Reducing uncertainty in local climate projections

Saïd Qasmi (said.qasmi@meteo.fr)
Centre National de Recherches Météorologiques https://orcid.org/0000-0002-1825-3896
Aurélien Ribes
Centre National de Recherches Météorologiques https://orcid.org/0000-0001-5102-7885

Article

Keywords: climate change, climate projections, adaptation, temperature

Posted Date: February 16th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-364943/v3

Reducing uncertainty in local climate projections

Saïd Qasmi*,[1] and Aurélien Ribes[1]

[1]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

*Corresponding author: said_qasmi@meteo.fr

Abstract

Planning for adaptation to climate change requires accurate climate projections. Recent studies have shown that the uncertainty in global mean surface temperature projections can be considerably reduced by using historical observations. However, the transposition of these new results to the local scale is not yet available. Here we adapt an innovative statistical method that combines the latest generation of climate model simulations, global observations, and local observations to reduce uncertainty in local temperature projections. By taking advantage of the tight links between local and global temperature, we can derive the local implications of global constraints. The model uncertainty is reduced by 30% up to 50% at any location worldwide, allowing to substantially improve the quantification of risks associated with future climate change. A rigorous evaluation of these results within a perfect model framework indicates a robust skill, leading to a high confidence in our constrained climate projections.

Main

As the global mean temperature keeps rising and climate change intensifies, there is a growing demand for local scale monitoring of current and future climate change. Assessing and planning the adaptation to the expected unprecedented impacts of climate change on humans activities, ecosystems and the biosphere as a whole, require an accurate local information with well calibrated uncertainties. This need relates to estimates of warming to date and the future warming in response to set of scenarios of future greenhouse gas emissions.

There is now clear evidence that the recent increase of the average Earth's temperature is mostly due to human activities [1, 2]. Concurrently, the anthropogenic influence is not detected everywhere at the local scale [3, 4]. Natural climate variability can blur the emergence of the anthropogenic signal for the next few years at high latitudes, while a significant warming is already reported in several tropical regions [5, 6]. Regarding climate projections, the IPCC concluded in its 5th assessment report (AR5) [7], that "Future [human-induced] warming trends cannot be predicted precisely, especially at local scales".

In the IPCC AR6 [2], a new generation of climate models [8] has been used to provide a range of projections in response to different socio-economic scenarios [9]. Based on this new dataset, various studies have recently shown that uncertainty in global mean warming can be considerably reduced by using the information provided by recent observed warming trends via so-called "constraint" methods [10, 11, 12, 13]. These studies consistently point towards a downward revision of the expected warming in

all emission scenarios [12, 10], with a decrease in model uncertainty of nearly 40% for end of century projections [11], and even more at shorter lead times. This is an important result as, until then, observations have failed to provide clear evidence in reducing the range of climate projections [14].

The next challenge is to transpose these new findings on global warming to regional and local scales. At the regional scale, a few studies have adopted the partitioning from the Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [15] (SREX) and have attempted to narrow model uncertainty with sophisticated techniques with promising results [16, 17]. But the SREX regions are typically continental-wide, and do not provide relevant information for local adaptation. At the local scale (defined as the size of a global climate model grid box of about 200 km), and to the best of our knowledge, only a few studies have attempted to narrow climate model uncertainty, by using weighting methods to account for inter-dependencies between models [18, 19], or by focusing on specific and limited areas [20]. In particular, although constrained GSAT projections are now used in the IPCC AR6 [2], local climate projections are still solely based on a raw ensemble of available climate models (https://interactive-atlas.ipcc.ch/), derived from global warming levels.

Here we assess how much uncertainty in local temperature projections can be reduced. We first take advantage of the tight links that exist between local climate and GMST [21, 22]. Specifically, we describe the local implications of the recent advances in the reduction of the uncertainties in GMST projections. We then provide a set of local-scale temperature projections, which encapsulate another source

of information: the observed local warming to date. If compared to the global mean temperature record, local observations are typically more affected by internal variability and measurement uncertainty. Yet, they still provide a useful source of information on both past and future trends, particularly over some specific regions. We discuss how much these two types of observations (global and local) narrow uncertainty on future warming ranges. Such a reduction is expected to provide more accurate information that becomes critical for policy-makers in the local climate risk management [23], as well as for the climate science community.

The Kriging for Climate Change (KCC) method used by Ribes et al. [11], is one of the statistical techniques that have led to a significant reduction of uncertainty in probabilistic projections of GMST by combining observations and models. This Bayesian method involves first of all the definition of a prior distribution, which is here based on the temperature response simulated by the climate models, after filtering out internal variability as much as possible. This prior is subsequently conditioned by temperature observations over the historical period to derive a posterior probability distribution of the forced response, *i.e.*, a constrained temperature response, for both past and future periods. Compared to many emergent constraint approaches, in which observations are often summarised into one single variable, the KCC method is able to take the full observed time-series into account. Here, we further extend this technique in order to account for multiple time series, and potential dependencies between them (see Methods). When applied to the GSAT time series simulated by the models from the Coupled Models Intercomparison Project phase 6 (CMIP6) [8] models and the Shared Socio-Economic Pathway (SSP) 5-8.5

scenario, the amplitude of the projected GSAT changes constrained by the observations is revised downwards by 0.5 °C by 2100, with a reduction in model uncertainty of nearly 40% [11]. Figure 1a offers an update of this result using GMST instead of GSAT (see Methods for the difference between GSAT and GMST): the warming of 5.3 °C projected by CMIP6 is in this case revised downward by 0.6 °C. Differences with Ribes et al. [11] are explained by the addition of several CMIP6 models, which affects the prior, and by the lower warming observed in GMST compared to GSAT [24, 25]. In the following, we consider the GMST metric to be consistent with the local observation dataset used (see Data).

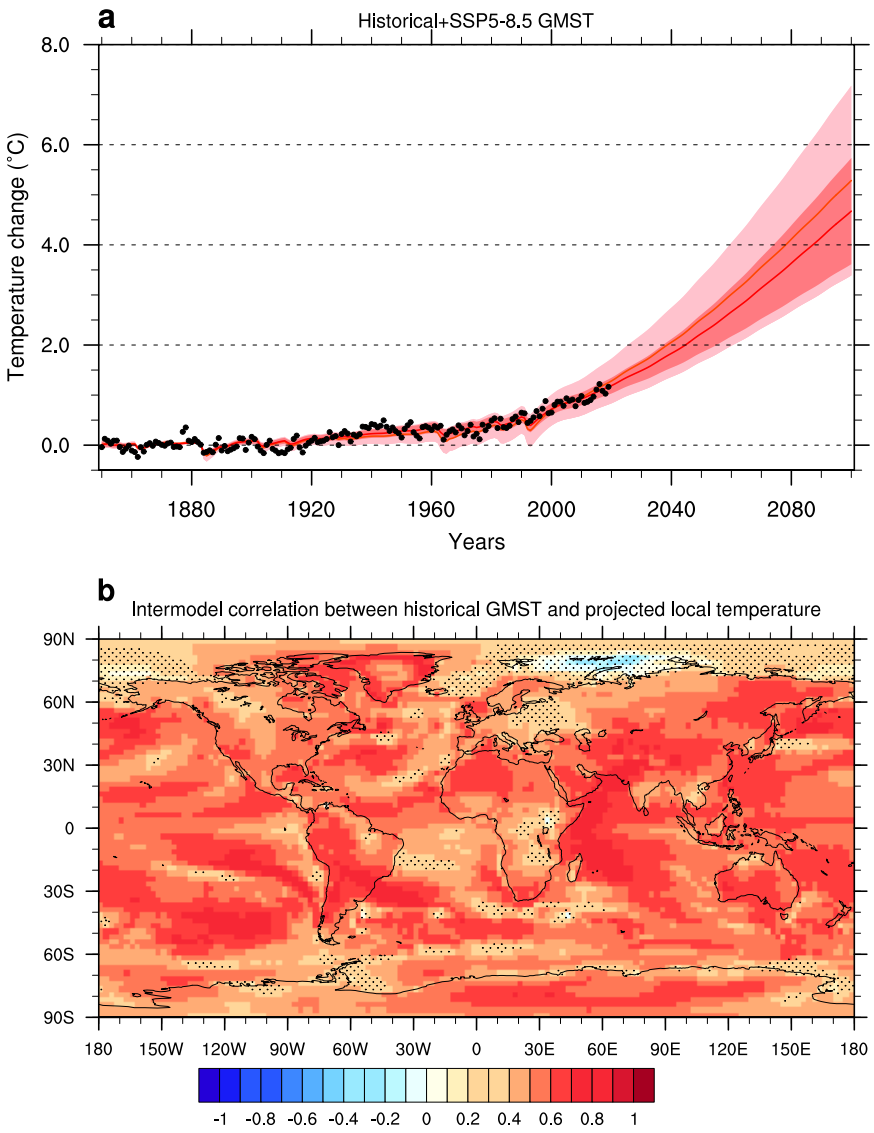

Fig. 1: **GMST time series and its correlation with local temperature.** (a) GMST annual observations from the Cowtan and Way [26] dataset (black points) are used to constrain concatenated historical and SSP5-8.5 scenario simulations of GMST. The unconstrained (pink) and constrained (red) ranges stand for the 5–95% confidence interval of the forced response as estimated from 27 CMIP6 models. The thick pink (red) line stands for the ensemble mean (best estimate). All values are anomalies with respect to the 1850–1900 period. (b) Intermodel correlation between simulated GMST trends over the 1850-2019 period and local temperature trends over the 2020-2100 period. Stippling indicates regions with non-significant correlation (p-value > 0.05 based on a two-sided Student's t-test).

Constrain local climate projections with global observations

Climate models exhibit a strong correlation between current GMST changes and future local warming over most regions of the globe (Figure 1b). To take such a relationship into account, we extend the KCC method to constrain local temperature projections. This is done by deriving the local warming conditional on the observed GMST record (hereafter the GMST-only case, see Methods). As an example, we consider the North American city of Dallas for which the simulated local temperature over the 2020-2100 period is significantly correlated with GMST (see the corresponding blue point 7 in Supplementary Figure 3). Consistent with the constrained GMST, the constrained local temperature range indicates a decrease in uncertainty of about 30% over the 2081-2100 period (Figure 2a), and a downward revision of the best estimate of local warming by 0.4 °C compared to the unconstrained projections (hereafter the unconstrained case).

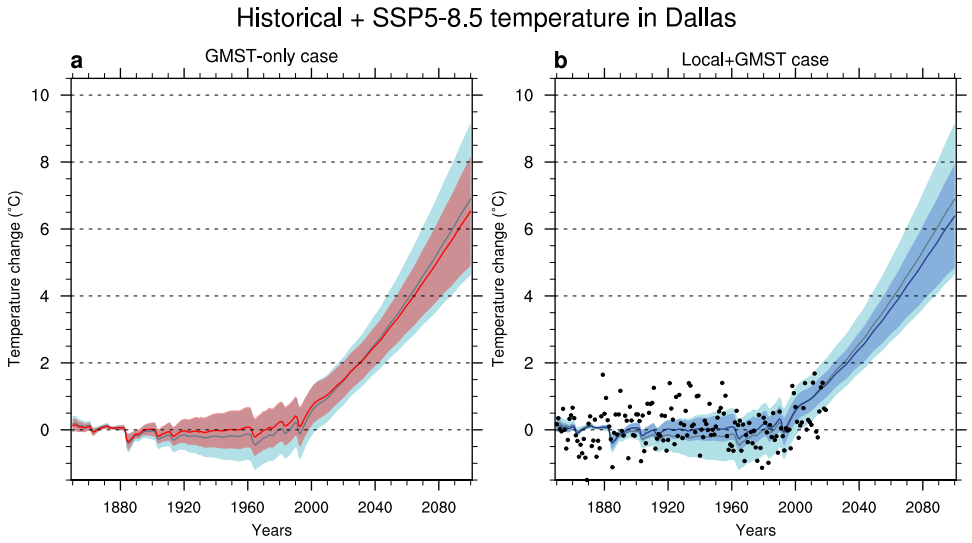

Fig. 2: **Observational constraints on historical and SSP5-8.5 local temperature changes in Dallas.** (a) Constrained local temperature for the grid point at [48.75 °N ; 11.25 °E] (see blue point 7 in Supplementary Figure 3) in the GMST-only case. The constrained (unconstrained) 5–95% spread of the simulated response to all external forcings is in red (light blue), the red (grey) line stands for the best estimate (ensemble mean). (b) Constrained local temperature in the Local+GMST case. The constrained (unconstrained) 5–95% spread of the simulated forced response is in blue (light blue), its best estimate (ensemble mean) is in dark blue (grey). Black points are the observations. Note that because single year observations are affected by internal variability, they often lie outside the 5–95% assessed ranges of the forced responses. Unconstrained simulated ranges are identical between panel a) and b). All values are anomalies with respect to the 1850–1900 period.

When the method is applied to any location worldwide, the results in the projected mean temperature and in model uncertainty depend on the level of correlation between the local temperature and GMST (Figure 3bd and 1b). The reduction of uncertainty in local projections is the highest at the locations where the correlation with GMST is the strongest. For these locations, e.g. over several continental regions, the North Pacific and the Indian Ocean, a reduction of the ensemble spread

of about 40% is obtained over the 2081-2100 period. In addition, the best estimate warming is revised downward between 0.5 °C to 1.5 °C (Figure 3ac). Conversely, for locations where the correlation is low, like in tropical Africa, the Barents Sea and the South Pacific gyre, the local temperature response is weakly constrained, with a reduction of the model uncertainty of 10% and a revision of the best estimate by 0.5 °C or less. These revised ranges lead to a warming pattern, at +2 °C of global warming, considerably different from the projections of the IPCC in its Special Report on Global Warming of 1.5 °C [27] (SR1.5) (Figure 4abd). For example, local temperatures over North America are expected to be 0.5 °warmer than in the unconstrained case under a global warming of 2 °C.

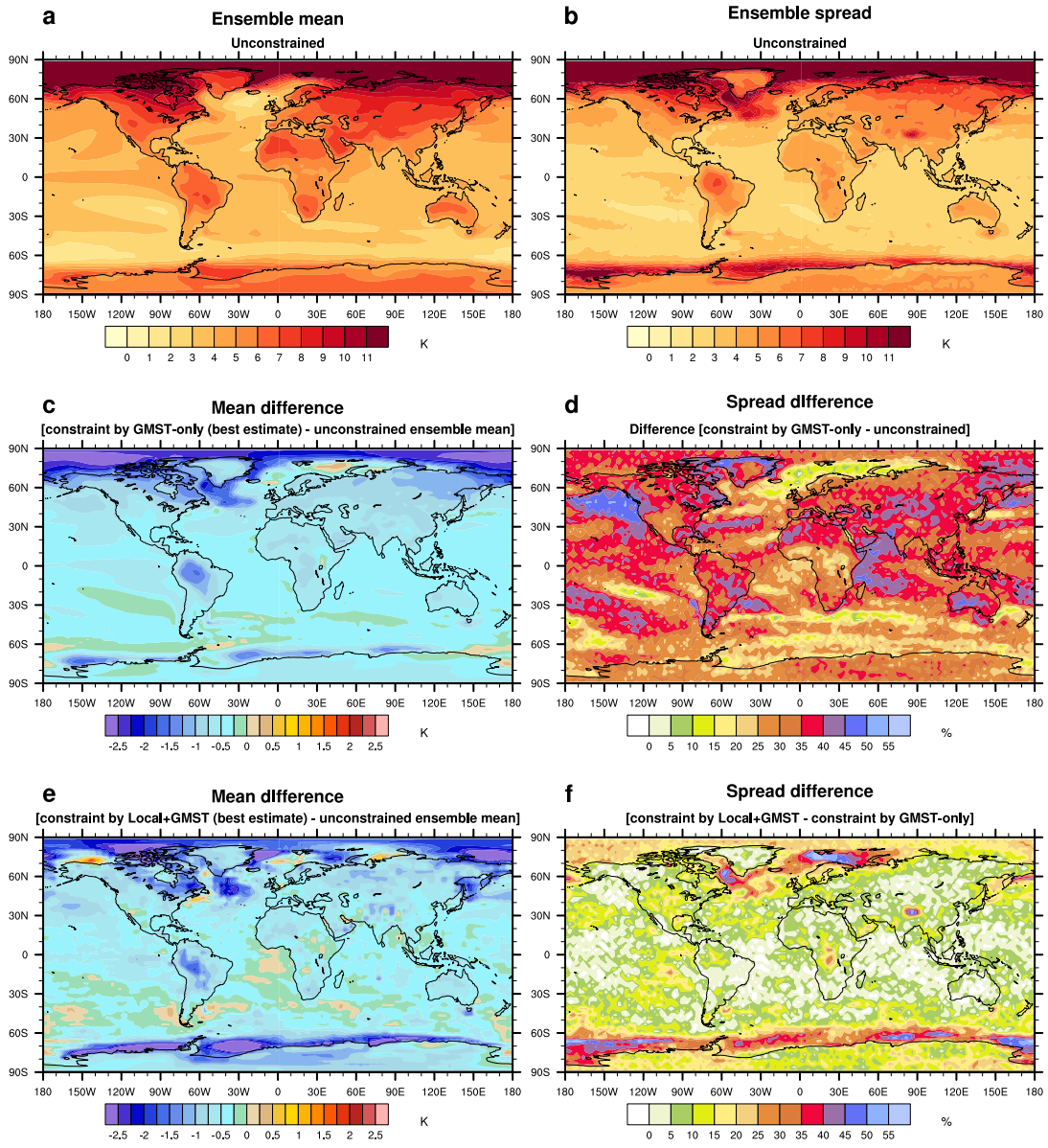

Fig. 3: **Unconstrained and constrained local temperature projections over the 2081-2100 period**. (a) Ensemble mean of the unconstrained local temperature changes. (b) Ensemble spread of the unconstrained local temperature changes, defined as the 5-95% confidence interval of the multimodel ensemble. (c) Difference of local temperature changes ensemble mean between the GMST-only case and the unconstrained case. (d) Relative difference of local temperature changes ensemble spread between the GMST-only case and the unconstrained case. (e) Same as (c) but for the Local+GMST case. (f) Relative difference of local temperature changes ensemble spread between the Local+GMST case and the GMST-only case, illustrating how much incorporating local observations narrows uncertainty. All values are anomalies with respect to the 1850–1900 period.

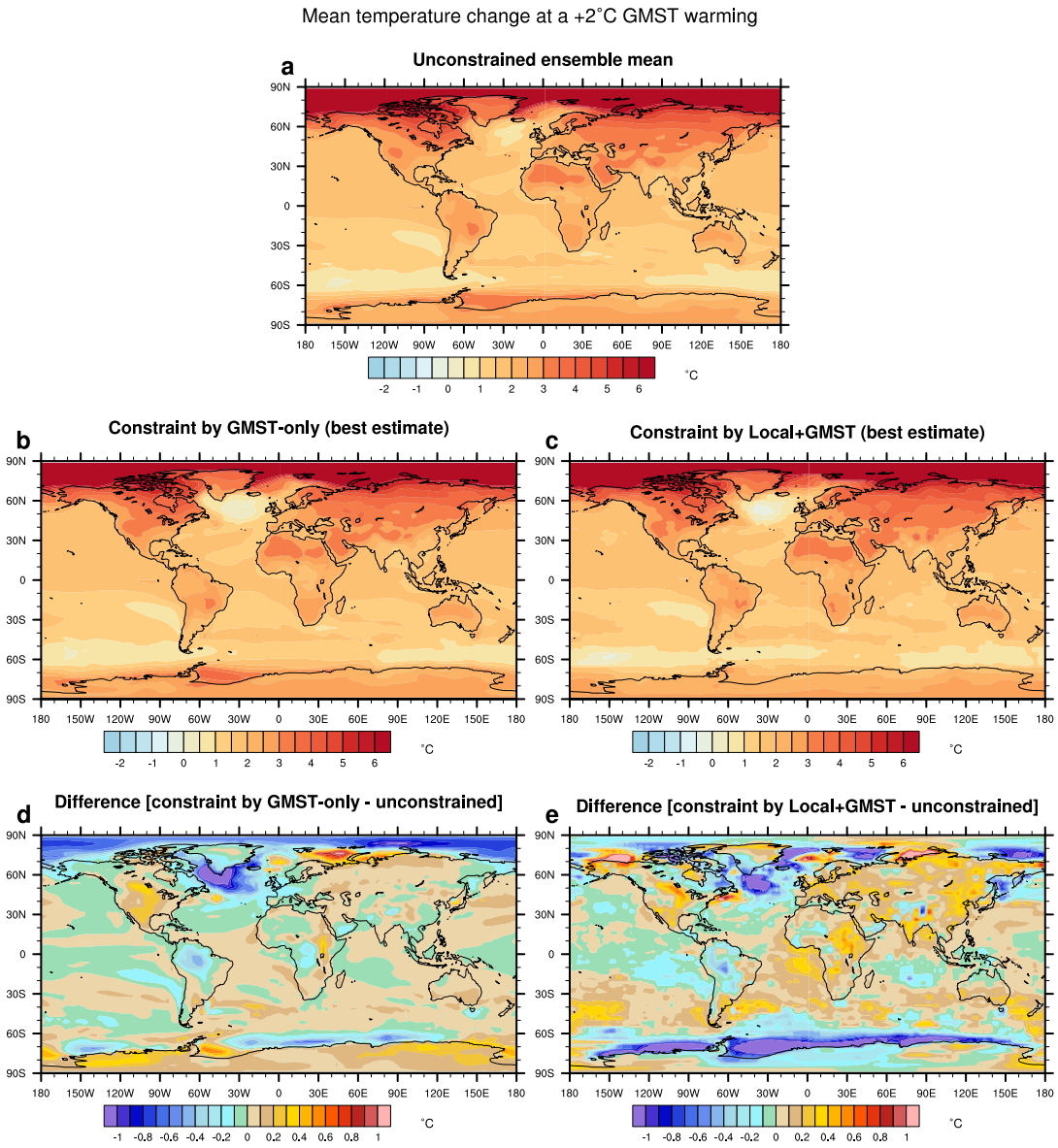

Fig. 4: **Warming pattern at +2 °C of GMST warming**. (a) Ensemble mean of the unconstrained local temperature changes. (b) Best estimate of the constrained local temperature changes in the GMST-only case. (c) Same as (b) but for the Local+GMST case. (d) Difference of local temperature changes between the GMST-only case (best estimate) and the unconstrained case (ensemble mean). (e) Difference of local temperature changes between the Local+GMST case (best estimate) and the unconstrained case (ensemble mean). All values in (a), (b) and (c) are anomalies with respect to the 1850-1900 period.

Added-value of local observations to the constraints

Beyond the useful information provided by the historical GMST time-series, one further very natural question involves the consistency between the expected local response (regardless whether observed GMST is accounted for) and local historical observations. Current and past warming is spatially heterogeneous, and some regions like the Arctic are warming faster than others [28] (Figure 4a). Evidence suggests that climate models have underestimated the Arctic temperature increase over the last few years [29]. This is in contradiction with the downward revision of the Beaufort Sea temperature change that is implied by the GMST-only constraint (Figure 3c). Therefore, it is very attractive to account for both GMST and local observations to provide local projections consistent with all available observations. Using recent local observations could particularly affect short term projections (typically over the 2020-2030 period) and could provide a different picture of the constrained temperature ranges.

In order to make such a calculation, we derive a posterior of the expected local warming given local historical observations in addition to the GMST observations (hereafter the Local+GMST case, see Methods). Following the example of Dallas considered in Figure 2a, the constrained local temperature ranges become more consistent with local observations, particularly over the 2000-2019 period (Figure 2b). Compared to the GMST-only case, the added value of local observations in the reduction of model uncertainty is limited in this example, with a decrease of about 10% of the confidence range width compared to the GMST-only case. Two reasons

contribute to this limited impact, and must be considered for any location. First, the local signal-to-noise ratio can be small. This may happen if local internal variability or measurement uncertainty is large (i.e., local observations provide little insight on the externally-forced response). Second, the global and local responses can be highly correlated with each other, so that they partly provide the same information, leading to a limited impact of local observations on uncertainty ranges. In both cases, the model uncertainty will be only marginally reduced by local historical data (see Methods).

The application of the Local+GMST constraint to all grid points worldwide results in a global mean projected warming which is about the same as the GMST-only case (Figure 3ce), but with regional differences. On the one hand, for several regions over the Arctic, especially the Beaufort and Barents Seas, the warming is revised upwards compared to the unconstrained case, making the projections more consistent with recent observations, and implying a much higher warming than predicted in the GMST-only case. On the other hand, a downward revision is slightly strengthened over Northern Central Asia, Eastern North America, the East Siberian Sea, and along the Antarctica coast. The added value of local observations in the reduction of model uncertainty is the largest over these regions where the correlation in Figure 1 is low (Figure 3f). Note that for both the GMST-only and the Local+GMST cases, the global mean of the constrained local ranges is very close to the constrained GMST ranges shown in Figure 1a (not shown). The addition of local information can also clearly modify the warming pattern at +2 °C of global warming (Figure 4ce). For example, while a downward revision of the temperature change of -0.2 °C

is obtained over Europe in the GMST-only case, an upward revision of 0.3 °C is obtained in the Local+GMST case. This change of sign is widespread over Eurasia. In the context of an urgent need of adaptation to the threat of climate change, our constrained warming pattern provides a revised and a more relevant information for local adaptation planning.

Evaluation of the constrained projections

The robustness of these promising results is quantified within a so-called perfect model framework, using a leave-one-out cross validation (see Methods). Each member of each model is considered as pseudo-observations over the 1850-2019 period. These are subsequently used to constrain the temperature projections, using all other models as a prior. The constrained temperature range is then compared to the warming simulated by the model from which pseudo-observations were taken. As making this evaluation for all the grid points is computationally expensive, this procedure is applied to several locations, considered as representative of the diversity of the worldwide climate (see Supplementary Figure 3). As for the real observations, we assess both the GMST-only and Local+GMST constraints. The continuous ranked probability skill score (CRPSS) [30] is used to measure the accuracy of the method, taking the unconstrained projections as a baseline as a first step (see Methods).

Figure 5 shows that the median of the CRPSS distribution based on all pseudo-observations and locations is positive in the GMST-only case with an improvement of about 30% over the 2081-2100 period. Depending on the location, the skill is

remarkably improved by 10% to 40%, except for one location (out of 55) where it comes close to 0 (Supplementary Figs. 4 to 8). In the Local+GMST case, the skill is also always positive, and lies between 10% and 50 % relative to the unconstrained case. These results clearly demonstrate the performance of the method. Moreover, the comparison between the Local+GMST and the GMST-only constraints indicates that the skill is slightly improved when adding local observations to constrain projections. Over the 55 tested locations, the CRPSS is positive for 35 locations, and is slightly negative for the remaining points. The significance of this result is assessed with a binomial test. Under the null hypothesis that adding local observations has no impact on the skill (i.e. that the GMST-only and Local+GMST cases are categories equally likely, such as a coin toss), the probability of getting this result by pure chance is 2.9%. This suggests that there is a slight but significant added value in considering the constrained ranges derived from the Local+GMST case relative to the GMST-only case. A third case for which we only use local observations (Local-only case) to constrain projections indicates lower scores than in the GMST-only and Local+GMST cases (Figure 5), and confirms that using the combination of global and local observations enhances the accuracy of the method.

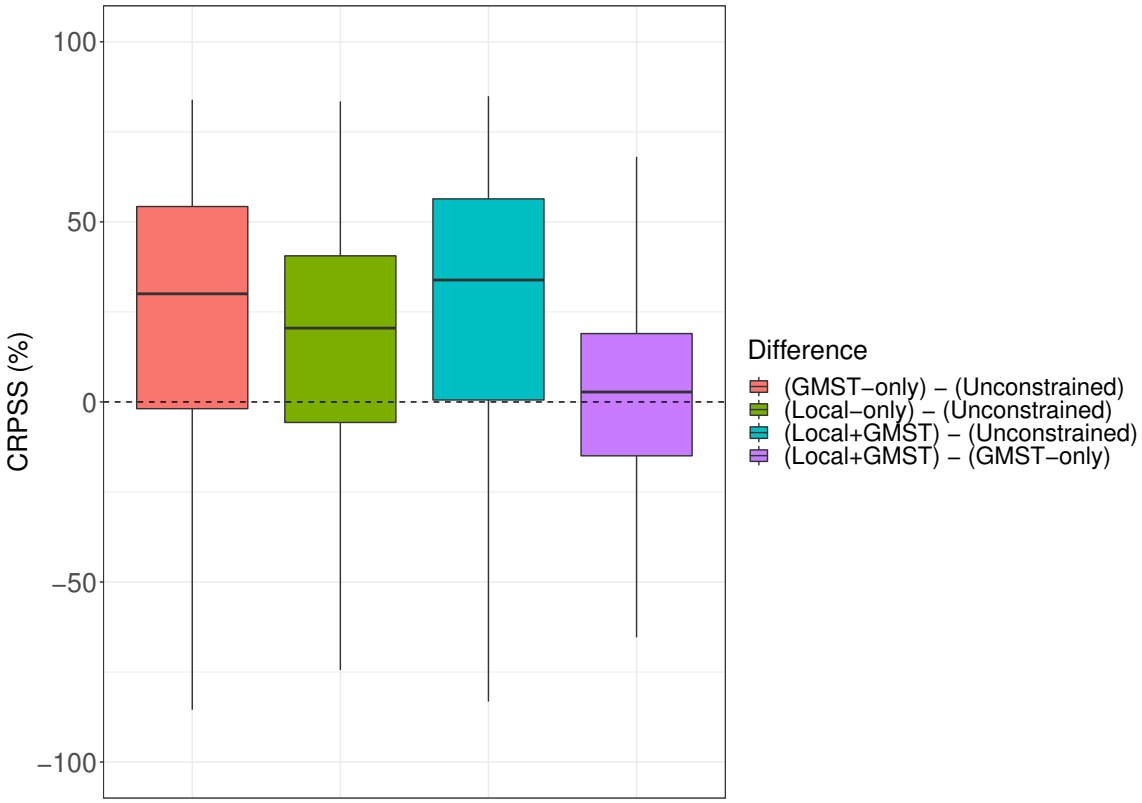

Fig. 5: **CRPSS for the constrained temperature projections within the perfect model framework.** The red, green and blue boxplots indicate the CRPSS distributions of the GMST-only, Local-only and Local+GMST constraints, respectively, compared to the unconstrained case. Calculation is made for all locations (see blue points in Supplementary Figure 3) and models, over the 2081-2100 period. The boxplot in magenta indicates the added value of the local observations and stand for the CRPSS distributions in the Local+GMST case compared to the GMST-only case. The top (bottom) of the box represents the 25th (75th) percentile of the distribution and the upper (lower) whisker represents the 95th (5th) percentile. Values are normalised by the number of members in each model. A CRPSS of 0 (dashed line) indicates the absence of added value of the method.

209 Note that models with large low-frequency variability tend to pull down the CRPSS

values (Supplementary Discussion). A second evaluation criterion of the method based on coverage probabilities leads to similar conclusions (Supplementary Discussion).

From all of these evaluation results, we retain the Local+GMST case to provide guidance in constraining local projections. The evaluation of the KCC method suggests that the constrained temperature ranges are reliable, and demonstrate that relying on unconstrained projections to assess the local future climate is no longer the best approach.

Implications for local pathways in a 2 °C warmer world

We have shown, using a statistical method combining the entire temperature observation records with model simulations, that uncertainty in local temperature projections can be substantially narrowed. Local projections constrained by both global and local observations exhibit a reduction of the uncertainty of 40% in average by 2100. This demonstrates the benefits of merging model simulations with observations to provide the best picture of future climate change. Figure 6 offers a complementary perspective to the IPCC SR1.5 [27] conclusions that were solely based on raw (unconstrained) projections. For each location, a temporal evolution from 1850 to 2100 of the constrained temperature and its uncertainty can be derived, with revised projections for the near and the long term time scales. This provides a considerable revision of the local exposure to the consequences of the on-going climate change [31]. An

online tool that implements the method and illustrates the constrained temperature ranges for every point over a horizontal grid of 2.5 degree resolution is available via the following demonstrator: `https://saidqasmi.shinyapps.io/KCC-shinyapp/`.

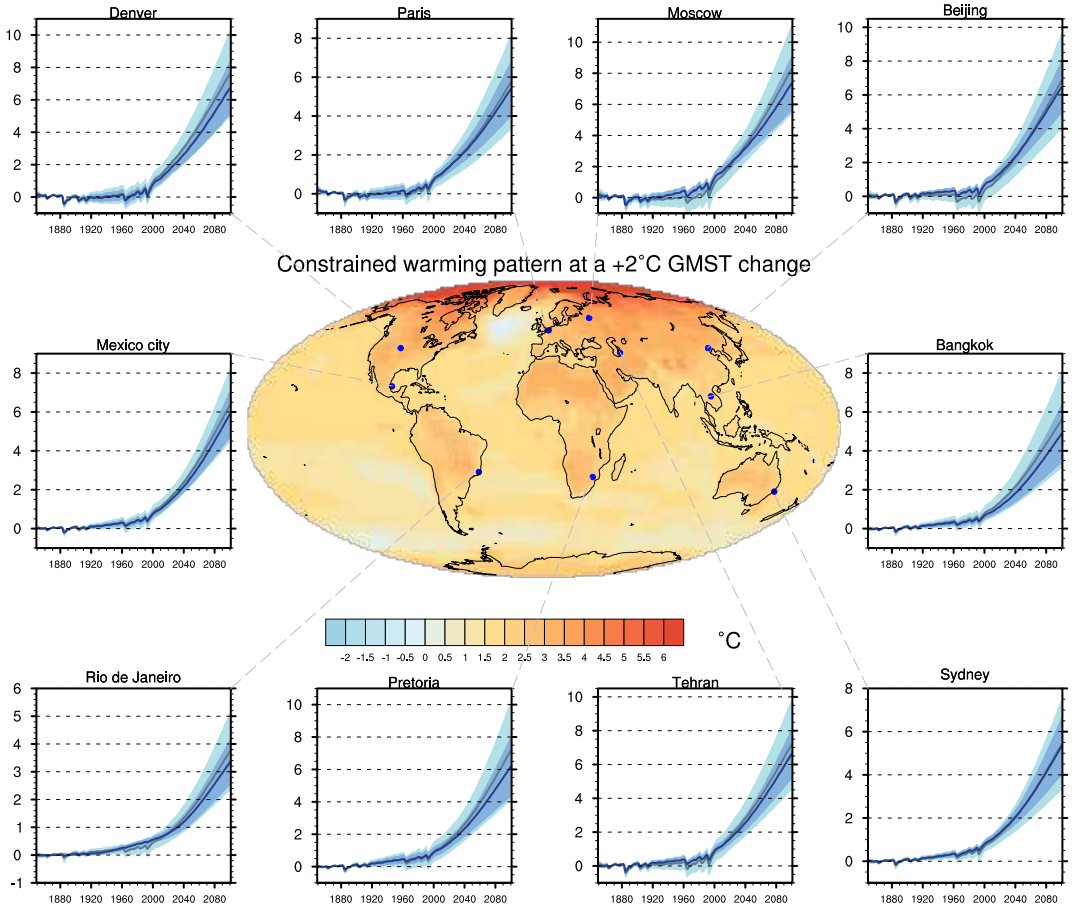

Fig. 6: **Mean temperature change at a +2 °C GMST warming**. Best estimate of the constrained local temperature changes in the Local+GMST case. Similarly to Figure 2b, the constrained and unconstrained temperature ranges are shown for several world capitals cities over the 1850-2100 period. All values are anomalies with respect to the 1850–1900 period.

Promising prospects exist to improve the constrained projections. CMIP5 [32] and

CMIP6 [8] ensembles sample model uncertainty in a probabilistic way by using all climate models as an "ensemble of opportunities" [33, 34]. This approach, in which our study fits, has several limitations that can bias the estimation of climate uncertainty [35]. One of them is that each model output is considered as independent and contributes equally to the multi-model ensemble. This "model democracy" paradigm has been largely used to summarise projection information in IPCC assessment reports [7], even though it can be criticised [36]. Therefore, using a subset of models qualified as independent a priori, or weighting the models in this way [37, 38], before applying the observational constraint, may provide even more reliable results.

Our results demonstrate that available observations offer valuable information to sharpen climate projections. As the climate system will continue to change over the next decades, observations will further constrain the local and the global responses to the increasing greenhouse effect. Therefore, it is critical to account for this new source of information and to regularly bridge the gap between monitoring recent changes and predicting future changes. This is particularly important as the spread among global climate models has not decreased over the last CMIP generations. With our new projection ranges, the storyline approach [39, 40, 41], which is increasingly adopted in the climate risk management, could be refined.

The KCC method itself can also be improved. Although it can be used on larger areas to easily derive constrained projections, eg on the SREX regions [16], the current implementation does not take into account the spatial dependence in the climate variability between locations. Taking the spatial dimension fully into account could bring additional useful information and would result in consistent uncertainties at

all spatial scales. In addition, generalising the method to other variables of high societal impacts, e.g. extreme precipitation, droughts, snow cover, some of which are also tightly related to GMST changes, would also be very relevant. In this way the climate science community could take a step forward towards a more accurate assessment of past and future human-induced climate change.

Methods

Observational dataset and models

The temperature observations are from the Cowtan and Way [26] (hereafter CW) dataset which complements the HadCRUT4 [42] data by filling in the missing data over the 1850-2019 period. The temperature field comes from a blending of near-surface temperature and sea surface temperature using land sea mask and sea ice concentration. The measurement uncertainty of the CW dataset is estimated from a ensemble of 100 equiprobable realisations. Most of the other observational products are included in the temperature range estimated by this ensemble, which confirms our choice to consider the CW dataset as a reference.

CMIP6 models are selected according to the availability of the following data: at least 200 years of pre-industrial control simulation; at least one member of a historical simulation and one member of a projection simulation for the SSP5-8.5 scenario. In order to constrain the simulated temperatures at a grid point scale in a consistent way, a blended temperature field T_{blend} is computed in each CMIP6 model based on the formulation of Cowtan et al. [24]:

$$\begin{cases} w_{\text{air}} & = (1 - f_{\text{ocean}}) + f_{\text{ocean}} f_{\text{ice}}, \\ T_{\text{blend}} & = w_{\text{air}} T_{\text{air}} + (1 - w_{\text{air}}) T_{\text{ocean}}, \end{cases} \tag{1}$$

where T_{air}, T_{ocean}, f_{ice}, and f_{ocean} are for each grid point near-surface air temperature, sea surface temperature, sea ice concentration and sea area fraction. The 27 models for which these variables are available and which satisfy the above criteria are listed in Table 1.

We define the global surface air temperature (GSAT) as the global mean of T_{air} and the global mean surface temperature (GMST) as the global mean of T_{blend}. Several studies have shown that GMST and GSAT significantly differ as GMST warms significantly less than GSAT [24, 25].

Models are interpolated on a common horizontal grid of 2.5 ° resolution before calculating blended temperatures and applying the constraining method. This choice is motivated by a compromise between the different resolutions of the CMIP6 models (between 1.5 ° and 2.5 °). Note that the KCC method can be applied to finer resolutions if observations are available at this scale. For temperature, for which the spatial autocorrelation is high, the reduction in the uncertainty is expected to be the similar as for the 2.5 ° resolution.

Model \ Simulation	historical + SSP5-8.5
ACCESS-CM2	1
ACCESS-ESM1-5	3
BCC-CSM2-MR	1
CanESM5-CanOE	3
CanESM5	50
CESM2	2
CESM2-WACCM	1
CNRM-CM6-1	6
CNRM-CM6-1-HR	1
CNRM-ESM2-1	5
EC-Earth3	7
EC-Earth3-Veg	3
FGOALS-g3	1
FIO-ESM-2-0	3
GFDL-ESM4	1
HadGEM3-GC31-LL	1
INM-CM4-8	1
INM-CM5-0	1
IPSL-CM6A-LR	6
MIROC6	3
MIROC-ES2L	1
MPI-ESM1-2-HR	2
MPI-ESM1-2-LR	10
MRI-ESM2-0	1
NESM3	1
NorESM2-MM	1
UKESM1-0-LL	5
27 models	121 members

Table 1: **List of the available CMIP6 Models and the associated number of members in the historical and SSP5-8.5 simulations used to constrain temperature projections.**

Statistical Method

The statistical method is based on the same one used by Ribes et al. [11], whose formulation and principle is similar to kriging, which is a method originally developed to interpolate geophysical data based on prior covariances. In Ribes et al., this method is applied to the analysis of time series from climate simulations of CMIP5 and CMIP6 models, and is used for several purposes: (i) reducing model uncertainty on past and future global warming estimated by CMIP and ScenarioMIP [9] simulations, (ii) reducing uncertainty on warming attributed to several external forcings via DAMIP [43] models, (iii) complete or statistically reconstruct missing simulations from other physically relevant simulations (e.g. using the so-called 1%CO_2 simulations, in which the CO_2 concentration increases by 1% each year, to reconstruct DAMIP historical simulations in which GHGs follow their historical concentrations, while other forcings are kept constant). Here we apply this method of Kriging for Climate Change (KCC) to reduce the model uncertainty in the past and future temperature changes simulated by CMIP6 models at each grid point. Note that a confusion can be made with techniques based on so-called emergent constraints methods [44, 45]. Emergent constraints would usually consider the sole observed global warming trend (a single scalar); e.g. over the 1980-2019 period, to constrain the simulated temperature changes in the future. The KCC method has several advantages compared to this approach. Instead of simply constraining a trend over a sub-period, it uses the entire observed time series of temperature, which avoids ignoring useful information. In addition, the method takes into account the model temporal pattern uncertainty and provides confidence ranges specifically for

the forced response, while many other studies also include internal variability.

For a given grid point, we define $\boldsymbol{y}^*_{\text{loc}}$ as the yearly time series of the real (and unknown) temperature response to external forcings over the 1850-2019 period, and $\boldsymbol{y}_{\text{loc}}$ as the observed yearly temperature time series over the same period. Similarly, we define for the GMST the vectors $\boldsymbol{y}^*_{\text{glo}}$ and $\boldsymbol{y}_{\text{glo}}$ as the unknown response to external forcings and the observed time series, respectively. They constitute the following $\boldsymbol{y}^*$ and $\boldsymbol{y}$ vectors, both of size $2n_y$ where $n_y = 170$:

$$\boldsymbol{y}^* = \begin{pmatrix} \boldsymbol{y}^*_{\text{loc}} \\ \boldsymbol{y}^*_{\text{glo}} \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_{\text{loc}} \\ \boldsymbol{y}_{\text{glo}} \end{pmatrix}. \tag{2}$$

Assuming that the observed temperature total variability can be decomposed as the sum of a term of forced variability and a term including both internal variability and measurement errors, $\boldsymbol{y}$ takes the following form:

$$\boldsymbol{y} = \boldsymbol{y}^* + \boldsymbol{\epsilon}, \tag{3}$$

where $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_{\text{loc}}, \boldsymbol{\epsilon}_{\text{glo}})$ is a vector of size $2n_y$, and corresponds to the local and global terms of measurement errors and observed internal variability. Further assuming that models are indistinguishable from the truth, *i.e.*, that observations and models are exchangeable [46, 47, 48], observations $\boldsymbol{y}$ can be rewritten:

$$\begin{cases} \boldsymbol{x} &= \begin{pmatrix} \boldsymbol{x}_{\text{loc}} \\ \\ \boldsymbol{x}_{\text{glo}} \end{pmatrix}, \\ \\ \boldsymbol{y} &= \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\epsilon}, \end{cases} \tag{4}$$

where $\boldsymbol{x}_{\text{loc}}$ and $\boldsymbol{x}_{\text{glo}}$ are the yearly time series over the 1850-2100 period of the local and global temperature responses to external forcings estimated in CMIP6 models, respectively, i.e. vectors of size $n_x = 251$. $\boldsymbol{H}$ is an observation operator of size $2n_y \times 2n_x$, which extracts the part of $\boldsymbol{x}$ that is observed in $\boldsymbol{y}$, *i.e.*, the forced response from 1850 to 2019, and whose form depends on the type of the applied constraint (using only GMST observations or both GMST and local observations, see equation 21 in Supplementary Information). Note that the assumption of exchangeability between observations and models has been suggested as well supported by observations, especially for temperature [46, 49].

For a given CMIP6 model m listed in Table 1, we choose to estimate the simulated response to all external forcings $\boldsymbol{x}_{m,\text{glo}}$, by decomposing the simulated GMST over 1850-2100 into an anthropogenic response $\boldsymbol{x}_{m,\text{ant,glo}}$, and a natural response $\boldsymbol{x}_{m,\text{nat,glo}}$. Therefore, after averaging all available members of the model m, the simulated GMST time series over 1850-2100 $\boldsymbol{x}_{m,\text{glo}}$ writes:

$$\boldsymbol{x}_{m,\text{glo}} = \boldsymbol{x}_{m,\text{nat,glo}} + \boldsymbol{x}_{m,\text{ant,glo}} + \boldsymbol{\epsilon}_m, \tag{5}$$

where $\boldsymbol{\epsilon}_m$ is a random term for internal variability.

To estimate $\boldsymbol{x}_{m,\text{nat,glo}}$ and $\boldsymbol{x}_{m,\text{ant,glo}}$ in the model m, we use a Generalized Additive Model (GAM) to compute the response to all external forcings, $\boldsymbol{x}_{m,\text{glo}}$ (recall that $\boldsymbol{x}_{m,\text{glo}}$ follows equation 4):

$$\begin{cases} \boldsymbol{x}_{m,\text{glo}} & = \boldsymbol{x}_{m,\text{all,glo}} + \boldsymbol{\epsilon}_m, \\ \boldsymbol{x}_{m,\text{all,glo}} & = \underbrace{\beta_m \boldsymbol{e}}_{\boldsymbol{x}_{m,\text{nat,glo}}} + \underbrace{f(\boldsymbol{t})}_{\boldsymbol{x}_{m,\text{ant,glo}}}, \end{cases} \tag{6}$$

where β_m is an unknown scaling factor. $\boldsymbol{e}$ is the multi-model mean of the temperature response from the EBM using the EBM parameters from Geoffroy et al [50], and is a vector of size n_x. $f(\boldsymbol{t})$ is a time series (with $\boldsymbol{t} = (1850, \ldots, 2100)$) and refers to an assumed smoothed response of GMST to the anthropogenic forcings. The function f corresponds to smoothing splines to filter out part of internal variability, with 6 degrees of freedom (a value which was selected as a bias-variance trade-off).

We apply the exact same procedure to estimate the local forced responses as simulated by each CMIP model. For each grid point from the model m, we consider $\boldsymbol{x}_{m,\text{loc}}$, the average of all available members, to estimate the local forced response, $\boldsymbol{x}_{m,\text{all,loc}}$.

We assume that the local natural response scales linearly with the globally-averaged natural forcings time series, as the EBM response $\boldsymbol{e}$ used to calculate $\boldsymbol{x}_{m,\mathrm{nat,glo}}$ is also used when we fit the GAM to compute the local natural response $\boldsymbol{x}_{m,\mathrm{nat,loc}}$. Thus $\boldsymbol{x}_{m,\mathrm{nat,glo}}$ and $\boldsymbol{x}_{m,\mathrm{nat,loc}}$ only differ by their scaling factor β_m. We believe that our results are not sensitive to this choice given the reduced strength of, and uncertainty in, the natural response compared to the anthropogenic response.

The multimodel ensemble of the local and global simulated responses to all external forcings is used to derive a distribution of $\boldsymbol{x}$, noted $\Pi(\boldsymbol{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathrm{mod}})$, built from all $\boldsymbol{x}_{m,\mathrm{glo}}$ and $\boldsymbol{x}_{m,\mathrm{loc}}$. $\boldsymbol{\mu} = (\boldsymbol{\mu}_{\mathrm{loc}}, \boldsymbol{\mu}_{\mathrm{glo}})$ is a vector of size $2n_x$ and is the multimodel ensemble mean of the concatenated local and global forced responses. $\boldsymbol{\Sigma}_{\mathrm{mod}}$ is a variance-covariance matrix of size $2n_x \times 2n_x$ that describes the model spread, with the following form:

$$\boldsymbol{\Sigma}_{\mathrm{mod}} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{\mathrm{mod,loc}} & \boldsymbol{\Sigma}_{\mathrm{mod,dep}} \\ \hline \boldsymbol{\Sigma}'_{\mathrm{mod,dep}} & \boldsymbol{\Sigma}_{\mathrm{mod,glo}} \end{array}\right], \tag{7}$$

where $\boldsymbol{\Sigma}_{\mathrm{mod,loc}}$ and $\boldsymbol{\Sigma}_{\mathrm{mod,glo}}$ are the sample covariance matrices of size $n_x \times n_x$ modelling local and global model spread within $\boldsymbol{x}_{\mathrm{loc}}$ and $\boldsymbol{x}_{\mathrm{glo}}$, respectively. $\boldsymbol{\Sigma}_{\mathrm{mod,dep}}$ is the covariance matrix modelling the dependence between $\boldsymbol{x}_{\mathrm{loc}}$ and $\boldsymbol{x}_{\mathrm{glo}}$.

In our Bayesian framework, $\Pi(\boldsymbol{x})$ is a first (probabilistic) estimate of $\boldsymbol{x}$, which makes no use of observations, and is only based on climate models. We want to update this estimate by incorporating the observational evidence provided by $\boldsymbol{y}$. Following the

Bayesian theory, the calculation of the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$ is required. A pre-requisite is to define the observational uncertainty, i.e., the covariance matrix associated with $\boldsymbol{y}$.

Modelling of observational uncertainty

Given equation 4, we assume that $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\text{obs}})$, where $\boldsymbol{\Sigma}_{\text{obs}} = \boldsymbol{\Sigma}_{\text{meas.}} + \boldsymbol{\Sigma}_{\text{iv}}$ is the observation error covariance matrix. $\boldsymbol{\Sigma}_{\text{meas.}}$ and $\boldsymbol{\Sigma}_{\text{iv}}$ are both of size $2n_y \times 2n_y$ and describe the measurement error and internal variability, respectively. $\boldsymbol{\Sigma}_{\text{meas.}}$ is estimated as the sample covariance matrix over the 100-member ensemble of the CW dataset [26].

$\boldsymbol{\Sigma}_{\text{iv}}$ is estimated by using observed annual time series of global and local temperature over the 1850-2019 period. First, we compute the global observational residuals by subtracting the CMIP6 response to all external forcings $\boldsymbol{\mu}_{\text{glo}}(1, \ldots, n_y)$ to the observations $\boldsymbol{y}_{\text{glo}}$. Similarly, we derive local residuals by subtracting $\boldsymbol{\mu}_{\text{loc}}(1, \ldots, n_y)$ to $\boldsymbol{y}_{\text{loc}}$. These residuals constitute an estimate of global and local internal variability.

We define $\boldsymbol{\Sigma}_{\text{iv}}$ as a matrix of size $2n_y \times 2n_y$ of the following form:

$$\boldsymbol{\Sigma}_{\text{iv}} = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{\text{iv,loc}} & \boldsymbol{\Sigma}_{\text{iv,dep}} \\ \hline \boldsymbol{\Sigma}'_{\text{iv,dep}} & \boldsymbol{\Sigma}_{\text{iv,glo}} \end{array}\right], \tag{8}$$

where $\boldsymbol{\Sigma}_{\text{iv,loc}}$ and $\boldsymbol{\Sigma}_{\text{iv,glo}}$ are the covariance matrices of size $n_y \times n_y$ modelling local and global internal variability within $\boldsymbol{y}_{\text{loc}}$ and $\boldsymbol{y}_{\text{glo}}$, respectively. $\boldsymbol{\Sigma}_{\text{iv,dep}}$ is the covariance

matrix modelling the dependence between global and local temperature.

To compute $\mathbf{\Sigma}_{\text{iv}}$, we take into account decadal internal variability that exists in the global [51], regional [52], and even local [53] observations, by using a mixture of two autoregressive processes or order 1 (AR1), hereafter MAR, as done by Ribes et al. [11]. The MAR formulation includes a fast (f) and a slow (s) components such that global internal variability $\boldsymbol{\epsilon}_{\text{iv,glo}}$ within the GMST residuals writes at a time t:

$$\begin{cases} \epsilon_{\text{iv,glo}}(t) & = \epsilon_{\text{iv,f,glo}}(t) + \epsilon_{\text{iv,s,glo}}(t), \\ \epsilon_{\text{iv,f,glo}}(t) & = \alpha_{\text{f,glo}}\epsilon_{\text{iv,f,glo}}(t-1) + Z_{\text{f,glo}}(t), \\ \epsilon_{\text{iv,s,glo}}(t) & = \alpha_{\text{s,glo}}\epsilon_{\text{iv,s,glo}}(t-1) + Z_{\text{s,glo}}(t), \end{cases} \tag{9}$$

where the parameters $\alpha_{\text{s,glo}}$ and $\alpha_{\text{f,glo}}$ are the lag 1 coefficients of the AR1 processes, and $\alpha_{\text{s,glo}} \geq \alpha_{\text{f,glo}}$ by convention. $Z_{\text{s,glo}}(t) \sim \mathcal{N}(0, \sigma^2_{\text{s,glo}})$ and $Z_{\text{f,glo}}(t) \sim \mathcal{N}(0, \sigma^2_{\text{f,glo}})$ are white noises associated with the two AR1. The slow component is able to generate a dependence on time scales of typically one decade, while the fast component accounts for interannual variability. Following the principle of parsimony, only 4 coefficients $(\sigma^2_{\text{f,glo}}, \alpha_{\text{f,glo}}, \sigma^2_{\text{s,glo}}, \alpha_{\text{s,glo}})$ are thus needed to characterize internal variability at the global scale. We fill the covariance matrix $\mathbf{\Sigma}_{\text{iv,glo}}$ following the calculations of each of its coefficients, as detailed in Supplementary Equation 8. In practice, we apply a maximum likelihood procedure to the local and global residuals according to the statistical model from equation 9. Uncertainty related to these coefficients is not taken into account. Then, we make the same assumptions, and estimate 4 other

parameters, $(\sigma^2_{\mathrm{f,loc}}, \alpha_{\mathrm{f,loc}}, \sigma^2_{\mathrm{s,loc}}, \alpha_{\mathrm{s,loc}})$, to characterize fast and slow components in local internal variability $\boldsymbol{\epsilon}_{\mathrm{iv,loc}}$ and to compute $\boldsymbol{\Sigma}_{\mathrm{iv,loc}}$. The autocorrelations from this MAR model suggest that our statistical representation of internal variability effectively captures decadal variability (typically between lag 5 and lag 10) in the GMST and local temperature time series, e.g., for the Atlantic, African and South American regions (Fig. 7). We are aware that initial condition large ensembles and long pre-industrial control (piControl) simulations provide a nice sampling of internal variability, and could also be used to estimate this variability. However, we choose to not directly rely on it because of the huge discrepancies between models in terms of their simulated internal variability [51]. Supplementary Figs. 16 to 26 illustrate this aspect with the piControl simulations from the CMIP6 models, including those used to build large ensembles. In all cases, the models do not converge to a consistent estimate of internal variability. For instance, over the Atlantic ocean, many models exhibit clear pseudo-periodic low frequency variability, while other models do not simulate decadal variability.

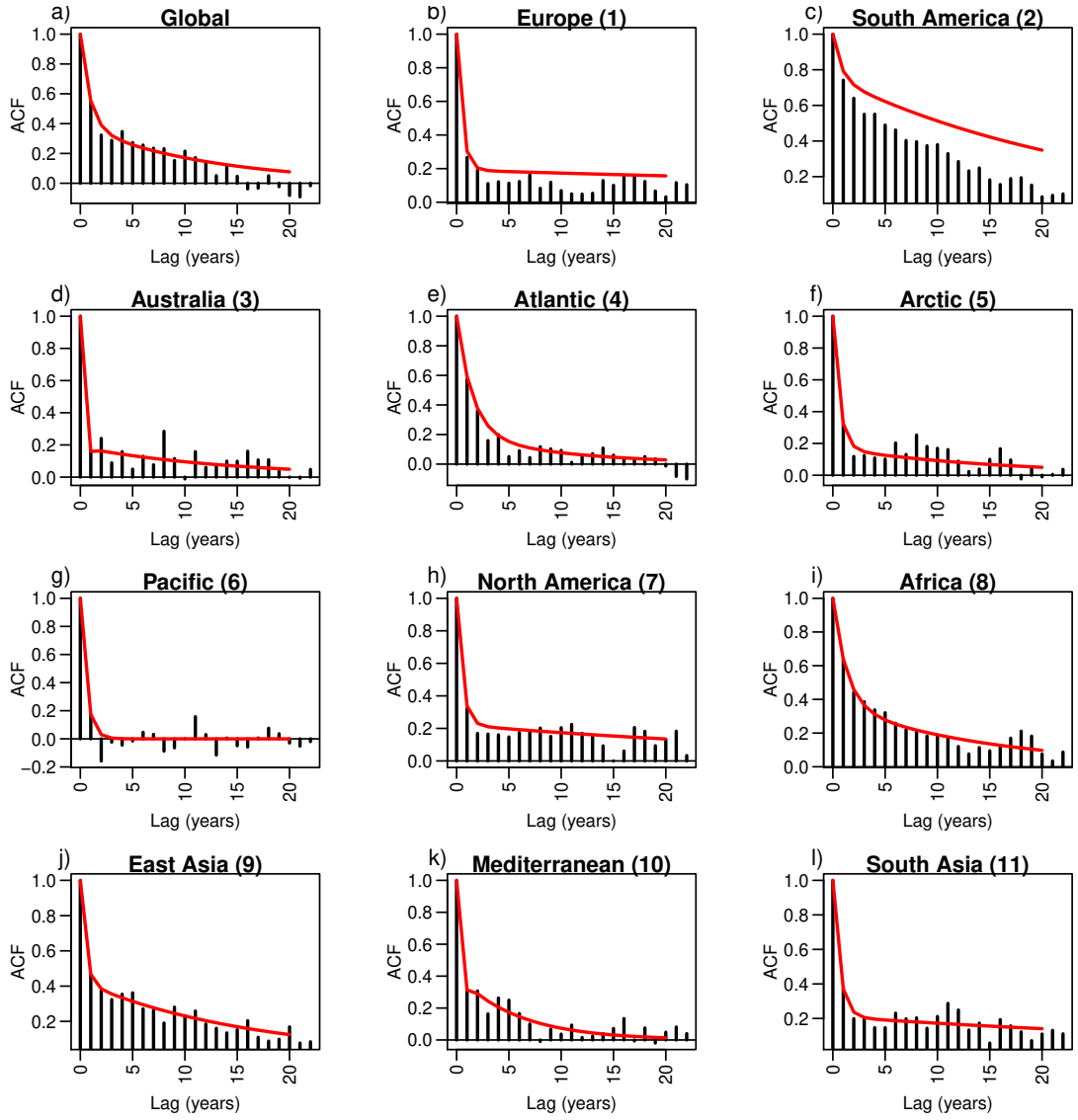

Fig. 7: **Estimation of observed internal variability.** (a) Autocorrelation function (ACF) of the residuals from GMST observations (bars) after subtracting the CMIP6 multimodel mean forced response to the observed time-series. The ACF of the MAR processes fitted to the same residual time series according to equation 9 is in red. (b-l) Same as (a) but for local residuals from 11 selected points (see Supplementary Fig. 3).

Modelling of the dependence between local and global internal variability

As impacts from Pacific and Atlantic decadal variability (and potential other modes of variability) on GMST have been reported over the historical period [54, 55], we need to allow a potential dependence between global and local internal variability in $\mathbf{\Sigma}_{\mathrm{iv,dep}}$. Therefore, finding a simple and parsimonious dependence model that is compatible with the MAR structure is required. Allowing the covariances $\mathrm{Cov}[\epsilon_{\mathrm{s,glo}}(t), \epsilon_{\mathrm{s,loc}}(t)]$ and $\mathrm{Cov}[\epsilon_{\mathrm{f,glo}}(t), \epsilon_{\mathrm{f,loc}}(t)]$ to be non zero is not trivial, and these terms need to be quantified to fill the covariance matrix $\mathbf{\Sigma}_{\mathrm{iv,dep}}$. Note that the fast and slow components remain always independent, and that $\mathbf{\Sigma}_{\mathrm{iv}}$ is computed for each location separately, as the spatial dependence among various locations is not considered in the method. To compute $\mathbf{\Sigma}_{\mathrm{iv,dep}}$, we introduce a nineth parameter λ accounting for some correlation between the local vs global components in the MAR modelling. The formulation of the covariances is slightly different in this case, and the calculations are detailed in the Supplementary Information.

Calculation of $p(\boldsymbol{x}|\boldsymbol{y})$

As $\Pi(\boldsymbol{x})$ and $\boldsymbol{\epsilon}$ are assumed to follow normal distributions, the Gaussian conditioning theorem is applicable to derive the posterior, or the "constrained" distribution $p(\boldsymbol{x}|\boldsymbol{y})$. Its formulation (detailed in equation 23 in the Supplementary Information) indicates that the method is conservative: the uncertainty in $p(\boldsymbol{x}|\boldsymbol{y})$ is never larger than that in $\Pi(\boldsymbol{x})$. Therefore, if observed internal variability is very large, then the model uncertainty in $p(\boldsymbol{x}|\boldsymbol{y})$ will remain very close to that in $\Pi(\boldsymbol{x})$.

Perfect model evaluation

We evaluate the performance of the KCC method within a perfect model framework, following a leave-one-out cross-validation:

1. For a given model, we consider a single member as pseudo-observations $\boldsymbol{y}$ over the 1850–2019 period (the historical simulation is extended by the SSP5-8.5 simulation over the 2015-2019 period).

2. We use the other 26 models to derive the prior $\Pi(\boldsymbol{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\text{mod}})$.

3. As there is no measurement uncertainty in models, $\boldsymbol{\Sigma}_{\text{meas.}}$ is null, therefore $\boldsymbol{\Sigma}_{\text{obs}} = \boldsymbol{\Sigma}_{\text{iv}}$. As done with the real observations, internal variability within the pseudo-observations is estimated from the difference between the pseudo-observations time series and the forced temperature response estimated by the ensemble mean of the 26 other models. $\boldsymbol{\Sigma}_{\text{iv}}$ is then derived from the MAR fitted on the obtained residuals.

4. We apply the KCC method using the inputs $\boldsymbol{y}$, $\boldsymbol{\Sigma}_{\text{obs}}$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}_{\text{mod}}$ to calculate projected changes constrained by pseudo-observations.

5. These 4 steps are repeated for each available member of the considered model, and for all available models.

Continuous Ranked Probability Score

We use the continuous ranked probability score [30] (CRPS) to quantify the performance of the KCC method. It is defined as the quadratic measure of discrepancy

between (i) $\mathbb{1}(x \geq y_{\text{pobs}})$, the empirical cumulative distribution function (CDF) of a scalar pseudo-observation y_{pobs} simulated by one model and averaged over the 2081-2100 period, and (ii) the projected CDF G_{cons} of $p(\boldsymbol{x}|\boldsymbol{y})$ (derived from all of the other models) over the same period:

$$\text{CRPS}_{\text{cons}}(G_{\text{cons}}, y_{\text{pobs}}) = \int_{\mathbb{R}} [G_{\text{cons}}(x) - \mathbb{1}(x \geq y_{\text{pobs}})]^2 dx, \tag{10}$$

where $\mathbb{1}$ is the indicator function (note that x is here a bound variable in the integral, different from the vector $\boldsymbol{x}$ in equation 4). Similarly, we define a reference CRPS, CRPS_{ref} based on G_{ref}, the CDF of $\Pi(\boldsymbol{x})$, the unconstrained distribution, and y_{pobs}. We can compute the continuous ranked probability skill score (CRPSS), which quantifies the performance of the KCC method if compared to the reference:

$$\text{CRPSS} = 1 - \frac{\text{CRPS}_{\text{cons}}}{\text{CRPS}_{\text{ref}}} \tag{11}$$

The CRPSS is computed over all available pseudo-observations (121 values, see Table 1). $\text{CRPS}_{\text{cons}}$ is calculated in both GMST-only and Local+GMST cases. Therefore, the quantity $1 - \frac{\text{CRPS}_{\text{cons}}(\text{Local+GMST})}{\text{CRPS}_{\text{cons}}(\text{Global-only})}$ allows to quantify the added value from local observations compared to the sole use of GMST observations. A positive (negative) value, indicates an improvement (deterioration). The higher the CRPSS (bounded at 1), the better the performance.

Data availability

The datasets generated during and/or analysed during the current study are available at: https://gitlab.com/saidqasmi/KCC_notebook.

Code availability

All required programs to run the statistical method are in the associated KCC R package, which is available under a GNU General Public License, version 3 (GPLv3), at https://doi.org/10.5281/zenodo.5233947 [56].

A reproducible example of application is provided in the Jupyter Notebook available at: https://gitlab.com/saidqasmi/KCC_notebook.

Acknowledgements

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme in the framework of the EUCP project (Grant Agreement 776613) and Météo-France. The authors thank Hervé Douville for fruitful discussions about this work. The authors thank the climate modeling groups involved in CMIP6 exercises for producing and making available their simulations. The authors thank Kevin Cowtan for providing the blended temperature data, and the ETH Zurich for providing CMIP6 data through their cmip6-ng interface (http://dx.doi.org/10.5281/zenodo.3734128). The analyses and figures were produced with the R software (https://www.R-project.org/) and the NCAR Command

Language Software (http://dx.doi.org/10.5065/D6WD3XH5).

Author information

S.Q. and A.R. contributed to the design of the study, the interpretation of the results and the writing of the manuscript. S.Q. processed the data and produced the figures.

Competing interests

The authors declare no competing interests.

References

[1] Bindoff, N. L. *et al.* Chapter 10 - Detection and attribution of climate change: From global to regional. In *Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5* (Cambridge University Press, Cambridge, 2013). URL http://pure.iiasa.ac.at/id/eprint/10552/.

[2] Lee, J., Marotzke, J., Bala, G., Cao, L. & Corti, S. Chapter 4 - 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* (Cambridge University Press, Cambridge, 2021).

[3] Hawkins, E. & Sutton, R. Time of emergence of climate signals. *Geophysical Research Letters* **39** (2012). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011GL050087. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011GL050087.

[4] Hawkins, E. *et al.* Observed Emergence of the Climate Change Signal: From the Familiar to the Unknown. *Geophysical Research Letters* **47**, e2019GL086259 (2020). URL http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086259. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL086259.

[5] Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. Communication of the role of natural variability in future North American climate. *Nature Climate Change* **2**, 775–779 (2012). URL http://www.nature.com/articles/nclimate1562. Number: 11 Publisher: Nature Publishing Group.

[6] Mahlstein, I., Knutti, R., Solomon, S. & Portmann, R. W. Early onset of significant local warming in low latitude countries. *Environmental Research Letters* **6**, 034009 (2011). URL https://doi.org/10.1088/1748-9326/6/3/034009. Publisher: IOP Publishing.

[7] Collins, M. *et al.* Chapter 12 - Long-term climate change: Projections, commitments and irreversibility. In IPCC (ed.) *Climate Change 2013: The Physical Science Basis. IPCC Working Group I Contribution to AR5* (Cambridge University Press, Cambridge, 2013). URL http://pure.iiasa.ac.at/id/eprint/10551/.

[8] Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9**, 1937–1958 (2016). URL https://www.geosci-model-dev.net/9/1937/2016/.

[9] O'Neill, B. C. *et al.* The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development* **9**, 3461–3482 (2016). URL https://gmd.copernicus.org/articles/9/3461/2016/. Publisher: Copernicus GmbH.

[10] Tokarska, K. B. *et al.* Past warming trend constrains future warming in CMIP6 models. *Science Advances* **6**, eaaz9549 (2020). URL https://advances.sciencemag.org/content/6/12/eaaz9549. Publisher: American Association for the Advancement of Science Section: Research Article.

[11] Ribes, A., Qasmi, S. & Gillett, N. P. Making climate projections conditional on historical observations. *Science Advances* **7**, eabc0671 (2021). URL https://advances.sciencemag.org/content/7/4/eabc0671. Publisher: American Association for the Advancement of Science Section: Research Article.

[12] Brunner, L. *et al.* Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics* **11**, 995–1012 (2020). URL https://esd.copernicus.org/articles/11/995/2020/. Publisher: Copernicus GmbH.

[13] Nijsse, F. J. M. M., Cox, P. M. & Williamson, M. S. Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS)

from historical warming in CMIP5 and CMIP6 models. *Earth System Dynamics* **11**, 737–750 (2020). URL https://esd.copernicus.org/articles/11/737/2020/. Publisher: Copernicus GmbH.

[14] Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C. & Eyring, V. Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics* **11**, 1233–1258 (2020). URL https://esd.copernicus.org/articles/11/1233/2020/. Publisher: Copernicus GmbH.

[15] Field, C. B., Barros, V. & Dokken, D. J. *Managing the risks of extreme events and disasters to advance climate change adaption: special report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, New York, NY, 2012). OCLC: ocn796030880.

[16] Brunner, L. *et al.* Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework. *Journal of Climate* **33**, 8671–8692 (2020). URL https://journals.ametsoc.org/jcli/article/33/20/8671/348736/Comparing-Methods-to-Constrain-Future-European. Publisher: American Meteorological Society.

[17] Brunner, L., Lorenz, R., Zumwald, M. & Knutti, R. Quantifying uncertainty in European climate projections using combined performance-independence weighting. *Environmental Research Letters* **14**, 124010 (2019). URL https://doi.org/10.1088%2F1748-9326%2Fab492f. Publisher: IOP Publishing.

[18] Abramowitz, G. & Bishop, C. H. Climate Model Dependence and the Ensemble Dependence Transformation of CMIP Projections. *Journal of Climate* **28**, 2332–

2348 (2015). URL https://journals.ametsoc.org/view/journals/clim/28/6/jcli-d-14-00364.1.xml. Publisher: American Meteorological Society Section: Journal of Climate.

[19] Bishop, C. H. & Abramowitz, G. Climate model dependence and the replicate Earth paradigm. *Climate Dynamics* **41**, 885–900 (2013). URL https://doi.org/10.1007/s00382-012-1610-y.

[20] Borodina, A., Fischer, E. M. & Knutti, R. Emergent Constraints in Climate Projections: A Case Study of Changes in High-Latitude Temperature Variability. *Journal of Climate* **30**, 3655–3670 (2017). URL https://journals.ametsoc.org/view/journals/clim/30/10/jcli-d-16-0662.1.xml. Publisher: American Meteorological Society Section: Journal of Climate.

[21] Sutton, R., Suckling, E. & Hawkins, E. What does global mean temperature tell us about local climate? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **373**, 20140426 (2015). URL https://royalsocietypublishing.org/doi/10.1098/rsta.2014.0426. Publisher: Royal Society.

[22] Tebaldi, C. & Arblaster, J. M. Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Climatic Change* **122**, 459–471 (2014). URL https://doi.org/10.1007/s10584-013-1032-9.

[23] Viner, D. *et al.* Understanding the dynamic nature of risk in climate change assessments—A new starting point for discussion. *Atmospheric Science Letters* **21**, e958 (2020). URL https://rmets.

onlinelibrary.wiley.com/doi/abs/10.1002/asl.958. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/asl.958.

[24] Cowtan, K. *et al.* Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures. *Geophysical Research Letters* **42**, 6526–6534 (2015). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL064888. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015GL064888.

[25] Richardson, M., Cowtan, K. & Millar, R. J. Global temperature definition affects achievement of long-term climate goals. *Environmental Research Letters* **13**, 054004 (2018). URL https://iopscience.iop.org/article/10.1088/1748-9326/aab305.

[26] Cowtan, K. & Way, R. G. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society* **140**, 1935–1944 (2014). URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2297. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2297.

[27] Hoegh-Guldberg, O. *et al.* Impacts of 1.5^{o}C global warming on natural and human systems. In *Global Warming of 1.5°C* (IPCC, 2018).

[28] Screen, J. A. & Simmonds, I. The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature* **464**, 1334–1337 (2010). URL https://www.nature.com/articles/nature09051. Number: 7293 Publisher: Nature Publishing Group.

[29] Jansen, E. *et al.* Past perspectives on the present era of abrupt Arctic climate change. *Nature Climate Change* **10**, 714–721 (2020). URL http://www.nature.com/articles/s41558-020-0860-7. Number: 8 Publisher: Nature Publishing Group.

[30] Gneiting, T., Raftery, A. E., Westveld, A. H. & Goldman, T. Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Review* **133**, 1098–1118 (2005). URL https://journals.ametsoc.org/mwr/article/133/5/1098/67504/Calibrated-Probabilistic-Forecasting-Using. Publisher: American Meteorological Society.

[31] Schleussner, C.-F. *et al.* Differential climate impacts for policy-relevant limits to global warming: the case of 1.5 °C and 2 °C. *Earth System Dynamics* **7**, 327–351 (2016). URL https://esd.copernicus.org/articles/7/327/2016/.

[32] Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society* **93**, 485–498 (2012). URL https://journals.ametsoc.org/doi/10.1175/BAMS-D-11-00094.1.

[33] Tebaldi, C. & Knutti, R. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**, 2053–2075 (2007). URL https://royalsocietypublishing.org/doi/10.1098/rsta.2007.2076. Publisher: Royal Society.

[34] Sanderson, B. M., Knutti, R. & Caldwell, P. A Representative Democracy to Reduce Interdependency in a Multimodel Ensemble. *Journal of Climate* **28**, 5171–5194 (2015). URL https://journals.ametsoc.org/jcli/article/28/13/5171/106259/A-Representative-Democracy-to-Reduce. Publisher: American Meteorological Society.

[35] Eyring, V. *et al.* Taking climate model evaluation to the next level. *Nature Climate Change* **9**, 102–110 (2019). URL https://www.nature.com/articles/s41558-018-0355-y. Number: 2 Publisher: Nature Publishing Group.

[36] Sanderson, B. M. & Knutti, R. On the interpretation of constrained climate model ensembles. *Geophysical Research Letters* **39** (2012). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052665. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL052665.

[37] Knutti, R. *et al.* A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters* **44**, 1909–1918 (2017). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL072012. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL072012.

[38] Sanderson, B. M., Wehner, M. & Knutti, R. Skill and independence weighting for multi-model assessments. *Geoscientific Model Development* **10**, 2379–2395 (2017). URL https://gmd.copernicus.org/articles/10/2379/2017/. Publisher: Copernicus GmbH.

[39] Shepherd, T. G. Storyline approach to the construction of regional climate change information. *Proceedings. Mathematical, Physical, and Engineering Sciences* **475** (2019). URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6545051/.

[40] Sutton, R. T. Climate Science Needs to Take Risk Assessment Much More Seriously. *Bulletin of the American Meteorological Society* **100**, 1637–1642 (2019). URL https://journals.ametsoc.org/view/journals/bams/100/9/bams-d-18-0280.1.xml. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society.

[41] Zappa, G. Regional Climate Impacts of Future Changes in the Mid–Latitude Atmospheric Circulation: a Storyline View. *Current Climate Change Reports* **5**, 358–371 (2019). URL https://doi.org/10.1007/s40641-019-00146-7.

[42] Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres* **117** (2012). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD017187. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2011JD017187.

[43] Gillett, N. P. *et al.* The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6. *Geoscientific Model Development* **9**, 3685–3697 (2016). URL https://gmd.copernicus.org/articles/9/3685/2016/. Publisher: Copernicus GmbH.

[44] Hall, A. & Qu, X. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophysical Research Letters* **33** (2006). URL http://agupubs.onlinelibrary.wiley.com/doi/10.1029/2005GL025127.

[45] Hall, A., Cox, P., Huntingford, C. & Klein, S. Progressing emergent constraints on future climate change. *Nature Climate Change* **9**, 269–278 (2019). URL http://www.nature.com/articles/s41558-019-0436-6.

[46] Annan, J. D. & Hargreaves, J. C. Reliability of the CMIP3 ensemble. *Geophysical Research Letters* **37** (2010). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL041994. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL041994.

[47] Ribes, A., Zwiers, F. W., Azaïs, J.-M. & Naveau, P. A new statistical approach to climate change detection and attribution. *Climate Dynamics* **48**, 367–386 (2017). URL https://doi.org/10.1007/s00382-016-3079-6.

[48] Rougier, J., Goldstein, M. & House, L. Second-Order Exchangeability Analysis for Multimodel Ensembles. *Journal of the American Statistical Association* **108**, 852–863 (2013). URL https://doi.org/10.1080/01621459.2013.802963. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2013.802963.

[49] Oldenborgh, G. J. v., Reyes, F. J. D., Drijfhout, S. S. & Hawkins, E. Reliability of regional climate model trends. *Environmental Research Letters* **8**, 014055 (2013). URL https://doi.org/10.1088/1748-9326/8/1/014055. Publisher: IOP Publishing.

[50] Geoffroy, O. *et al.* Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate* **26**, 1841–1857 (2013). URL https://journals.ametsoc.org/view/journals/clim/26/6/jcli-d-12-00195.1.xml. Publisher: American Meteorological Society Section: Journal of Climate.

[51] Parsons, L. A., Brennan, M. K., Wills, R. C. J. & Proistosescu, C. Magnitudes and Spatial Patterns of Interdecadal Temperature Variability in CMIP6. *Geophysical Research Letters* **47**, e2019GL086588 (2020). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086588. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL086588.

[52] Qasmi, S., Cassou, C. & Boé, J. Teleconnection Between Atlantic Multidecadal Variability and European Temperature: Diversity and Evaluation of the Coupled Model Intercomparison Project Phase 5 Models. *Geophysical Research Letters* **44**, 11,140–11,149 (2017). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL074886.

[53] Laepple, T. & Huybers, P. Ocean surface temperature variability: Large model–data differences at decadal and longer periods. *Proceedings of the National Academy of Sciences* **111**, 16682–16687 (2014). URL https://www.pnas.org/content/111/47/16682. Publisher: National Academy of Sciences Section: Physical Sciences.

[54] Trenberth, K. E. & Fasullo, J. T. An apparent hiatus in global

warming? *Earth's Future* **1**, 19–32 (2013). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013EF000165. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013EF000165.

[55] Douville, H., Voldoire, A. & Geoffroy, O. The recent global warming hiatus: What is the role of Pacific variability? *Geophysical Research Letters* **42**, 880–888 (2015). URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL062775. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014GL062775.

[56] Qasmi, S. & Ribes, A. Kriging for Climate Change Package (2021). URL https://github.com/saidqasmi/KCC.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- paperSI.pdf