Showcasing MESMER-X: Spatially resolved emulation of annual maximum temperatures of Earth System Models

Yann Quilcaille[1], Lukas Gudmundsson[2], Lea Beusch[3], Mathias Hauser[2], and Sonia I. Seneviratne[1]

[1]ETH Zurich
[2]ETH Zürich
[3]Federal Office of Meteorology and Climatology MeteoSwiss and ETH Zurich

November 24, 2022

Abstract

Emulators of Earth System Models (ESMs) are complementary to ESMs by providing climate information at lower computational costs. Thus far, the emulation of spatially resolved climate extremes has only received limited attention, even though it is one of the most impactful aspects of climate change. Here, we propose a method for the emulation of local annual maximum temperatures, with a focus on reproducing essential statistical properties such as correlations in space and time. We test different emulator configurations and find that driving the emulations with global mean surface temperature offers an optimal compromise of model complexity and performance. We show that the emulations can mimic the temporal evolution and spatial patterns of the underlying climate model simulations and are able to reproduce their natural variability. The general design and the good performance for annual maximum temperatures suggests that the proposed methodology can be applied to other climate extremes.

Posted on 24 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10511207.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Showcasing MESMER-X: Spatially resolved emulation of annual maximum temperatures of Earth System Models

Y. Quilcaille[1], L. Gudmundsson[1], L. Beusch[1,*], M. Hauser[1], and S.I. Seneviratne[1]

[1] Institute for Atmospheric and Climate Science, Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland.

[*] Now at: Center for Climate Systems Modeling (C2SM), ETH Zurich, Zurich, Switzerland and MeteoSwiss, via ai Monti 146, 6605 Locarno-Monti, Switzerland

Corresponding author: Yann Quilcaille (yann.quilcaille@env.ethz.ch)

Key Points:

- We present a method for the emulation of spatially resolved annual maximum temperatures from Earth System Models.
- The emulator reproduces statistical properties and correlations in space and time and can be extended to other extreme indicators.
- The method exhibits good performance for annual maximum temperatures.

(The above elements should be on a title page)

Abstract

Emulators of Earth System Models (ESMs) are complementary to ESMs by providing climate information at lower computational costs. Thus far, the emulation of spatially resolved climate extremes has only received limited attention, even though it is one of the most impactful aspects of climate change. Here, we propose a method for the emulation of local annual maximum temperatures, with a focus on reproducing essential statistical properties such as correlations in space and time. We test different emulator configurations and find that driving the emulations with global mean surface temperature offers an optimal compromise of model complexity and performance. We show that the emulations can mimic the temporal evolution and spatial patterns of the underlying climate model simulations and are able to reproduce their natural variability. The general design and the good performance for annual maximum temperatures suggests that the proposed methodology can be applied to other climate extremes.

Plain Language Summary

Climate models are invaluable tools for studying climate change but take a very long time to run, even on modern super computers. Emulators of climate models are statistical tools that can be calibrated to mimic the behaviour of complex climate models with a much reduced computational demand. However, they are typically not made for reproducing climate extremes, despite the fact that extreme climate events belong to the most impactful aspects of climate change. In this study, we propose a method for the emulation of annual maximum temperature over time and space. This method also reproduces the natural variability of climate models, even though it is driven only by global mean surface temperature. We show that the emulations are very similar to the data created by climate models. In an example application, we use the emulator to examine the extreme temperatures for different climate scenarios.

TOTAL WORDS: 3892 / 4000

1 Introduction

The impacts of climate change will affect the entire social and economical system (IPCC, 2014, 2021). In particular, changes in climate extremes count among the most impactful consequences of climate change. Climate extremes are substantially affected by human-induced climate change (Seneviratne et al., 2021). For example, the annual average losses to weather-related disasters were USD168 billions per year over 2001-2010 and have increased to USD248 billions per year over 2011-2020 (Aon, 2021). Climate extremes affect numerous economical sectors, for instance agriculture (Sivakumar et al., 2005; Vogel et al., 2019) or the energy sector (Schaeffer et al., 2012; Perera et al., 2020). Not only do climate extremes have direct consequences on food or energy security (Hasegawa et al., 2021), but they may also have indirect impacts on societies due to feedbacks with societal drivers (Raymond et al., 2020). Even if climate change is limited to 1.5°C, changes in climate extremes remain a crucial issue (Seneviratne et al., 2018), and society will be impacted in many aspects (IPCC, 2018).

Traditionnaly, Earth System Models (ESMs) are used to derive climate change projections and the associated climate extremes (Flato et al., 2013; Collins et al., 2013; Lee et al., 2021). These outputs are crucial to assess what consequences climate extremes would have on society (Rosenzweig et al., 2017). However, ESMs require detailed scenarios to simulate climate change and have very high computational cost, solving a very large number of equations on several grids. These requirements make ESMs expensive tools and hinder their use to explore new scenarios and to characterize the internal climate variability.

ESM emulators have been developed for a quicker assessment of climate change in response to given scenario pathways. A large class of emulators, termed “simple climate models” or “reduced complexity models” provide projections of key variables of the Earth system such as global mean temperature (Nicholls et al., 2020; Nicholls et al., 2021), however they do not provide local information which is essential for studying climate impacts. A second class of emulators derives spatially resolved climate responses from global mean temperature trajectories (“spatially resolved emulators”), such as the recently developed Modular Earth System Model Emulator with spatially Resolved output (MESMER) (Beusch et al., 2020) that this work builds upon. Spatially resolved emulators usually rely on some version of pattern scaling to derive local responses from global variables (Mitchell, 2003; Fordham et al., 2012; Herger et al., 2015; Lynch et al., 2017; Alexeeff et al., 2018). While other approaches exist (Castruccio et al., 2014; Holden et al., 2014), pattern scaling shows good performances in spite of its simplicity (Tebaldi and Arblaster, 2014; Tebaldi and Knutti, 2018). For the representation of natural variability, there is no single most established method. Some emulators resample actual ESM fields (McKinnon et al., 2017; Alexeeff et al., 2018), some resample principle components with perturbed phases (Link et al., 2019), and others rely on autoregressive processes with spatially correlated innovations (Beusch et al., 2020; Nath et al., 2021). Almost all currently available spatially resolved emulation approaches have been developed to emulate mean quantities, but to better assess the impacts of climate change for diverse emission pathways, emulation of climate extremes is needed too. A first step in this direction has been made by (Tebaldi et al., 2020), using pattern scaling to emulate the average evolution of climate extremes, but does not consider natural variability. Thus, an emulator that reproduces the full distribution of the climate extremes is still lacking.

In this paper, we propose a new method for the emulation of climate extremes that accounts for both the spatio-temporal structure and their internal variability. Building on the MESMER emulator, the presented approach is referred as MESMER-X. The statistical framework of the method is introduces in Section 3. We use annual maximum temperature data from the 6th phase of the Coupled Model Intercomparison Project (CMIP6, (Eyring et al., 2016)) to illustrate our method (Section 4). Finally, we discuss the potential of this method for extension to other climate extremes (Section 5).

2 Data

Simulations from 18 ESMs contributing to Scenario Model Intercomparison Project (ScenarioMIP, (O'Neill et al., 2016)) of CMIP6 are considered (listed in Supplementary table S.1). We use the ESMs which provide data for concentration-driven historical (Meinshausen et al., 2017) and for at least two of the five scenarios SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 (Meinshausen et al., 2020). As another condition, we retain only the ESMs providing the daily maximum near-surface air temperature, the near-surface air temperature and the downward surface sensible heat flux over the ocean.

All simulations are interpolated to the same 2.5° x 2.5° grid using second-order conservative remapping for the two temperatures and inverse distance-weighted average remapping for the heat flux (Brunner et al., 2020a). Spatially resolved local annual maximum temperature (TXx) is calculated as the annual maximum of the daily maximum temperature. The anomaly of the local annual maximum temperature is defined by substracting the 1850-1900 mean.The global mean surface air temperature (GSAT) is derived by first averaging annual mean near-surface temperature, then its anomaly is also calculated by subtracting the 1850-1900 mean. The same operations are performed to obtain the global downward heat flux in sea water (HFDS).

In Section 4, both the global trend and global variability of GSAT and HFDS are used to identify adequate drivers for the emulations. These two components are decomposed using a locally weighted scatterplot smoothing, accounting for volcanic eruptions as explained in (Beusch et al., 2020). For the sake of clarity, this paper shows mostly results with the global trend of GSAT, but the full results are shown in supplementary information.

In this paper, some results are aggregated to sub-continental regions defined for the 6th Assessment Report of IPCC regions (Iturbide et al., 2020).

3 A method for the emulation of climate extremes

3.1 Statistical distribution of local climate extremes

Climate variables can be characterized by stochastic processes, and climate extremes are rare values or events of these climate variables, in the tail of their probability distribution (Wilks, 2011; Storch and Zwiers, 1999). This definition implies that changes in the distribution of climate variables will also affect the distribution of climate extremes. For instance, if the local annual surface temperature increases, it is likely that the local annual maximum surface temperature will increase as well. Regional anomalies of climate extremes have been found to scale linearly with anomalies in GSAT (Seneviratne et al., 2016; Wartenburger et al., 2017; Seneviratne et al., 2018; Tebaldi et al., 2020). Here, the principle is extended: instead of having the regional mean anomalies of climate extremes scaled with anomalies in GSAT, we scale the distribution of the local anomalies of climate extremes with anomalies in GSAT.

For clarity, this method is explained for TXx, but this method is designed to be applicable to other climate extremes as discussed in Section 5. We write $\Delta X_{s,t}$ the local anomaly of TXx at each point in space s and timestep t. We assume here that $\Delta X_{s,t}$ follows a Generalized Extreme Value (GEV) distribution, because TXx is a block maxima (Coles, 2001; Wilks, 2011) and we note that the GEV has been successfully used to model TXx elsewhere (Hauser et al., 2016; Huang et al., 2016; Kim et al., 2020). We further assume that the location, scale and shape parameters of the GEV are point- and timestep-dependent, written as $\mu_{s,t}$, $\sigma_{s,t}$ and $\xi_{s,t}$. More precisely, we disentangle these dependencies by assuming that these parameters follow the functions f_s, g_s and h_s, taking a matrix of covariates $\Delta \boldsymbol{V_t}$ as input. This matrix is defined as timeseries of the anomalies in global climate variables such as GSAT. We define the emulator configuration E as the set of equations (1). Examples are shown in Section 4.1.

$$E:\begin{cases}\Delta X_{s,t} \sim GEV\left(\mu_{s,t}, \sigma_{s,t}, \xi_{s,t}\right) \\ \mu_{s,t} = f_s(\Delta \boldsymbol{V_t}) \\ \sigma_{s,t} = g_s(\Delta \boldsymbol{V_t}) \\ \xi_{s,t} = h_s(\Delta \boldsymbol{V_t})\end{cases} \quad (1)$$

For each ESM, the coefficients in the functions f_s, g_s and h_s are estimated by minimizing the negative log likelihood over scenarios and available ensemble members. To ensure the convergence of the fit, the local first guess of the coefficients for the parameters is optimized using an adapted method of moments as described in the supplementary information.

3.2 Spatio-temporal coherent sampling of climate extremes

The first step of our emulation method provides the local statistical properties of the climate extremes and their evolution with external covariates. For approximating internal climate variability, we aim at devising a stochastic model that produces spatially and temporaly correlated samples of TXx. To this end, we follow previous work which parameterizes internal climate variability in annual mean temperature anomalies using a local auto-regressive processes with spatially correlated innovations (Beusch et al., 2020). A key assumption of this approach is that the variability is stationary in time and approximately normally distributed. This is however not the case for residuals of the model mentioned in equation (1). Instead, we propose an approach that exploits the model to transform TXx to a standard normal distribution using the probability integral transform (Angus, 1994; Gneiting et al., 2007; Gudmundsson et al., 2012). For the emulator configuration defined in Section 3.1, the GEV of TXx and its cumulative distribution function $\mathcal{F}_{GEV}\left(\Delta X_{s,t} | \Delta \boldsymbol{V_t}, f_s, g_s, h_s\right)$ are known over the full training dataset. We define $\mathcal{F}_{\mathcal{N}}^{-1}$ as the quantile function of the standard normal distribution. Using these two functions, we transform $\Delta X_{s,t}$ to a standard normally distributed transformed TXx, that we write as $\Phi_{s,t}$.

$$\Phi_{s,t} = \mathcal{F}_{\mathcal{N}}^{-1}\left(\mathcal{F}_{GEV}\left(\Delta X_{s,t} | \Delta \boldsymbol{V_t}, f_s, g_s, h_s\right)\right) \quad (2)$$

While $\Delta X_{s,t}$ follows a non-stationary GEV distribution, $\Phi_{s,t}$ has a normal distribution stationary in time, thus respecting the required conditions (Humphrey and Gudmundsson, 2019; Beusch et al., 2020). Note that no information is lost in this transformation, because the GEV associated with $\Phi_{s,t}$ is known at each point s and timestep t, which will be used in Section 3.3. We train on $\Phi_{s,t}$ a local auto-regressive process of order 1 with parameters $\gamma_{s,0}$ and $\gamma_{s,1}$, with spatially correlated innovations $\upsilon_{s,t}$. These innovations are sampled from a multivariate normal distribution deduced from an empirically estimated and localized covariance matrix that represents spatial dependence between points as explained in (Beusch et al., 2020).

$$\Phi_{s,t} = \gamma_{s,0} + \gamma_{s,1}\Phi_{s,t-1} + \upsilon_{s,t} \quad (3)$$

3.3 Emulating spatio-temporally correlated climate extremes

The two steps described in section 3.1 and 3.2 form the full training of the emulator. Here, we explain how to emulate TXx under different scenarios. Any scenario can be emulated if it provides the covariates $\Delta\boldsymbol{V_t}$ that are timeseries of anomalies in global climate variables. Thanks to this scenario, the distribution of TXx is a direct result from equation 1.

Using the auto-regressive processes with spatially correlated innovations, we draw realizations $\Phi_{s,t,e}$ for all points s, timesteps t, and index of emulation e. These realizations are transformations of TXx onto a standard normal distribution, and independent from the scenario so far. Because the probability integral transformation can be reversed, we transform back the realizations $\Phi_{s,t,e}$ onto the distribution of TXx using its quantile function $\mathcal{F}_{GEV}^{-1}(p|\Delta\boldsymbol{V_t}, f_s, g_s, h_s)$, p being here a probability and the cumulative distribution function of the standard normal distribution $\mathcal{F}_{\mathcal{N}}$, leading to the emulations of TXx written $\Delta X_{s,t,e}$:

$$\Delta X_{s,t,e} = \mathcal{F}_{GEV}^{-1}\left(\mathcal{F}_{\mathcal{N}}\left(\Phi_{s,t,e}\right)|\Delta\boldsymbol{V_t}, f_s, g_s, h_s\right) \quad (4)$$

4 Emulating extreme temperatures under climate change

4.1 Evaluating and selecting emulator configurations

The method of Section 3 is applied and we test a set of different configurations (Figure 1), looking for a good compromise between simplicity and accuracy. For each of the 18 ESMs, we use its historical period over 1850-2014 and all available scenarios over 2015-2100 to calibrate the emulator configuration (Section 3.1) and the auto-regressive process with spatially-correlated innovations (Section 3.2). We then draw 1000 realizations that we back-transform into emulations of all available scenarios. (Section 3.3). For each ESM and emulator configuration, we evaluate the ability of the emulations to reproduce the ESM's TXx anomaly distribution using the Continuous Rank Probabiliy Score (CRPS) and the CRPS Skill Score (CRPSS), commonly used in climate sciences (Wilks, 2011; Jolliffe and Stephenson, 2012). The CRPS measures the quadratic discrepancy between the cumulative distribution function of the emulations to the one of the ESM. We calculate this score for each point of the sample. The CRPSS is defined as one minus the ratio of the CRPS to another CRPS used as a reference, thus expressing the decrease in the CRPS relative to the reference. Both scores are then averaged globally for the sake of clarity.

We show a selection of emulator configurations in Figure 1, using the decomposition of the GSAT anomaly into the global trend ΔT^{GT} and the global variability ΔT^{GV} (Beusch et al., 2020). The global trend ΔT^{GT} is meant to capture the signal from global warming while the global variability ΔT^{GV} would rather capture interannual variability processes. We use here linear evolutions of covariates, for simplicity and given their observed linearity with global mean temperature (Seneviratne et al., 2016; Wartenburger et al., 2017; Tebaldi et al., 2020). In Figure 1, the configurations are distinguished into two groups: the first row corresponds to a primitive configuration, with no covariates, used for benchmarking of the second group.

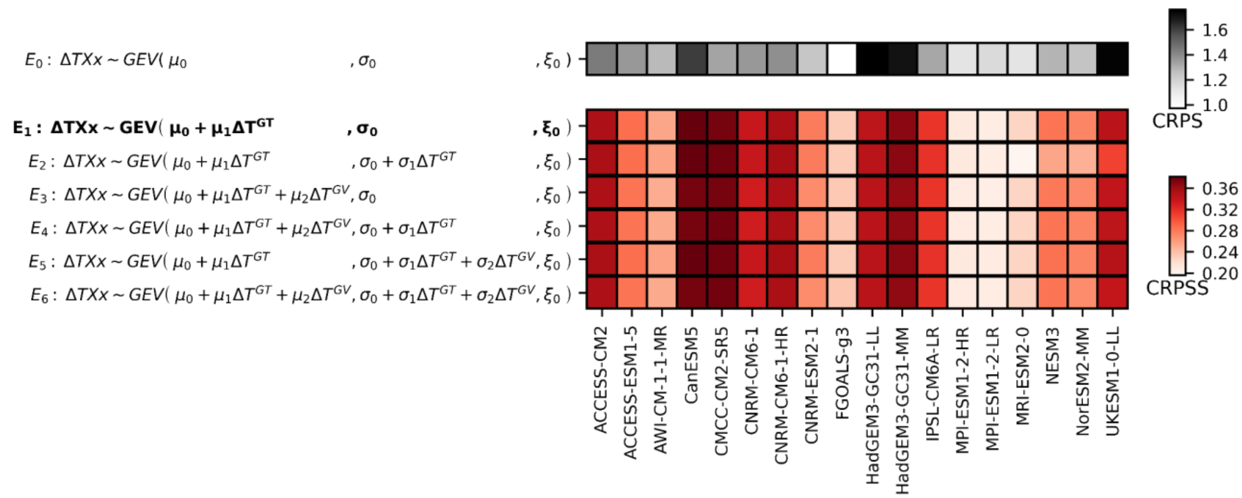

Figure 1. Selection of an emulator configuration. The first row shows the CRPS (lower is better) for the emulator configuration E_0 used as a reference. On the following rows, the CRPSS (higher is better) with reference to the emulator configuration E_0 show the respective global performance of the different emulator configurations for different ESMs.

On the first row of Figure 1, the emulator configuration has its GEV with constant parameters over time, despite a changing climate. On the second row of Figure 1, the configuration E_1 has only its location covarying linearly with ΔT^{GT}. Compared to E_0, it reduces the CRPS on average by about 28%. The ESMs with a low CRPS in E_0 (eg FGOALS-g3) have their TXx less influenced by climate change than those with a higher CRPS such as HadGEM3-GC31-LL, HadGEM3-GC31-MM and UKESM1-0-LL. Those ESMs with a low CRPS have a low CRPSS as well, because the new emulator configuration brings little improvement. However, those with a higher CRPS benefit from a stronger reduction in their CRPS by including a dependency of the GEV to climate change.

On the following rows, different combinations are tried to further improve E_1. However, these more complex models have only marginal gains, or even lead to a reduction in the performances (e.g. E_2). These results are confirmed by comparing the global distribution of CRPS using Mann-Whitney U tests: adding additional terms for the emulation of TXx either brings no significant improvement, or slightly reduces the quality of the emulations. It would suggest that it would only overfit the data.

We observe that the emulator configurations E_2 to E_6 bring improvement only in some regions of the Earth (not shown), while they hamper the fit in many others, which is consistent with (Kharin and Zwiers, 2005; Kim et al., 2020). In our framework, E_1 is sufficient to capture the evolution of the distribution of TXx in a changing climate. We observe that with the combination of ΔT^{GT} and the emulated ΔT^{GV}, the location and scale parameters vary over broader domains than those of the ESM, and the scale parameters may even become negative. Because ΔT^{GV} can be characterized as an independent stochastic process, using it as a covariate hinders the emulation.

In Figures S.1 to S.7, we show other emulator configurations. We use ΔH^{GT}, the global trend of the anomaly in HFDS, to disentangle contributions with different timescales (Geoffroy et al., 2013; King et al., 2020). It shows that ΔH^{GT} does not bring the desired improvement: the differences in transient and equilibrium TXx appear mostly at the end of low-warming scenarios. Using the extensions of scenarios up to 2300 may help the algorithm in seizing this signal. A logistic regression is also tried on the shape parameter, to limit the range of its evolution and to account for changes in albedo, for instance due to reduction in snow cover.

This analysis shows that the emulator configuration E_1 provides the best compromise of simplicity and quality for emulations of TXx. The results in the rest of the paper will therefore use E_1, i.e. with the only the location parameter of the GEV varying linearly with ΔT^{GT}.

4.2 Example of emulations

Figure 2 shows an example of our results for MPI-ESM1-2-HR, one of the 18 trained ESMs. We compare the maps of the anomaly in TXx of the ESM (topmost row) with 3 of the 1000 emulations for this ESM. We show the years 2014 and 2100, the end of the historical scenario and the end of SSP5-8.5 to illustrate the performance under current and high warming conditions.

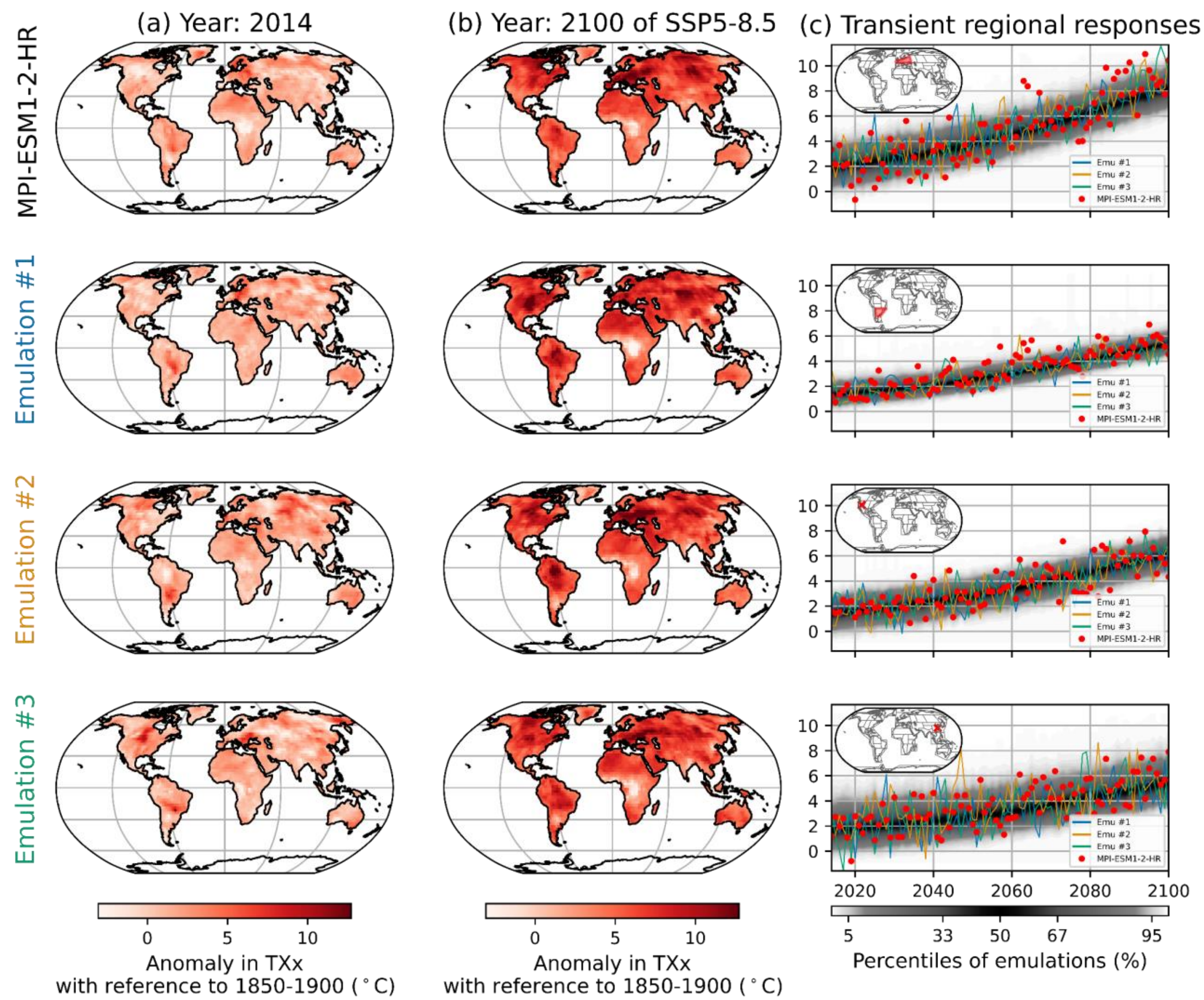

Figure 2. Example of emulations. The local anomalies in TXx are shown for MPI-ESM1-2-HR and three emulations for the years 2014 and 2100, in columns (a) and (b), respectively. The transient regional response from 2014 to 2100 is shown in column (c) for selected regions and points. The 1st and 2nd rows of column (c) are respectively the regions West & Central Europe and the South-East of South America. The 3rd and 4th rows of column (c) are two points located in the United States and in China respectively. It features the values from MPI-ESM1-2-HR, the same three emulations shown in maps and the density of the 1000 emulations drawn for this emulator configuration.

The emulations capture the general spatial features in TXx well, be it in 2014 or in 2100, but no exact match to the ESM simulation can be expected since they include a representation of natural variability. For example, both the emulations and the ESM simulate the positive anomaly over Eastern Europe and the center of South America or the lower anomaly over Central Africa. Because each emulation includes natural variability, some features are more pronounced than others, such as the high anomaly in the center of North America. The strongest differences to MPI-ESM1-2-HR are in the South-East of South America and in the center of North America.

To further investigate the similarities and discrepancies, we represent the transient response in two specific regions and two specific points as detailed in Figure 2. Overall, the emulations show a good agreement with the ESM. The ensemble of emulations correctly encompasses the realization by MPI-ESM1-2-HR.

Figures S.8 to S.25 show the same results for the 17 other ESMs employed in this study. They highlight that the emulator captures the spatial and temporal features of these models as well, even though the ESMs present different mean warming, internal variability, and spatial patterns of TXx anomalies.

4.3 Evaluation of regional performance

We have selected the emulator configuration on the basis of its global performance in Section 4.1, and verified that the emulations are visually convincing at different spatial scales in Section 4.2. Here, we want to quantify the performance on a regional level. To do so, we compare regional percentiles of the emulations to the ESMs following the same approach as (Beusch et al., 2020). For each ESM and each emulation, the anomalies in TXx are averaged over the AR6 regions. Next, we calculate the 95%, 50% and 5% percentiles of the regional emulations. We count how often the regional values of the ESM exceed these thresholds. We determine the deviations of the ESM to the percentiles of the emulation, hence how well we reproduce the dispersion of the ESM.

Figure 3 shows the regional deviation in the quantiles. Panel (a) shows that the 95% quantile of the emulations is generally too low, while panel (c) shows that the 5% quantile of the emulations is mostly too high. This means that the emulation is underdispersive, a feature expected for emulations (Beusch et al., 2020). The performance of the emulator is lowest in South-East Asia and in the Sahara, but overall the performance remains good: the regional deviations are below 5% in most of the cases (for 93%, 99% and 92% of the model-region combinations for the quantiles 95%, 50% and 5%, respectively). The average of the regional deviations across regions and ESMs is -2.4%, -0.3% and 2.9%.

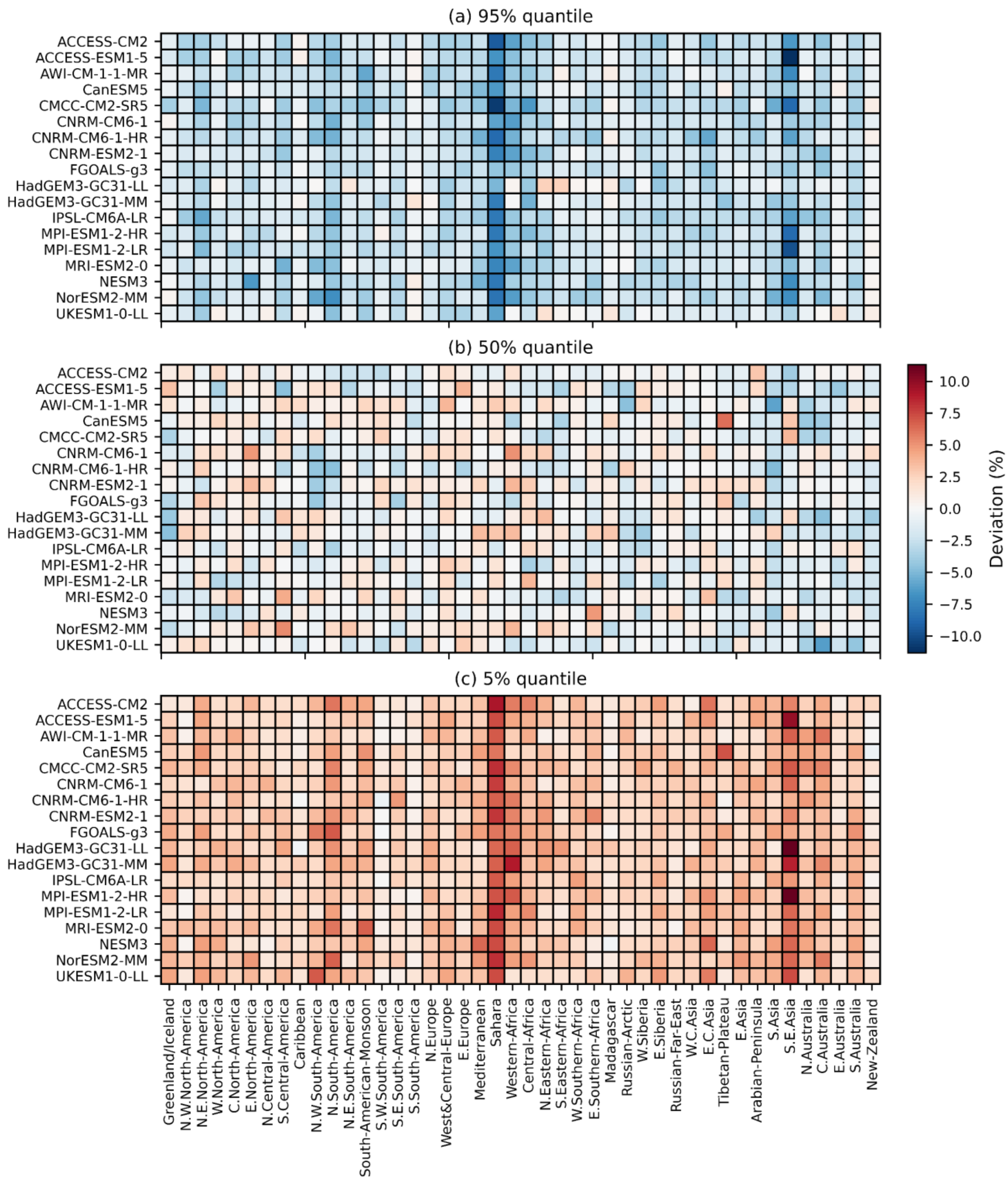

Figure 3. Regional deviations of ESMs from the 5%. 50%, 95% quantiles of the emulations, respectively in panels (a), (b) and (c). Red (blue) indicates that the quantile of the emulations is higher (lower) than the one of ESM, because the ESM is more frequently below (above) the quantile than expected.

4.4 Example application

In Sections 4.1 to 4.3, we evaluate the performance of the emulator using training data. However, this method is not only meant to reproduce training data, but also to emulate other scenarios. For instance, some ESMs run only a subset of the scenarios SSP1-1.9, SS1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5, which hinders the evaluation of a distribution of anomalies in TXx based on all ESMs. Here, for each ESM, we use the emulator trained on available scenarios from Sections 4.2 to 4.3 to calculate all these SSPs.

In the selected configuration, MESMER-X can emulate scenarios if timeseries of the smoothed anomaly in GSAT (ΔT^{GT}) are provided. For each of the scenarios (Section 2), we average ΔT^{GT} over all ESMs that have run the scenario. These averaged ΔT^{GT} are used as common drivers to create emulations for all ESMs for every scenario. For each of the 18 ESMs, we calculate an ensemble of 1000 realizations which combines two sources of dispersion: the local variability in TXx modeled by the ESM and the uncertainty in this modeling by ESMs, also termed “regional climate sensitivity” (Seneviratne and Hauser, 2020). Yet, it does not encompass the global uncertainty due to the different global climate sensitivities of the ESMs. Additionally, we are not weighting ESMs according to their performances nor accounting for ESM-interdepecincies (Abramowitz et al., 2019; Brunner et al., 2020b). Here, we solely aim to show the capacity of this emulator by synthetizing differences in the modeling of TXx in the ESMs. Using the emulations, we calculate the distributions of the anomaly in TXx for any point in space and time, as illustrated in the right panel of Figure 4. From these emulations, we deduce the return periods in 2100 for each ESM and scenario. Then we deduce the mean and standard deviation of these return periods, corresponding to the uncertainty induced by the different ESMs' different representation of natural variability, as shown in left panel of Figure 4.

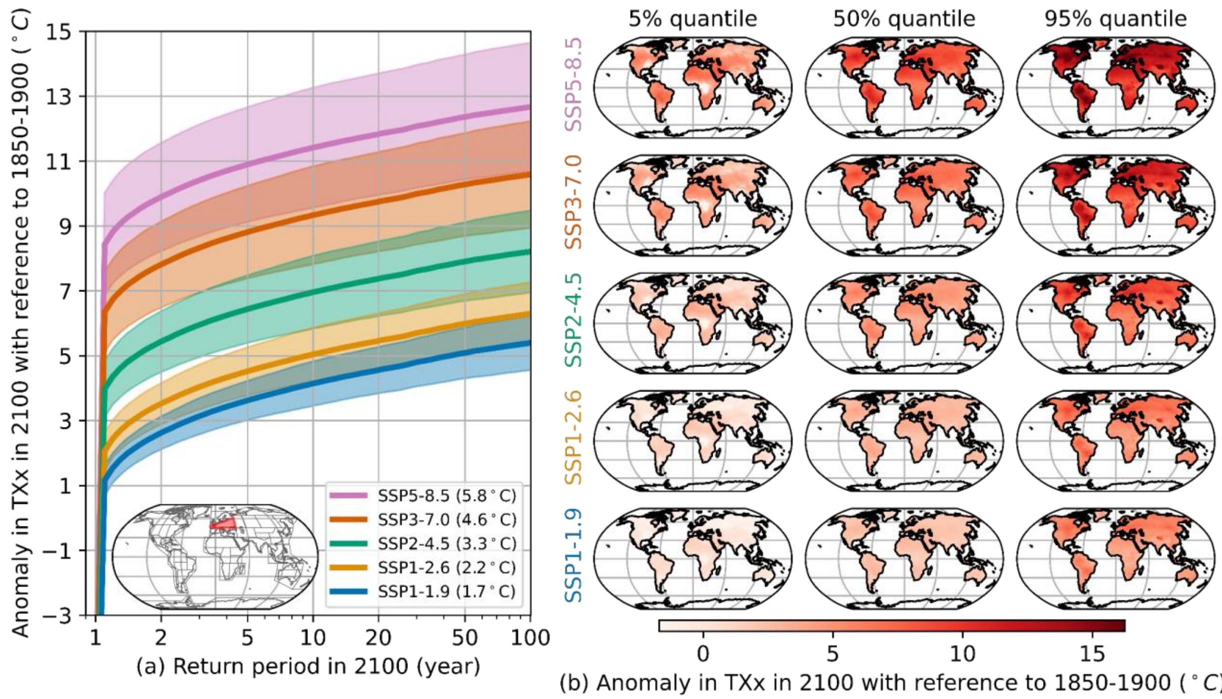

Figure 4. Illustration of the ensemble formed by 1000 emulations of the 18 trained ESMs, applied over common scenarios. The return periods in 2100 in West & Central Europe of each ESM are shown in panel (a) through their mean and one standard deviation range. In the legend, the anomaly in GSAT in 2100 of each scenario is provided. Panel (b) shows the local 5%, 50% and 95% quantiles in 2100, all ESMs being pooled together. Each row corresponds to a different scenario.

In the left panel of Figure 4, we notice that in West & Central Europe, an anomaly of 5°C would happen about once in 40 years in 2100 under SSP1-1.9, but every 10 years under SSP1-2.6 and every 1 or 2 years under SSP2-4.5. This result is consistent with how climate extremes are projected for 1.5°C (Seneviratne et al., 2018) and the change from 1.5°C to 2°C (Hoegh-Guldberg et al., 2018).

In the right panel of Figure 4, we show the maps in 2100 for selected quantiles. Here, all emulations and ESMs are pooled together, which implies that both the natural local variability in TXx and the uncertainty in this modeling by ESMs contribute to this range. For the median, the regions with the highest anomalies of TXx are Central North America, Central South America and the Meditarenean region. Those with the lowest anomalies are Greenland, South Asia and Central Africa. These results are even more distinct when considering the 95% quantile. The 95% quantile of SSP1-1.9 seems overall only slightly higher than the 5% quantile of SSP5-8.5.

Broadly speaking, it would suggest that anomalies in TXx that had only 5% of chances to occur or be exceeded in SSP1-1.9 in 2100, would have their probability increase to 95% in SSP5-8.5.

5 Discussion and conclusions

This paper has introduced a method for the emulation of climate extremes under climate change, used to extend the MESMER emulator (Beusch et al., 2020) to MESMER-X. This method does not only reproduce the mean evolution of climate extremes but also their distribution. Besides, it accounts for their spatial and temporal features.

Fits of non-stationnary GEV for TXx have already been performed using different covariates on the location (Zwiers et al., 2011; Hauser et al., 2016; Wehner et al., 2020; Wehner, 2020). Here, we leverage this approach to model the distribution of TXx at each point conditional on global covariates. The proposed method is improved in its greater versatility in the use of covariates and in its sampling of stochastic realizations of timeseries fields. We show that the emulator mimics well the local annual maximum temperature of the ESMs, with an underdispersion below 5% for most regions and ESMs.

This method is designed to be directly applied to other indicators of climate extremes, as long as their distribution can be parametrized by a GEV. Moreover, the framework can be easily adapted to different distributions which be more appropriate for other indicators, such as a Poisson distribution for counting extreme events (Wilks, 2011) or a generalized Pareto distribution for climate extremes based on peak-over-threshold exceedances (Coles, 2001; Naveau et al., 2005). The parameters of these distributions may vary with any combination of global drivers to improve the quality of the emulator configuration.

Similar to MESMER (Beusch et al., 2022a; Beusch et al., 2022b), MESMER-X could be coupled to a SCM in future work to gain the ability to transform any emission scenario into local annual climate extremes in a fast and probabilistic way. Such an emulator chain could be used to provide detailed climate information into integrated assessment models, for instance to assess how climate extremes affect different transformation pathways.

(All figures and tables should be cited in order. For initial submission, please embed figures, tables, and their captions within the main text near where they are cited. At revision, figures should be uploaded separately, as we need separate files for production. Tables and all captions should be moved to the end of the file.)

Acknowledgments

We acknowledge the World Climate Research Program for the coordination and promotion of CMIP6. We thank the climate modeling groups for producing their model outputs and making them available. We acknowledge the Earth System Grid Federation for archiving the data and providing access. We thank Urs Beyerle for the download and curation of CMIP6 data, and Lukas Brunner for the processing into the CMIP6-ng archive. We acknowledge the European Research Council (ERC) for the ERC Proof Of Concept Grant MESMER-X (Grant Agreement ID 964013) that funded this work.

Open Research

Data from CMIP6 is available at https://esgf-node.llnl.gov/projects/esgf-llnl/ (last access: 3 April 2022). Code from MESMER is available at https://github.com/MESMER-group/mesmer.

References

Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., Lorenz, R., Pincus, R., and Schmidt, G. A.: ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing, Earth Syst. Dynam., 10, 91-105, 10.5194/esd-10-91-2019, 2019.

Alexeeff, S. E., Nychka, D., Sain, S. R., and Tebaldi, C.: Emulating mean patterns and variability of temperature across and within scenarios in anthropogenic climate change experiments, Climatic Change, 146, 319-333, 10.1007/s10584-016-1809-8, 2018.

Angus, J. E.: The probability integral transform and related results, SIAM review, 36, 652-654, 1994.

AON: Weather, climate & catastrophe insight, 2020 Annual Report, 2021.

Beusch, L., Gudmundsson, L., and Seneviratne, S. I.: Emulating Earth system model temperatures with MESMER: from global mean temperature trajectories to grid-point-level realizations on land, Earth Syst. Dynam., 11, 139-159, 10.5194/esd-11-139-2020, 2020.

Beusch, L., Nauels, A., Gudmundsson, L., Gütschow, J., Schleussner, C.-F., and Seneviratne, S. I.: Responsibility of major emitters for country-level warming and extreme hot years, Communications Earth & Environment, 3, 7, 10.1038/s43247-021-00320-6, 2022a.

Beusch, L., Nicholls, Z., Gudmundsson, L., Hauser, M., Meinshausen, M., and Seneviratne, S. I.: From emission scenarios to spatially resolved projections with a chain of computationally efficient emulators: coupling of MAGICC (v7.5.1) and MESMER (v0.8.3), Geosci. Model Dev., 15, 2085-2103, 10.5194/gmd-15-2085-2022, 2022b.

Brunner, L., Hauser, M., Lorenz, R., and Beyerle, U.: The ETH Zurich CMIP6 next generation archive: technical documentation (v1.0-final), Zenodo, https://doi.org/10.5281/zenodo.3734128, 2020a.

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., and Knutti, R.: Reduced global warming from CMIP6 projections when weighting models by performance and independence, Earth Syst. Dynam. Discuss., 2020, 1-23, 10.5194/esd-2020-23, 2020b.

Castruccio, S., McInerney, D. J., Stein, M. L., Liu Crouch, F., Jacob, R. L., and Moyer, E. J.: Statistical Emulation of Climate Model Projections Based on Precomputed GCM Runs, Journal of Climate, 27, 1829-1844, 10.1175/JCLI-D-13-00099.1, 2014.

Coles, S.: An introduction to statistical modeling of extreme values, Springer, London; New York2001.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029-1136, 2013.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937-1958, 2016.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.:

Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 741-866, 2013.

Fordham, D. A., Wigley, T. M. L., Watts, M. J., and Brook, B. W.: Strengthening forecasts of climate change impacts with multi-model ensemble averaged projections using MAGICC/SCENGEN 5.3, Ecography, 35, 4-8, https://doi.org/10.1111/j.1600-0587.2011.07398.x, 2012.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., and Tytéca, S.: Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs, Journal of Climate, 26, 1859-1876, 10.1175/JCLI-D-12-00196.1, 2013.

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic forecasts, calibration and sharpness, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69, 243-268, https://doi.org/10.1111/j.1467-9868.2007.00587.x, 2007.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, 3383-3390, 10.5194/hess-16-3383-2012, 2012.

Hasegawa, T., Sakurai, G., Fujimori, S., Takahashi, K., Hijioka, Y., and Masui, T.: Extreme climate events increase risk of global food insecurity and adaptation needs, Nature Food, 2, 587-595, 10.1038/s43016-021-00335-4, 2021.

Hauser, M., Orth, R., and Seneviratne, S. I.: Role of soil moisture versus recent climate change for the 2010 heat wave in western Russia, Geophysical Research Letters, 43, 2819-2826, https://doi.org/10.1002/2016GL068036, 2016.

Herger, N., Sanderson, B. M., and Knutti, R.: Improved pattern scaling approaches for the use in climate impact studies, Geophysical Research Letters, 42, 3486-3494, https://doi.org/10.1002/2015GL063569, 2015.

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., and Zhou, G.: Impacts of 1.5ºC global warming on natural and human systems, in: Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, edited by: Masson-Delmotte, V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T., In Press, https://www.ipcc.ch/sr15/download/, 2018.

Holden, P. B., Edwards, N. R., Garthwaite, P. H., Fraedrich, K., Lunkeit, F., Kirk, E., Labriet, M., Kanudia, A., and Babonneau, F.: PLASIM-ENTSem v1.0: a spatio-temporal emulator of future climate change for impacts assessment, Geosci. Model Dev., 7, 433-451, 10.5194/gmd-7-433-2014, 2014.

Huang, W. K., Stein, M. L., McInerney, D. J., Sun, S., and Moyer, E. J.: Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions, Adv. Stat. Clim. Meteorol. Oceanogr., 2, 79-103, 10.5194/ascmo-2-79-2016, 2016.

Humphrey, V. and Gudmundsson, L.: GRACE-REC: a reconstruction of climate-driven water storage changes over the last century, Earth Syst. Sci. Data, 11, 1153-1170, 10.5194/essd-11-1153-2019, 2019.

IPCC, Field, C. B., Barros, V. R., Dokken, D. J., Mach, K. J., Mastrandrea, M. D., Bilir, T. E., Chatterjee, M., Ebi, K. L., Estrada, Y. O., Genova, R. C., Girma, B., Kissel, E. S., Levy, A. N., MacCracken, S., Mastrandrea, P. R., White, L. L., and (eds.) (Eds.): Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2014.

IPCC: Global warming of 1.5C. An IPCC Special Report on the impacts of global warming of 1.5C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, In Press2018.

IPCC, Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (Eds.): Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press2021.

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cerezo-Mota, R., Cimadevilla, E., Cofiño, A. S., Di Luca, A., Faria, S. H., Gorodetskaya, I. V., Hauser, M., Herrera, S., Hennessy, K., Hewitt, H. T., Jones, R. G., Krakovska, S., Manzanas, R., Martínez-Castro, D., Narisma, G. T., Nurhati, I. S., Pinto, I., Seneviratne, S. I., van den Hurk, B., and Vera, C. S.: An update of IPCC climate reference regions for subcontinental analysis of climate model data: definition and aggregated datasets, Earth Syst. Sci. Data, 12, 2959-2970, 10.5194/essd-12-2959-2020, 2020.

Jolliffe, I. T. and Stephenson, D. B.: Forecast verification: a practitioner's guide in atmospheric science, John Wiley & Sons2012.

Kharin, V. V. and Zwiers, F. W.: Estimating Extremes in Transient Climate Change Simulations, Journal of Climate, 18, 1156-1173, 10.1175/JCLI3320.1, 2005.

Kim, Y.-H., Min, S.-K., Zhang, X., Sillmann, J., and Sandstad, M.: Evaluation of the CMIP6 multi-model ensemble for climate extreme indices, Weather and Climate Extremes, 29, 100269, https://doi.org/10.1016/j.wace.2020.100269, 2020.

King, A. D., Lane, T. P., Henley, B. J., and Brown, J. R.: Global and regional impacts differ between transient and equilibrium warmer worlds, Nature Climate Change, 10, 42-47, 10.1038/s41558-019-0658-7, 2020.

Lee, J. Y., Marotzke, J., Bala, G., Cao, L., Corti, S., Dunne, J. P., Engelbrecht, F., Fischer, E., Fyfe, J. C., Jones, C., Maycock, A., Mutemi, J., Ndiaye, O., Panickal, S., and Zhou, T.: Future Global Climate: Scenario Based Projections and Near-Term Information, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, 2021.

Link, R., Snyder, A., Lynch, C., Hartin, C., Kravitz, B., and Bond-Lamberty, B.: Fldgen v1.0: an emulator with internal variability and space–time correlation for Earth system models, Geosci. Model Dev., 12, 1477-1489, 10.5194/gmd-12-1477-2019, 2019.

Lynch, C., Hartin, C., Bond-Lamberty, B., and Kravitz, B.: An open-Access CMIP5 pattern library for temperature and precipitation: Description and methodology, Earth System Science Data, 9, 281-292, 2017.

McKinnon, K. A., Poppick, A., Dunn-Sigouin, E., and Deser, C.: An "Observational Large Ensemble" to Compare Observed and Modeled Temperature Trend Uncertainty due to Internal Variability, Journal of Climate, 30, 7585-7598, 10.1175/JCLI-D-16-0905.1, 2017.

Meinshausen, M., Nicholls, Z. R. J., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., Canadell, J. G., Daniel, J. S., John, A., Krummel, P. B., Luderer, G., Meinshausen, N., Montzka, S. A., Rayner, P. J., Reimann, S., Smith, S. J., van den Berg, M., Velders, G. J. M., Vollmer, M. K., and Wang, R. H. J.: The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500, Geosci. Model Dev., 13, 3571-3605, 10.5194/gmd-13-3571-2020, 2020.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geoscientific Model Development, 10, 2057-2116, 2017.

Mitchell, T. D.: Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates, Climatic Change, 60, 217-242, 10.1023/A:1026035305597, 2003.

Nath, S., Lejeune, Q., Beusch, L., Schleussner, C. F., and Seneviratne, S. I.: MESMER-M: an Earth System Model emulator for spatially resolved monthly temperatures, Earth Syst. Dynam. Discuss., 2021, 1-38, 10.5194/esd-2021-59, 2021.

Naveau, P., Nogaj, M., Ammann, C., Yiou, P., Cooley, D., and Jomelli, V.: Statistical methods for the analysis of climate extremes, Comptes Rendus Geoscience, 337, 1013-1022, https://doi.org/10.1016/j.crte.2005.04.015, 2005.

Nicholls, Z., Meinshausen, M., Lewis, J., Corradi, M. R., Dorheim, K., Gasser, T., Gieseke, R., Hope, A. P., Leach, N. J., McBride, L. A., Quilcaille, Y., Rogelj, J., Salawitch, R. J., Samset, B. H., Sandstad, M., Shiklomanov, A., Skeie, R. B., Smith, C. J., Smith, S. J., Su, X., Tsutsui, J., Vega-Westhoff, B., and Woodard, D. L.: Reduced Complexity Model Intercomparison Project Phase 2: Synthesising Earth system knowledge for probabilistic climate projections, Earth's Future, n/a, e2020EF001900, https://doi.org/10.1029/2020EF001900, 2021.

Nicholls, Z. R. J., Meinshausen, M., Lewis, J., Gieseke, R., Dommenget, D., Dorheim, K., Fan, C. S., Fuglestvedt, J. S., Gasser, T., Golüke, U., Goodwin, P., Kriegler, E., Leach, N. J., Marchegiani, D., Quilcaille, Y., Samset, B. H., Sandstad, M., Shiklomanov, A. N., Skeie, R. B., Smith, C. J., Tanaka, K., Tsutsui, J., and Xie, Z.: Reduced complexity model intercomparison project phase 1: Protocol, results and initial observations, Geosci. Model Dev. Discuss., 2020, 1-33, 10.5194/gmd-2019-375, 2020.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J. F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461-3482, 2016.

Perera, A. T. D., Nik, V. M., Chen, D., Scartezzini, J.-L., and Hong, T.: Quantifying the impacts of climate change and extreme climate events on energy systems, Nature Energy, 5, 150-159, 10.1038/s41560-020-0558-0, 2020.

Raymond, C., Horton, R. M., Zscheischler, J., Martius, O., AghaKouchak, A., Balch, J., Bowen, S. G., Camargo, S. J., Hess, J., Kornhuber, K., Oppenheimer, M., Ruane, A. C., Wahl, T., and White, K.: Understanding and managing connected extreme events, Nature Climate Change, 10, 611-621, 10.1038/s41558-020-0790-4, 2020.

Rosenzweig, C., Arnell, N. W., Ebi, K. L., Lotze-Campen, H., Raes, F., Rapley, C., Smith, M. S., Cramer, W., Frieler, K., Reyer, C. P. O., Schewe, J., van Vuuren, D., and Warszawski, L.: Assessing inter-sectoral climate change risks: the role of ISIMIP, Environmental Research Letters, 12, 010301, 10.1088/1748-9326/12/1/010301, 2017.

Schaeffer, R., Szklo, A. S., Pereira de Lucena, A. F., Moreira Cesar Borba, B. S., Pupo Nogueira, L. P., Fleming, F. P., Troccoli, A., Harrison, M., and Boulahya, M. S.: Energy sector vulnerability to climate change: A review, Energy, 38, 1-12, https://doi.org/10.1016/j.energy.2011.11.056, 2012.

Seneviratne, S. I. and Hauser, M.: Regional Climate Sensitivity of Climate Extremes in CMIP6 Versus CMIP5 Multimodel Ensembles, Earth's Future, 8, e2019EF001474, https://doi.org/10.1029/2019EF001474, 2020.

Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R., and Wilby, R. L.: Allowable CO2 emissions based on regional and impact-related climate targets, Nature, 529, 477-483, 10.1038/nature16542, 2016.

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., and Zhou, B.: Weather and Climate Extreme Events in a Changing Climate, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. e., Cambridge University Press, 2021.

Seneviratne, S. I., Wartenburger, R., Guillod, B. P., Hirsch, A. L., Vogel, M. M., Brovkin, V., van Vuuren, D. P., Schaller, N., Boysen, L., Calvin, K. V., Doelman, J., Greve, P., Havlik, P., Humpenöder, F., Krisztin, T., Mitchell, D., Popp, A., Riahi, K., Rogelj, J., Schleussner, C.-F., Sillmann, J., and Stehfest, E.: Climate extremes, land–climate feedbacks and land-use forcing at 1.5°C, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376, 20160450, 10.1098/rsta.2016.0450, 2018.

Sivakumar, M. V., Motha, R. P., and Das, H. P.: Natural disaster and extreme events in agriculture, Springer2005.

Storch, H. v. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, DOI: 10.1017/CBO9780511612336, 1999.

Tebaldi, C. and Arblaster, J. M.: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations, Climatic Change, 122, 459-471, 10.1007/s10584-013-1032-9, 2014.

Tebaldi, C. and Knutti, R.: Evaluating the accuracy of climate change pattern emulation for low warming targets, Environmental Research Letters, 13, 055006, 10.1088/1748-9326/aabef2, 2018.

Tebaldi, C., Armbruster, A., Engler, H. P., and Link, R.: Emulating climate extreme indices, Environmental Research Letters, 15, 074006, 10.1088/1748-9326/ab8332, 2020.

Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N., and Frieler, K.: The effects of climate extremes on global agricultural yields, Environmental Research Letters, 14, 054010, 10.1088/1748-9326/ab154b, 2019.

Wartenburger, R., Hirschi, M., Donat, M. G., Greve, P., Pitman, A. J., and Seneviratne, S. I.: Changes in regional climate extremes as a function of global mean temperature: an interactive plotting framework, Geoscientific Model Development Discussions, 1-30, 2017.

Wehner, M., Gleckler, P., and Lee, J.: Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 1, model evaluation, Weather and Climate Extremes, 30, 100283, https://doi.org/10.1016/j.wace.2020.100283, 2020.

Wehner, M. F.: Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 2, projections of future change, Weather and Climate Extremes, 30, 100284, https://doi.org/10.1016/j.wace.2020.100284, 2020.

Wilks, D. S.: Statistical methods in the atmospheric sciences, 2011.

Zwiers, F. W., Zhang, X., and Feng, Y.: Anthropogenic Influence on Long Return Period Daily Temperature Extremes at Regional Scales, Journal of Climate, 24, 881-892, 10.1175/2010JCLI3908.1, 2011.

Showcasing MESMER-X: Spatially resolved emulation of annual maximum temperatures of Earth System Models

Supplementary Information

Y. Quilcaille[1], L. Gudmundsson[1], L. Beusch[1,*], M. Hauser[1], and S.I. Seneviratne[1]

[1] Institute for Atmospheric and Climate Science, Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland.

[*] Now at: Center for Climate Systems Modeling (C2SM), ETH Zurich, Zurich, Switzerland and

MeteoSwiss, via ai Monti 146, 6605 Locarno-Monti, Switzerland

ESM	Simulations available	Ensemble member used
ACCESS-CM2	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
ACCESS-ESM1-5	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
AWI-CM-1-1-MR	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
CanESM5	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
CMCC-CM2-SR5	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
CNRM-CM6-1	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f2
CNRM-CM6-1-HR	*historical*, *ssp126*, *ssp585*	r1i1p1f2
CNRM-ESM2-1	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f2
FGOALS-g3	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
HadGEM3-GC31-LL	*historical*, *ssp126*, *ssp245*, *ssp585*	r1i1p1f3
HadGEM3-GC31-MM	*historical*, *ssp126*, *ssp585*	r1i1p1f1
IPSL-CM6A-LR	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
MPI-ESM1-2-HR	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
MPI-ESM1-2-LR	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
MRI-ESM2-0	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
NESM3	*historical*, *ssp126*, *ssp245*, *ssp585*	r1i1p1f1
NorESM2-MM	*historical*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f1
UKESM1-0-LL	*historical*, *ssp119*, *ssp126*, *ssp245*, *ssp370*, *ssp585*	r1i1p1f2

Table S.1: ESMs selected for emulation, based on the availability of data.

Optimization of the first guess for the fit of the GEV with covariates:

In section 3.1 of the main text of this paper, we describe how a distribution is fitted for the climate extreme, using covariates on parameters. As written in section 3.1, $\Delta X_{s,t}$ corresponds to the sample of the climate extreme and $\boldsymbol{\Delta C_{t,k}}$ to the vector of covariates. We assume immediately that we are on a given gridpoint s to drop the index. In this section, we note ΔX_t the full sample of the climate extreme, historical and scenarios together.

The objective is to identify coefficients for the emulator configuration. We illustrate this method with a GEV here of location μ, scale σ and shape ξ. We write in equation (A.1) the objective. The coefficients μ_0, σ_0 and ξ_0 are constant terms. We separate the i coefficients $\mu_{lin,i}$ on linear covariates from the j coefficients $\mu_{other,j}$ on non-linear covariates, for all parameters.

$$\begin{cases} (\mu_0, \dots \mu_{lin,i} \cdots, \dots \mu_{other,j} \cdots) \\ (\sigma_0, \dots \sigma_{lin,k} \cdots, \dots \sigma_{other,l} \cdots) \\ (\xi_0, \dots \xi_{lin,m} \cdots, \dots \xi_{other,n} \cdots) \end{cases} \qquad (A.1)$$

The general idea of this method is to propose a first guess of the constant terms for the location, scale and shape of the distribution using the analytical expressions of the mean, variance and skewness. By optimizing a first evaluation of these constant terms to the observed moments of the distribution, we obtain an optimized first guess.

Step 1:

To begin with, the sample of the climate extreme ΔX_t is detrended using ordinary least squares, and only with the terms on the location that were assumed linear in the emulator configuration. The constant term is noted $\mu_{fg1,0}$, while the coefficients on the i linear terms are written $\mu_{fg,lin,i}$.

Step 2:

From the detrended climate extremes, we deduce the residuals. From these residuals, we calculate the mean M, the variance V and the skewness S of the full sample.

Step 3:

The support of a GEV is defined as shown in equation (A.2). In our data, we observe that the shape is mostly negative, pointing at an upper limit in ΔX_t.

$$\begin{cases} \Delta X_t \in \ [\mu - \sigma/\xi \, , +\infty[\quad when\ \xi > 0 \\ \Delta X_t \in \ ]-\infty, +\infty[\quad when\ \xi = 0 \\ \Delta X_t \in \ ]-\infty, \mu - \sigma/\xi] \quad when\ \xi < 0 \end{cases} \qquad (A.2)$$

An initial value ξ_{raw} for the shape is calculated using this support and an ad-hoc value, as shown in equation (A.3). This value will not be the first guess for the shape of the GEV. This ξ_{raw} is meant to ensure that all points of the sample are within the support of the GEV.

$$\xi_{raw} = max\left(-0.25, \frac{V}{M - \max(\Delta X_t)} + 0.1\right) \qquad (A.3)$$

<u>Step 4:</u>

We write a first set of coefficients, shown in equation (A.4). The coefficients $\mu_{fg,other,j}$ are written so that the ensuing evolutions would be small compared to the constant. For instance, using notations from Figure 1, the logistic terms are set to $\xi_{\lambda,1}$= 0.1 yr^{-1} and $\xi_{\delta,1} = 0.01\ \xi_0$.

$$\begin{cases} (M, \ldots \mu_{fg,lin,i} \cdots, \ldots \mu_{fg,other,j} \cdots) \\ (\sqrt{V}, \ldots 0 \ldots, \ldots \sigma_{fg,other,l} \cdots) \\ (\xi_{raw}, \ldots 0 \ldots, \ldots \xi_{fg,other,n} \cdots) \end{cases} \qquad (A.4)$$

The mean M_{GEV}, the variance V_{GEV} and the skewness S_{GEV} of a GEV of location μ, scale σ and shape ξ can be written as shown in equations (A.5). We write γ as the Euler's constant, Γ as the Gamma function, sgn as the sign function and ζ as the Riemann's zeta function.

$$\begin{cases} M_{GEV} = \begin{cases} \mu + \sigma\,(g_1 - 1)/\xi \quad when\ \xi \neq 0, \xi < 1 \\ \mu + \sigma\gamma \quad when\ \xi = 0 \\ \infty \quad when\ \xi \geq 1 \end{cases} \\ V_{GEV} = \begin{cases} \sigma^2\,(g_2 - g_1^2)/\xi^2 \quad when\ \xi \neq 0, \xi < 1/2 \\ \sigma^2\,\pi^2/6 \quad when\ \xi = 0 \\ \infty \quad when\ \xi \geq 1/2 \end{cases} \\ S_{GEV} = \begin{cases} sgn(\xi)\dfrac{g_3 - 3g_2 g_1 + 2g_1^3}{(g_2 - g_1^2)^{\frac{3}{2}}} \quad when\ \xi \neq 0, \xi < \dfrac{1}{3} \\ 12\sqrt{6}\,\zeta(3)/\pi^3 \quad when\ \xi = 0 \end{cases} \\ g_k = \Gamma(1 - k\xi) \end{cases} \qquad (A.5)$$

We optimize now the constant coefficients (μ_c, σ_c, ξ_c) with starting values $(M, \sqrt{V}, \xi_{raw})$ from (A.4), by minimization of the differences to the moments of the GEV deduced from (A.5). This process is illustrated in equation (A.6), and the solution is noted $(\mu_{fg2,0}, \sigma_{fg,0}, \xi_{fg,0})$.

$$(\mu_{fg2,0}, \sigma_{fg,0}, \xi_{fg,0}) = min_{(\mu_c,\sigma_c,\xi_c)}^{constraints}\begin{pmatrix} (M_{GEV}(\mu_c, \sigma_c, \xi_c) - M)^2 \\ +(V_{GEV}(\mu_c, \sigma_c, \xi_c) - V)^2 \\ +(S_{GEV}(\mu_c, \sigma_c, \xi_c) - S)^2 \end{pmatrix} \qquad (A.6)$$

Equation (A.6) shows that the minimization is performed with constraints. For every set of values (μ_c, σ_c, ξ_c), the evolutions of the parameters (μ_t, σ_t, ξ_t) of the GEV are computed. To do so, the covariates are used along the coefficients from equation (A.4), although values

$\left(M, \sqrt{V}, \xi_{raw}\right)$ are replaced by the current values $\left(\mu_{fg1,0} + \mu_c, \sigma_c, \xi_c\right)$. We pinpoint that the actual mean for the calculation of the evolution of the coefficients was $\mu_{fg1,0} + \mu_c$, not only μ_c. This is due to the dependency of the mean of the GEV to its scale and shape, as shown in equation (A.5), and the linear detrend used in step 1.

The computation of the evolutions of the parameters allow the verification of conditions, as shown in equation (A.7). The first condition verifies that the sample falls within the support of the current tested GEV, and is a direct consequence of equation (A.2). The second condition is meant to avoid problematic values on the shape. The low and high thresholds on the shape were respectively set to $-\infty$ and $1/3$, to avoid an infinite skewness, as shown in equation (A.5). The third condition simply answers to obvious mathematical and physical grounds. The fourth condition is meant to avoid spurious evolutions of coefficients in ill-defined emulator configurations, causing a trend in coefficients, almost compensating in the evolutions of parameters. This second low threshold were set to -2, this value were observed to provide good results. The last condition actually corresponds to other mathematical conditions on coefficients, such as the time constant in logistic evolutions that are meant to be positive.

$$\begin{cases} \begin{cases} \Delta X_t \in \left[\mu_t - \sigma_t/\xi_t, +\infty\right[& when\ \xi_t > 0 \\ \Delta X_t \in \left]-\infty, \mu_t - \sigma_t/\xi_t\right] & when\ \xi_t < 0 \end{cases} \\ \xi_t \in \left[\xi_{threshold,low}, \xi_{threshold,high}\right] \\ \sigma_t > 0 \\ \xi_c > \xi_{threshold,low,2} \\ \xi_\lambda > 0, \dots \end{cases} \quad (A.7)$$

Step 5:

Thanks to the former optimization, better values for the constant terms have been found. By feeding the result of (A.6) in (A.4), we calculate the negative log likelihood of the current solution, a first optimized first guess.

Then we repeat step 4, although by removing the term on the mean. The second optimized first guess is then used to calculate the negative log likelihood.

We deduce the first guess by taking the one with the lower negative log likelihood. Equation (A.8) shows the optimal first guess used for the fit of the distribution from section 3.1. We pinpoint that the conditions (A.7) are used as well during the fit of the distribution.

$$\begin{cases} \left(\mu_{fg1,0} + \mu_{fg2,0}, \dots \mu_{fg,lin,i} \dots, \dots \mu_{fg,other,j} \dots\right) \\ \left(\sigma_{fg,0}, \dots 0 \dots, \dots \sigma_{fg,other,l} \dots\right) \\ \left(\xi_{fg,0}, \dots 0 \dots, \dots \xi_{fg,other,n} \dots\right) \end{cases} \quad (A.8)$$

Comparison of emulator configuration over each scenario individually:

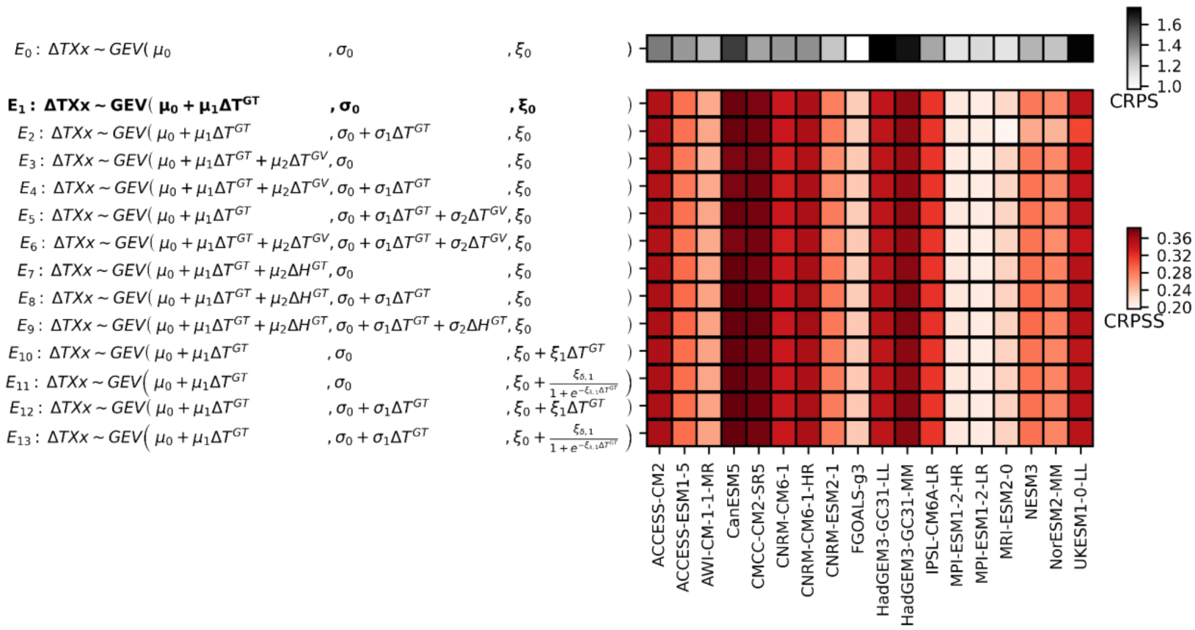

Figure S.1. Emulator configurations over historical and all scenarios. The first row shows the CRPS (lower is better) for E_0 used as a reference. On the following rows, the CRPSS (higher is better) with reference to the emulator configuration E_0 show the respective global performance of the different emulator configurations for different ESMs.

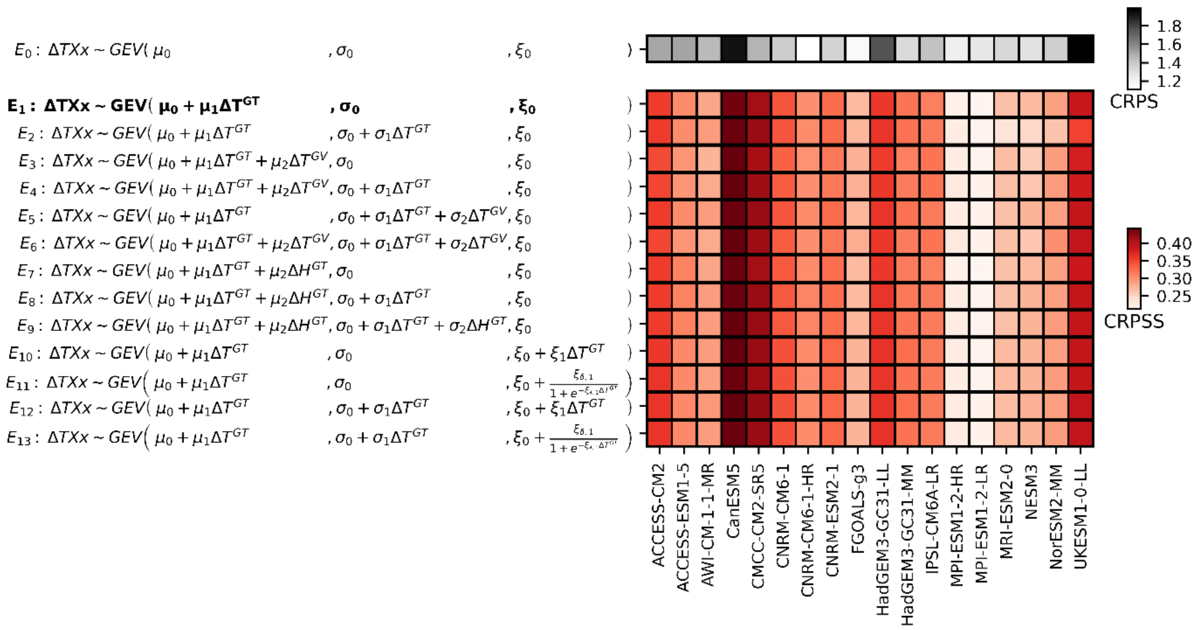

Figure S.2. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the historical (1850-2014).

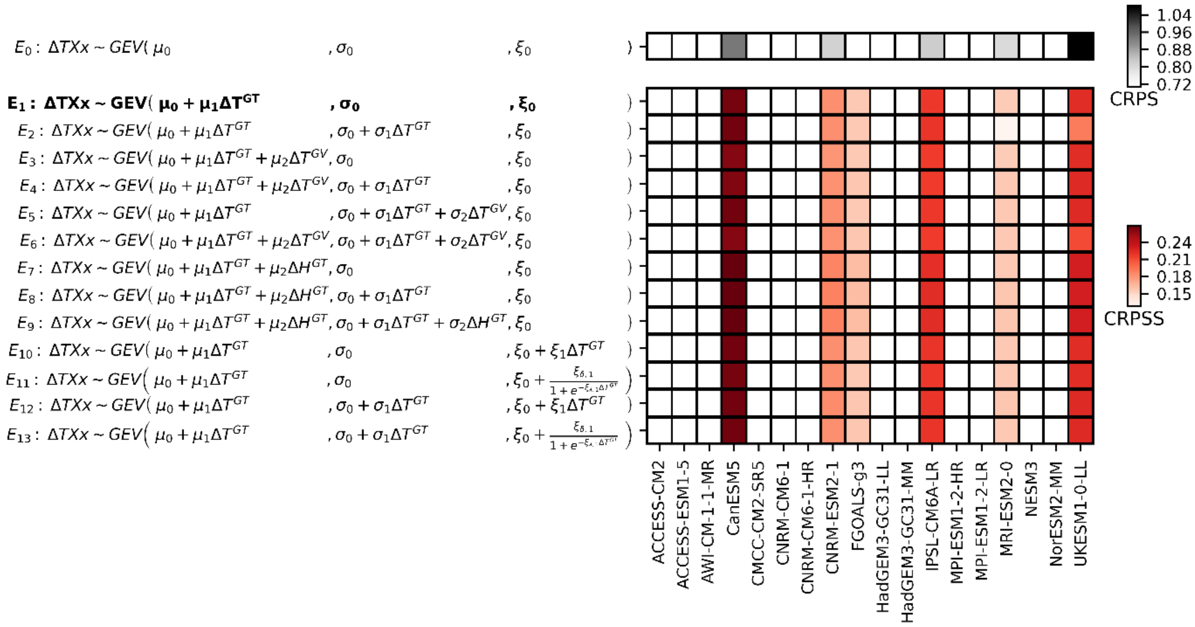

Figure S.3. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the *ssp119* (2015-2100).

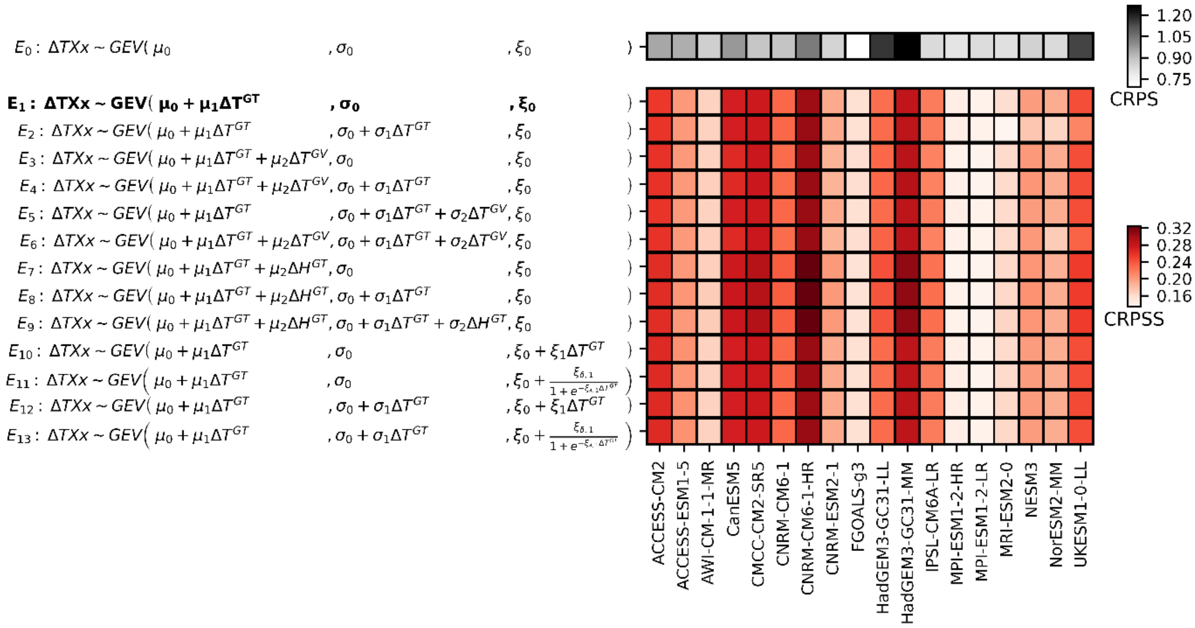

Figure S.4. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the *ssp126* (2015-2100).

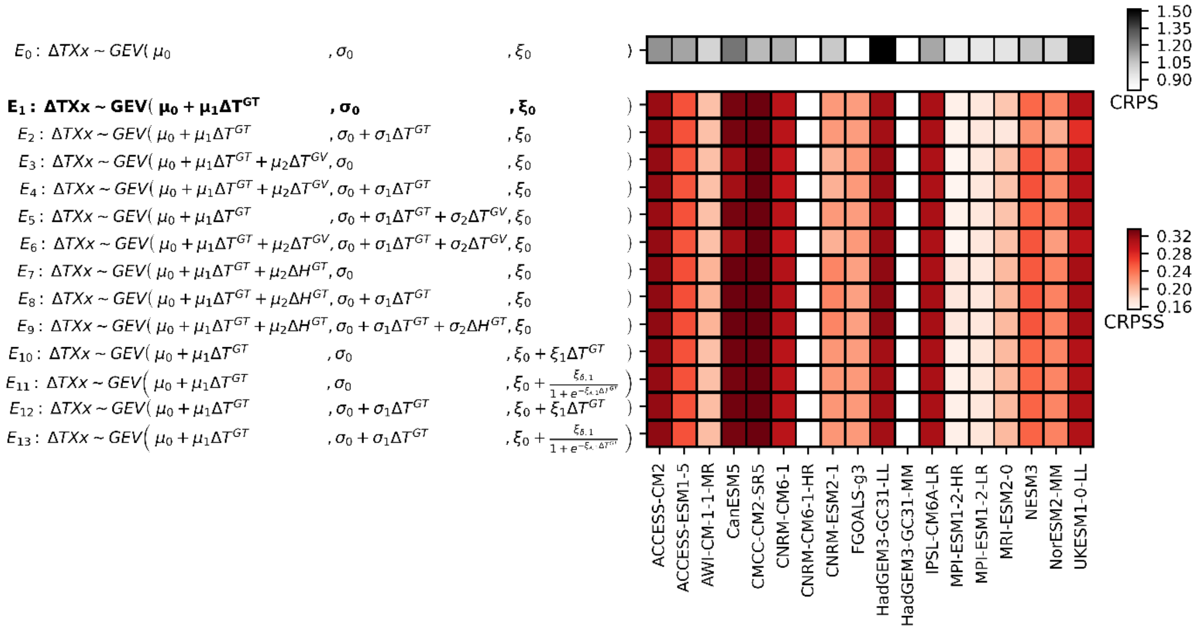

Figure S.5. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the *ssp245* (2015-2100).

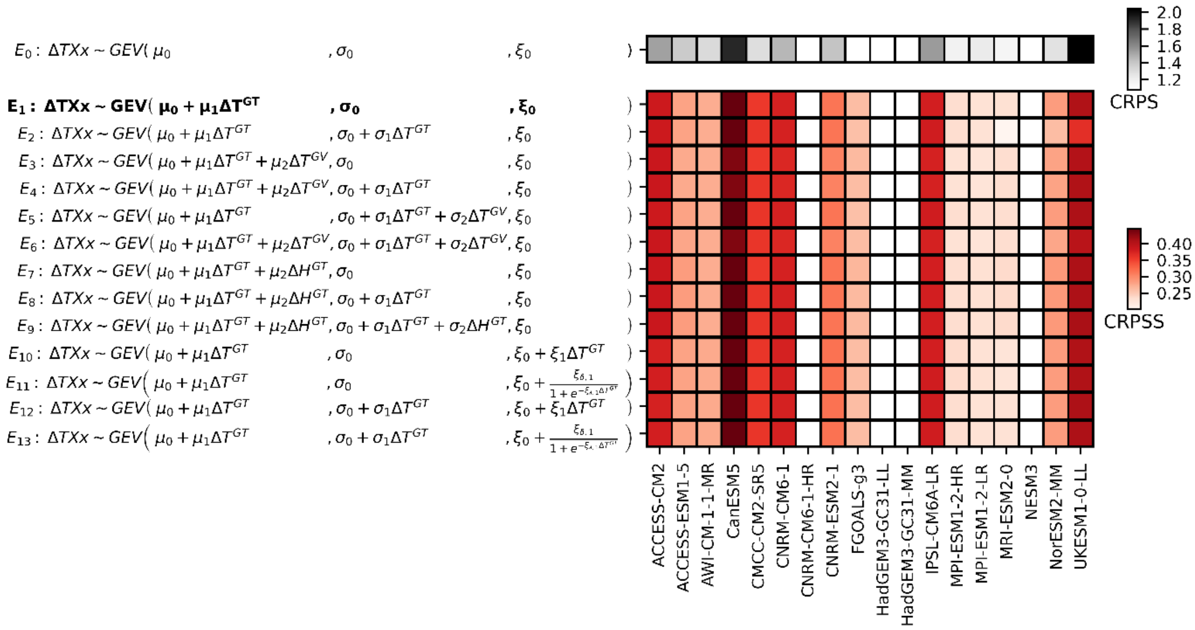

Figure S.6. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the *ssp370* (2015-2100).

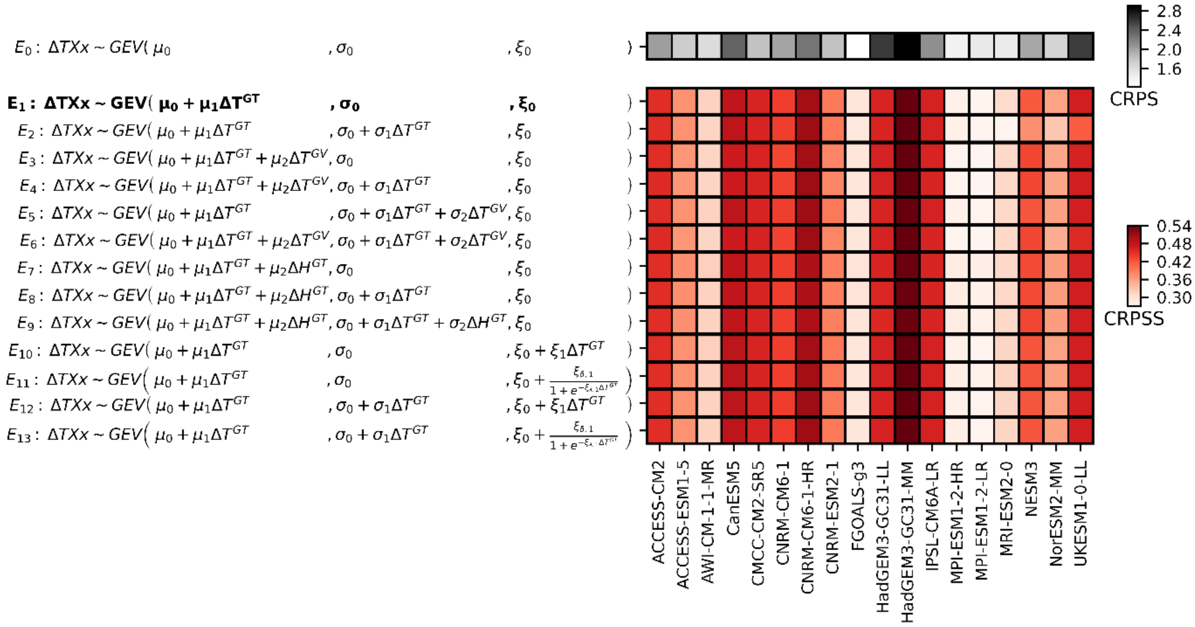

Figure S.7. Same as Figure S.1, with training over all available scenarios, but evaluation solely over the *ssp585* (2015-2100).

Examples of emulations under each ESM:

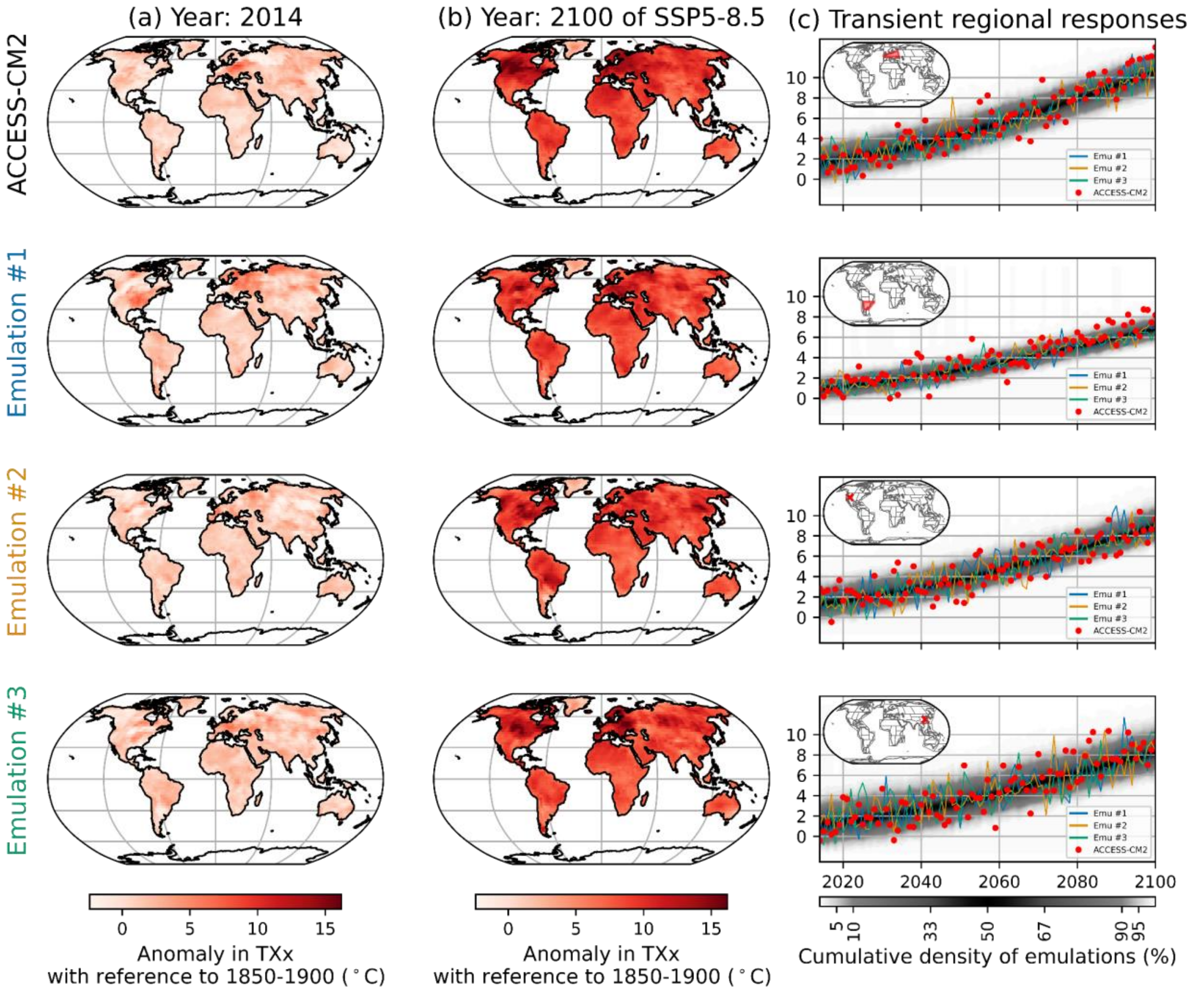

Figure S.8. Example of emulations for ACCESS-CM2 and three of its emulations in 2014 and 2100, in columns (a) and (b), respectively. The transient regional response from 2014 to 2100 are shown in column (c) for selected regions and grid points. It features the values from ACCESS-CM2, the same three emulations shown in maps and the density of the 1000 emulations drawn for this emulator configuration.

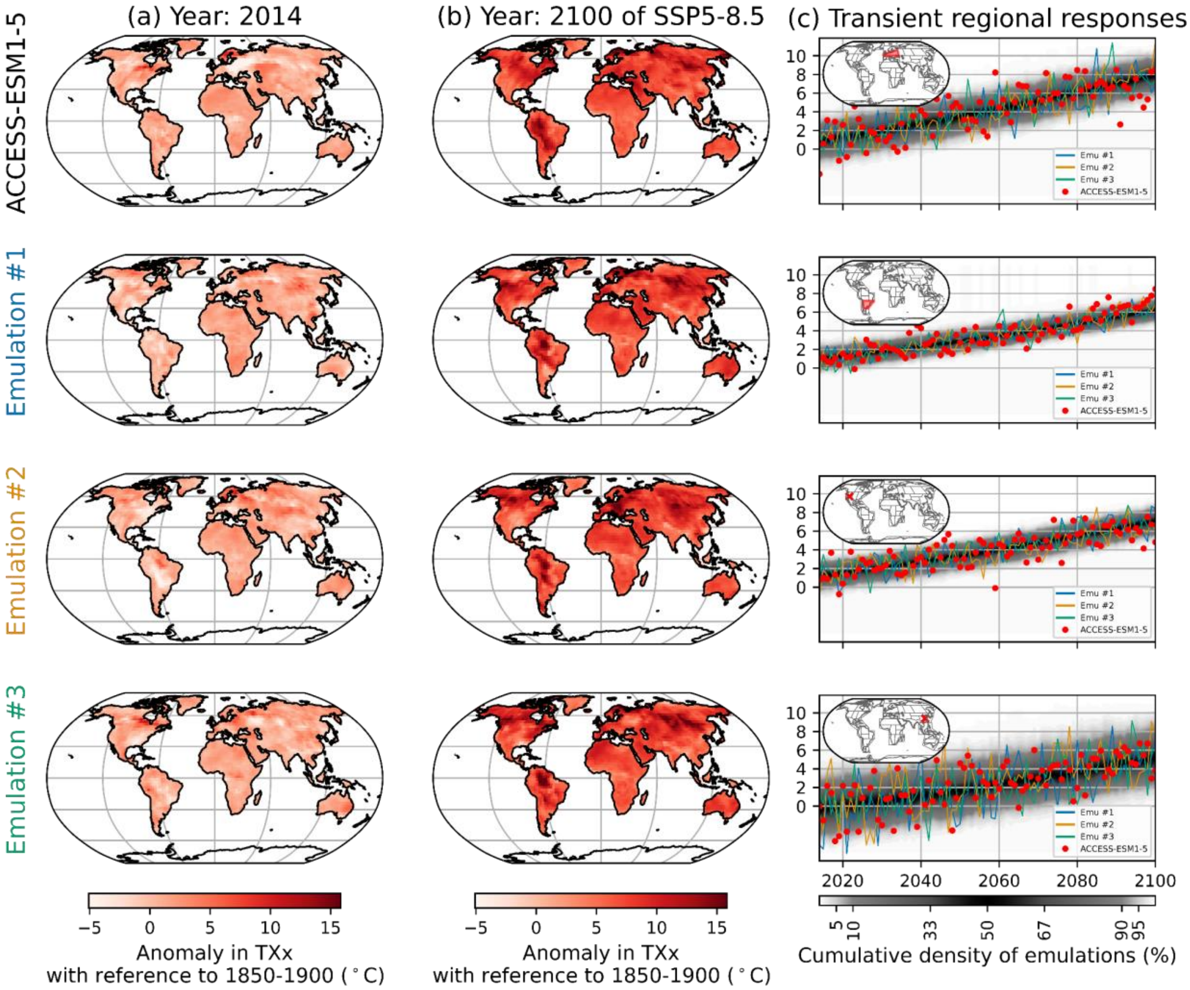

Figure S.9. Same as Figure S.8, but with ACCESS-ESM1-5.

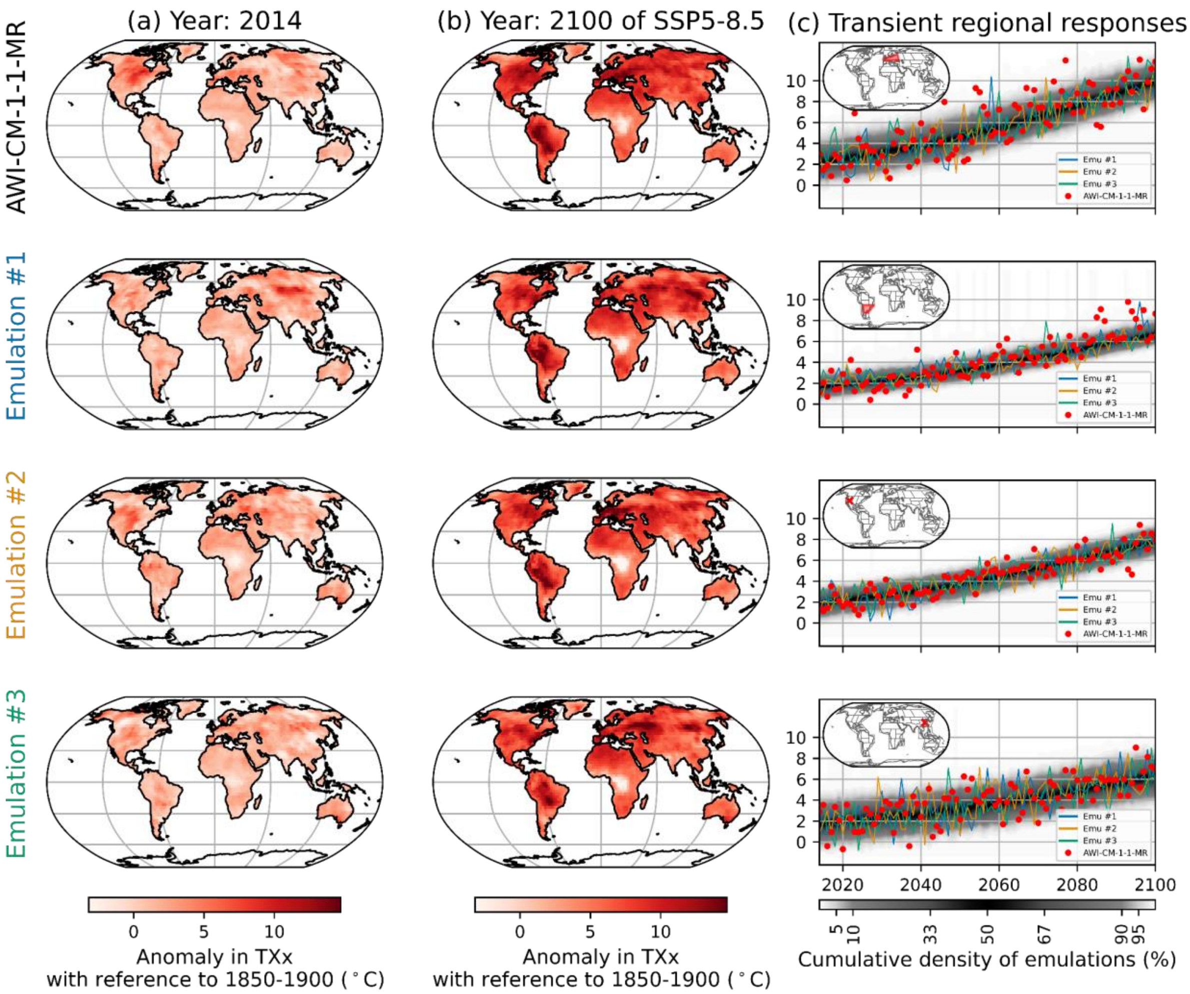

Figure S.10. Same as Figure S.8, but with AWI-CM-1-1-MR.

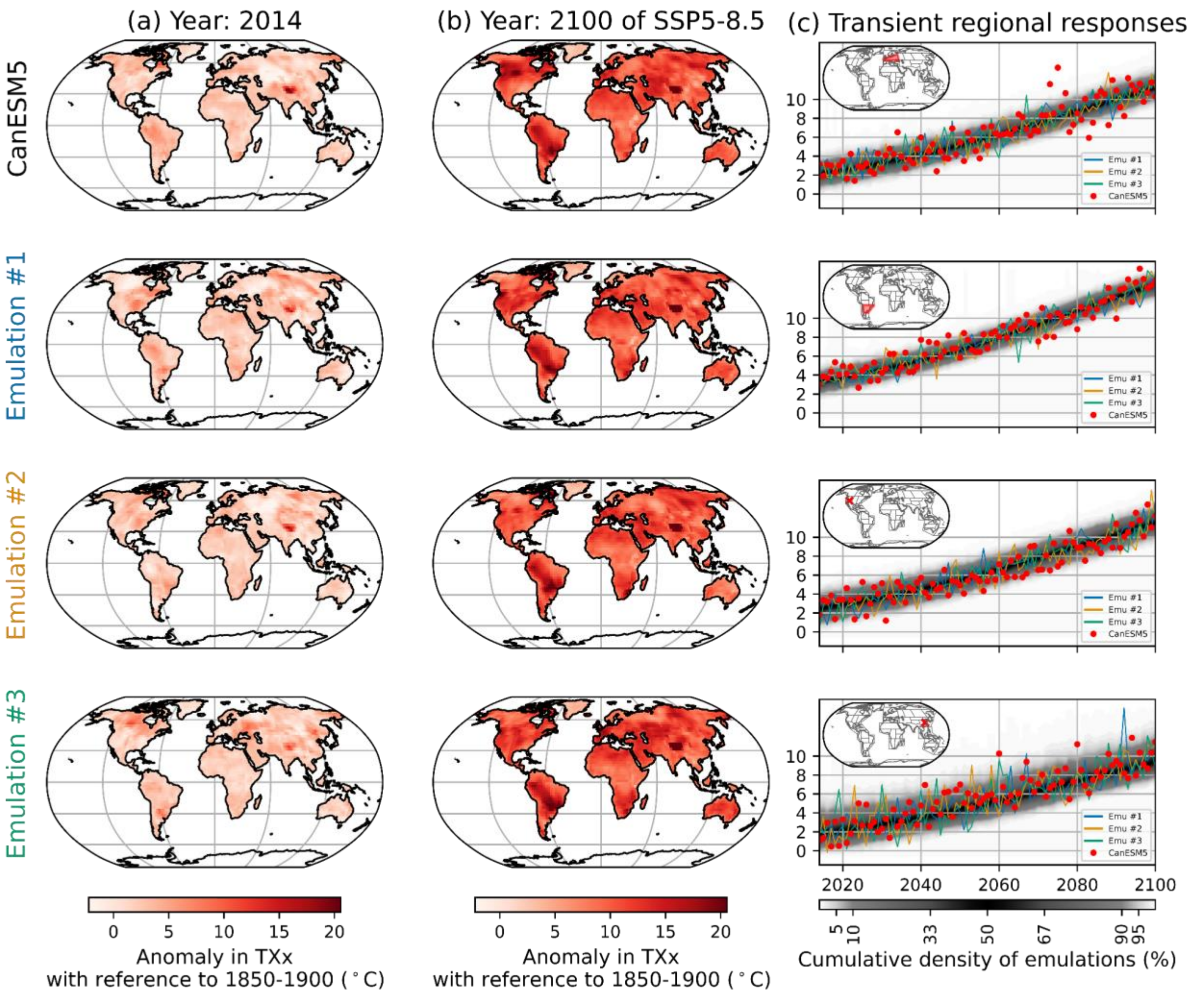

Figure S.11. Same as Figure S.8, but with CanESM5.

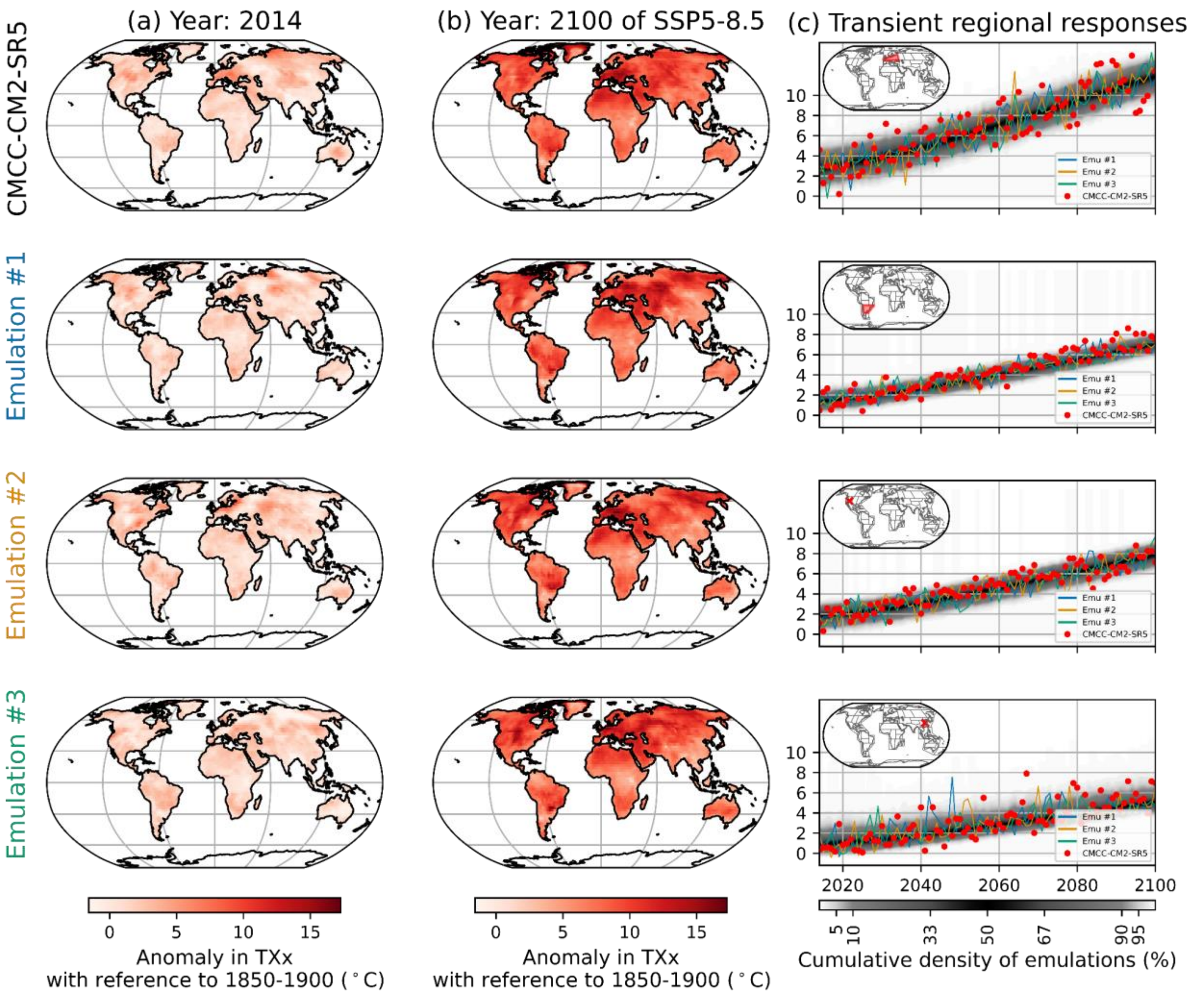

Figure S.12. Same as Figure S.8, but with CMCC-CM2-SR5.

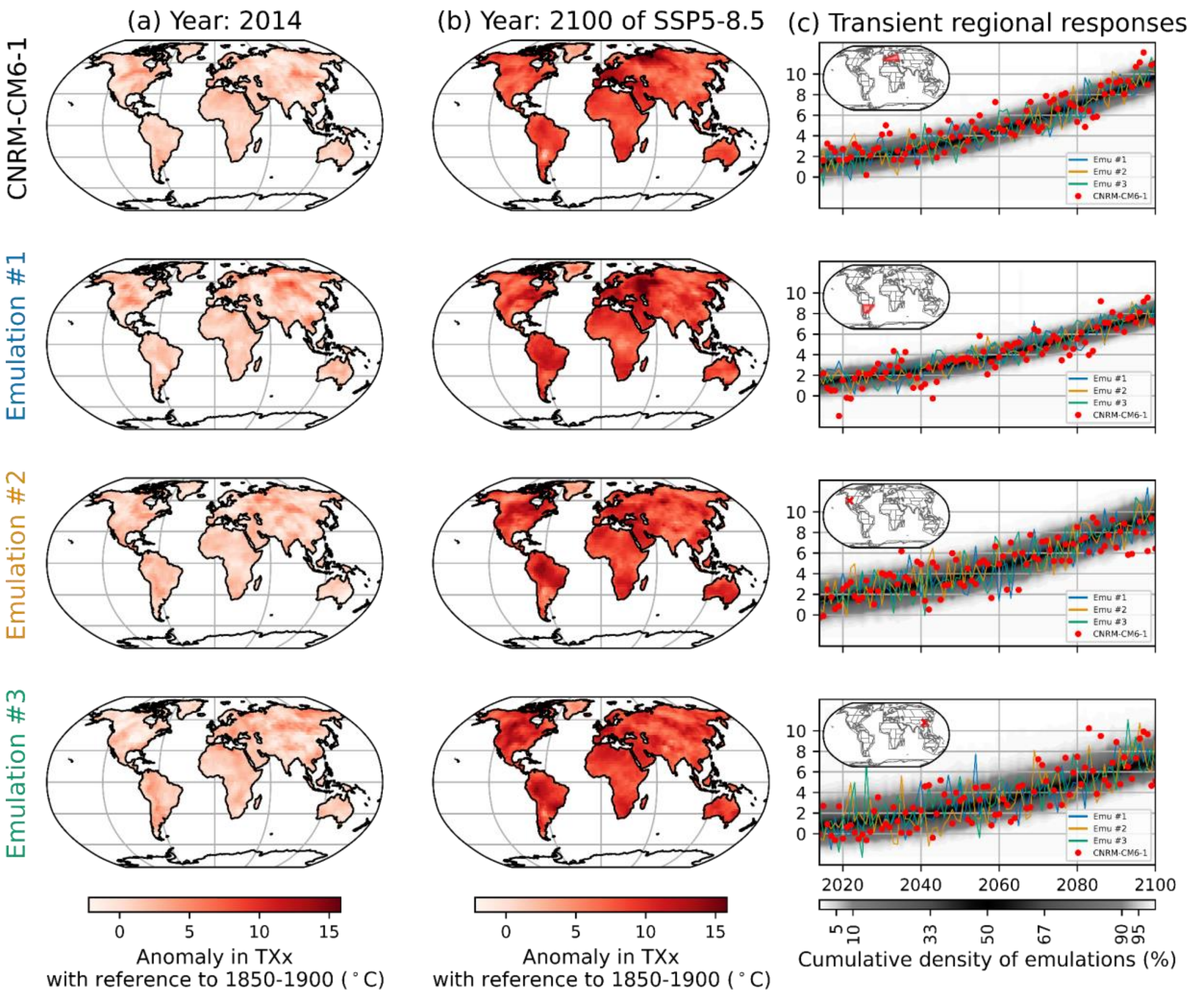

Figure S.13. Same as Figure S.8, but with CNRM-CM6-1.

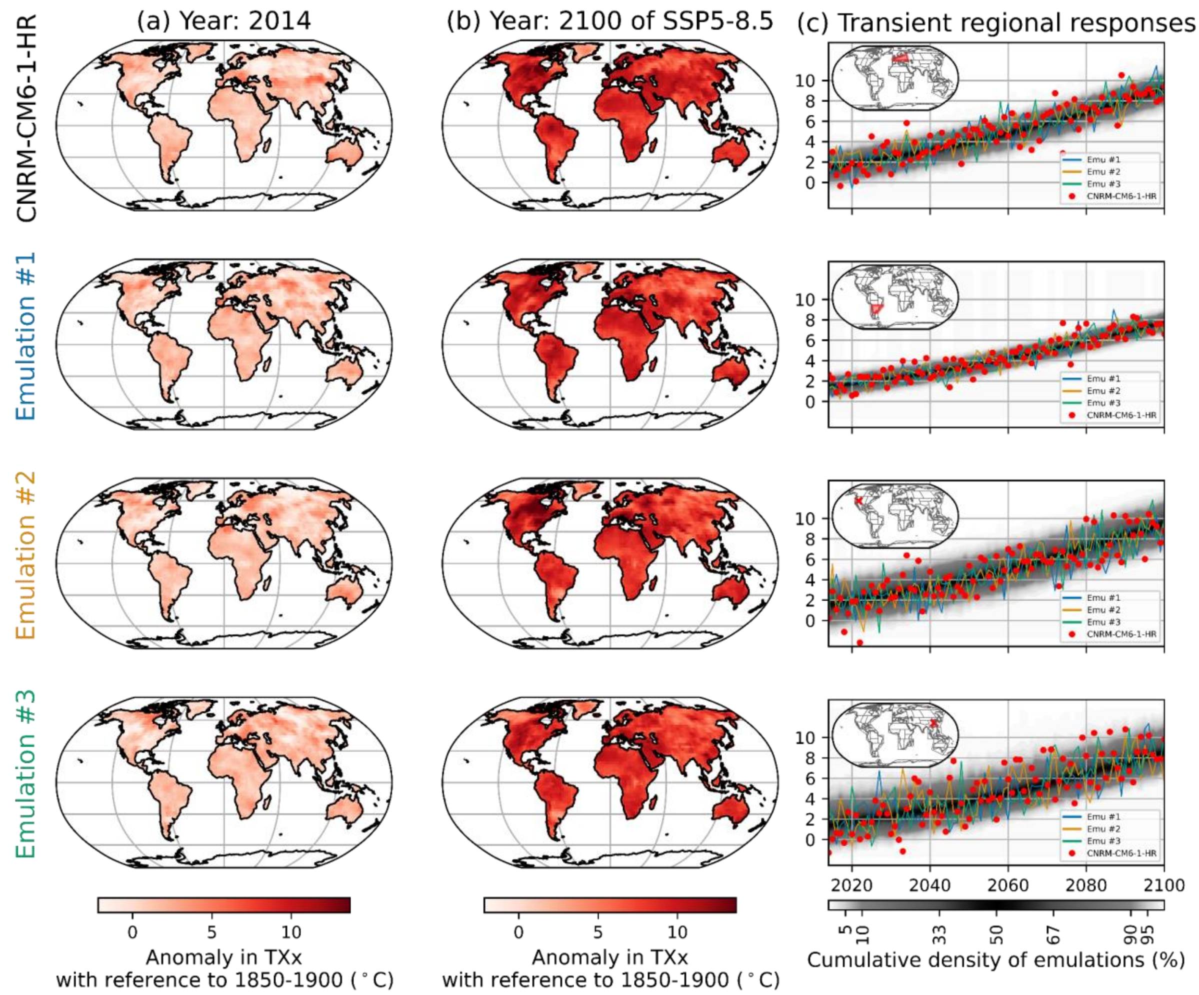

Figure S.14. Same as Figure S.8, but with CNRM-CM6-1-HR.

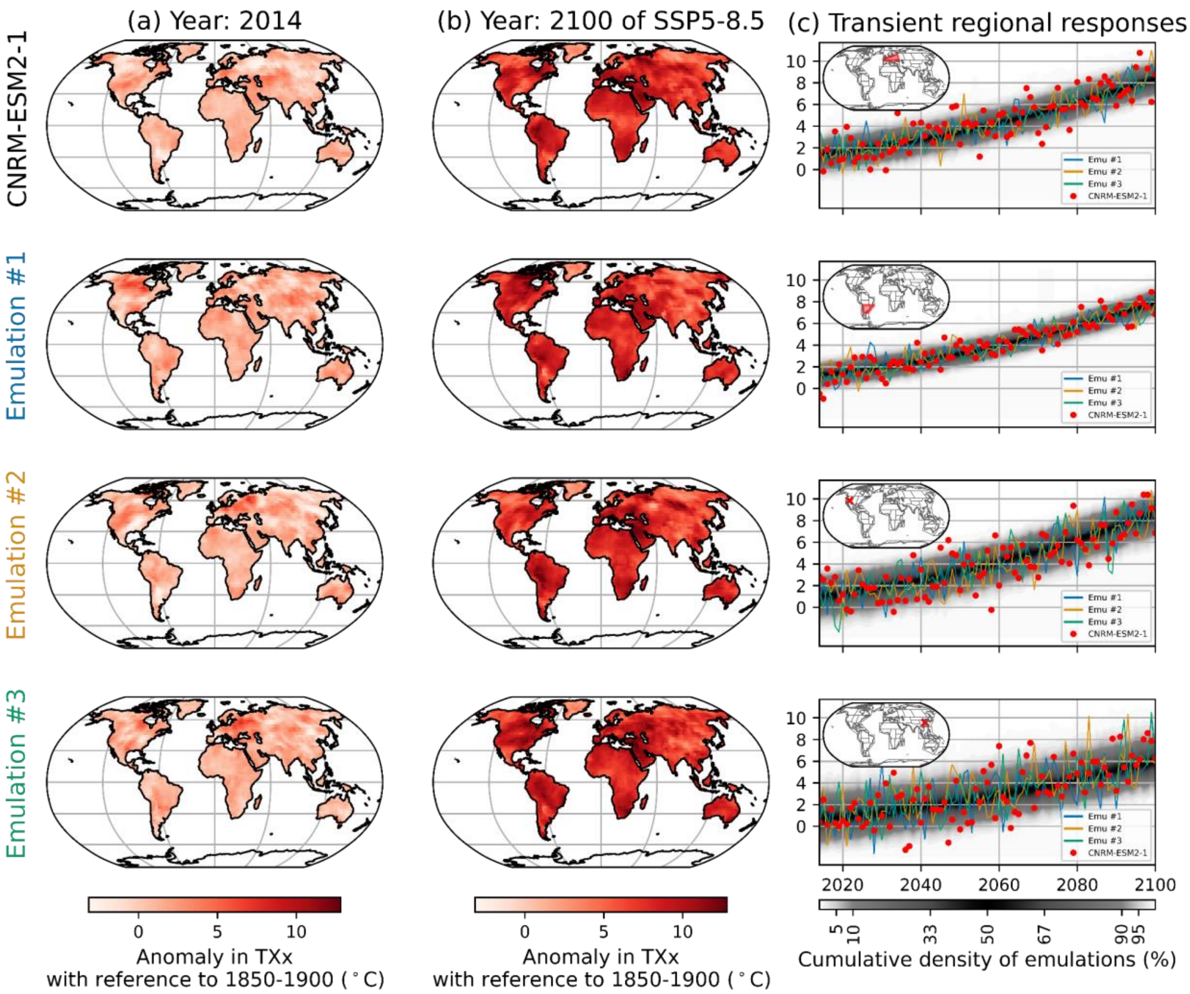

Figure S.15. Same as Figure S.8, but with CNRM-ESM2-1.

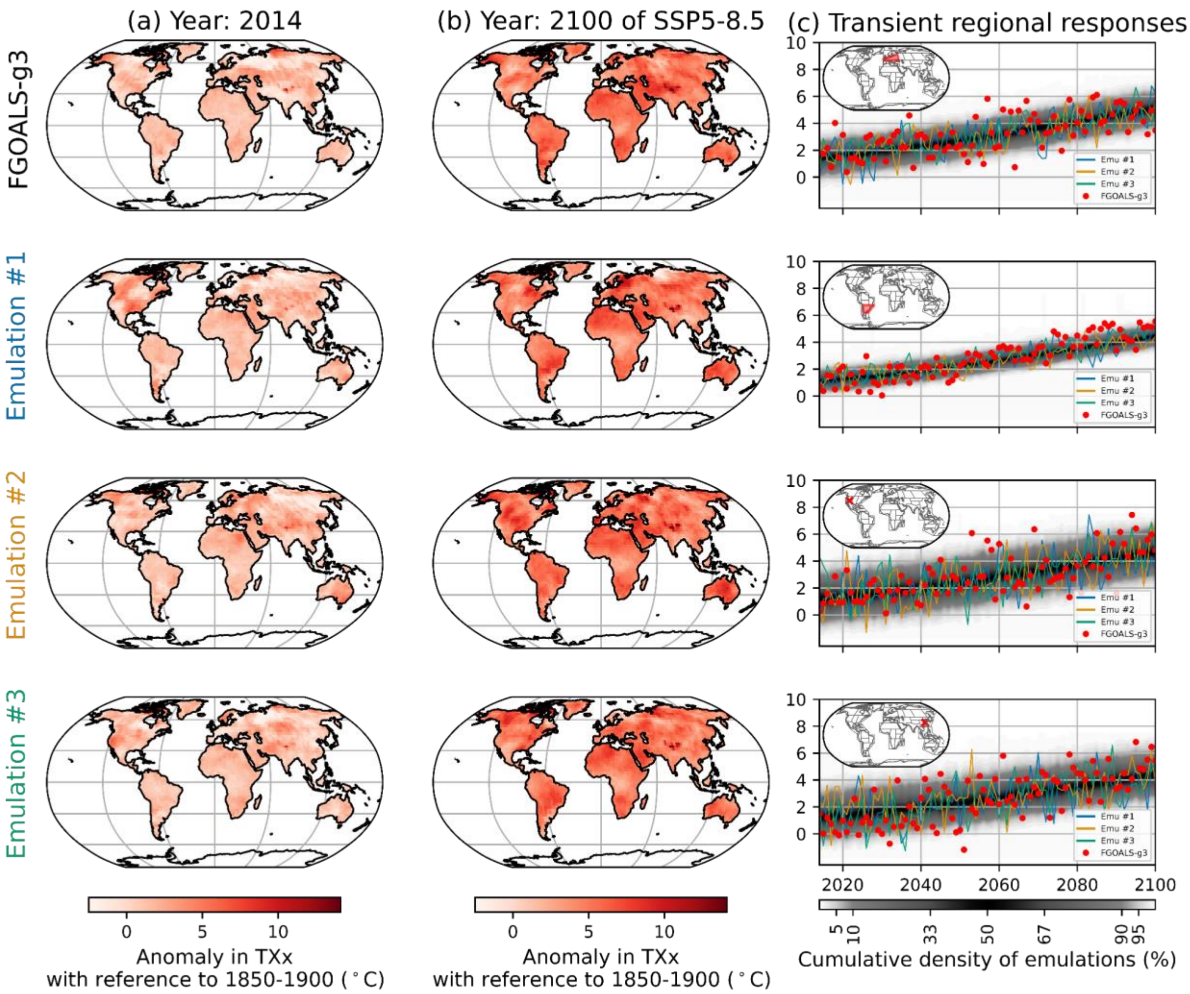

Figure S.16. Same as Figure S.8, but with FGOALS-g3.

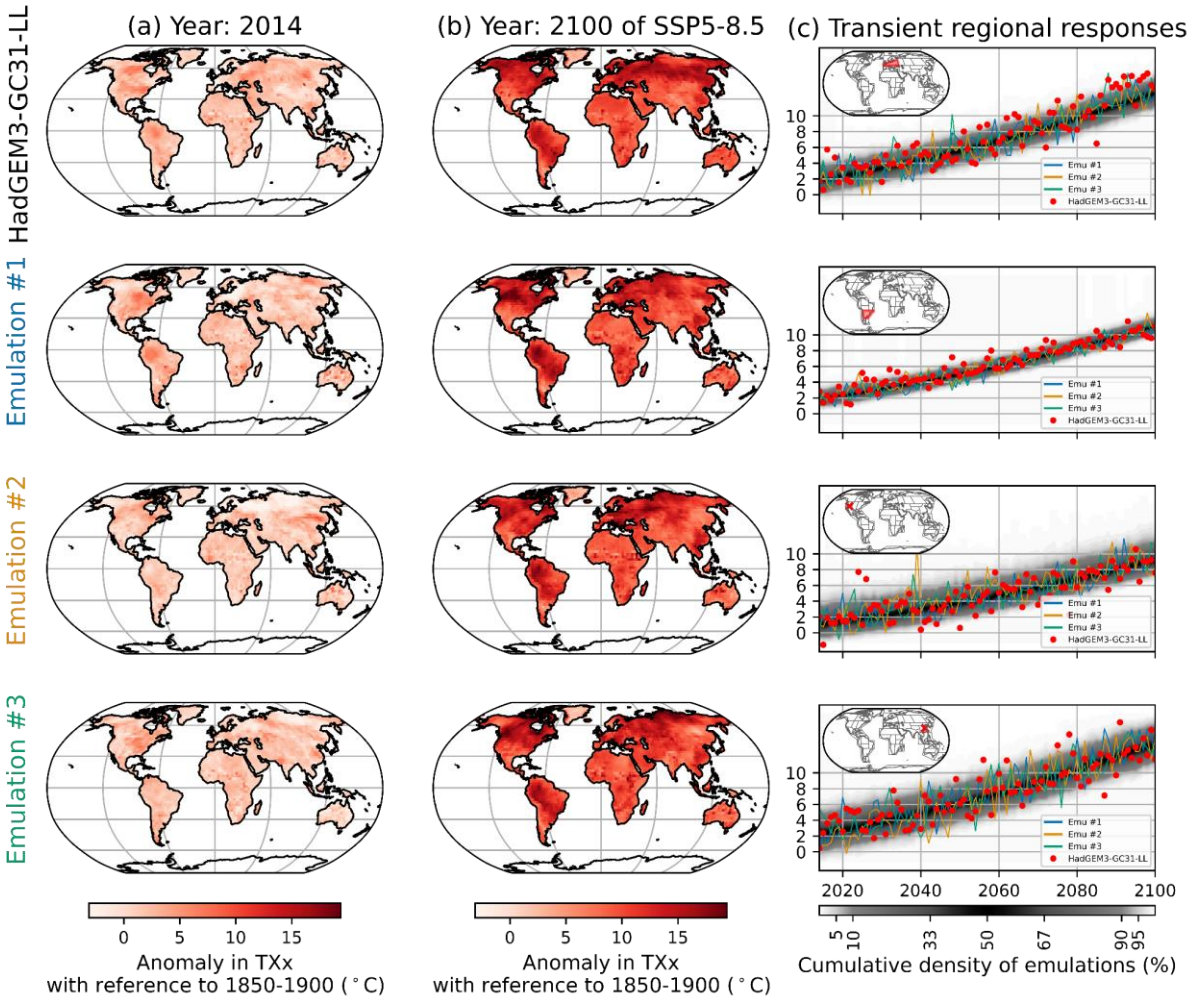

Figure S.17. Same as Figure S.8, but with HadGEM3-GC31-LL.

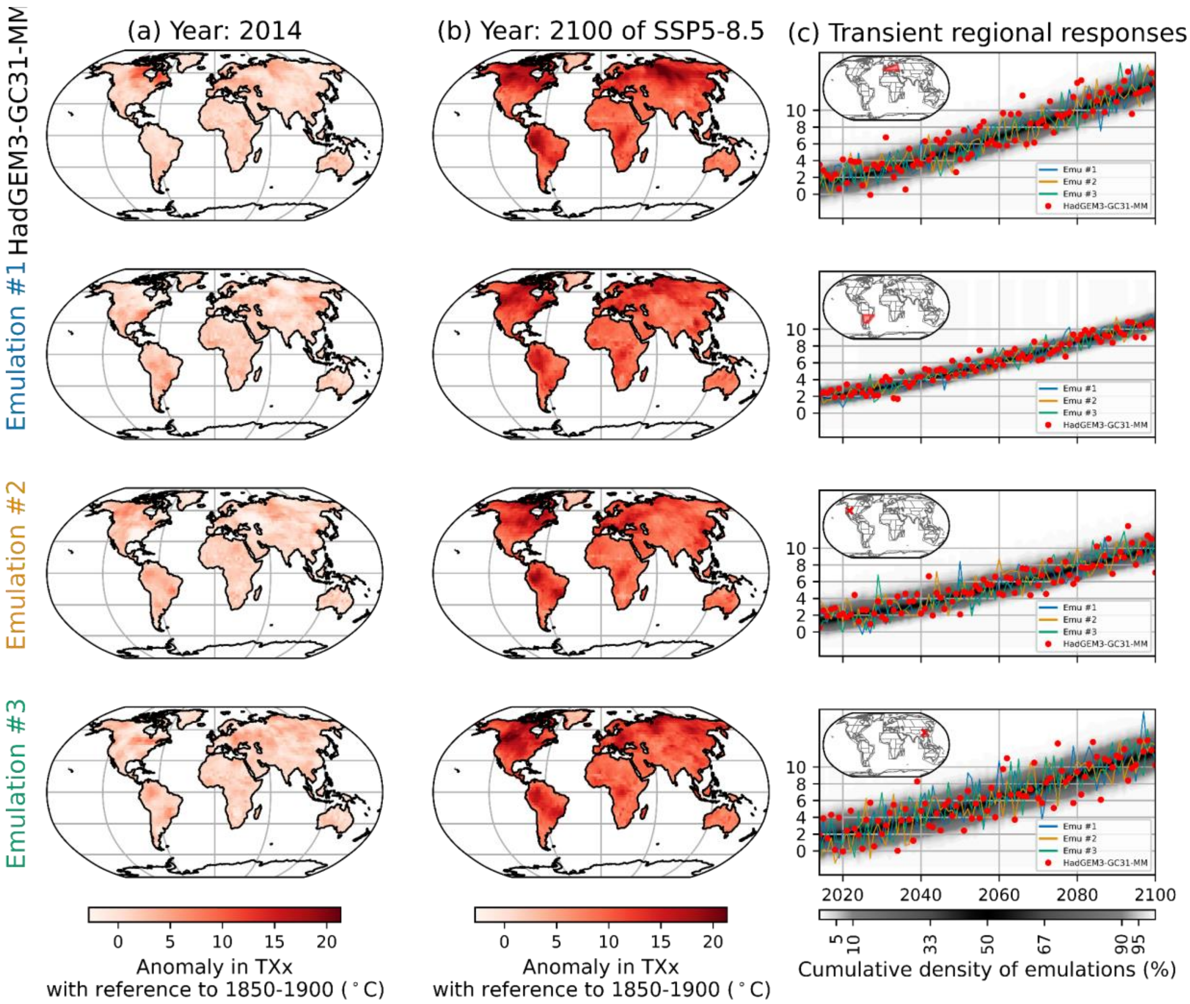

Figure S.18. Same as Figure S.8, but with HadGEM3-GC31-MM.

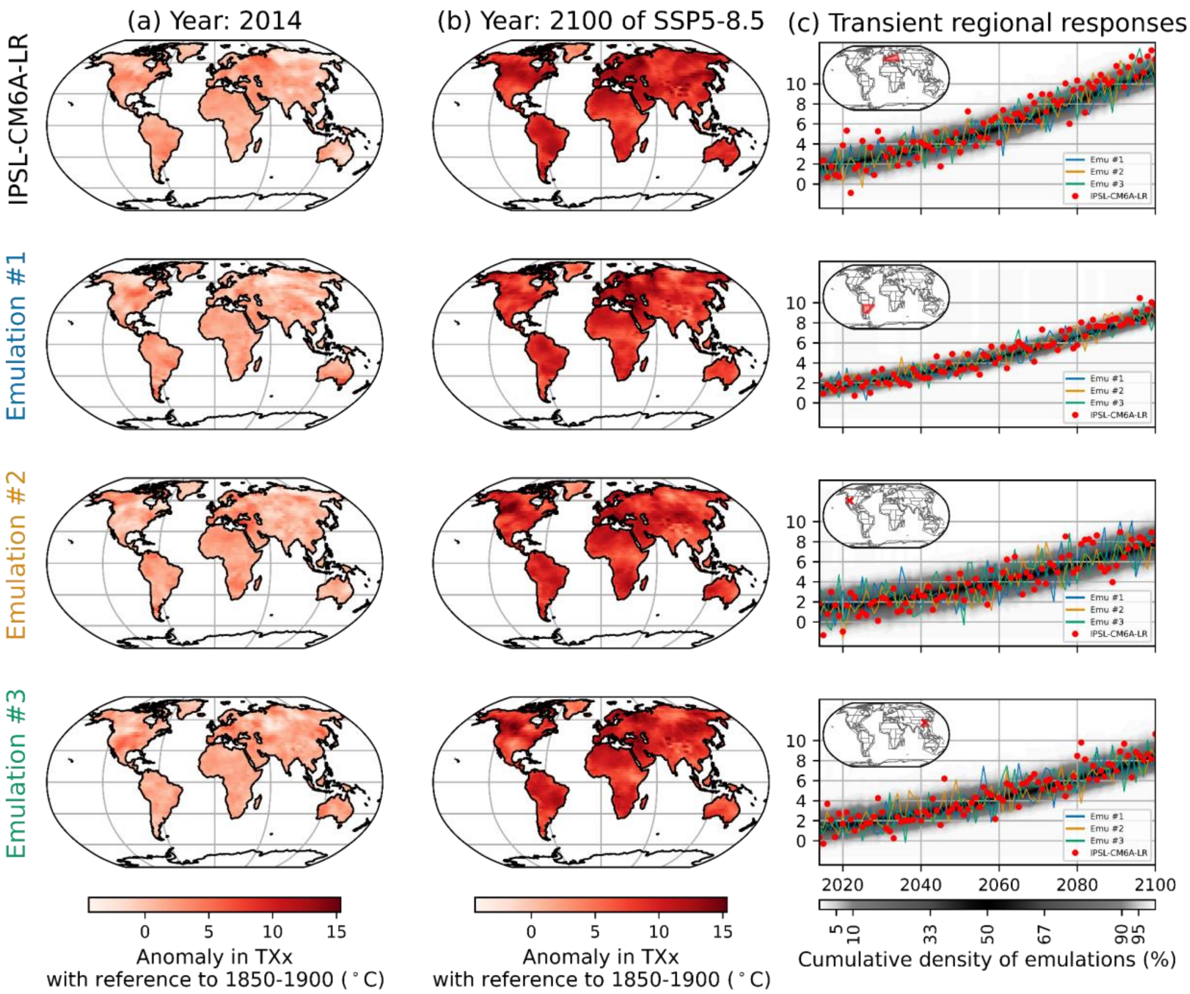

Figure S.19. Same as Figure S.8, but with IPSL-CM6A-LR.

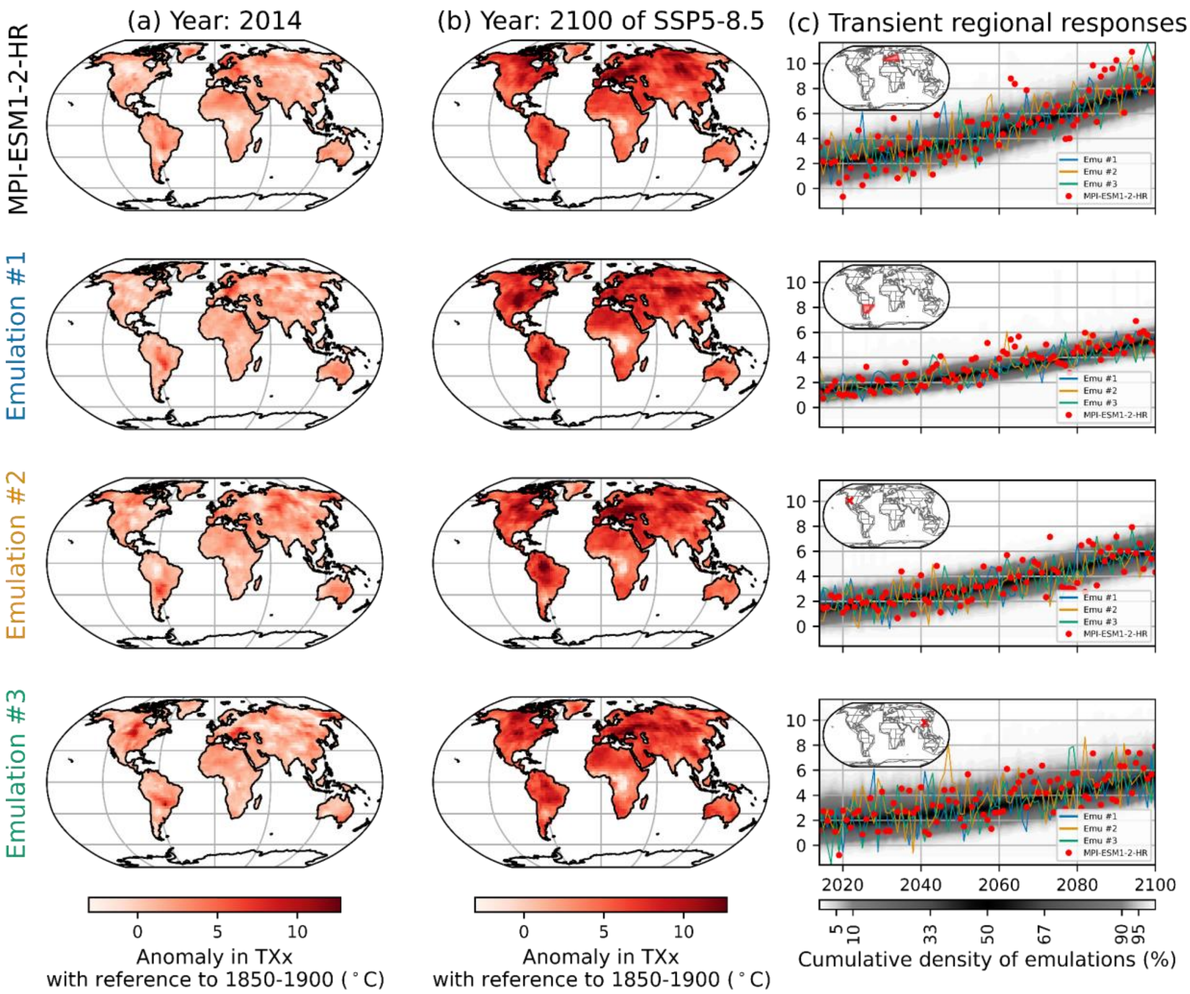

Figure S.20. Same as Figure S.8, but with MPI-ESM1-2-HR.

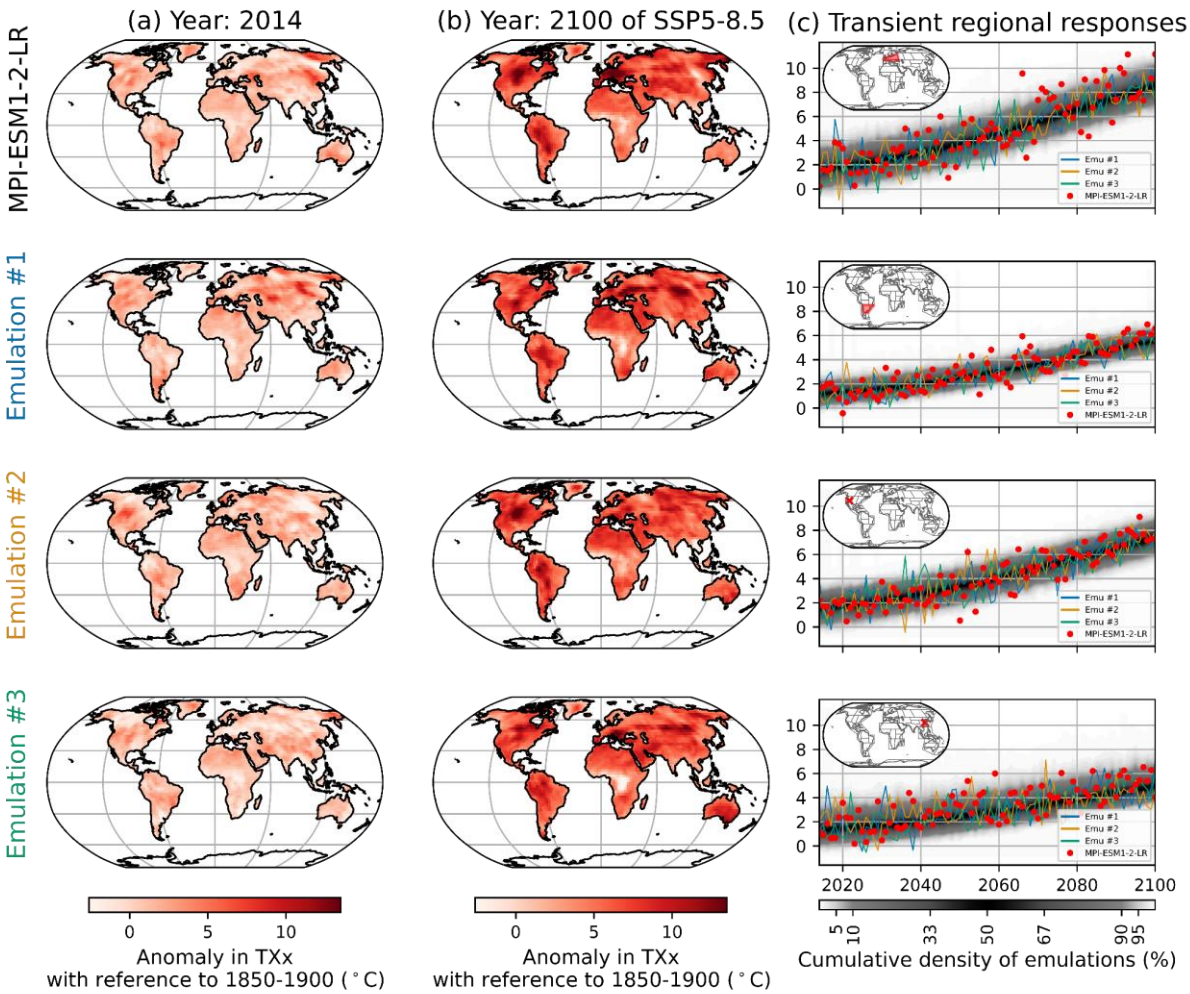

Figure S.21. Same as Figure S.8, but with MPI-ESM1-2-LR.

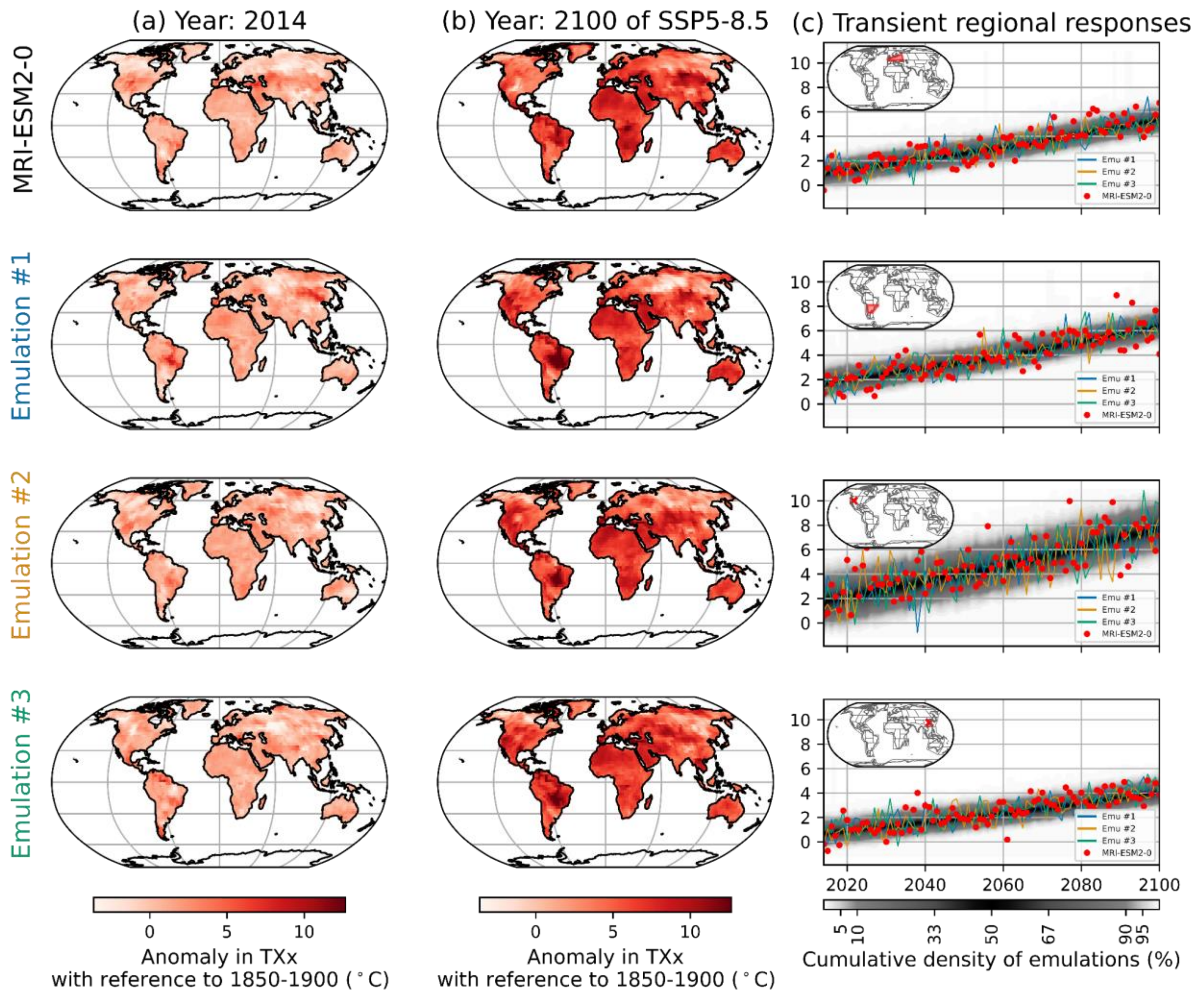

Figure S.22. Same as Figure S.8, but with MRI-ESM2-0.

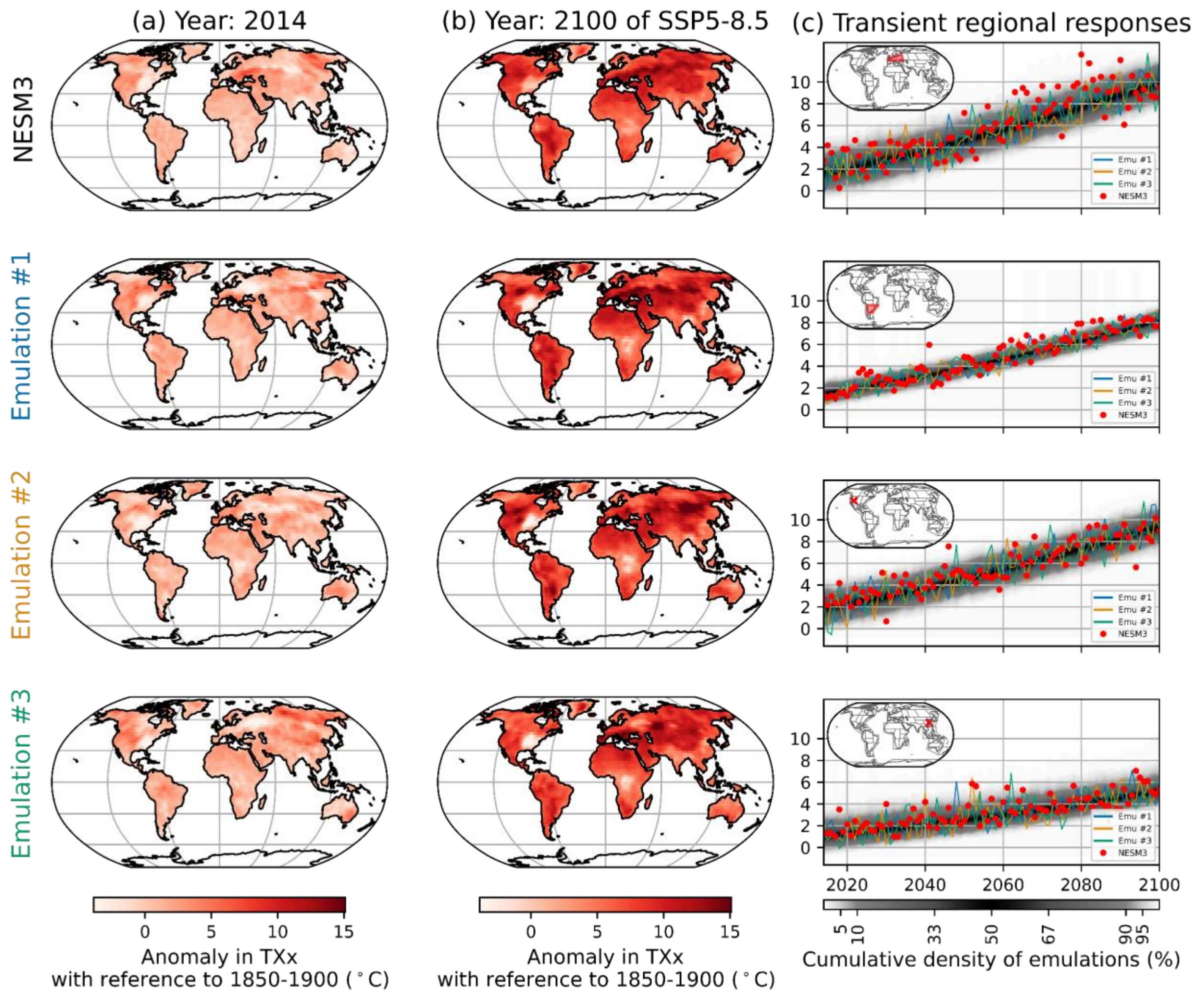

Figure S.23. Same as Figure S.8, but with NESM3.

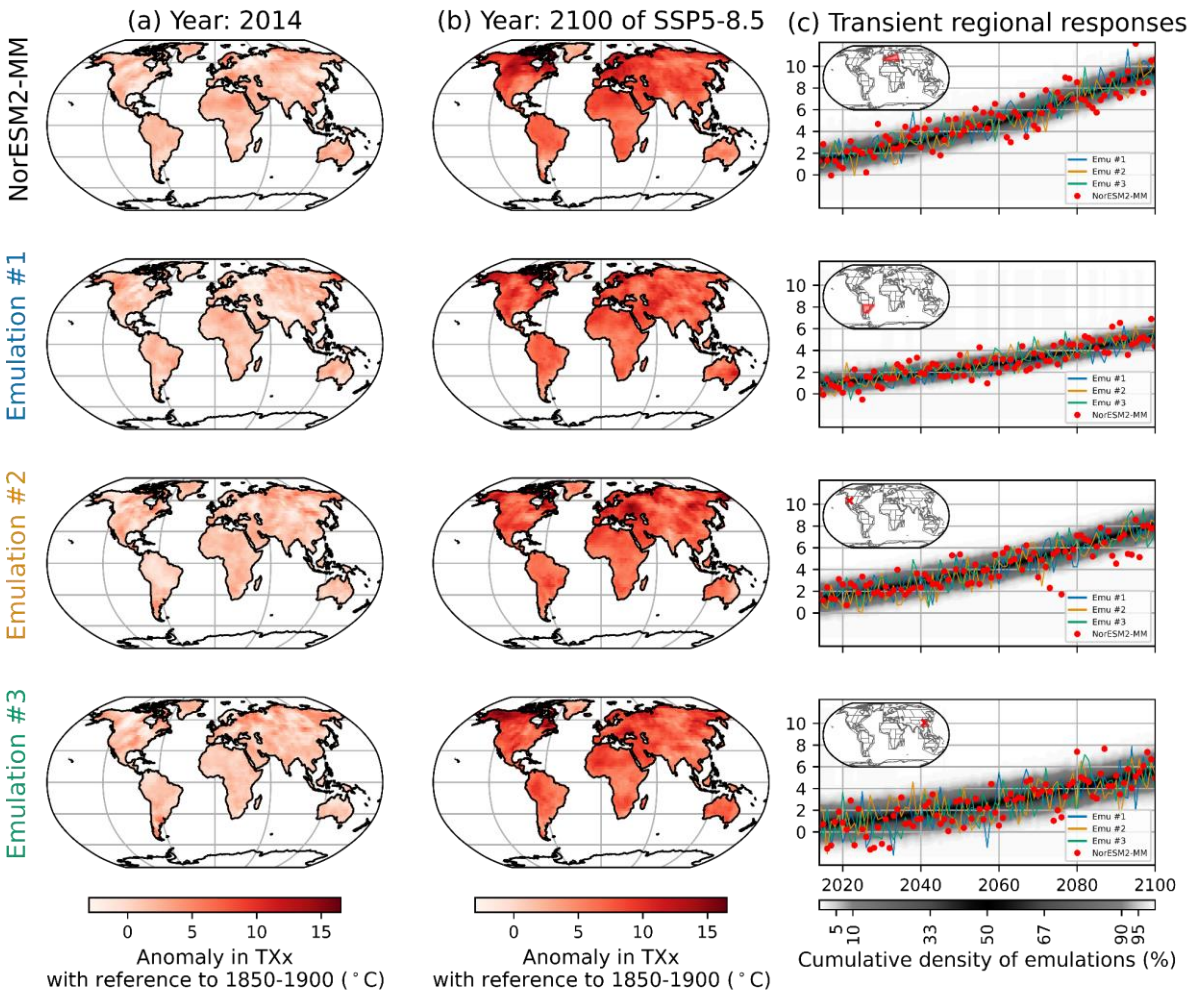

Figure S.24. Same as Figure S.8, but with NorESM2-MM.

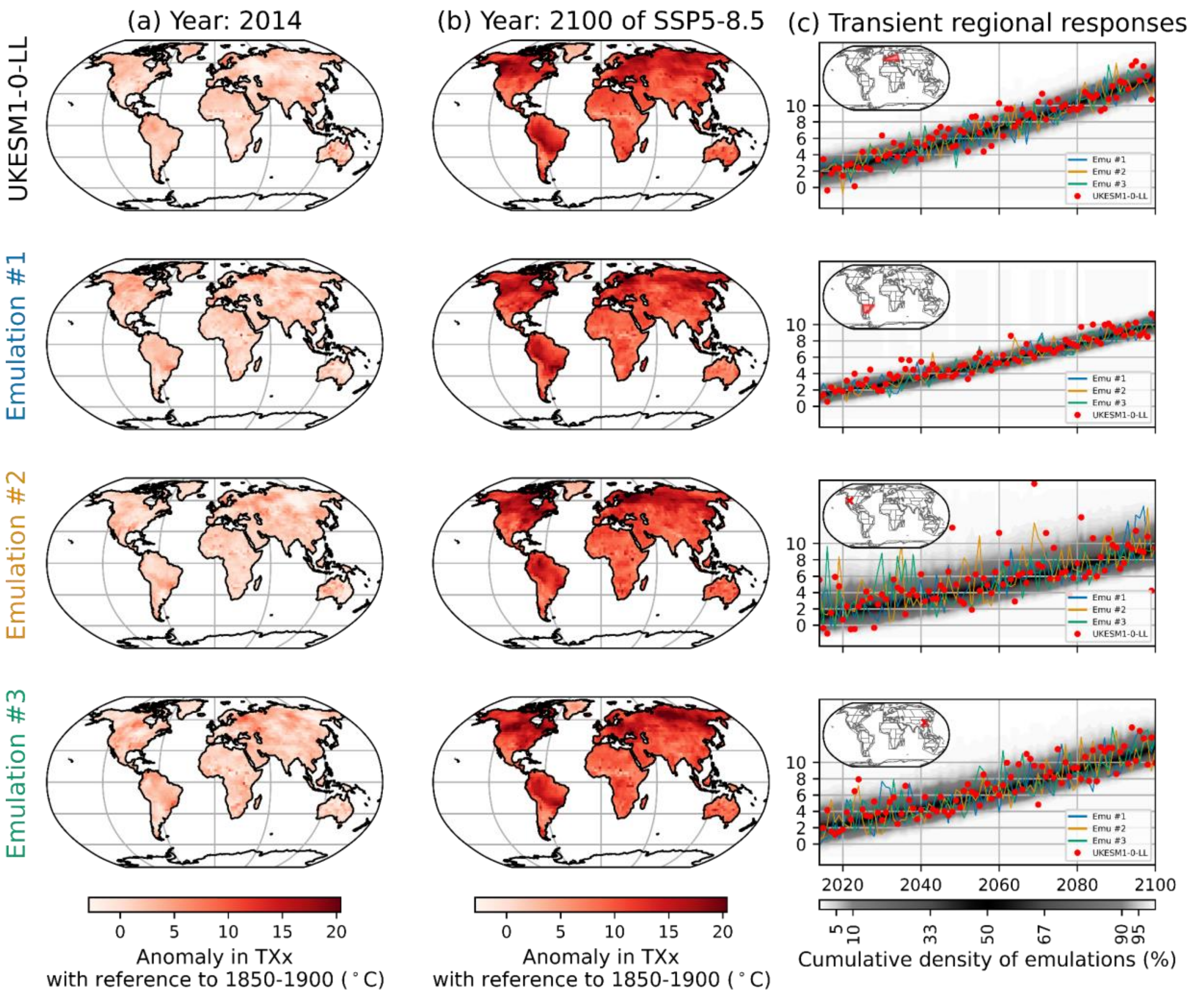

Figure S.25. Same as Figure S.8, but with UKESM1-0-LL.