Supplement for "The role of internal variability in global climate projections of extreme events"

Mackenzie L. Blanusa[1*], Carla J. López-Zurita[2] and Stephan Rasp[2,3]

[1]Department of Marine Sciences, University of Connecticut, Avery Point, 1084 Shennecossett Rd, Groton, 06340, CT, U.S..
[*]Work done during an internship at ClimateAi.
[2]ClimateAi, San Francisco.
[3]Now at Google Research.

*Corresponding author(s). E-mail(s): mackenzie.blanusa@uconn.edu;
Contributing authors: carla@climate.ai; raspstephan@gmail.com;

S1 Sensitivity to hyper-parameters in the default extreme definition

The extreme definition used in the main text, herein referred to as the default extreme definition, relies on estimating extreme event occurrences over an aggregation period based on a quantile threshold. There are several hyper-parameters in the default extreme definition including the quantile, coarsening the data, consecutive days, aggregation period, and post processing. Sensitivity analyses have been performed for each hyper-parameter and are discussed in the subsequent sections below. The default values for each hyper-parameter are: quantile/return period = 10 years[1], coarsening = 1 day, consecutive days = 1, aggregation period = 10 years. For each sensitivity analysis, only one parameter was altered at a time and the rest of the hyper-parameters were set to their default values, unless stated otherwise. We use the same three example locations as in the main text. The model percent contribution to the total uncertainty (neglecting scenario uncertainty), $M\%$, is the variable used to assess the sensitivity of each hyper-parameter.

$$\text{Model \% Contribution} = (M(t,l)/(M(t,l) + I(t,l))) * 100 \tag{1}$$

S1.1 Quantile/"Return Period"

The "return period" was set to 5, 10, 20, 50, and 100 years while keeping all other parameters at their default values. To get from the "return period" to the quantile, we first compute the expected number of events, E, in the length of the reference period:

$$E = \#\text{ years in reference period} \,/\, \text{return period} \tag{2}$$

The quantile, q, can then be computed as

$$q = 1 - E \,/\, \#\text{ days in reference period} \tag{3}$$

Generally, barring some noise, larger "return periods" lead to a smaller fraction of model uncertainty (Fig. S1). This is the result of larger internal variability attached to rarer events. Seattle and Montreal daily precipitation and Seattle maximum temperature do not show a clear pattern of the quantile effecting $M\%$ due to the already large contribution of internal variability at small return periods (e.g. Fig. 2). Overall, regions where model uncertainty dominates will be more sensitive to the quantile. While there is certainly a noticeable effect of changing the quantile, the differences are small enough to not drastically change the qualitative results.

[1] As described in the main text, since we are using daily values rather than yearly maximums, our extremes do not correspond to the typical definition of return periods. However, for convenience, we still use the terminology here.

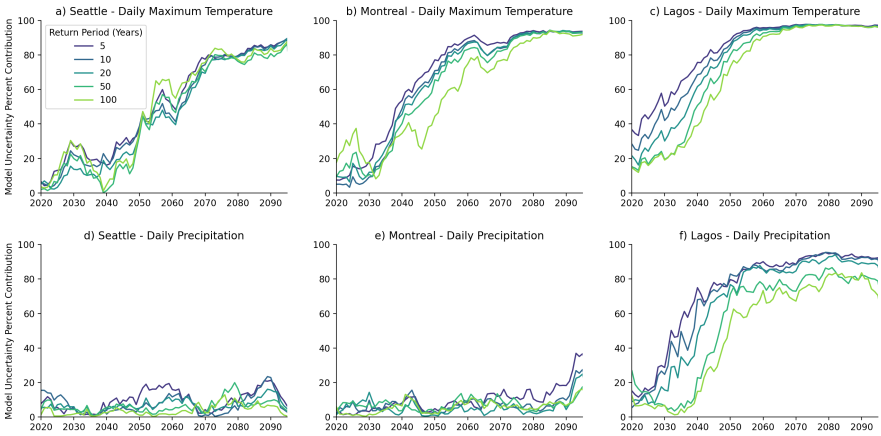

Fig. S1 Model % contribution from 2020-2095 for various "return periods". Return periods of 5, 10, 20, 50, and 100 years were tested. Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

S1.2 Days Coarsened

To avoid counting consecutive days with extreme events as separate events, we can coarsen the data, so that the coarsened period is True if one or more days within indicate an extreme event. We investigate the effects of coarsening the data by setting the parameter to 1, 3, 5 and 7 days. Fig. S2 clearly shows that $M\%$ is insensitive to the choice of days coarsened for both maximum temperature and precipitation. Seattle maximum temperature shows the most spread, with values converging at the end of the century.

S1.3 Consecutive Days

We also test the impact of using multi-day averages instead of individual daily values. This is motivated by extreme events, especially for temperature, lasting several days in many cases. Consecutive days of 1, 3, 5, and 7 were tested and all other parameters were set to their default values except the coarsening period, which was set to 7 days, since doubly counting of events is potentially a bigger issue for consecutive day extremes (co-varying both parameters ended up having a small effect in the end, however). Generally, consecutive days was found to have negligible effects on $M\%$ (Fig. S3). However, there are notable effects for Seattle maximum temperature that show less $M\%$ over time for greater consecutive days. Differences for Seattle may be attributed to the larger contributions of internal variability and scenario uncertainty that persist throughout the century (Fig. 2).

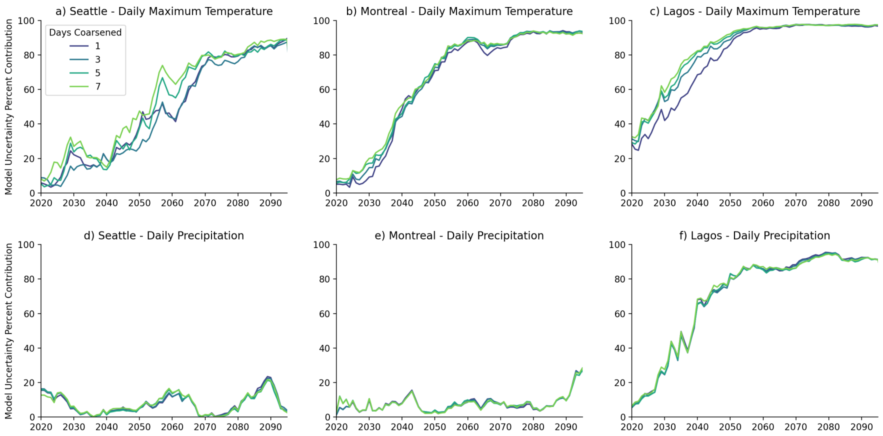

Fig. S2 Model % contribution from 2020-2095 for various coarsening values. Coarsening periods of 1, 3, 5, and 7 days were tested. Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

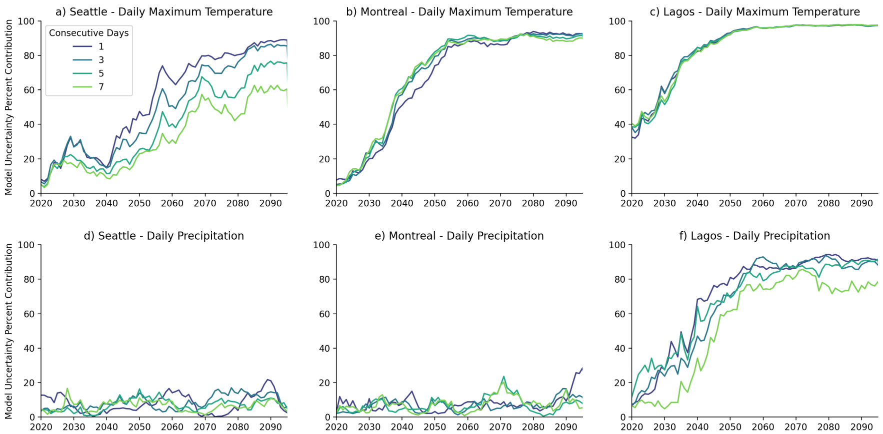

Fig. S3 Model % contribution from 2020-2095 for various consecutive periods. Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

S1.4 Aggregation Period

We look at the average occurrence of extreme events over an aggregation period, with a default value of 10 years. As expected, this hyper-parameter has significant effects on $M\%$, where a greater aggregation period results in greater $M\%$ (Fig. S4). When increasing the aggregation period, the data becomes smoother, which reduces the contribution of internal variability. This sensitivity is to be expected. The default value of 10 years was selected to be consistent with previous studies (Hawkins and Sutton, 2009, 2011; Lehner et al, 2020) and our own experience from delivering climate projections to customers.

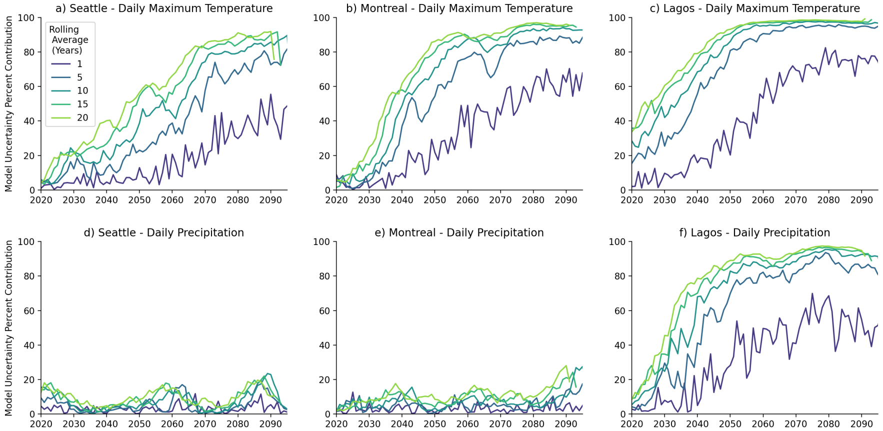

Fig. S4 Model % contribution from 2020-2095 for various aggregation periods. Rolling averages of 1, 5, 10, 15, and 20 years were tested. Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

S2 Sensitivity to post-processing

Computing the quantile value for each model separately performs an implicit quantile bias correction. In Fig. S5 we compare this to taking a single quantile value across all models ("No implicit post-p"). Naturally, model variability increases drastically since climate models have large regional biases.

We further test a state-of-the-art postprocessing method, quantile delta mapping (QDM) (Cannon et al, 2015). It was implemented taking a 30 year window as well as 3-month windows to ensure seasonality, as in the original reference. The QDM has an impact on the variance partitioning but the effect are non-systematic. This suggests that further analysis would be required to make any conclusive statements. We did not perform the QDM globally, because the algorithm is quite computationally expensive.

S3 Temperature variable

We tested the impact of using mean temperature instead of maximum temperature on $M\%$ (Fig. S6). Overall, the differences are negligible and since the study is focused on extremes we decided to utilize daily maximum temperature.

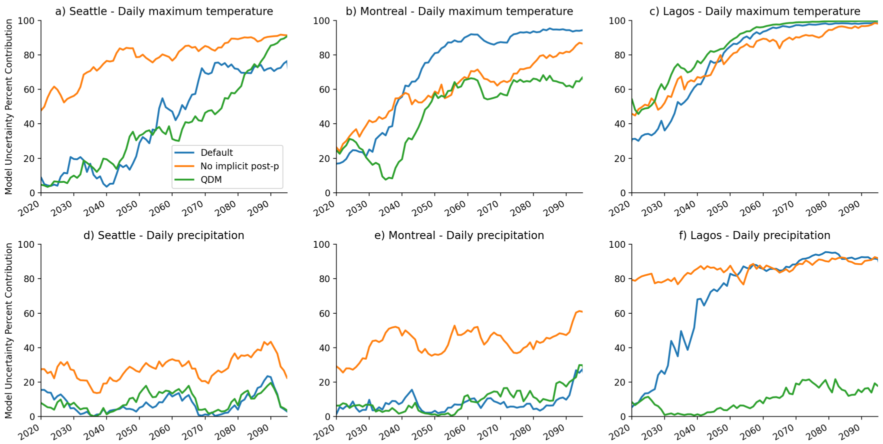

Fig. S5 Model % contribution from 2020-2095 for default data (blue), default data with no implicit post-processing (orange), and post-processed data using quantile delta mapping (greeen) . Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

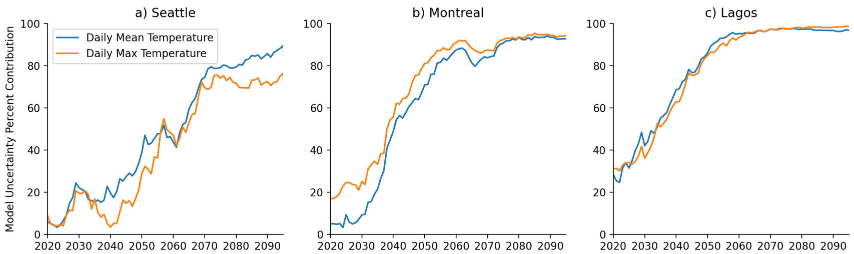

Fig. S6 Model % contribution from 2020-2095 for daily mean temperature (blue) and daily maximum temperature (orange). Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

S4 Sensitivity to extreme event definition

We also investigated different extreme event definitions from the daily quantile definition used in the main paper.

S4.1 T/PXx return

First we can compute the return period based on annual maximum of maximum temperature (TXx) and daily precipitation (PXx) values rather than daily data. This is consistent with the usual definition of return periods. One problem with this is that, in a high emission scenario, this metric will quickly saturate, as all years will have a larger maximum temperature as a 1-in-10 year event in the past. Hence, the model uncertainty contribution flattens out quickly for temperature and for Lagos precipitation (Fig. S7). This metric therefore is not appropriate for our context. One would have to look at much longer return periods to obtain reasonable results.

S4.2 T/PXx

One can also look at the variability of T/PXx itself, rather than the associated event occurrence. In other words, x would then be the annual maximum in °C and mm. In figure S7 we see that this has a non-negligible effect for the three example locations, going in different directions. For TXx, results stay similar in Seattle but the contribution of model uncertainty decreases in Montreal while it increases in Lagos. For PXx, the results a qualitatively unchanged for Seattle and Montreal, with a larger change in Lagos. Looking at global figures (not shown) suggests that for most locations the general picture is unchanged, however.

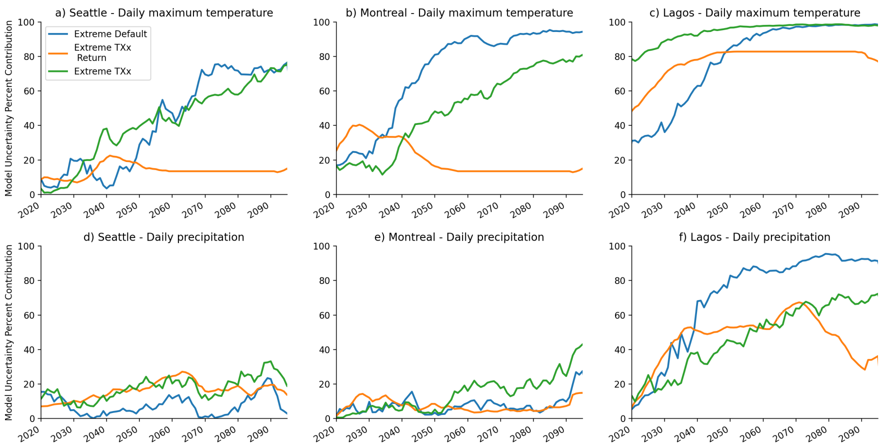

Fig. S7 Model % contribution from 2020-2095 for three extreme event definitions: default (blue), extreme TXx return (orange), and extreme TXx (green). Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f).

S5 Sensitivity to partitioning variability method

We can compare the large ensemble method of estimating uncertainty partitioning with the fitting method introduced by HS09. For this we follow L20 and use the first ensemble member to estimate a fourth-order polynomial fit (Fig. S8). The fit generally agrees quite well with the ensemble mean for temperature, but for precipitation the fit method picks up some noise as forced response.

We can also apply the fitting method for our larger sample of 14 single realization CMIP6 models used for estimating the scenario uncertainty. The results for I and M for the different methods are shown in Figure S9. The biggest drawback of the fitting method is that it estimates a constant I, which is clearly not appropriate for our extreme event definition. For temperature, M estimates from the large ensemble and fitting methods generally agree in magnitude and trend. However, for precipitation large differences can be seen, a result of the fitting method picking up signal as noise. When using the fitting method on CMIP6 models larger differences can be seen, for example in Seattle temperature. We did not further investigate this behavior but also note that our quantile estimation method might be unreliable for single realization models due to low sample size (only 2 events expected in a 20 year reference period). Internal variability estimates seem to agree between the large ensemble and CMIP6 fits but are obviously not comparable to the time-changing I in the large ensemble method for most locations.

Finally, we can look at the I estimates for each of the large ensembles separately (Fig. S10). Here we see differences up to a factor of 5 between models but also general agreement on the order of magnitude compared to M.

This analysis shows that in most cases the methods agree on the order of magnitude of I and M but also hints at large uncertainties in the method as already described in Section 4.1 in the main text.

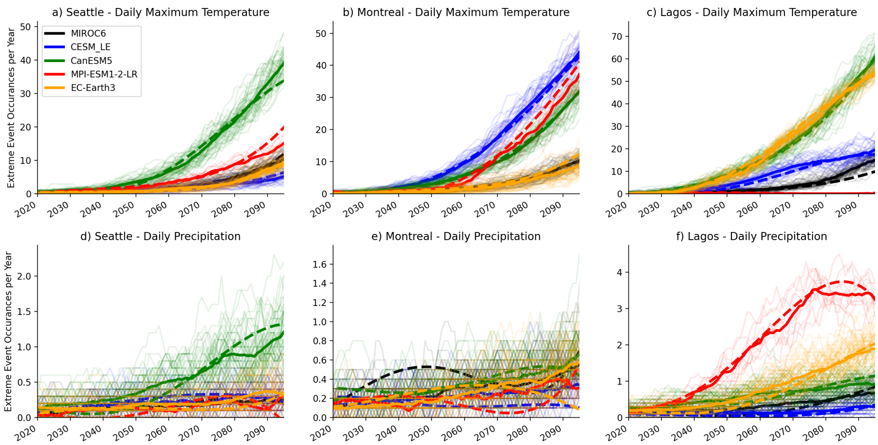

Fig. S8 Extreme event occurrence (days) per year from 2020-2095 as estimated from five large ensembles. Occurrences were calculated from a 10 year rolling average and divided by 10 to obtain a yearly estimate. Top row: daily maximum temperature for three locations; Seattle, U.S.A (a), Montreal, Canada (b), Lagos, Nigeria (c). Bottom row: daily precipitation for the same three regions (d-f). Dashed lines are 4th order polynomial fit for the first ensemble member of each model.

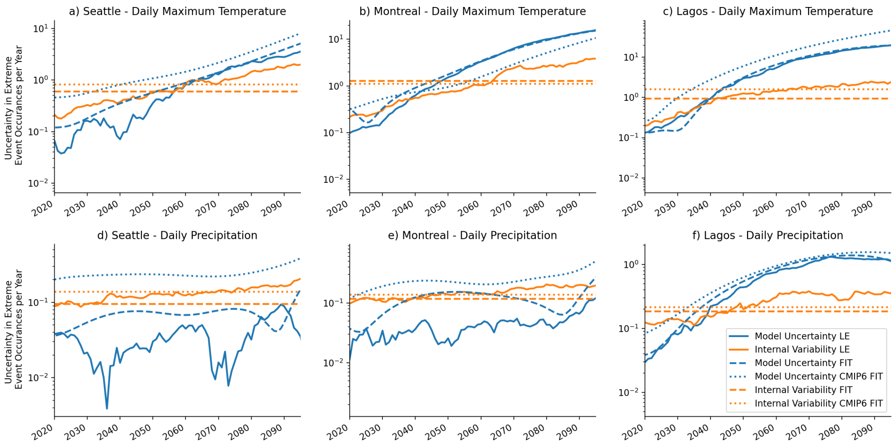

Fig. S9 Model uncertainty (blue) and internal variability (orange) from 2020-2095 for the large ensemble method (solid), fit method using large ensemble models (dashed), and fit method using CMIP6 single realization models (dotted). Top row: daily maximum temperature and bottom row: daily precipitation for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), and Lagos, Nigeria (c,f). Uncertainties were calculated with a 10 year rolling average and divided by 10 to obtain a yearly estimate.

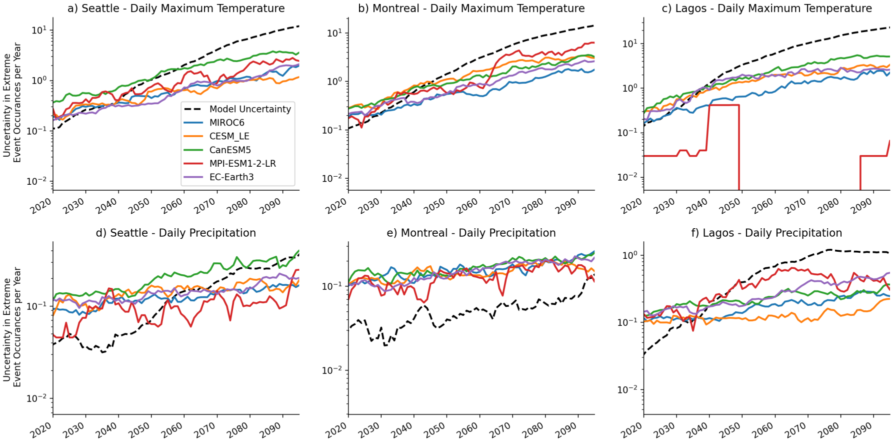

Fig. S10 The spread of internal variability as estimated from the five large ensembles and model uncertainty (black, dashed) from 2020-2095 for three locations: Seattle, U.S.A (a), Montreal, Canada (b), Lagos, Nigeria (c). Bottom row: daily precipitation for the same three regions (d-f). Uncertainties were calculated with a 10 year rolling average and divided by 10 to obtain a yearly estimate.

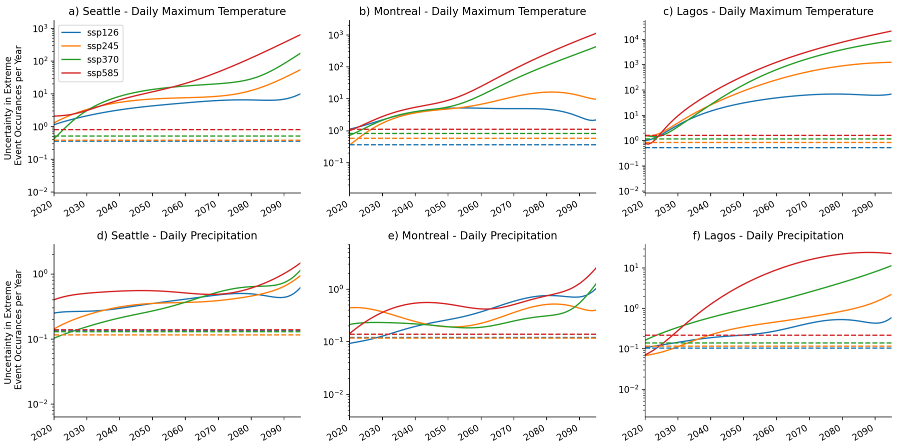

Fig. S11 Sensitivity analysis for the effects of scenario forcing on model uncertainty and internal variability. Each panel shows model uncertainty (solid) and internal variability (dashed) for four SSP scenarios: SSP1-2.6 (blue), SSP2-4.5 (orange), SSP3-7.0 (green), and SSP5-8.5 (red). Top row is for daily maximum temperature for three locations: Seattle, U.S.A (a), Montreal, Canada (b), and Lagos, Nigeria (c). Bottom row (d-f) is for daily precipitation for the same three locations. Uncertainties were calculated with a 10 year rolling average and divided by 10 to obtain a yearly estimate.

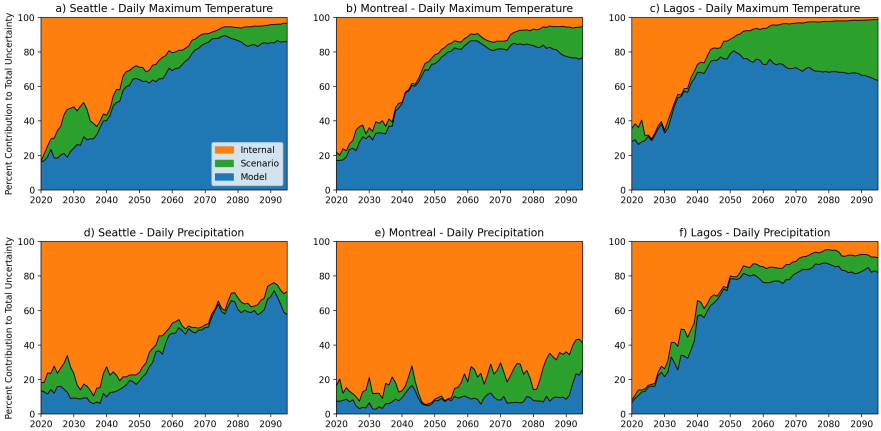

Fig. S12 Percent contribution to total uncertainty including CanESM5 model for model uncertainty, scenario uncertainty, and internal variability from 2020-2095 for three locations: Seattle, U.S.A (a,d), Montreal, Canada (b,e), Lagos, Nigeria (c,f). Top row (a-c) is for daily maximum temperature and bottom row (d-f) for daily precipitation.

References

Cannon AJ, Sobie SR, Murdock TQ (2015) Bias correction of gcm precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? Journal of Climate 28(17):6938 – 6959. https://doi.org/10.1175/JCLI-D-14-00754.1, URL https://journals.ametsoc.org/view/journals/clim/28/17/jcli-d-14-00754.1.xml

Hawkins E, Sutton R (2009) The Potential to Narrow Uncertainty in Regional Climate Predictions. Bulletin of the American Meteorological Society 90(8):1095–1108. https://doi.org/10.1175/2009BAMS2607.1, URL https://journals.ametsoc.org/view/journals/bams/90/8/2009bams2607_1.xml, publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society

Hawkins E, Sutton R (2011) The potential to narrow uncertainty in projections of regional precipitation change. Climate Dynamics 37(1-2):407–418. https://doi.org/https://doi.org/10.1007/s00382-010-0810-6, URL https://www.proquest.com/docview/874300071/abstract/41E2758CCFAA4901PQ/1, num Pages: 407-418 Place: Heidelberg, Netherlands Publisher: Springer Nature B.V.

Lehner F, Deser C, Maher N, et al (2020) Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. Earth System Dynamics 11(2):491–508. https://doi.org/10.5194/esd-11-491-2020, URL https://esd.copernicus.org/articles/11/491/2020/, publisher: Copernicus GmbH