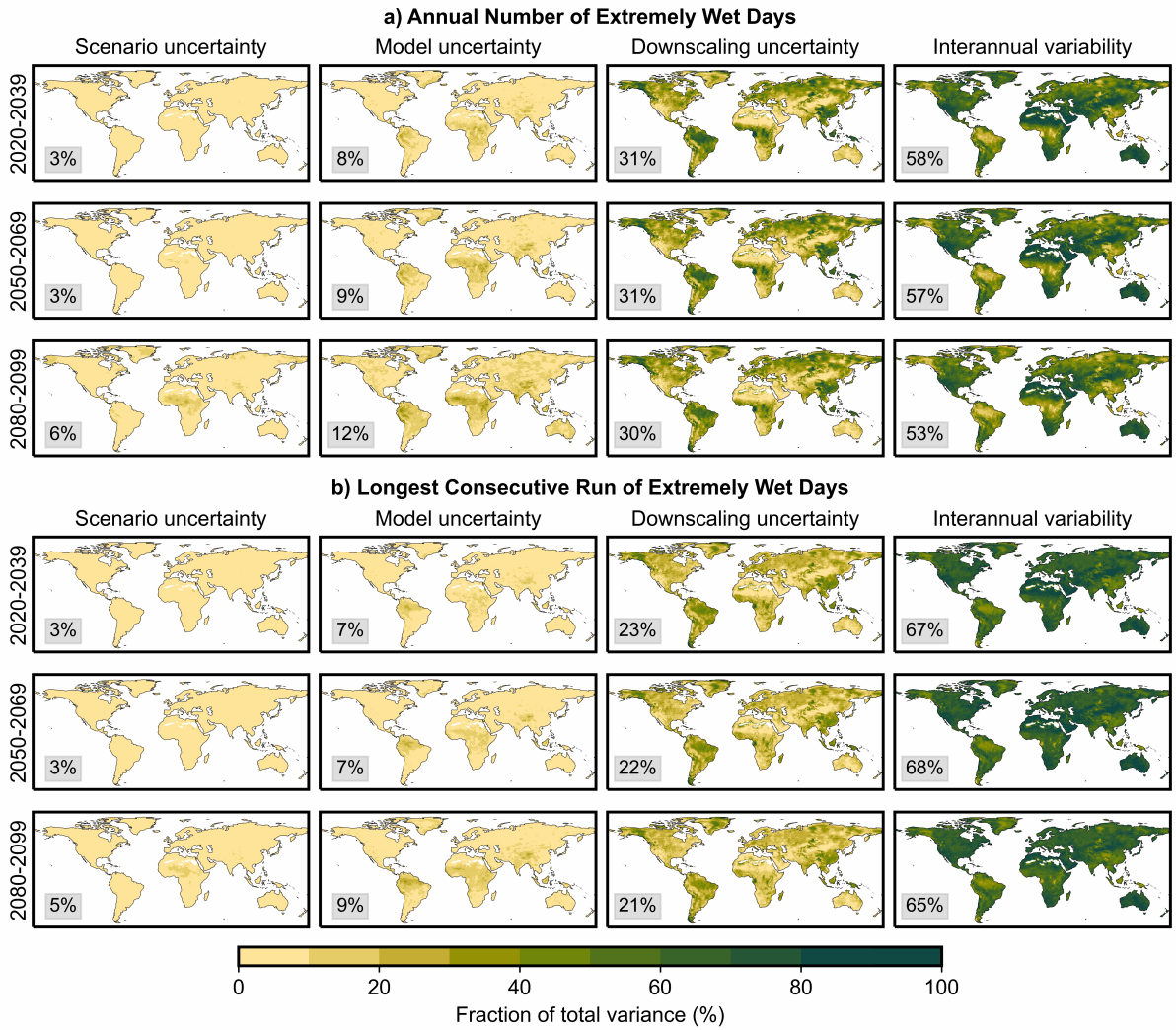
Supplementary Figure 51: **Methodological checks:** Variance decomposition for extremely hot days (defined as daily maximum temperature exceeding the 99th percentile from GMFD) using the Brekke & Barsugli definition for scenario uncertainty.



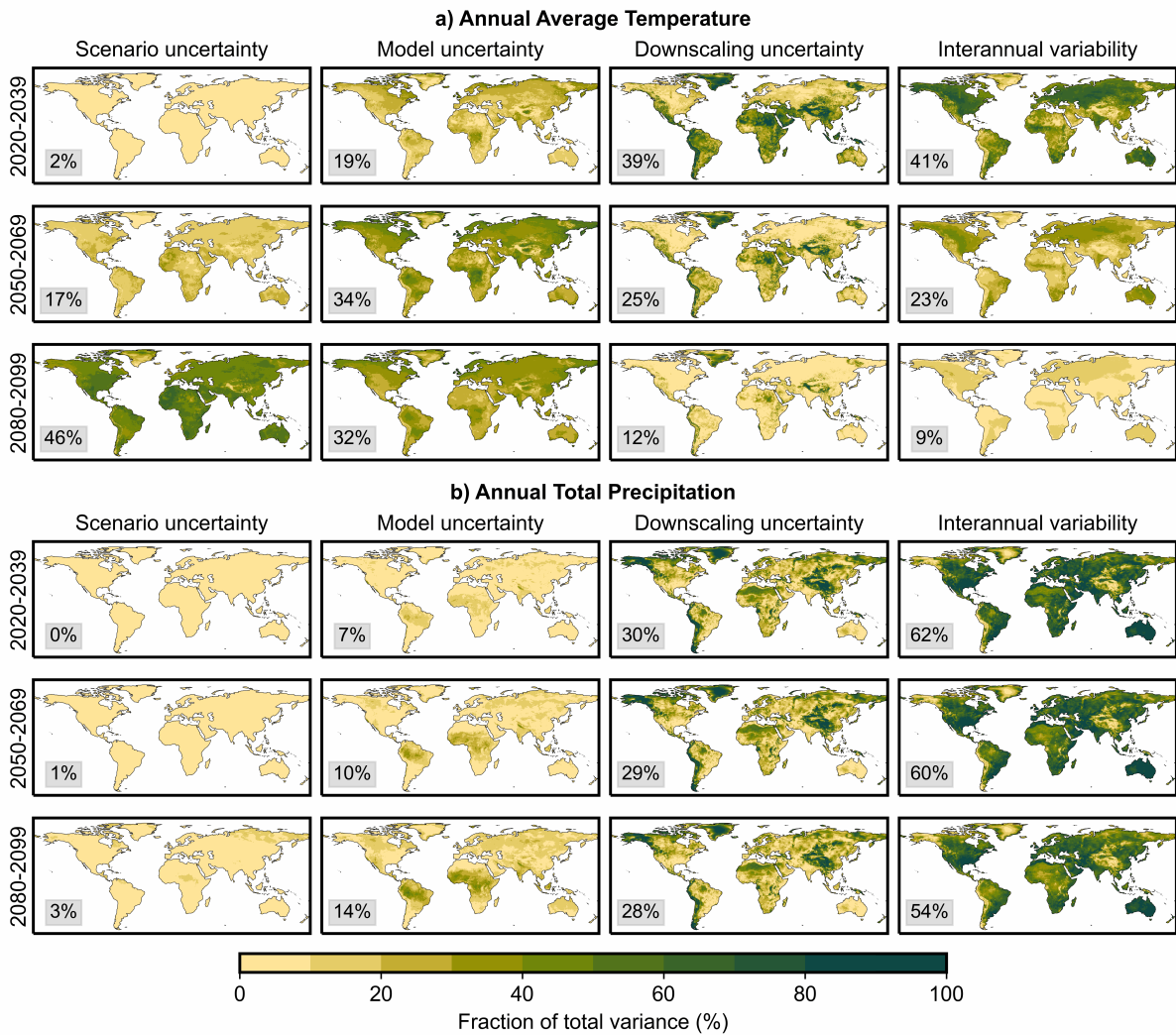
Supplementary Figure 52: **Methodological checks:** Variance decomposition for dry days using the Brekke & Barsugli definition for scenario uncertainty.



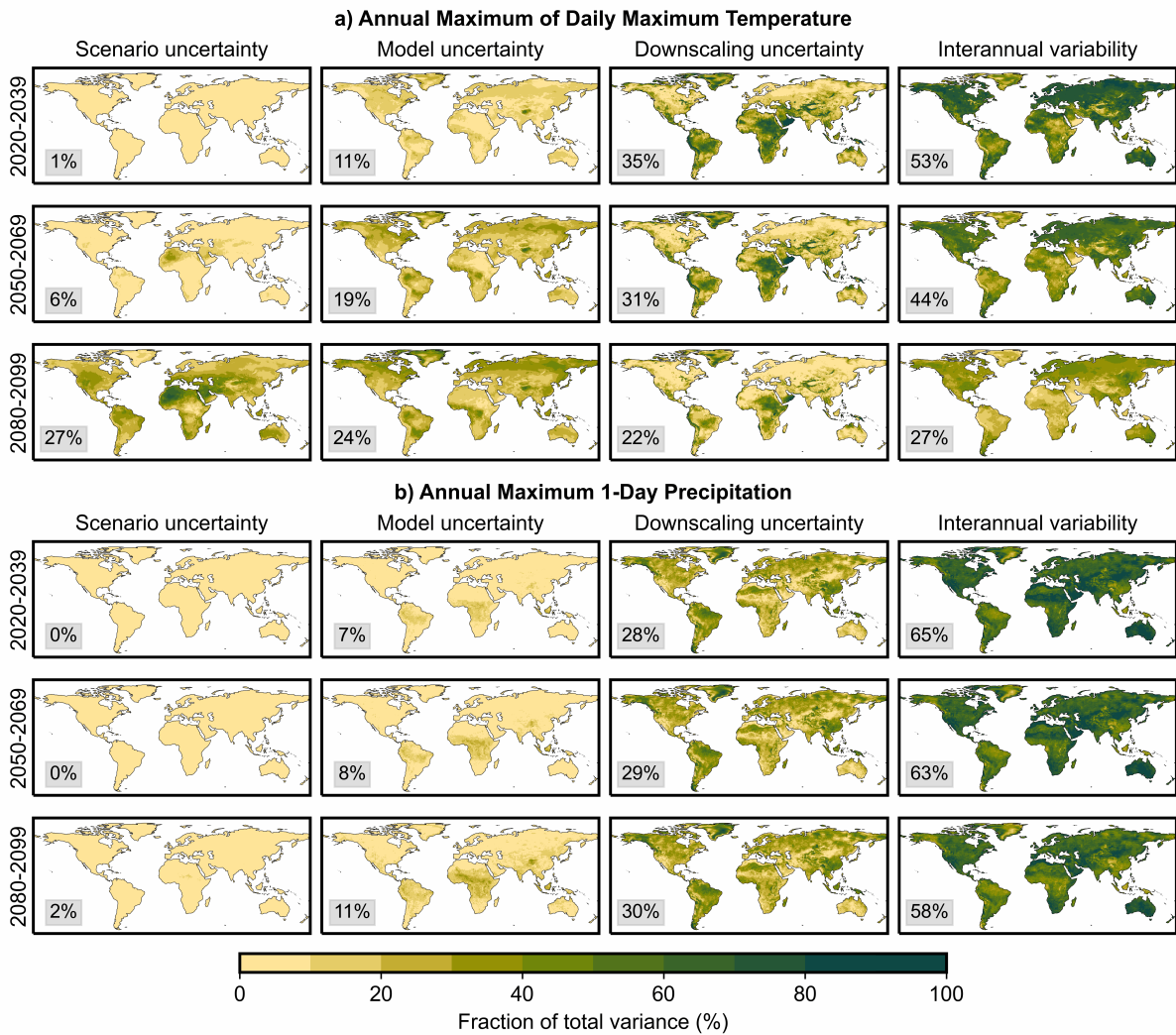
Supplementary Figure 53: **Methodological checks:** Variance decomposition for extremely hot & days (defined as daily maximum temperature exceeding the 99th percentile from GMFD and daily precipitation simultaneously less than 1mm) using the Brekke & Barsugli definition for scenario uncertainty.



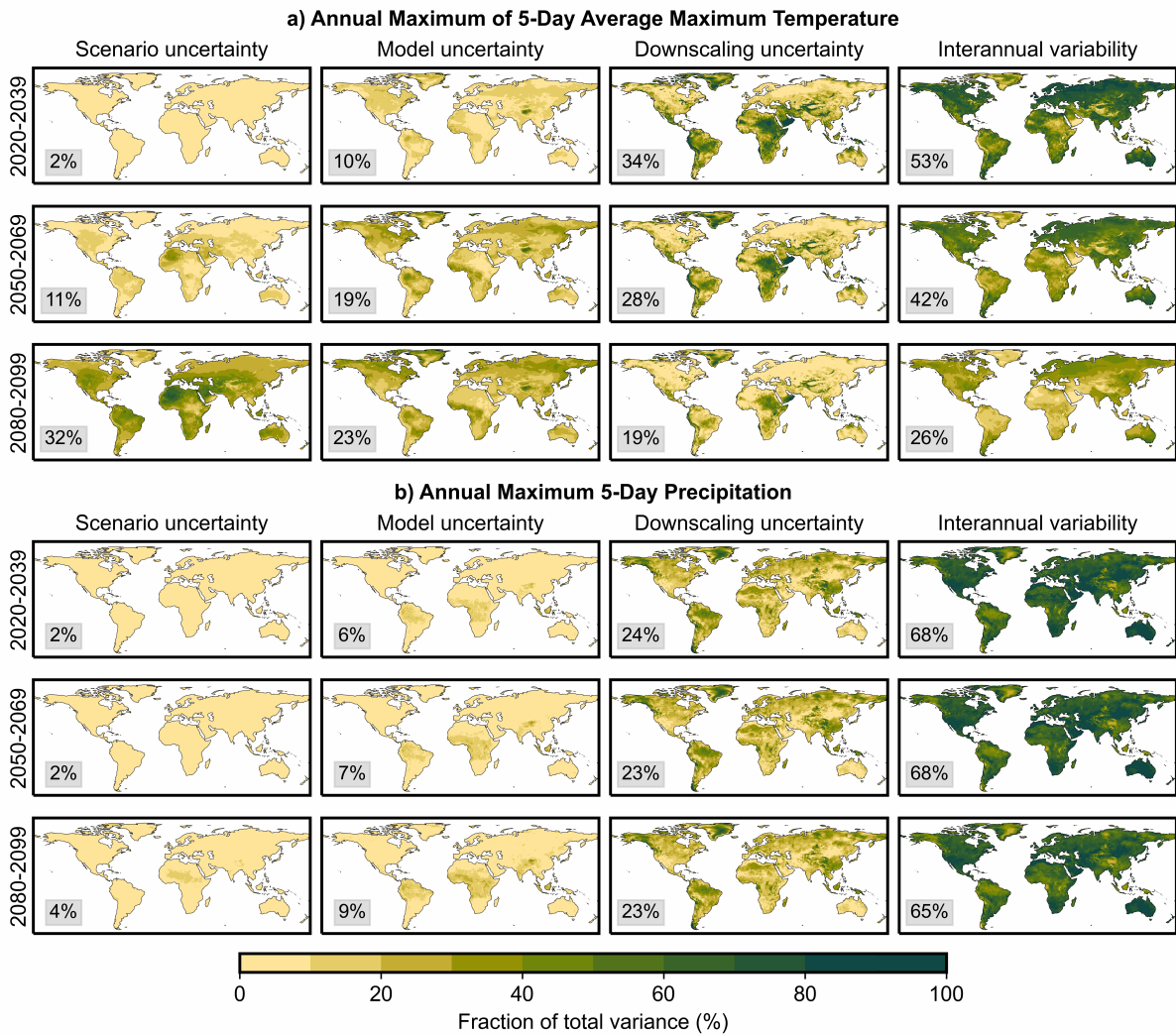
Supplementary Figure 54: **Methodological checks:** Variance decomposition for extremely wet days (defined as daily precipitation exceeding the 99th percentile from GMFD) using the Brekke & Barsugli definition for scenario uncertainty.



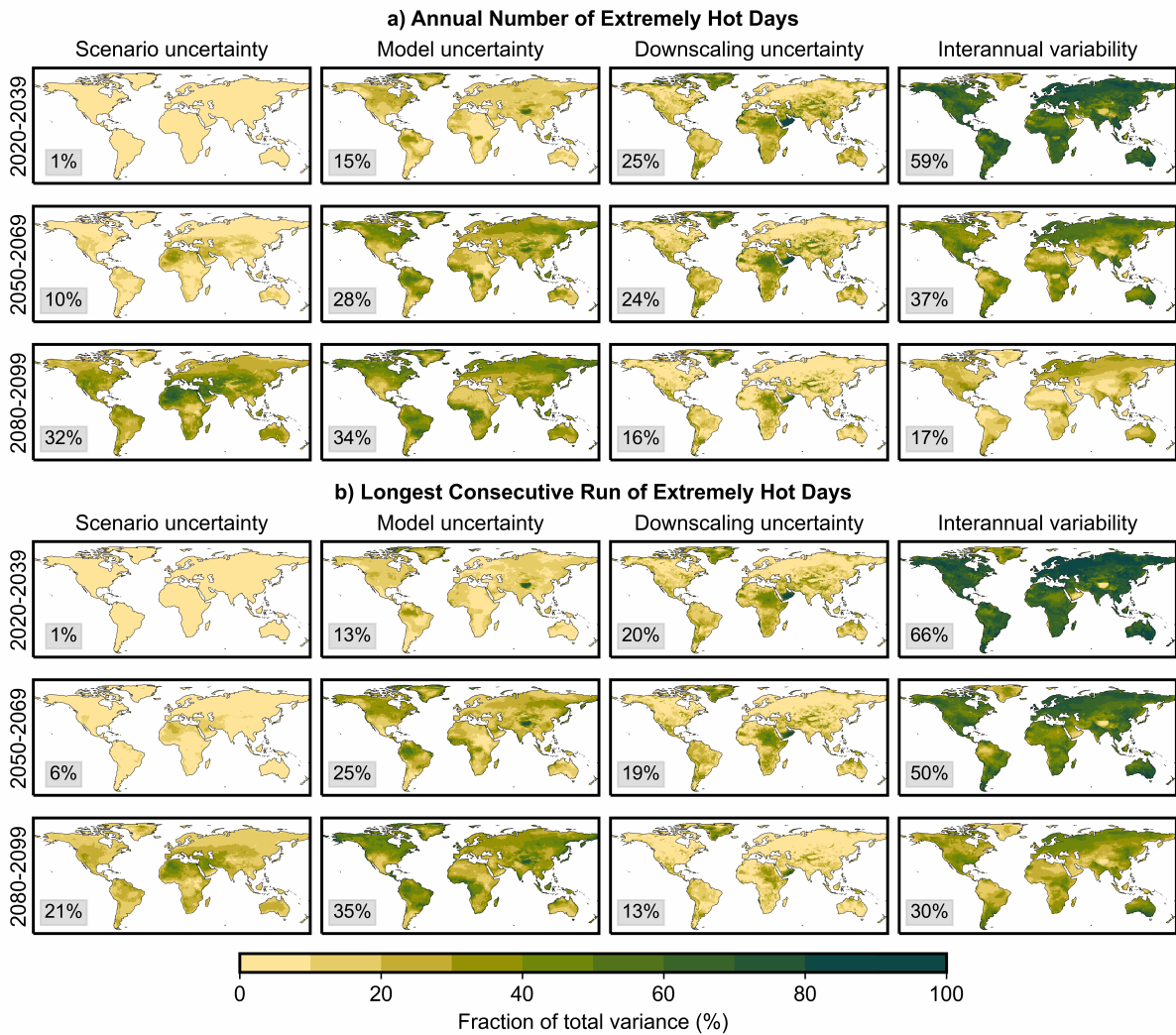
Supplementary Figure 55: **Methodological checks:** Variance decomposition for (a) annual average temperature and (b) annual total precipitation calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



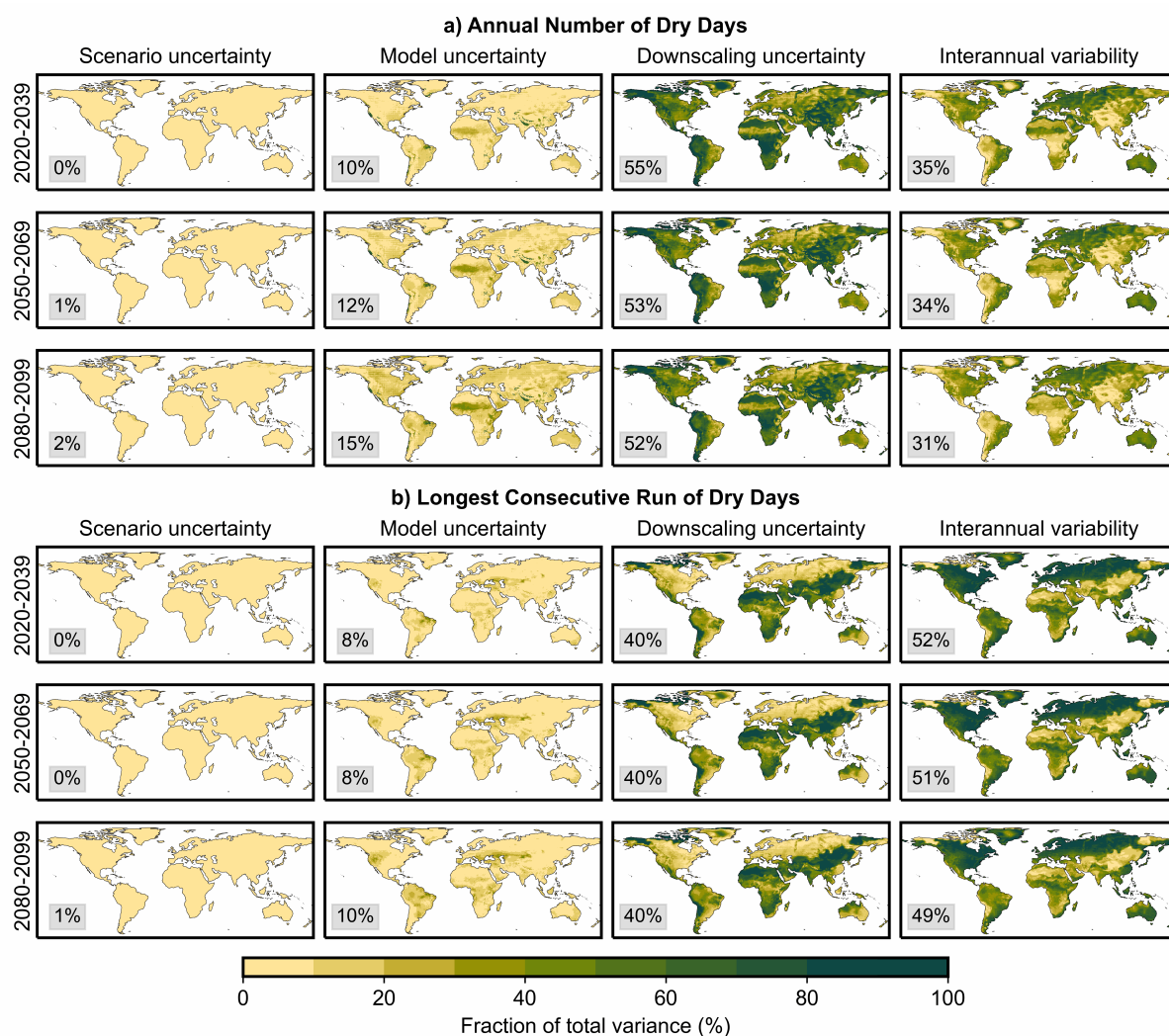
Supplementary Figure 56: **Methodological checks:** Variance decomposition for annual 1-day maxima of (a) daily maximum temperature and (b) daily precipitation calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



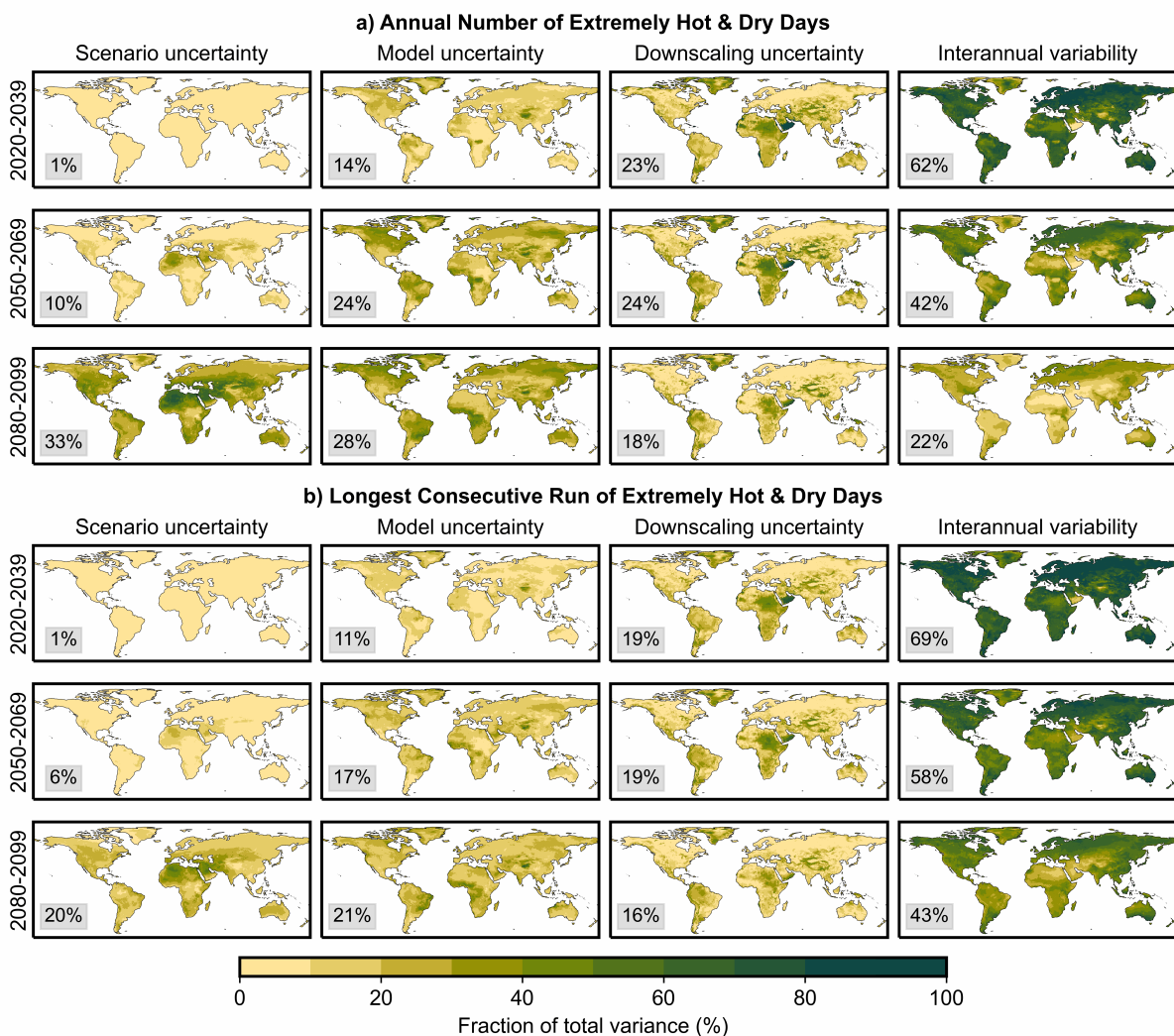
Supplementary Figure 57: **Methodological checks:** Variance decomposition for annual 5-day maxima of (a) daily maximum temperature and (b) daily precipitation calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



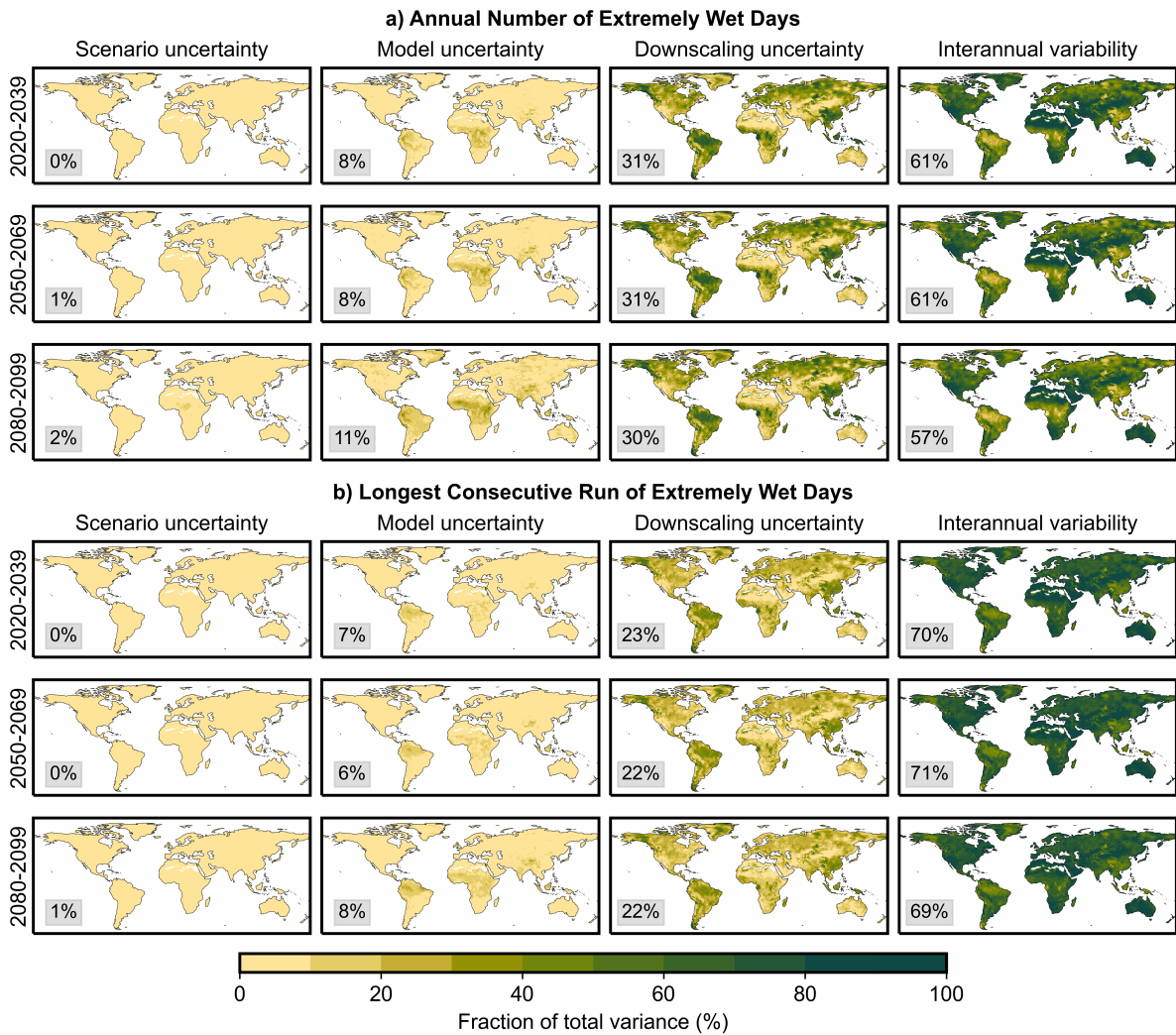
Supplementary Figure 58: **Methodological checks:** Variance decomposition for extremely hot days (defined as daily maximum temperature exceeding the 99th percentile from GMFD) calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



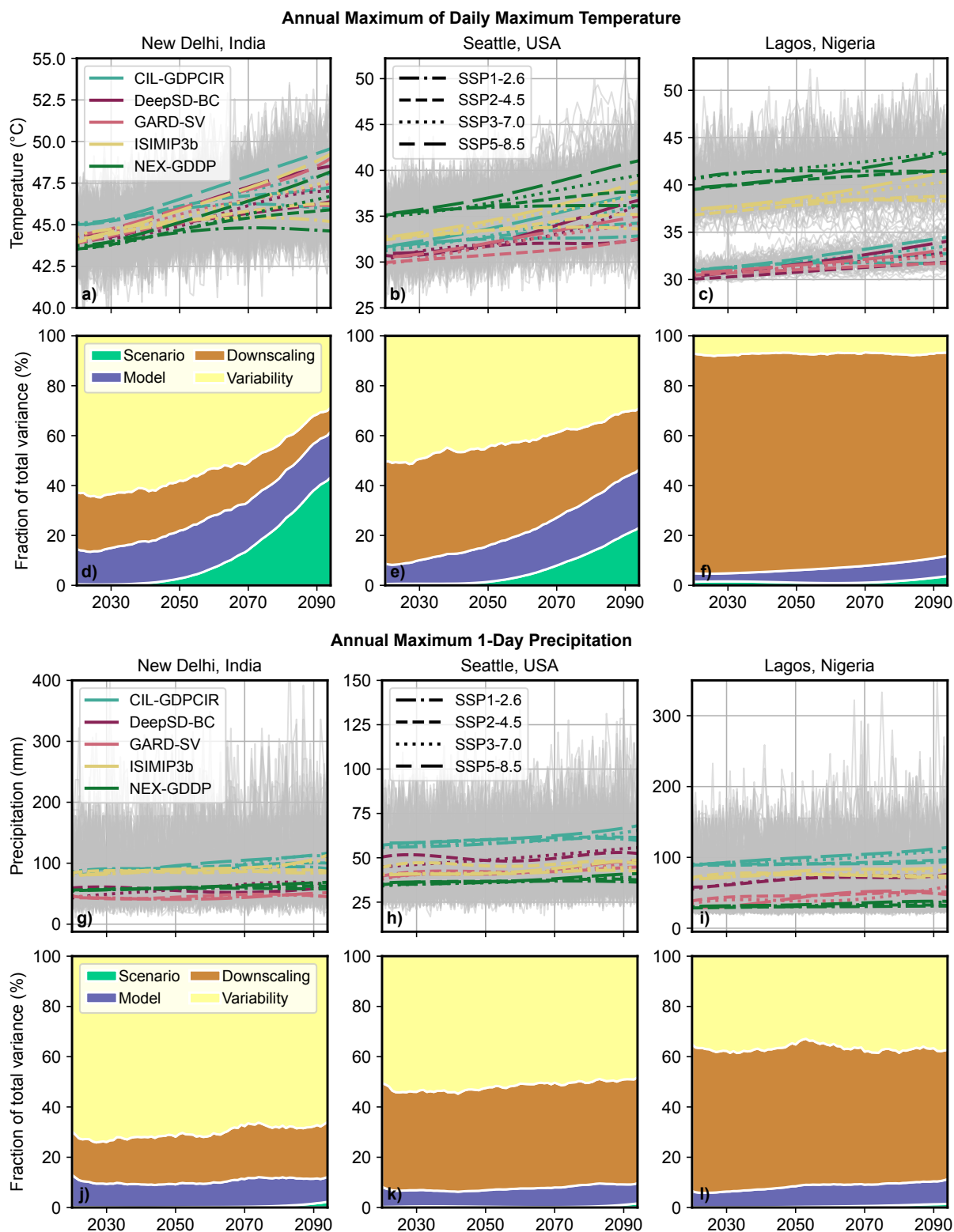
Supplementary Figure 59: **Methodological checks:** Variance decomposition for dry days calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



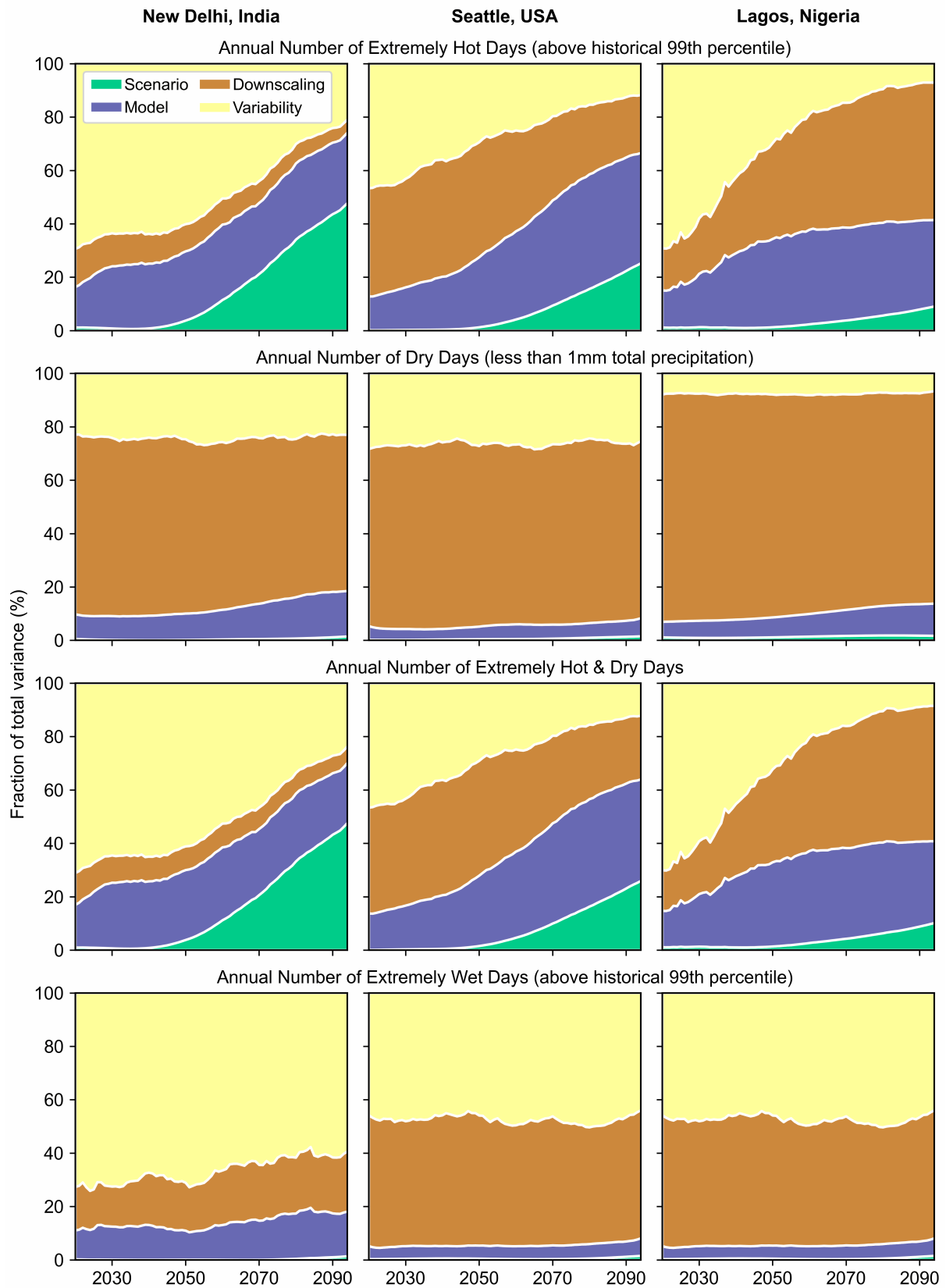
Supplementary Figure 60: **Methodological checks:** Variance decomposition for extremely hot & days (defined as daily maximum temperature exceeding the 99th percentile from GMFD and daily precipitation simultaneously less than 1mm) calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).



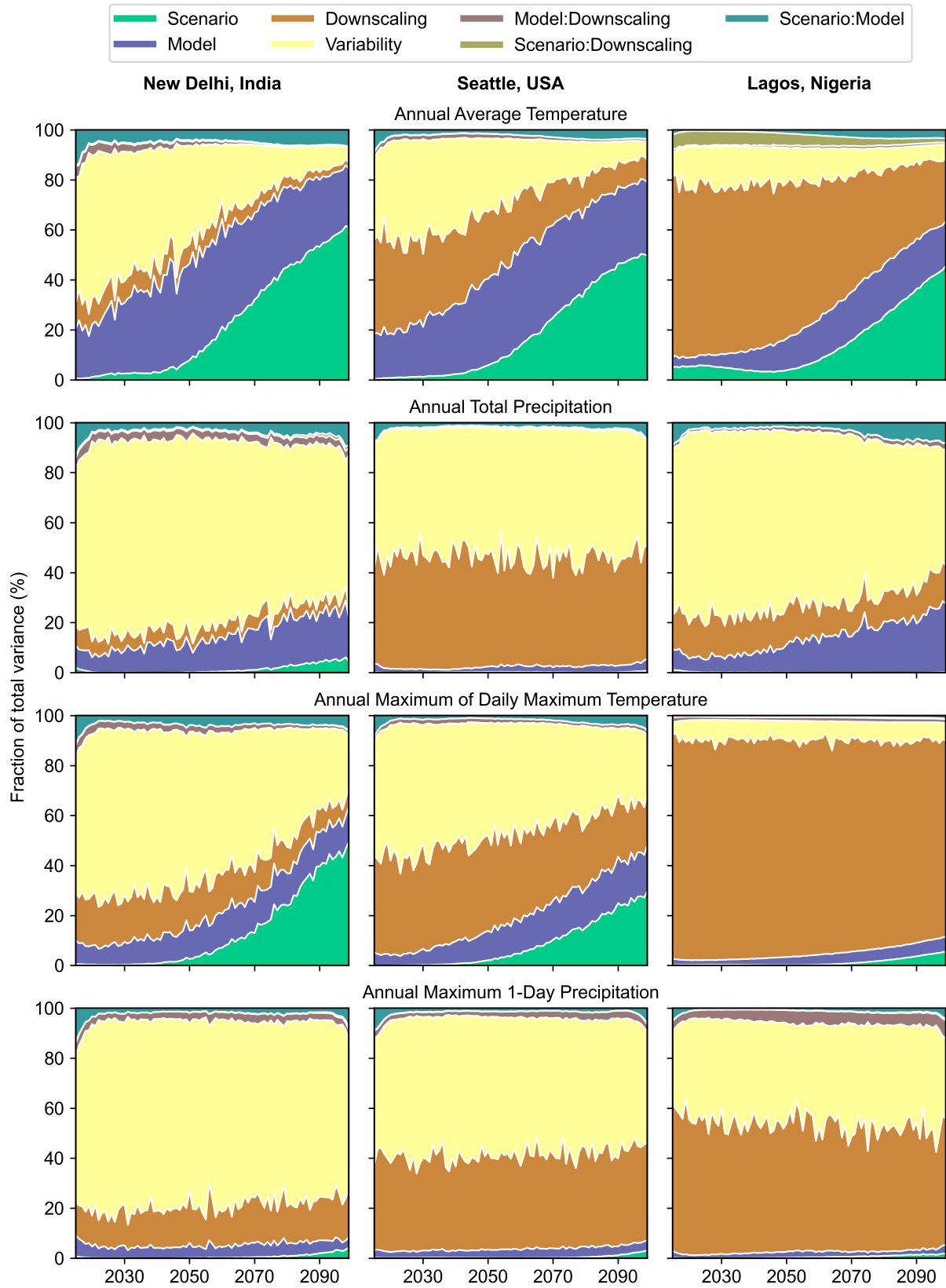
Supplementary Figure 61: **Methodological checks:** Variance decomposition for extremely wet days (defined as daily precipitation exceeding the 99th percentile from GMFD) calculated by taking a simple average of each individual variance estimate (i.e., without weighting to account for model/scenario/downscaled output availability).

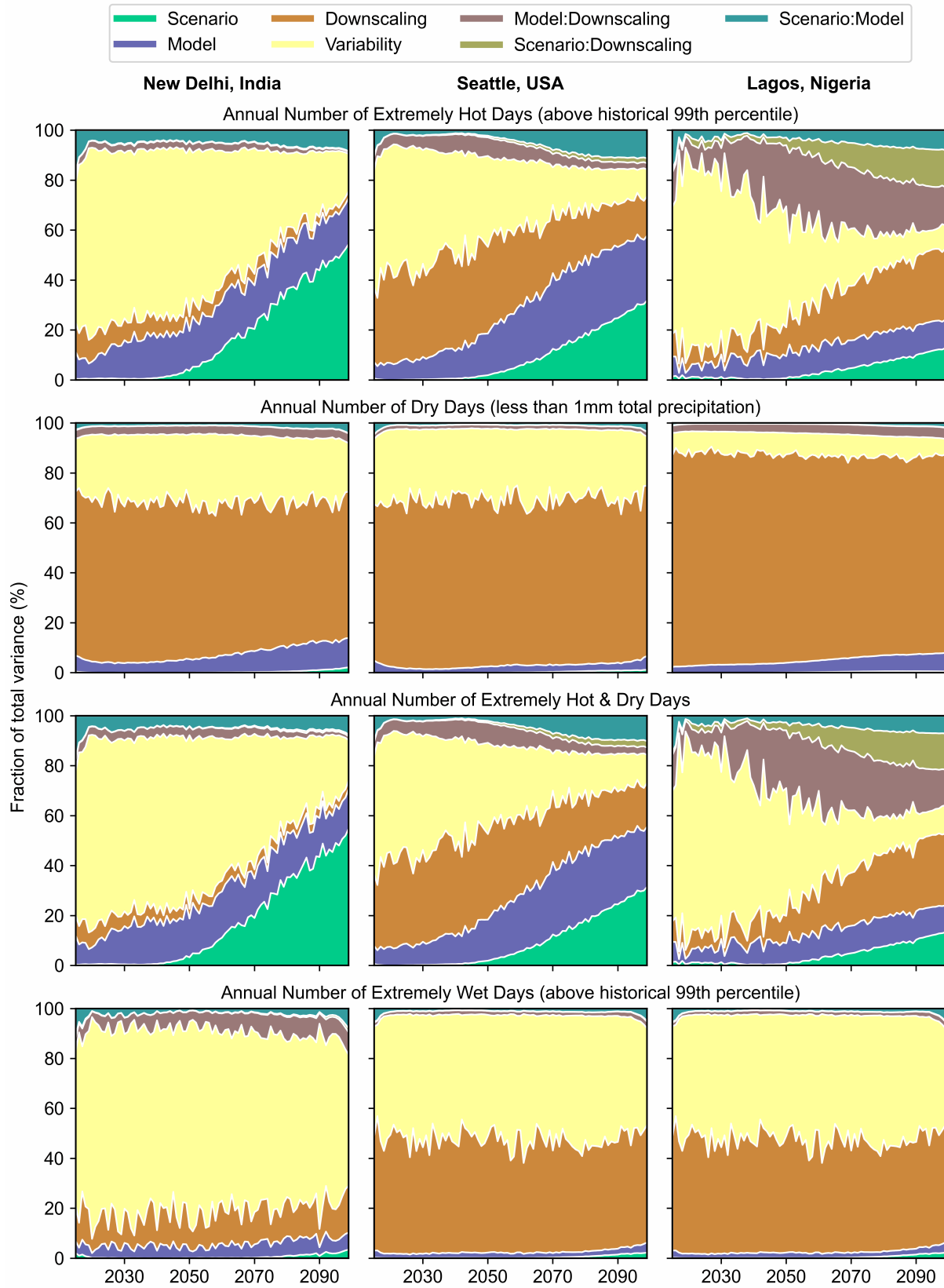


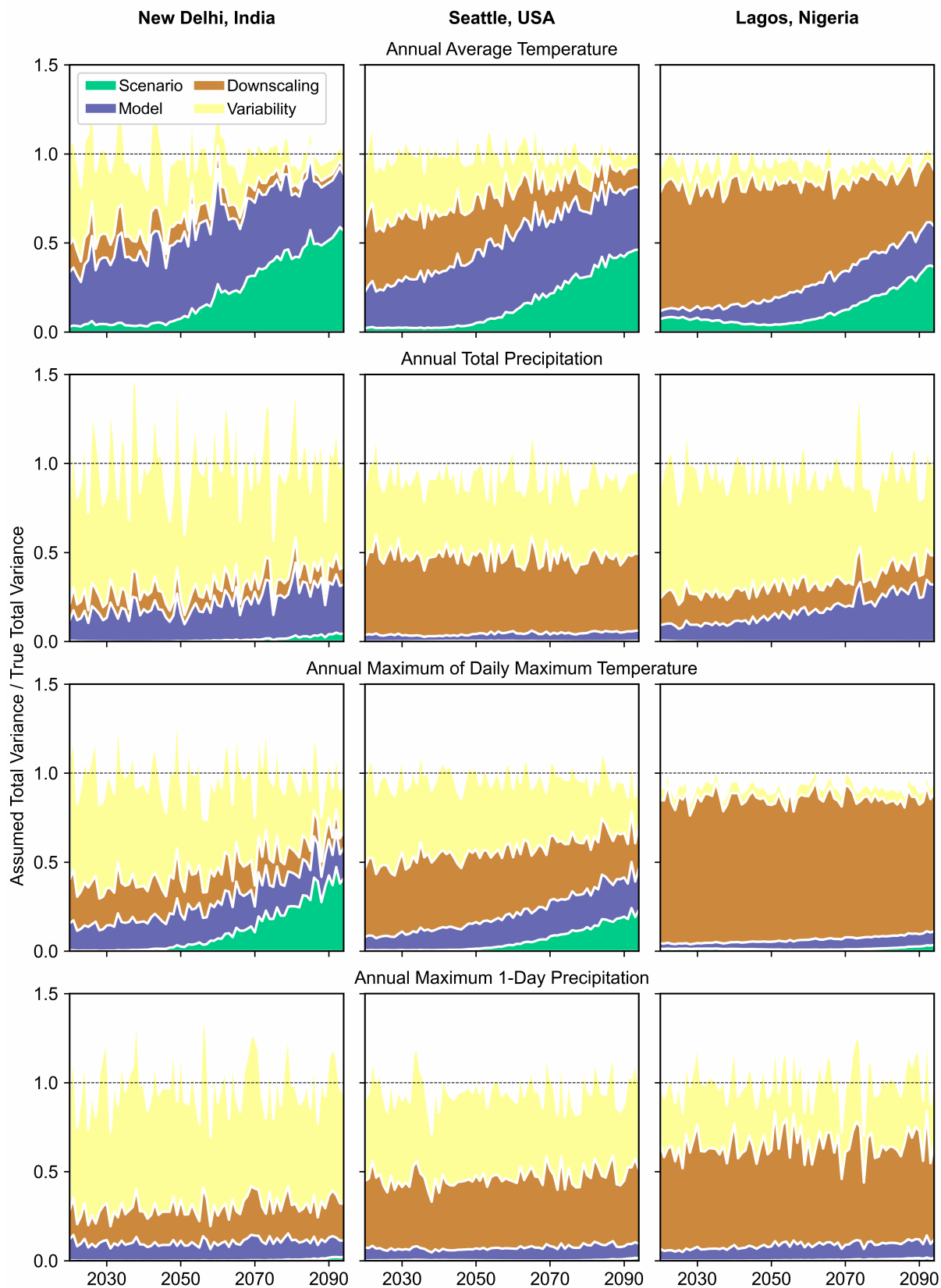
Supplementary Figure 62: **Methodological checks:** Timeseries and resulting variance decomposition in New Delhi, Seattle, Lagos for annual maxima of daily maximum temperature and daily precipitation. This figure can be used for reference against Fig. 64.



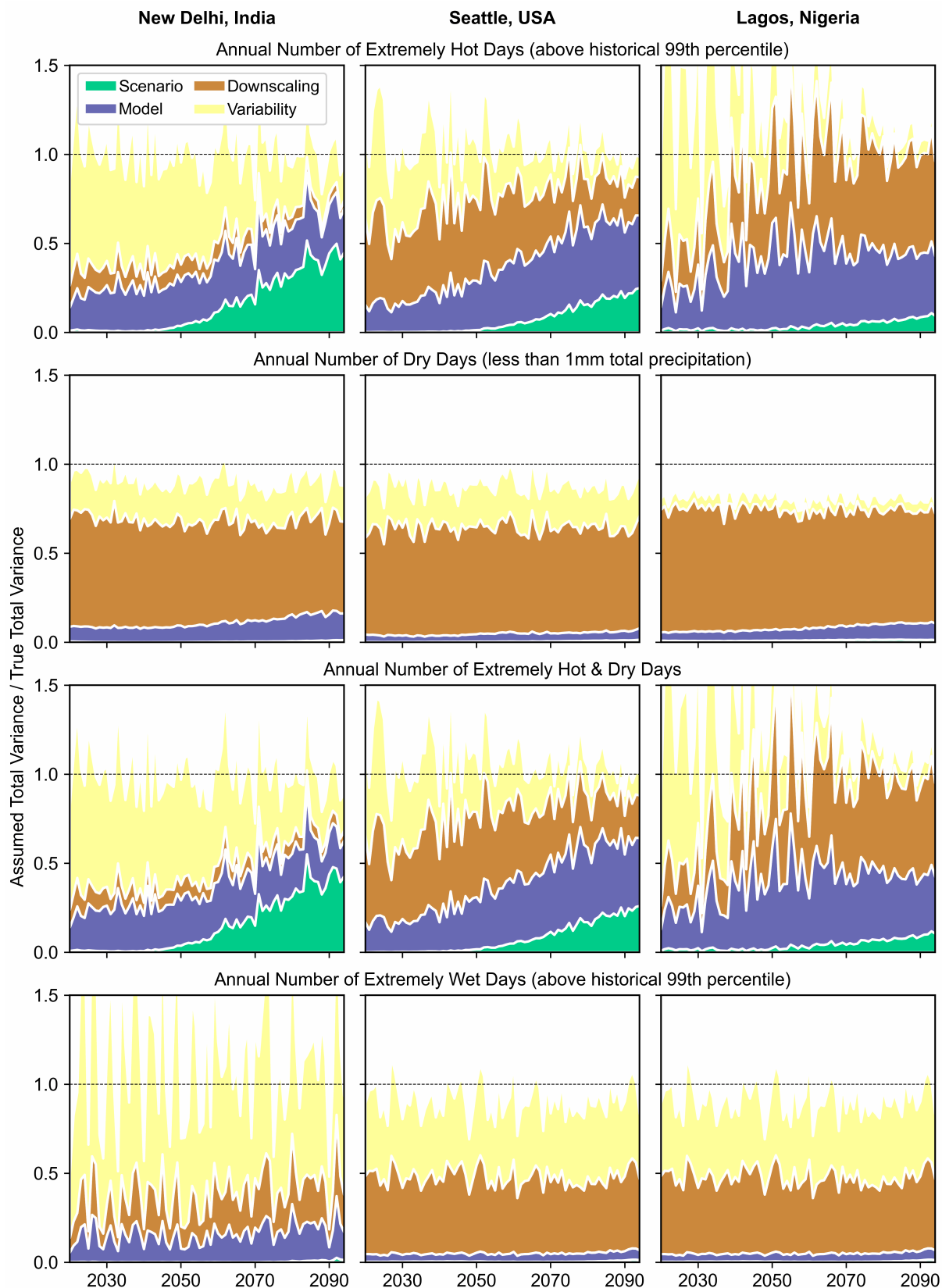
Supplementary Figure 63: **Methodological checks:** Variance decomposition in New Delhi, Seattle, Lagos for all threshold indices of extremes, following the methodology of our main results. This figure can be used for reference against Fig. 65.



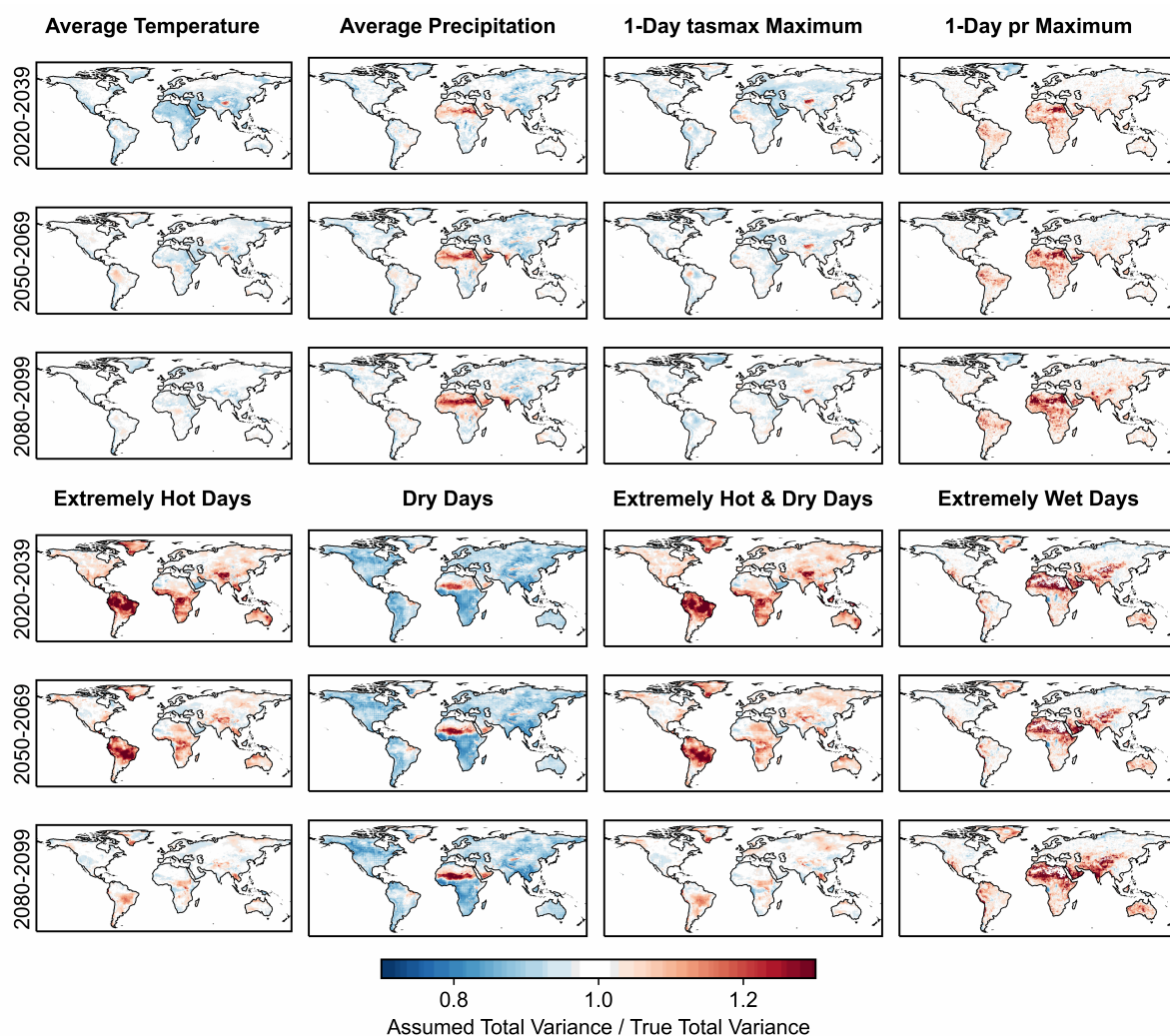




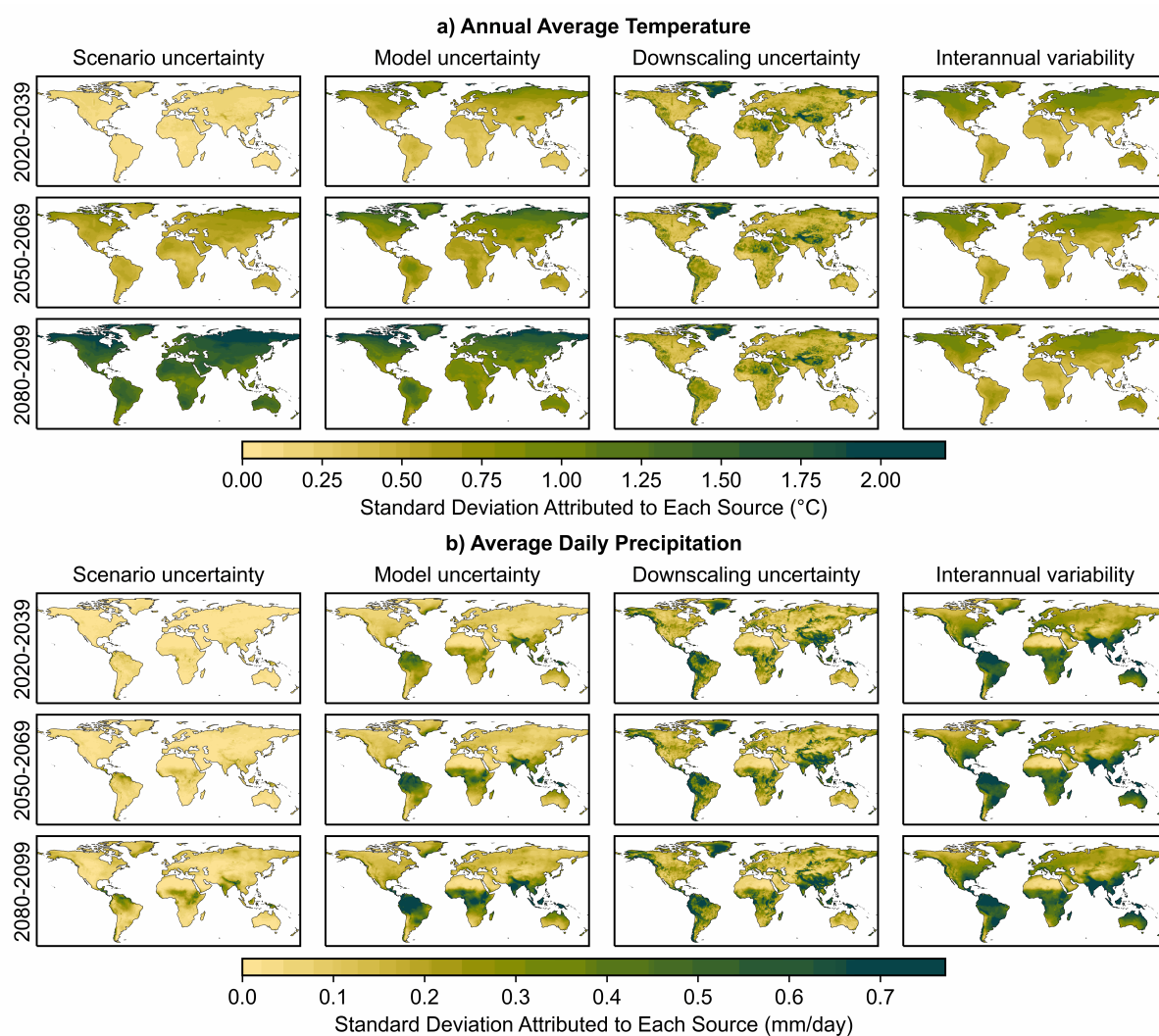
Supplementary Figure 66: **Methodological checks:** Ratio of the assumed total variance (sum of each individual contribution) to the true total variance for annual averages and 1-day maxima in New Delhi, Seattle, Lagos. Large departures from unity indicate the presence of interaction effects that we neglected in the main results.



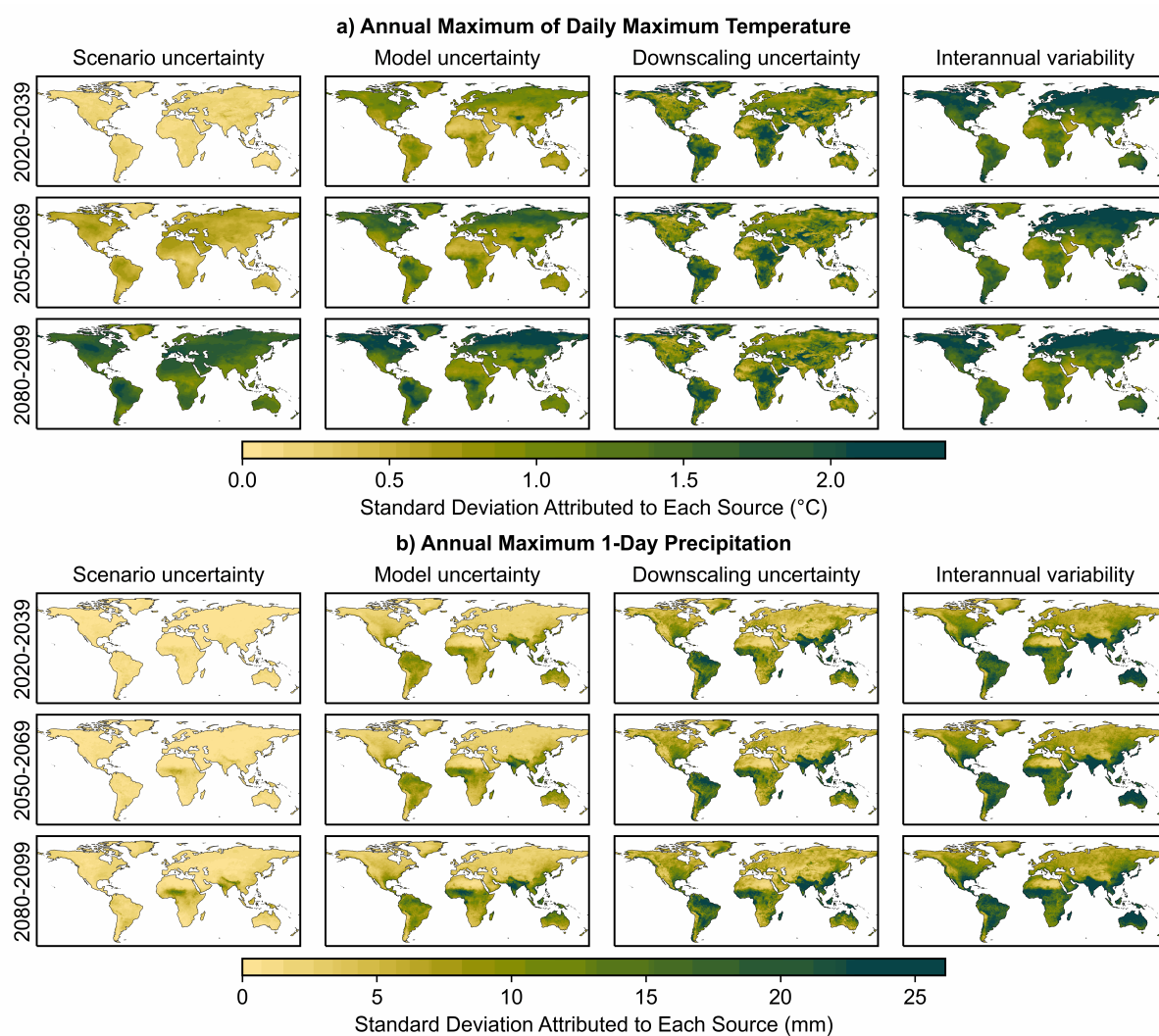
Supplementary Figure 67: **Methodological checks:** Ratio of the assumed total variance (sum of each individual contribution) to the true total variance for threshold indices in New Delhi, Seattle, Lagos. Large departures from unity indicate the presence of interaction effects that we neglected in the main results.



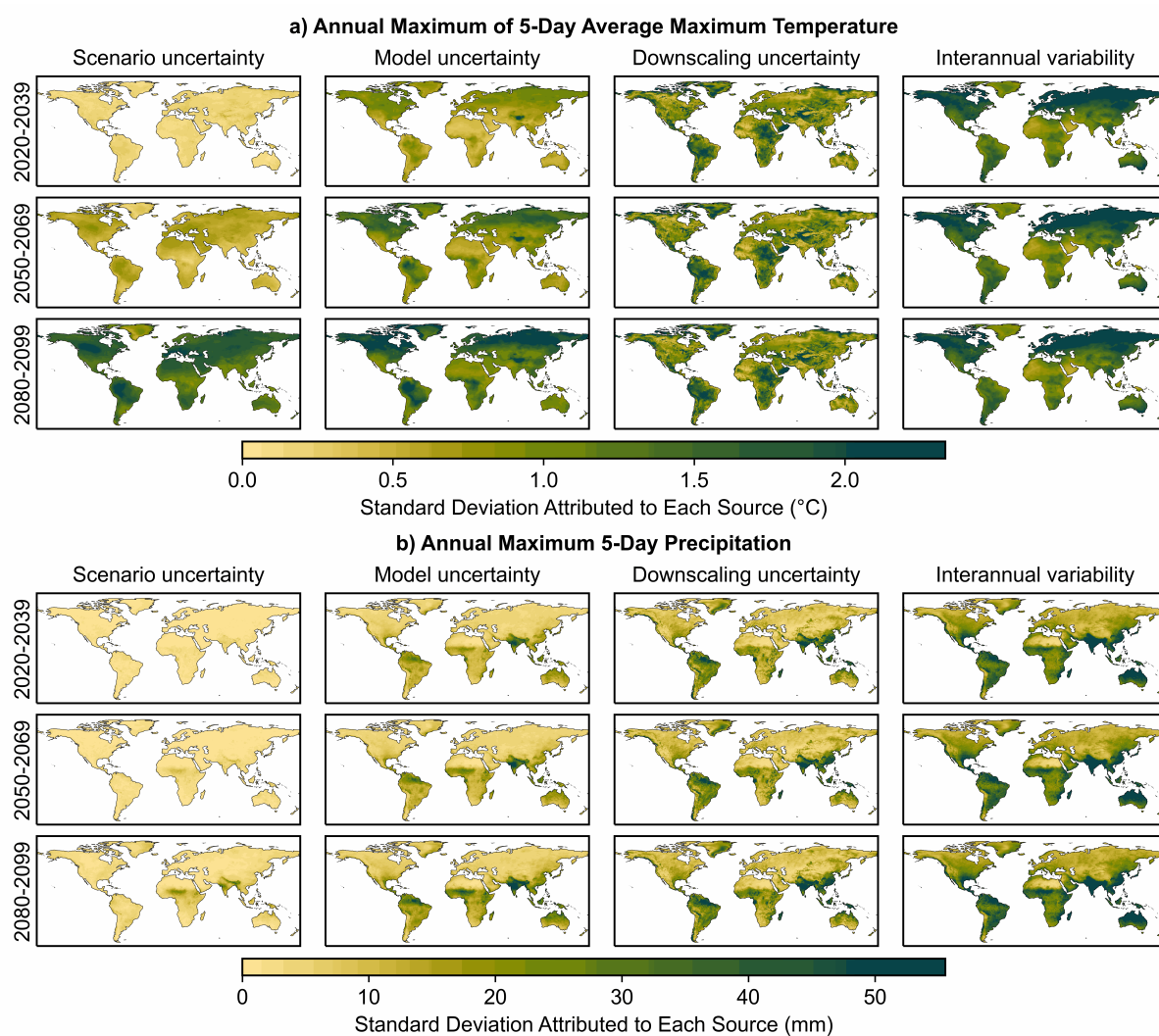
Supplementary Figure 68: **Methodological checks:** Ratio of the assumed total variance (sum of each individual contribution) to the true total variance for all metrics globally. Large departures from unity indicate the presence of interaction effects that we neglected in the main results.



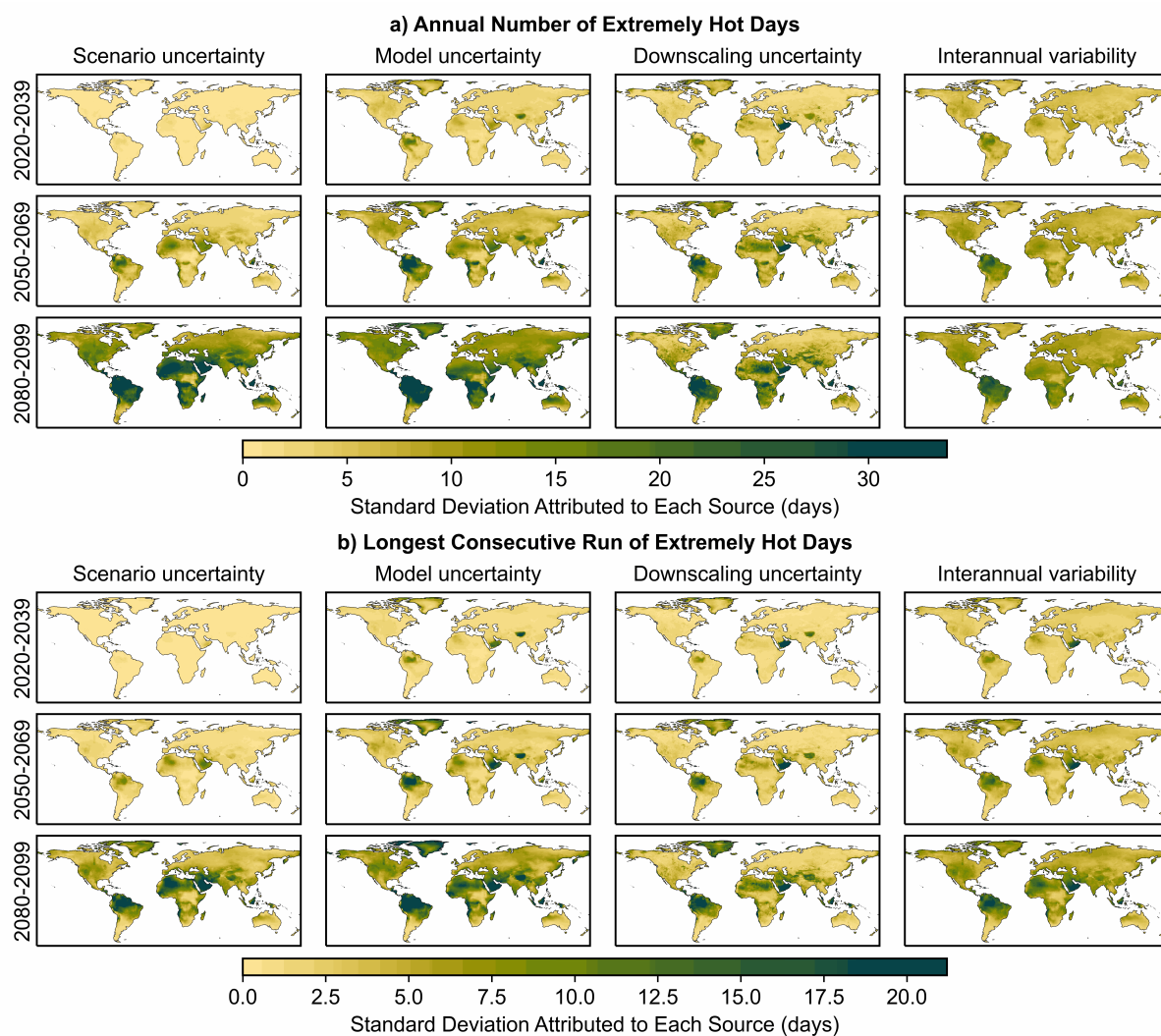
Supplementary Figure 69: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for annual averages.



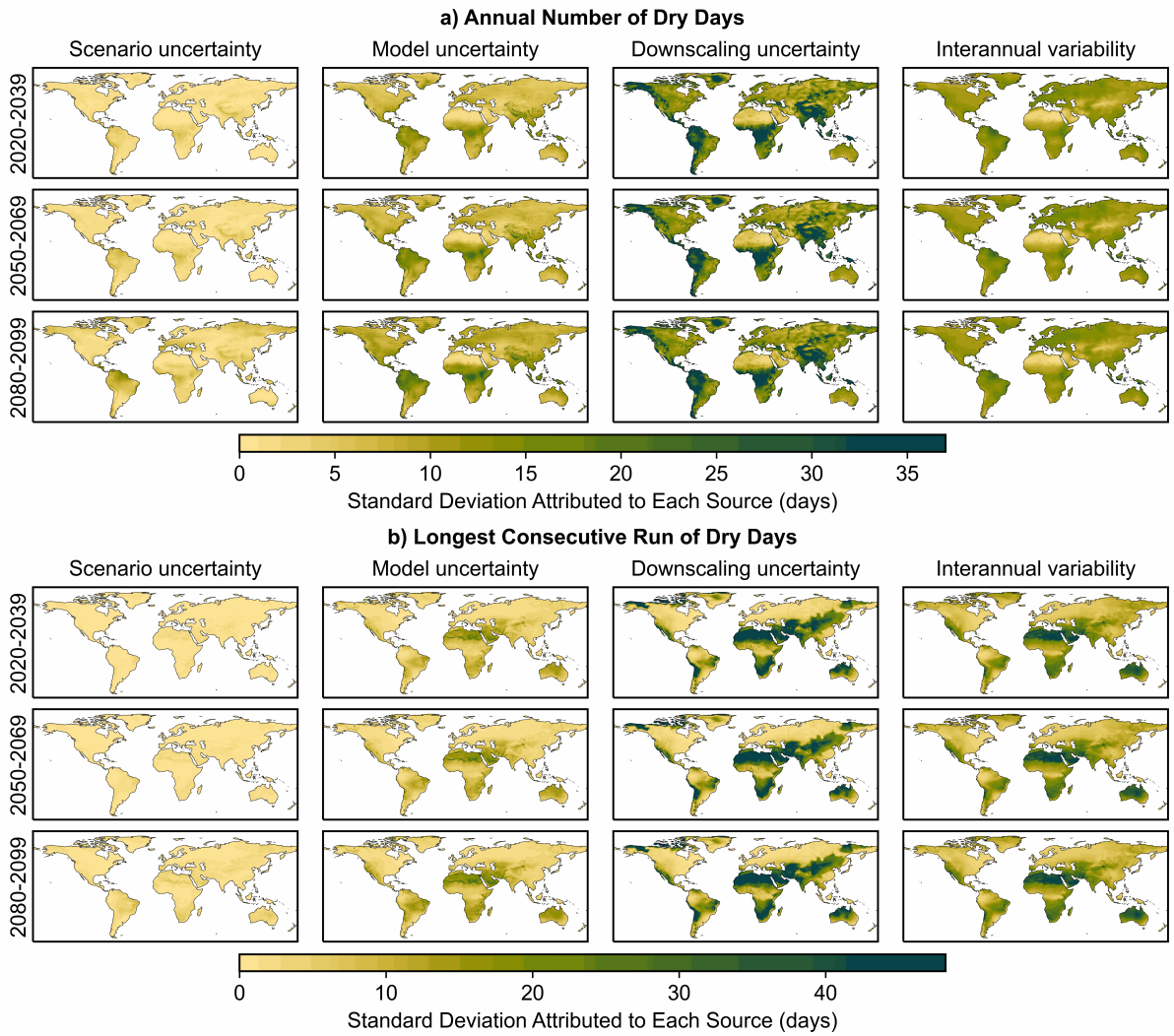
Supplementary Figure 70: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for 1-day maxima.



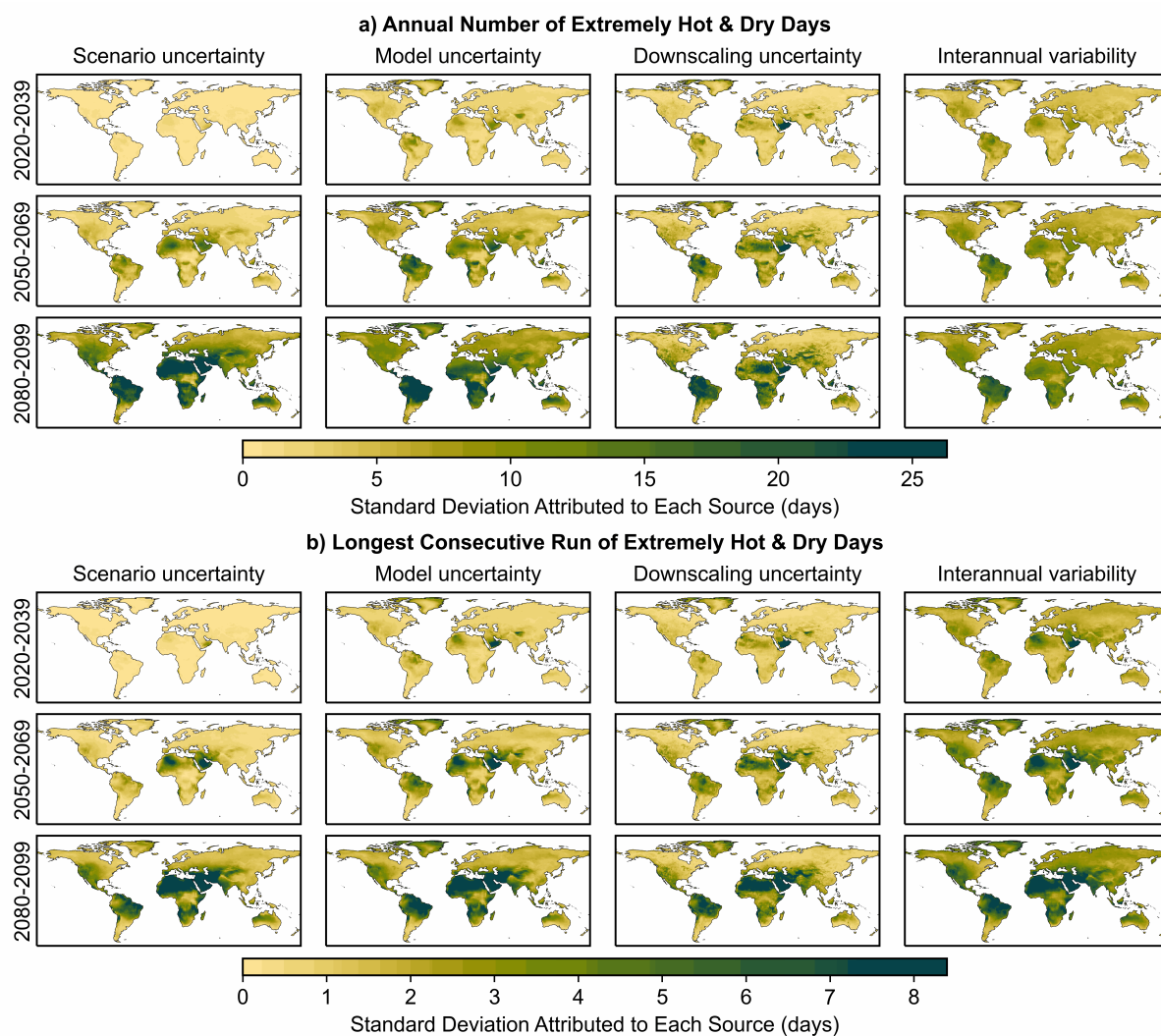
Supplementary Figure 71: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for 5-day maxima.



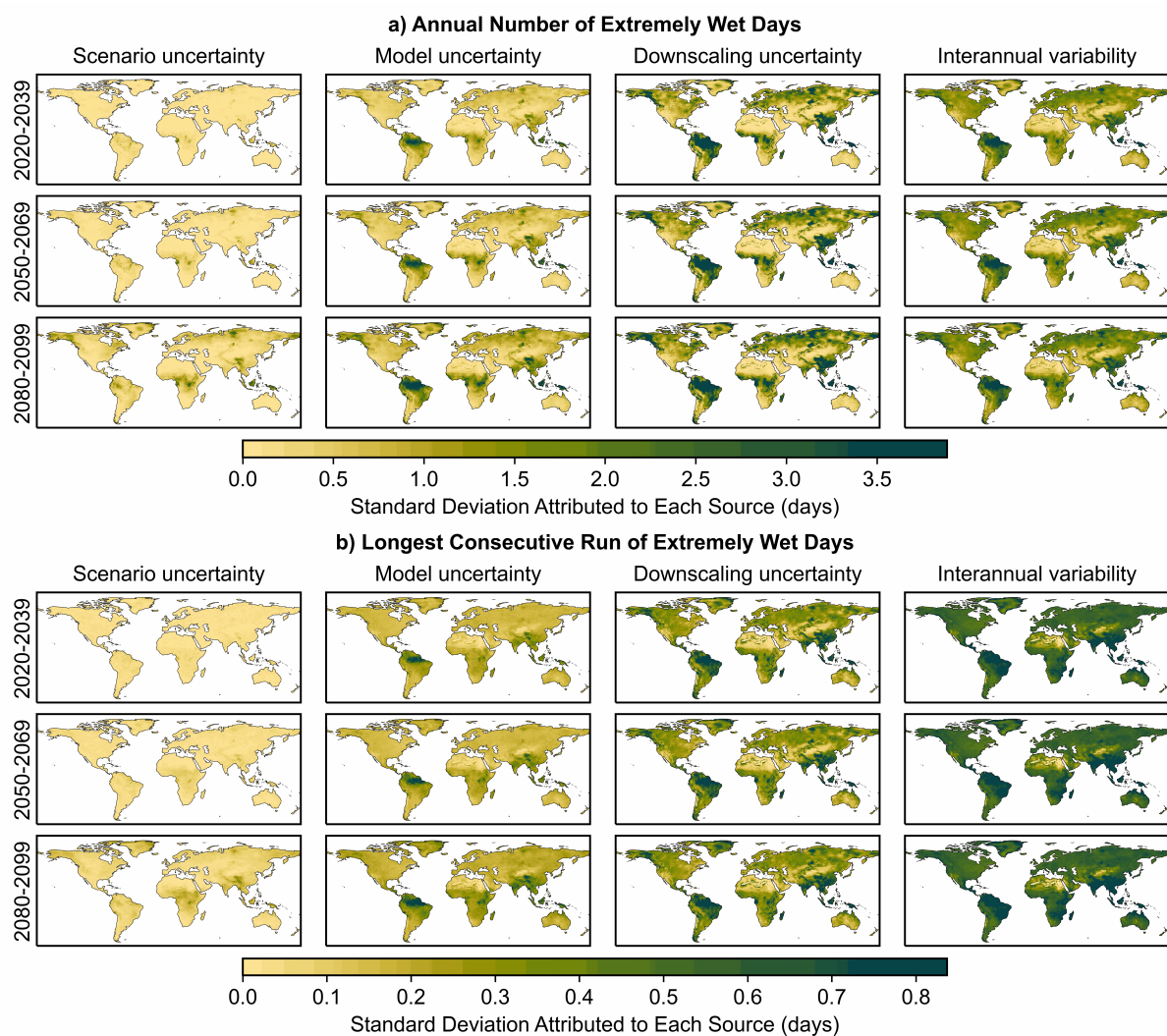
Supplementary Figure 72: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for extremely hot days (defined as daily maximum temperature exceeding the 99th percentile from GMFD).



Supplementary Figure 73: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for dry days (defined as daily precipitation less than 1mm).



Supplementary Figure 74: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for extremely hot & dry days (defined as daily maximum temperature exceeding the 99th percentile from GMFD and daily precipitation simultaneously less than 1mm).



Supplementary Figure 75: **Absolute uncertainty comparison:** The absolute uncertainty (measured by the standard deviation) contributed by each uncertainty source over each time horizon for extremely wet days (defined as daily precipitation exceeding the 99th percentile from GMFD).

Supplementary References

- [1] Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. *Journal of Climate* **19**, 3088–3111 (2006). URL <https://doi.org/10.1175/JCLI3790.1>.
- [2] Hersbach, H. *et al.* The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **146**, 1999–2049 (2020). URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803>.
- [3] Lange, S. *et al.* WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0) (2021). URL <https://doi.org/10.48364/ISIMIP.342217>.
- [4] Cucchi, M. *et al.* WFDE5: bias-adjusted ERA5 reanalysis data for impact studies. *Earth System Science Data* **12**, 2097–2120 (2020). URL <https://essd.copernicus.org/articles/12/2097/2020/>.
- [5] Sheridan, S. C., Lee, C. C. & Smith, E. T. A Comparison Between Station Observations and Reanalysis Data in the Identification of Extreme Temperature Events. *Geophysical Research Letters* **47**, e2020GL088120 (2020). URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL088120>.
- [6] Brekke, L. & Barsugli, J. *Uncertainties in Projections of Future Changes in Extremes*, vol. 65, 309–346 (Springer, Dordrecht, 2013). URL https://doi.org/10.1007/978-94-007-4479-0_11.
- [7] Fox, J. & Weisberg, S. *An R Companion to Applied Regression* (Sage, Thousand Oaks CA, 2019), third edn. URL <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- [8] Thrasher, B. *et al.* NASA Global Daily Downscaled Projections, CMIP6. *Scientific Data* **9**, 262 (2022).
- [9] Thrasher, B. & Weile, W. NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6) Technical Note. URL https://www.nccs.nasa.gov/sites/default/files/NEX-GDDP-CMIP6-Tech_Note.pdf.

- [10] Gergel, D. R. *et al.* Global downscaled projections for climate impacts research (GDPCIR): preserving extremes for modeling future climate impacts. *EGUsphere* **2023**, 1–35 (2023). URL <https://egusphere.copernicus.org/preprints/2023/egusphere-2022-1513/>.
- [11] Malevich, B. *et al.* Climateimpactlab/downscalecmip6: (2022). URL <https://doi.org/10.5281/zenodo.6403794>.
- [12] Lange, S. ISIMIP3BASD (2022). URL <https://doi.org/10.5281/zenodo.7151476>.
- [13] Lange, S. Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0). *Geoscientific Model Development* **12**, 3055–3070 (2019). URL <https://gmd.copernicus.org/articles/12/3055/2019/>.
- [14] Lange, S. ISIMIP3b bias adjustment fact sheet (2021). URL https://www.isimip.org/documents/413/ISIMIP3b_bias_adjustment_fact_sheet_Gnsz7C0.pdf.
- [15] Chegwidan, O. *et al.* Open data and tools for multiple methods of global climate downscaling (2022). URL <https://carbonplan.org/research/cmip6-downscaling-explainer>.
- [16] carbonplan. CMIP6-Downscaling: Downscaling methods (2022). URL <https://docs.carbonplan.org/cmip6-downscaling/downscaling-methods>.
- [17] Gutmann, E. D. *et al.* En-GARD: A Statistical Downscaling Framework to Produce and Test Large Ensembles of Climate Projections. *Journal of Hydrometeorology* **23**, 1545–1561 (2022).
- [18] Matwin, S. *et al.* DeepSD: Generating High Resolution Climate Change Projections through Single Image Super-Resolution. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1663–1672 (2017).