Monte Carlo Drift Correction – Quantifying the Drift Uncertainty of Global Climate Models

Benjamin S. Grandey[1], Zhi Yang Koh[1], Dhrubajyoti Samanta[2], Benjamin P. Horton[2,3], Justin Dauwels[4], and Lock Yue Chew[1]

[1]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
[2]Earth Observatory of Singapore, Nanyang Technological University, Singapore
[3]Asian School of the Environment, Nanyang Technological University, Singapore
[4]Department of Microelectronics, Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology (TU Delft), The Netherlands

Correspondence: Benjamin S. Grandey (benjamin.grandey@ntu.edu.sg)

Abstract. Global climate models are susceptible to drift, causing spurious trends in output variables. Drift is often corrected using data from a control simulation. However, internal climate variability within the control simulation introduces uncertainty to the drift correction process. To quantify this drift uncertainty, we develop a probabilistic technique: Monte Carlo drift correction (MCDC). MCDC involves random sampling of the control time series. We apply MCDC to an ensemble of global

5   climate models from the Coupled Model Intercomparison Project Phase 6 (CMIP6). We find that drift correction partially addresses a problem related to drift: energy non-conservation. Nevertheless, the energy balance of several models remains suspect. We quantify the drift uncertainty of global quantities associated with energy balance and thermal expansion of the ocean. When correcting drift in a cumulatively-integrated energy flux, we find that it is preferable to integrate the flux before correcting the trend: an alternative method would be to correct the bias before integrating the flux, but this alternative method

10   amplifies the drift uncertainty by up to an order of magnitude. We find that drift uncertainty is often smaller than other sources of uncertainty: for thermosteric sea-level rise projections for the 2090s, ensemble-mean drift uncertainty (9 mm) is an order of magnitude smaller than scenario uncertainty (138 mm) and model uncertainty (98 mm). However, drift uncertainty may dominate time series that have weak trends: for historical thermosteric sea-level rise since the 1850s, ensemble-mean drift uncertainty is 15 mm, which is of comparable magnitude to the impact of omitting volcanic forcing in control simulations.

15   Therefore, drift uncertainty may influence comparisons between historical simulations and observation-based estimates of thermosteric sea-level rise. When evaluating and analysing global climate model data that are susceptible to drift, researchers should consider drift uncertainty.

Short summary

Global climate models are susceptible to spurious trends known as *drift*. Fortunately, drift can be corrected when analysing data

20   produced by the models. To explore the uncertainty associated with drift correction, we develop a new method: *Monte Carlo*

drift correction. For historical simulations of thermosteric sea-level rise, drift uncertainty is relatively large. When analysing data susceptible to drift, researchers should consider drift uncertainty.

1 Introduction

Global climate models are susceptible to *drift* (Sen Gupta et al., 2013; Irving et al., 2020), causing spurious trends in output

25   variables such as thermosteric sea-level rise. Drift is not caused by transient external forcing (Sen Gupta et al., 2013). Instead, drift in long-term centennial-scale simulations is caused by two primary factors: insufficient spin-up and model errors (Hobbs et al., 2016).

First, drift may occur due to insufficient spin-up of equilibrium simulations (Sen Gupta et al., 2013; Hobbs et al., 2016). Following a perturbation to equilibrium forcing, the deep ocean responds slowly, taking millennia to approach a quasi-equilibrium

30   state (Stouffer, 2004; Stouffer et al., 2004; Eyring et al., 2016). This slow response of the deep ocean is consistent with the physics of the real climate system. This source of drift can be reduced by increasing the spin-up length of equilibrium simulations.

Second, drift may arise due to errors – including biases – in the climate model (Hobbs et al., 2016; Brunetti and Vérard, 2018; Irving et al., 2020). Even when climate models are spun-up for long periods under equilibrium forcing, modelled variables may

35   be inconsistent with a true quasi-equilibrium state – spurious trends remain (Hobbs et al., 2016). In contrast to the real climate system, global climate models often fail to conserve energy, mass, and salt (Irving et al., 2020). Of particular relevance here is the problem of *energy non-conservation*: unphysical sinks and sources of energy contribute to inconsistent biases in the top-of-atmosphere radiative flux and the sea-surface heat flux (Sect. 4.1; Lucarini and Ragone, 2011; Hobbs et al., 2016). Such biases indicate defects in the tuning procedure and the parameterisations included in the global climate model (Brunetti and

40   Vérard, 2018).

Variables influenced by the deep ocean are especially susceptible to drift (Sen Gupta et al., 2013). Derived climate indices may also be sensitive to drift (Sanderson, 2020).

Fortunately, drift can often be corrected when analysing global climate model data (Sen Gupta et al., 2013). Correcting drift may also address the problem of energy non-conservation (Hobbs et al., 2016; Irving et al., 2020).

45   Most drift correction methods rely on *control* simulations (Sen Gupta et al., 2013). These control simulations are forced by equilibrium forcing consistent with pre-industrial conditions, using 1850 as the reference year (Eyring et al., 2016). Such control simulations provide initial conditions for historical simulations.

A commonly used drift correction method involves fitting a linear trend to the entire control time series (Sen Gupta et al., 2013). This linear trend can then be subtracted from transient time series – such as time series produced by historical or pro-

50   jection simulations – to correct the drift. However, internal climate variability in the control simulation introduces uncertainty to the drift correction process (Sen Gupta et al., 2013; Sanderson, 2020).

In this paper, we quantify the *drift uncertainty* associated with the drift correction of global climate model data. Sen Gupta et al. (2013) recommend fitting a linear trend to the *entire* control time series to minimise the influence of internal variability.

In contrast, rather than seeking to *minimise* the influence of the internal variability, we seek to *quantify* the influence of this variability.

To do this, we propose a Monte Carlo drift correction (MCDC) framework. We explore two alternative flavours of MCDC: trend-method MCDC (Sect. 3.2) and integrated-bias-method MCDC (Sect. 3.3).

Using MCDC, we produce drift-corrected time series of excess system energy, excess ocean heat, and thermosteric sea-level rise. Relationships between these variables (Appendix A) illumine a problem related to drift: the problem of energy non-conservation. Using the drift-corrected time series, we quantify the drift uncertainty associated with estimates of the fraction of energy absorbed by the ocean (β), the expansion efficiency of heat (ϵ), and thermosteric sea-level rise (ΔZ).

Our overarching aim is to *quantify drift uncertainty*. To support this aim, we ask three primary research questions:

1. Do trend-method MCDC and integrated-bias-method MCDC produce similar measurements of drift uncertainty? If not, which method minimises drift uncertainty?

2. What is the contribution of drift uncertainty to estimates of β, estimates of ϵ, and ΔZ?

3. How large is the drift uncertainty compared with other sources of uncertainty?

2   Data

We use global climate model data produced for the Coupled Model Intercomparison Project Phase 6 (CMIP6; Eyring et al., 2016) and the CMIP6-endorsed Scenario Model Intercomparison Project (ScenarioMIP; O'Neill et al., 2016). CMIP6 includes contributions from many climate modeling groups globally.

Our MCDC framework uses the *control* simulations (Eyring et al., 2016). These control simulations follow equilibrium forcing representative of the year 1850.

We apply MCDC to the *historical* simulations (Eyring et al., 2016). Each historical simulation is initialised using data from a control simulation. The historical simulations follow transient forcing for the period 1850–2014. The 1850–2014 transient forcing includes both natural forcing (including volcanic eruptions) and anthropogenic forcing (including greenhouse gas concentrations and aerosol emissions).

We also apply MCDC to ScenarioMIP's four Tier 1 scenarios for the period 2015–2100: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 (O'Neill et al., 2016). These *projection* scenarios are based on the narratives of the Shared Socioeconomic Pathways (SSPs; Riahi et al., 2017).

We analyse global variables that shed light on energy balance and thermosteric sea-level rise (Appendix A). We derive these variables from the CMIP6 diagnostic variables as follows:

1. Net downward *top-of-atmosphere radiative flux* (R) is calculated as $R = rsdt - rsut - rlut$, where $rsdt$ is the incoming shortwave radiation, $rsut$ is the outgoing shortwave radiation, and $rlut$ is the outgoing longwave radiation. Each top-of-atmosphere flux ($rsdt$, $rsut$, and $rlut$) is multiplied by the atmosphere grid-cell area ($areacella$) then summed globally.

2. Net downward *sea-surface heat flux* (H) corresponds to the variable $hfds$, assuming there is no flux correction (Griffies et al., 2016). One model – MRI-ESM2-0 – has flux correction ($hfcorr$) data, so we calculate H as $H = hfds + hfcorr$ for this model. Each sea-surface flux ($hfds$ and $hfcorr$) is multiplied by the ocean grid-cell area ($areacello$) then summed globally.

3. Thermosteric sea-level rise (ΔZ) corresponds to the variable $zostoga$, a global-mean diagnostic.

Excess system energy ($\int R$) is calculated by integrating R cumulatively (Eq. A1). *Excess ocean heat* ($\int H$) is calculated by integrating H cumulatively (Eq. A2). The *fraction of energy absorbed by the ocean* (β) is calculated as the linear regression coefficient of $\int H$ versus $\int R$ (Eq. A3). The *expansion efficiency of heat* (ϵ) is calculated as the linear regression coefficient of ΔZ versus $\int H$ (Eq. A4). The coefficients β and ϵ are both estimated using ordinary least squares (Appendix B).

The projection simulations are essentially continuations of the historical simulation, beginning where the historical simulation ends. Therefore, when processing the projection time series, we prepend the historical period (1850–2014) before the projection period (2015–2100). In other words, each projection time series incorporates the contribution of the earlier historical time series. For example, the cumulative integration that produces the $\int R$ SSP5-8.5 time series begins in 1850 (not 2015). The $\int R$, $\int H$, and ΔZ time series are all initialised at zero in 1850 (or the first year, in the case of the control time series). We subsequently remove the historical period from the projection time series before we conduct further analysis – including plotting figures and calculating β and ϵ – to distinguish the projection scenarios from one another and from the historical time series.

Within the CMIP6 ensemble, we search for model variants that have monthly data for the required diagnostic variables ($rsdt$, $rsut$, $rlut$, $areacella$, $hfds$, $areacello$, and $zostoga$) and scenario (control, historical, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5). It is common practice to select only one ensemble member per model (e.g. Jevrejeva et al., 2021; Hermans et al., 2021). For each model, we use only the first available "r1i1" variant – in practice, this means we use either the "r1i1p1f1" or "r1i1p1f2" variant, depending on the model (Table S1). These constraints provide an ensemble of sixteen models (Table S1).

We focus on one model as an illustrative example: the *UK Earth System Model* (UKESM1; Sellar et al., 2019; Tang et al., 2019; Good et al., 2019). With a high equilibrium climate sensitivity of 5.4 K (Sellar et al., 2019), UKESM1 is a "hot model" that may overestimate future warming (Hausfather et al., 2022). Nevertheless, as a state-of-the-art global climate model, UKESM1 is well-suited for our present purpose of exploring drift uncertainty. Furthermore, UKESM1 has the longest control time series length (1100 yr) of any model within the ensemble (Table S1). Many of the figures presented in this manuscript and the supplement show results for UKESM1 (Figs. 1–7 and S1–S2). Nevertheless, our analysis also includes the larger ensemble (Figs. S3–S8; Tables 1 and S2).

3 Drift correction methods

3.1 Conventional drift correction using a single estimate of drift

Sen Gupta et al. (2013) compared drift correction methods, including linear, quadratic, and cubic drift correction approaches. Many recent studies focusing on the earth's energy budget or sea-level change use the linear drift correction approach (e.g. Hobbs et al., 2016; Jevrejeva et al., 2016; Slangen et al., 2017; Palmer et al., 2018, 2020; Cuesta-Valero et al., 2021; Hamlington et al., 2021; Lambert et al., 2021). Other studies use the quadratic drift correction approach (e.g. Gleckler et al., 2016; Lyu et al., 2020, 2021; Harrison et al., 2021; Hermans et al., 2021; Jevrejeva et al., 2021). To apply a quadratic or cubic drift correction to a transient time series, the *branch time* must be known: Which part of the control time series parallels the historical time series? Often, this branch-time information may be either unavailable or unreliable (Gleckler et al., 2016; Flynn and Mauritsen, 2020). Furthermore, a quadratic or cubic drift correction approach risks overfitting any curvature in the control time series (Hobbs et al., 2016). On the other hand, such curvature may contain valuable information about non-linearity in the drift. Here, we adopt a linear approach.

Sen Gupta et al. (2013) also considered whether to use all or only part of the control time series. They recommended that drift be estimated using the entire control time series, "to minimize the contamination of the drift estimate by internal variability" (Sen Gupta et al., 2013, p.8597).

To derive a single best estimate of drift, it makes sense to calculate the linear trend of the entire control time series, as recommended by Sen Gupta et al.. However, it would also be beneficial to estimate the *uncertainty* associated with drift correction. For example, we could then ask, How does drift uncertainty influence projections of thermosteric sea-level rise? To answer questions like this, we need to go beyond Sen Gupta et al.'s recommendation.

3.2 Trend-method Monte Carlo drift correction

To quantify drift uncertainty, we propose a Monte Carlo drift correction (MCDC) framework. To estimate the uncertainty associated with variability in the control time series, MCDC randomly draws segments of length 150 yr from the control time series. Within each 150 yr segment, MCDC also samples the standard error of the estimated trend or bias. MCDC can be applied in two alternative ways: *trend-method* MCDC (described below) and *integrated-bias-method* MCDC (described in Sect. 3.3).

Trend-method MCDC proceeds as follows:

1. Randomly draw (with replacement) a 150 yr segment from the control time series.

2. For each variable ($\int R$, $\int H$, and ΔZ), (i) use ordinary least squares to estimate the linear trend of the segment and the standard error of the estimated trend, (ii) construct a Gaussian distribution with mean equal to the estimated trend and standard deviation equal to the standard error, then (iii) randomly draw a *trend sample* from the Gaussian distribution. The Gaussian distribution from which the trend sample is drawn corresponds to the specific 150 yr segment drawn in step 1 above.

3. For each variable and scenario of interest (e.g. historical), subtract the trend sample from the time series to produce a *trend-corrected* time series. In other words, detrend the time series using the trend sample derived from the control time series.

4. Repeat the above steps 500 times, drawing 500 trend samples for each variable. This produces 500 trend-corrected time series for each variable and scenario.

For each global climate model, trend-method MCDC produces the following data for $\int R$: (i) 500 trend samples derived from the control simulation, and (ii) 500 trend-corrected time series for each scenario of interest. Trend-method MCDC also produces corresponding data for $\int H$ and ΔZ.

This MCDC approach is informed by several assumptions. Four important assumptions are described below.

First, we assume that branch-time metadata – indicating which part of the control simulation parallels the historical simulation – is either unavailable, unreliable, or irrelevant. In other words, we have no reason to favour one segment of the control time series over another. Therefore, we treat all segments of the control time series equally, using information from the entire length (in agreement with Sen Gupta et al., 2013). This assumption limits our investigation to linear estimates of drift, because quadratic and cubic drift correction approaches rely on branch-time information. If reliable branch-time data were to be taken into account, then it might be possible to reduce the drift uncertainty. Therefore, our assumption may lead us to overestimate the drift uncertainty. However, the risk of overestimation is not necessarily a disadvantage: When seeking to quantify uncertainty, is it not preferable to aim for an upper-bound estimate, to reduce the risk of underestimating the uncertainty?

Second, we assume that 150 yr segments of the control time series represent an appropriate period over which to derive our drift estimates. 150 yr is long enough to average over decadal variability, so any differences between 150 yr segments may be indicative of a genuine signal in the underlying model drift. 150 yr is also much shorter than the available length of most control simulations (Table S1), enabling sampling of different parts of the control time series. Such sampling may capture low-frequency variability, such as the multi-century oscillations reported in UKESM1's surface temperature and sea-ice cover (Sellar et al., 2019), although such low-frequency variability should not technically be categorised as drift (Sen Gupta et al., 2013). Additionally, 150 yr is of comparable length to the historical time series (165 yr), a target period of interest.

Third, we assume that the standard error of the trend should be included via random sampling of a Gaussian distribution. For any given 150 yr segment, the standard error of the estimated trend provides an estimate of the drift uncertainty within that segment.

Fourth, we assume that 500 trend samples are sufficient to characterise the drift uncertainty. When analysing our results, we quantify drift uncertainty using the 2nd–98th inter-percentile range of the drift-corrected data. If we were to use more extreme percentiles instead, then additional trend samples might be required.

When analysing relationships between variables, the different variables should be sampled consistently using the same 150 yr segments. For example, the first $\int R$ trend sample, $\int H$ trend sample, and ΔZ trend sample should all be derived from the same 150 yr segment of the control time series. Strong correlations exist between the trend samples for the different variables (Fig. S1a and c). These correlations should be preserved when analysing $\int R$–$\int H$–ΔZ relationships.

180   When estimating the linear trend of each segment (step 2i), we use ordinary least squares. Are our results sensitive to this choice of estimation method? To test this, we have re-run our analysis using a robust linear model with an M-estimator (Huber, 1964, 1981) instead of ordinary least squares. The results are very similar. We conclude that trend-method MCDC is not sensitive to the choice of estimation method.

3.3   Integrated-bias-method Monte Carlo drift correction

185 Variables derived from cumulatively-integrated fluxes – such as $\int R$ and $\int H$ – can also be corrected using an alternative method, *integrated-bias-method* MCDC:

1. Randomly draw (with replacement) a 150 yr segment from the control time series.

2. For each flux variable (R and H), (i) calculate the mean of the segment and the standard error of the mean, (ii) construct a Gaussian distribution with mean equal to the mean of the segment and standard deviation equal to the standard error, (iii) randomly draw a *bias sample* from the Gaussian distribution.

3. For each flux variable and scenario of interest, subtract the bias sample from the time series to produce a *bias-corrected* time series.

4. Repeat the above steps 500 times, drawing 500 bias samples for each flux variable. This produces 500 bias-corrected time series for each flux variable and scenario. (Steps 1–4 could be described as Monte Carlo *bias* correction.)

5. Integrate each bias-corrected time series cumulatively. This produces 500 *integrated-bias-corrected* time series for each cumulatively-integrated flux variable ($\int R$ and $\int H$) and scenario.

    For each global climate model, integrated-bias-method MCDC produces the following data for R: (i) 500 bias samples derived from the control time series, and (ii) 500 bias-corrected time series for each scenario of interest. Additionally, integrated-bias-method MCDC produces the following data for $\int R$: 500 integrated-bias-corrected time series for each scenario. Integrated-200 bias-method MCDC also produces corresponding data for H and $\int H$.

4   Results

4.1   Uncorrected time series

An example of drift is illustrated in Fig. 1. For the control time series, the top-of-atmosphere radiative flux (R) is close to zero (Fig. 1a), consistent with quasi-equilibrium radiative balance at the top of the atmosphere. The UKESM1 team have achieved 205 their stated goal of keeping R close to zero when tuning the model (Sellar et al., 2019). Therefore, the drift in excess system energy ($\int R$) is small (Fig. 1c).

    However, the sea-surface heat flux (H) has a negative bias (Fig. 1b), showing that the ocean loses heat spuriously. When we integrate the flux cumulatively, the negative bias in H integrates into a negative trend in excess ocean heat ($\int H$; Fig. 1d). This

210   flux (Hobbs et al., 2016).

In this example, the drift in $\int H$ also dominates the historical time series, obscuring the anthropogenic warming signal during the 20th century (Fig. 1d). Anthropogenic forcing only becomes strong enough to offset the negative bias in H towards the end of the 20th century (Fig. 1b).

If energy is conserved within the modelled climate system, the $\int R$–$\int H$ relationship should be approximately linear (Eq. A3).
215   In contrast, for UKESM1's historical simulation, the drift in $\int H$ drives a strange $\int R$–$\int H$ relationship (Fig. 1f): for most of the historical period, $\int R$ remains close to zero, but $\int H$ decreases. This reveals energy non-conservation. Furthermore, the inconsistency between $\int R$ and $\int H$ reveals that the energy leakage occurs somewhere between the top of the atmosphere and the sea surface (i.e. outside the ocean; Hobbs et al., 2016).

In this example, thermosteric sea-level rise (ΔZ) also exhibits drift (Fig. 1e). The drift in ΔZ obscures the anthropogenic
220   sea-level rise signal during the 20th century. The drift will also contaminate future projections. Furthermore, the $\int H$–ΔZ relationship (Fig. 1g) is inconsistent with Eq. (A4).

4.2   Drift-corrected time series

Figure 2a shows an example of estimating drift using a linear fit to the entire control time series (Fig. 2a "1100 yr trend" line). In this example, the spurious trend in $\int R$ is approximately 0.03 W m^{-2} (Fig. 2b "1100 yr trend" line). For all three variables,
225   the single drift estimate lies near the middle of the Monte Carlo drift estimates described below (Fig. 2b, f, and j).

When applying trend-method MCDC, 500 trend samples are produced by randomly sampling 150 yr segments of the control R time series (Fig. 2a). In this example, the trend samples range from below 0.00 W m^{-2} to above 0.05 W m^{-2}, with a median of approximately 0.03 W m^{-2} (Fig. 2b). Each trend sample is subtracted from the control time series, producing 500 trend-corrected time series (Fig. 2c). Each trend sample is also subtracted from the historical time series (Fig. 2d). The spread
230   between the trend-corrected time series reveals the drift uncertainty.

In this example, most of the $\int R$ trend samples are positive, but a few are negative (Fig. 2b). It is therefore plausible that there is no spurious trend. Accordingly, the uncorrected time series lies within the range of the trend-corrected time series (Fig. 2c and d).

The ΔZ trend samples also include values near zero (Fig. 2j). Accordingly, the uncorrected time series lies near the edge of
235   the range of the trend-corrected time series (Fig. 2k–l).

In contrast, the drift in $\int H$ is much larger than the variability within the control time series (Fig. 1e), so all trend samples lie far from zero (Fig. 1f). Accordingly, the uncorrected time series lies well outside the range of the trend-corrected time series (Fig. 2g–h).

An example of integrated-bias-method MCDC is shown in Fig. 3e–h. The R bias samples (Fig. 3f) cover a much wider
240   range than the $\int R$ trend samples (Fig. 3b). Therefore, the integrated-bias-corrected time series (Fig. 3g–h) also cover a much wider range than the trend-corrected time series (Fig. 3c–d). In other words, integrated-bias-method MCDC provides a much

weaker constraint on the drift correction: to minimise drift uncertainty, it is preferable to use trend-method MCDC. The same conclusion can be drawn from the corresponding $\int H$ results (Fig. S2).

Furthermore, even if only a single estimate of drift is used, the estimate derived using the integrated-bias method (Fig. 3f "1100 yr bias" line) differs slightly from that derived using the trend method (Fig. 3b "1100 yr trend" line). When correcting drift in cumulatively-integrated fluxes, researchers should carefully consider which approach is most appropriate in their specific context and then clarify their choice when reporting results.

The choice of method also influences the correlation between drift samples for different variables (Fig. S1). If the 150 yr segments are sampled consistently, a strong correlation exists between trend samples (Fig. S1a and c). However, the correlation between bias samples is much weaker (Fig. S1b and d).

We hypothesise that these differences can be ascribed to differences in the size of the standard error: the standard error of the trend of $\int R$ (or $\int H$) is generally smaller than the standard error of the mean of R (or H), because integrating cumulatively effectively averages over the substantial inter-annual variability in R (or H). For example, if we select the first 150 yr segment of UKESM1's control time series, the standard error of the trend of $\int R$ ($\pm$ 0.002 W m^{-2}) is an order of magnitude smaller than the standard error of the mean of R ($\pm$ 0.021 W m^{-2}). The same applies to the standard error of the trend of $\int H$ ($\pm$ 0.002 W m^{-2}) compared with the standard error of the mean of H ($\pm$ 0.020 W m^{-2}). When we sample the Gaussian distribution (step 2 in both methods), a wider distribution contributes to a larger spread of drift samples and also weakens the correlation between drift samples for different variables. This influences the drift uncertainty associated with relationships between variables, as we show below.

4.3 Fraction of excess energy absorbed by the ocean

If energy is conserved and if the fraction of excess energy absorbed by the ocean (β) is constant, then excess ocean heat ($\int H$) should be a linear function of excess system energy ($\int R$): $\int H = \beta \int R$ (Eq. A3).

However, as noted in Sect. 4.1 above, the uncorrected $\int R$–$\int H$ relationship for UKESM1's historical time series is neither linear nor consistent with energy conservation (Fig. 1f). Can drift correction address this problem?

After we apply trend-method MCDC, the trend-corrected $\int R$–$\int H$ relationships are linear (Fig. 4a–b). Estimates of the coefficient β are positive and less than 1.0 (Fig. 4c). Therefore, for UKESM1, the trend-corrected $\int R$–$\int H$ relationships are generally consistent with Eq. (A3) and with energy conservation.

When applying trend-method MCDC, we produce 500 trend-corrected time series for each variable and scenario. Therefore, for each scenario, we calculate 500 estimates of β. The β estimates are presented as histograms in Fig. 4c. The spread within each histogram indicates the *trend-method drift uncertainty*. The drift uncertainty in β is relatively large for the historical time series: the relatively weak global warming signal in the $\int R$ and $\int H$ historical time series (Fig. 2d and h) leaves them susceptible to drift uncertainty, which in turn influences the β estimates. The drift uncertainty in β is much smaller for the projection time series: global warming drives clear trends in the $\int R$ and $\int H$ projection time series, constraining the β estimates. For the UKESM1 projection simulations, the trend-method drift uncertainty is approximately 0.01 (Table S2).

275 The β estimates differ slightly between the different projection simulations, with the mean estimate ranging from 0.935 (SSP5-8.5) to 0.944 (SSP1-2.6; Fig. 4c). The spread between the projection simulations indicates the *scenario uncertainty*. We can quantify the scenario uncertainty using the inter-scenario range of mean β estimates. For the UKESM1 projection simulations, the scenario uncertainty is approximately 0.01, similar to the trend-method drift uncertainty (Table S2).

 For most models in the CMIP6 ensemble, the scenario uncertainty is similarly small (Table S2; Fig. S3). However, for a few
280 models, the scenario uncertainty is larger. For one model, the scenario uncertainty is 0.08 (Table S2; Fig. S3c).

 Alongside drift uncertainty and scenario uncertainty, a third source of uncertainty is *model uncertainty* (Tebaldi and Knutti, 2007; Knutti and Sedláček, 2013): estimates of β vary between different models within the ensemble (Fig. S3). We can estimate the model uncertainty using the inter-model range of mean β estimates for a given projection scenario (Table S2). The inter-model range is large, with estimates of β ranging from 0.88 (Fig. S3e and f) to 1.05 (Fig. S3h and m), with an ensemble mean
285 of approximately 0.97. The lower-end estimates of approximately 0.9 are consistent with our understanding of the climate system (von Schuckmann et al., 2020). In contrast, the estimates above 1.0 are unrealistic: even if the ocean were to absorb all excess energy entering the climate system, this would lead to a theoretical maximum of $\beta = 1.0$. In other words, even after drift correction, the energy balance of several CMIP6 models remains problematic.

 Four models (25 % of the ensemble) consistently fail to conserve energy after drift correction: β is consistently greater than
290 1.0 for all projection scenarios (Fig. S3h, k, l, and m). A further two models also have suspect energy balance that deserves further attention: β is greater than 1.0 for one projection scenario (Fig. S3c and d). For the remaining ten models, β is less than 1.0 for all scenarios, consistent with energy conservation.

 Overall, model uncertainty contributes the largest uncertainty to estimates of β (Table 1). The model uncertainty (0.17) is an order of magnitude larger than both the mean scenario uncertainty (0.02) and the mean trend-method drift uncertainty (0.01).
295 However, if integrated-bias-method MCDC is used instead, the *integrated-bias-method drift uncertainty* is relatively large (Figs. 5a–c and S4; Tables 1 and S2). Compared with trend-method MCDC, integrated-bias-method MCDC provides a much weaker constraint on the appropriate drift correction (Figs. 3, S2, and 5a–b). Consequently, the integrated-bias-corrected estimates of β exhibit a much wider spread (Fig. 5c and S4). When averaged across the ensemble, the integrated-bias-method drift uncertainty (0.11) is an order of magnitude larger than the trend-method drift uncertainty (0.01) and is of comparable size to
300 the model uncertainty (0.17; Table 1).

4.4 Expansion efficiency of heat

As the ocean warms, it expands, leading to thermosteric sea-level rise (ΔZ). We expect the relationship between $\int H$ and ΔZ to be approximately linear: $\Delta Z = \epsilon \int H$ (Eq. A4).

 As noted in Sect. 4.1 above, the uncorrected data from UKESM1's historical simulation are inconsistent with Eq. (A4)
305 (Fig. 1g). However, after we apply trend-method MCDC, the trend-corrected $\int H - \Delta Z$ relationships are approximately linear (Fig. 4d–e). Trend-method MCDC successfully establishes meaningful $\int H - \Delta Z$ relationships that can be interpreted using Eq. (A4).

However, in disagreement with Eq. (A4), the trend-corrected $\int H$–ΔZ relationships for the historical period exhibit hysteresis-like behaviour (Fig. 4d). $\int H$ and ΔZ are initialised at zero in the year 1850, so the $\int H$–ΔZ relationships begin at the origin.

310 $\int H$ and ΔZ then become negative for much of the historical period (Fig. 2h and 2l), driving the relationships towards the lower-left of Fig. 4d. $\int H$ and ΔZ become positive later in the historical period, driving the relationships towards the upper-right of Fig. 4d. However, $\int H$ and ΔZ do not necessarily both become positive at the same time, so the relationships do not necessarily pass through the origin, as demonstrated by the "Max intercept" and "Min intercept" lines in Fig. 4d. It is possible that the $\int R$–$\int H$–ΔZ relationships may depend on the sign of the forcing (c.f. Bouttes et al., 2013). However, the

315 hysteresis-like behaviour in Fig. 4d – and also in Figs. 4a, 5a, and 5d – does not reveal systemic hysteresis: for different MCDC samples, the intercept can be either negative or positive. Regardless of the explanations for this hysteresis-like behaviour, we use ordinary least squares with an intercept – not regression through the origin – to estimate the linear regression coefficient (Appendix B).

The linear regression coefficient is ϵ, the expansion efficiency of heat. For UKESM1, the estimates of ϵ depend on the

320 scenario, ranging from approximately 117 mm YJ^{-1} (SSP1-2.6) to 126 mm YJ^{-1} (SSP5-8.5; Fig. 4f). This scenario uncertainty of 9 mm YJ^{-1} is much larger than the trend-method drift uncertainty of 1 mm YJ^{-1} (Table S2), leading to distinct non-overlapping histograms for the projection scenarios (Fig. 4f).

Similar features are found for the other models in the CMIP6 ensemble. For all models, the scenario uncertainty is larger than the trend-method drift uncertainty (Table S2). The trend-corrected ϵ estimates are smallest for SSP1-2.6 and largest for

325 SSP5-8.5 (Fig. S5). Estimates of ϵ range from 108 mm YJ^{-1} (Fig. S5l SSP1-2.6) to 127 mm YJ^{-1} (Fig. S5h SSP5-8.5). The ensemble-mean estimate of 118 mm YJ^{-1} is comparable to Fox-Kemper et al. (2021)'s estimate of 121 mm YJ^{-1} for the period 1995–2014.

When averaged across the ensemble, the mean trend-method drift uncertainty is approximately 1 mm YJ^{-1} (Table 1). The mean scenario uncertainty is larger: 8 mm YJ^{-1}. The model uncertainty is even larger: 12 mm YJ^{-1}, an order of magnitude

330 larger than the trend-method drift uncertainty.

The integrated-bias-method drift uncertainty is much larger than the trend-method drift uncertainty (Figs. 5d–f and S6; Tables 1 and S2). This is because integrated-bias-method MCDC provides a much weaker constraint on the appropriate drift correction for $\int H$ (Fig. S2). (Integrated-bias-method MCDC cannot be applied to ΔZ, so trend-method MCDC is applied to ΔZ when estimating integrated-bias-corrected ϵ.) The mean integrated-bias-method drift uncertainty is approximately

335 10 mm YJ^{-1}, of comparable magnitude to the scenario uncertainty and the model uncertainty (Table 1).

4.5 Thermosteric sea-level rise

For the UKESM1 control simulation, the thermosteric sea-level rise (ΔZ) trend samples range from below -0.12 mm yr^{-1} to approximately 0.00 mm yr^{-1} (Fig. 2j). This drift uncertainty contributes substantial uncertainty to ΔZ trends during the historical period: between the 1850s and 1995–2014, the change in trend-corrected ΔZ ranges from approximately 0 mm to

340 20 mm (Fig. 6). Drift uncertainty may therefore influence the extent to which the historical time series agrees with estimates of ΔZ derived from observations, such as the estimate of Frederikse et al. (2020).

For future projections beyond the 2050s, the drift uncertainty is small compared with scenario uncertainty (Fig. 7). For the 2050s (relative to 1995–2014), UKESM1's scenario uncertainty is 33 mm, five times larger than the drift uncertainty of 6 mm (Table S2). For the 2090s, the scenario uncertainty is 176 mm, more than an order of magnitude larger than the drift uncertainty
345 of 10 mm (Table S2).

We can draw similar conclusions for other ensemble members (Figs. S7 and S8; Table S2). For the historical period (1850s relative to 1995–2014), the mean drift uncertainty is 15 mm (Table 1). For the 2050s, the mean drift uncertainty is 5 mm, smaller than the mean scenario uncertainty of 23 mm. For the 2090s, the mean drift uncertainty is 9 mm, more than an order of magnitude smaller than the mean scenario uncertainty of 138 mm.

350 Model uncertainty is substantial. For the 2050s, the model uncertainty of 45 mm represents the largest source of uncertainty (Table 1). For the 2090s, the model uncertainty of 98 mm is smaller than the scenario uncertainty of 138 mm. For longer time horizons, scenario uncertainty plays an increasingly dominant role in the uncertainty associated with future thermosteric sea-level rise.

5 Discussion

5.1 Energy balance

Following Hobbs et al. (2016) and Irving et al. (2020), we have shown that global climate models often fail to conserve energy. Inconsistent relationships between excess system energy ($\int R$) and excess ocean heat ($\int H$; Fig. 1f) are symptomatic of energy leakage outside the ocean domain (Hobbs et al., 2016).

In agreement with Hobbs et al. (2016) and Irving et al. (2020), we have shown that drift correction can partially address
360 this problem of energy non-conservation. Trend-corrected $\int R$–$\int H$ relationships are linear (Fig. 4a–b). For all sixteen models analysed here, the coefficient β – the fraction of excess energy absorbed by the ocean – is positive (Fig. S3). For most of these models, β is also less than 1.0, consistent with energy conservation.

However, even after drift correction, the energy balance of several models remains suspect: energy non-conservation is revealed by β estimates greater than 1.0. For four models, β is consistently greater than 1.0 for all projection scenarios. For a
365 further two models, β is greater than 1.0 for one projection scenario.

When evaluating the performance of global climate models, the coefficient β is a useful diagnostic. This diagnostic could be used alongside other diagnostics relating to energy balance, such as the diagnostics considered by Mayer et al. (2017) and Wild (2020).

5.2 Drift uncertainty

370 Using an ensemble of global climate models from CMIP6, we have explored and quantified three sources of uncertainty: (i) drift uncertainty, derived using Monte Carlo drift correction (MCDC), (ii) scenario uncertainty, quantified using the inter-

scenario range for a given model, and (iii) model uncertainty, quantified using the inter-model range for a given scenario. Our quantification of the former – drift uncertainty – is novel.

Sanderson (2020) previously explored the contribution of drift to uncertainty in global climate sensitivity indices. To do this,
375 Sanderson added drift (via spurious forcing) and inter-annual variability (via noise) to a simple two-timescale response model. In contrast, we have quantified drift uncertainty directly from global climate model data.

We have compared two alternative methods of correcting drift in cumulatively-integrated fluxes (such as $\int R$ and $\int H$). When applying *trend-method* MCDC, the flux is first integrated cumulatively, then the spurious trend is quantified and subtracted. When applying *integrated-bias-method* MCDC, the bias in the flux is first quantified and subtracted, then the bias-corrected
380 flux is integrated cumulatively. Conceptually, these two alternative approaches are similar: a constant bias in flux (e.g. R) will drive a spurious trend in cumulatively-integrated flux (e.g. $\int R$) (Hobbs et al., 2016). In practice, we find that these two alternative methods produce differing results, even if only a single estimate of drift is used (Fig. 3b and f). Researchers should clarify which method they use.

For derived coefficients such as β and ϵ, the integrated-bias-method drift uncertainty is substantial (Table 1). Trend-method
385 drift uncertainty is generally much smaller. In other words, compared with integrated-bias-method MCDC, trend-method MCDC provides a much stronger constraint on the appropriate drift correction to apply. Therefore, we recommend that trend-method MCDC should generally be preferred. Nevertheless, we recognise that there may be cases for which it makes sense to use integrated-bias-method MCDC: for example, when performing an analysis that involves both R and $\int R$, we may prefer to use integrated-bias-method MCDC for consistency.

390 Compared with other sources of uncertainty, trend-method drift uncertainty is generally small (Table 1). For β estimates, model uncertainty dominates. For ϵ estimates and ΔZ projections, model uncertainty and scenario uncertainty are both substantial. In these cases, it may be safe to ignore trend-method drift uncertainty, especially when analysing projections from model ensembles (such as CMIP6). It will often be sufficient to use a single estimate of drift, following the recommendations of Sen Gupta et al. (2013).

395 On the other hand, it may be important to quantify drift uncertainty when analysing time series with a relatively weak trend. In particular, historical time series may be susceptible to drift uncertainty (Fig. 6). For the historical period, the mean drift uncertainty in ΔZ is 15 mm (Table 1). This drift uncertainty is of comparable magnitude to the impact of omitting volcanic forcing in control simulations: Gregory et al. (2013) found that neglecting pre-industrial volcanic forcing leads to an underestimate of 5–30 mm over the period 1850–2000. Therefore, drift uncertainty deserves similar attention to that given to
400 volcanic forcing.

We hypothesise that drift uncertainty may influence the extent to which a historical simulation agrees with an observation-based estimate. Drift uncertainty should be considered when comparing historical simulations with observation-based estimates of thermosteric sea-level change (such as the estimate of Frederikse et al., 2020). Drift uncertainty should also be considered when using ocean-related variables – such as ocean heat content (Lyu et al., 2021) – as an emergent constraint for global
405 climate models.

We further hypothesise that drift uncertainty will be large for drift-susceptible time series that exhibit large internal variability, such as regional dynamic sea level (Richter et al., 2020; Hamlington et al., 2021). This hypothesis is informed by two considerations. First, large variability will lead to a large standard error for any given segment of the control time series. Second, large variability may lead to large differences between different segments. Therefore, drift uncertainty should be considered when analysing time series with large internal variability, provided that the time series is known to be susceptible to drift.

On the other hand, drift is often weak for time series unrelated to the deep ocean (Sen Gupta et al., 2013). Therefore, many analyses – such as multi-model mean hemispheric partitioning of upper ocean heat – may be insensitive to drift (Durack et al., 2014). In such cases, drift correction may be unnecessary and drift uncertainty may be ignored.

One limitation of our MCDC framework is that we ignore branch-time metadata. If branch-time metadata were to be both available and reliable, it might be possible to use this additional information to reduce the drift uncertainty. Branch-time metadata could be used in at least one of two ways. First, branch-time metadata enable higher degree polynomial fits (e.g. quadratic) to be applied when quantifying drift (Sen Gupta et al., 2013; Gleckler et al., 2016). Second, branch-time metadata provide a possible constraint on the appropriate segment of the control simulation to use, even when a linear trend is used to quantify drift. If branch-time metadata are available, a simplified estimate of drift uncertainty may be derived by comparing two linear trend estimates: (i) the trend of the appropriate segment of the control time series and (ii) the trend of the entire control time series.

In this study, we have focused on global variables. When correcting drift in multivariate regional contexts – such as might be done in preparation for dynamical downscaling – drift correction approaches may be informed by additional considerations to ensure consistency (Paeth et al., 2019).

Furthermore, we have focused on correcting drift in long-term climate simulations, which are dominated by external forcing. In contrast, shorter-term decadal simulations are further complicated by observation-based initialisation, hence decadal simulations require different drift correction methods (Choudhury et al., 2017; Hossain et al., 2021). The quantification of drift uncertainty in such contexts remains an open question.

6 Conclusions

We have developed a probabilistic technique: Monte Carlo drift correction (MCDC). MCDC has enabled us to quantify the drift uncertainty of global climate models.

We have also considered a problem related to drift: non-conservation of energy. Although energy non-conservation is partially addressed by drift correction, the energy balance of some CMIP6 models remains suspect. Following drift correction, a useful diagnostic for model evaluation is provided by the coefficient β (the fraction of excess energy absorbed by the ocean): Is $\beta \leq 1$?

We draw three conclusions about drift uncertainty:

1. When correcting drift in cumulatively-integrated fluxes (e.g. $\int R$, the excess system energy), it is generally preferable to integrate the fluxes before correcting the trends (trend-method MCDC), rather than correcting the bias before integrating the fluxes (integrated-bias-method MCDC). Trend-method drift uncertainty is smaller than integrated-bias-method drift uncertainty.

2. When analysing projections of thermosteric sea-level rise (ΔZ), drift uncertainty is much smaller than scenario uncertainty and model uncertainty. Therefore, a single estimate of drift will often be sufficient.

3. When analysing ΔZ during the historical period, however, drift uncertainty is relatively large. Between 1850–1859 and 1995–2014, the ensemble-mean drift uncertainty is 15 mm. This is comparable to the impact of omitting volcanic forcing in control simulations (Gregory et al., 2013).

We propose two hypotheses – each with an accompanying question – to be tested in future work:

1. Drift uncertainty will influence the extent to which a historical simulation agrees with observation-based estimates. In light of this, can ocean-related variables still be used as emergent constraints for the evaluation of global climate models?

2. Drift uncertainty will be large for regional variables with large internal variability, such as dynamic sea-level change. What are the implications for climate projections?

Finally, we offer one recommendation: When evaluating and analysing data that are prone to drift, researchers should consider the potential influence of drift uncertainty. Such drift uncertainty can be quantified using MCDC.

Code and data availability. The CMIP6 global climate model data can be downloaded from the Earth System Grid Federation (ESGF). The analysis code used to produce the figures and tables can be downloaded from https://doi.org/10.5281/zenodo.7488335 (Grandey, 2022).

Appendix A: Equations connecting energy fluxes, excess energy, and thermosteric sea-level rise

If the earth's climate system were in equilibrium, the incoming and outgoing energy fluxes would balance (Meyssignac et al., 2019; Wild, 2020). In reality, the energy fluxes vary due to internal variability, leading to a quasi-equilibrium state that approximates equilibrium when averaged over long timescales. When the earth's climate system experiences an external forcing – such as that caused by rising concentrations of greenhouse gases – the energy fluxes no longer balance.

A net downward *top-of-atmosphere radiative flux* (R) heats the earth's climate system (Melet and Meyssignac, 2015; von Schuckmann et al., 2020), increasing the *excess system energy* ($\int R$). Most of this excess system energy is transferred to the ocean via a net downward *sea-surface heat flux* (H), increasing the ocean heat content (Meyssignac et al., 2019) – here, we refer to the change in ocean heat content as the *excess ocean heat* ($\int H$). As the ocean warms, the water expands due to thermal expansion, causing *thermosteric sea-level rise* (ΔZ; Gregory et al., 2019).

To a first approximation, the relationships between these variables can be described using relatively simple equations. Following Melet and Meyssignac (2015), we describe the equations below.

We begin by noting that the total excess energy that enters the earth's climate system can be calculated by integrating the global-total top-of-atmosphere radiative flux (R) cumulatively over time:

470

$$\text{Excess system energy} \equiv \int R \equiv \int_{t_0}^{t} R(\tau)\,\mathrm{d}\tau \tag{A1}$$

where t_0 is the reference time. Expressed in global-mean units, R is estimated to be 0.9 W m^{-2} for the period 2010–2018 and is increasing (von Schuckmann et al., 2020).

On annual to decadal timescales and longer, most excess system energy is absorbed by the ocean (Palmer et al., 2011; Palmer and McNeall, 2014; Durack et al., 2018; Meyssignac et al., 2019). The change in ocean heat content can be calculated
475 by integrating the global-total sea-surface heat flux (H) cumulatively over time:

$$\text{Excess ocean heat} \equiv \int H \equiv \int_{t_0}^{t} H(\tau)\,\mathrm{d}\tau \tag{A2}$$

A geothermal heat flux of approximately 0.1 W m^{-2} (Davies and Davies, 2010) also contributes to ocean heat content. We ignore this geothermal heat flux – which is an order-of-magnitude smaller than R – to focus on the relationship between R and H.

480 If we assume that the fraction of excess energy absorbed by the ocean (β) is constant, then the excess ocean heat can be written as a linear function of the excess system energy:

$$\int H = \beta \int R \tag{A3}$$

For the period 2010–2018, β is estimated to be 0.90 (von Schuckmann et al., 2020).

As the ocean warms, the water expands, causing thermosteric sea-level rise (Gregory et al., 2019). The thermal expansion of
485 water varies between locations and depths, increasing with temperature, salinity, and pressure (Russell et al., 2000; Piecuch and Ponte, 2014). Remarkably, the thermal expansion coefficient is an order of magnitude larger in the warm low-latitude ocean than it is in the cold high-latitude ocean (Griffies and Greatbatch, 2012). Nevertheless, in practice, we can use a globally-representative coefficient: ϵ, the "expansion efficiency of heat" (Russell et al., 2000). Global-mean thermosteric sea-level rise (ΔZ) can then be written as a linear function of excess ocean heat (Kuhlbrodt and Gregory, 2012; Melet and Meyssignac,
490 2015):

$$\Delta Z = \epsilon \int H \tag{A4}$$

For the period 1995–2014, ϵ is estimated to be 121 $\pm$ 1 mm YJ^{-1} (Fox-Kemper et al., 2021).

In this manuscript, we express $\int R$ and $\int H$ in global-mean units of W m^{-2} yr instead of global-total units of YJ (where 1 YJ = 10^{24} J). We express ϵ in units of mm YJ^{-1} (equivalent to m RJ^{-1}, where 1 RJ = 10^{27} J). For the Gregorian and 365
495 day calendars used by all but one of the climate models (Table S1), 1 YJ $\approx$ 62 W m^{-2} yr integrated globally; conversely, 1 W m^{-2} yr $\approx$ 0.016 YJ. For the 360 day calendar used by UKESM1, 1 YJ $\approx$ 63 W m^{-2} yr.

Appendix B: Estimating β and ϵ using ordinary least squares

Informed by Eq. (A3) and Eq. (A4), we estimate β and ϵ using linear regression. Should we use ordinary least squares with an intercept or regression through the origin (Eisenhauer, 2003)?

Regression through the origin is controversial yet sometimes justifiable (Eisenhauer, 2003). In this particular case, Eq. (A3) and Eq. (A4) provide strong theoretical support for the application of regression through the origin. We would expect Eq. (A3) and Eq. (A4) to hold across the full range of the climate model time series, including the origin (because $\int R$, $\int H$, and ΔZ are initialised at zero in 1850).

However, the drift-corrected $\int R$–$\int H$ and $\int H$–ΔZ relationships often exhibit hysteresis-like behaviour during the historical period (Figs. 4a, 4d, 5a, and 5d, especially "Max intercept" and "Min intercept" lines; Sect. 4.4): when the time series transition from negative values to positive values, the $\int R$–$\int H$ and $\int H$–ΔZ relationships do not necessarily pass through the origin. Such hysteresis-like behaviour during the historical period influences the starting point of the projection time series. This undermines the appropriateness of regression through the origin.

Furthermore, we are interested primarily in linear relationships *within* the range of data for any given scenario. For example, when estimating β and ϵ for a projection scenario, we are interested in the relationship over the period 2015–2100. In light of these considerations, we use ordinary least squares with an intercept.

Author contributions. BSG: conceptualisation; data curation; formal analysis; investigation; methodology; software; visualisation; writing – original draft preparation; writing – review and editing. ZYK: validation; writing – review and editing. DS: validation; writing – review and editing. BPH: funding acquisition; supervision (supporting); writing – review and editing. JD: funding acquisition; supervision (supporting); writing – review and editing. LYC: funding acquisition; project administration; resources; supervision (lead); writing – review and editing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This Research/Project is supported by the National Research Foundation, Singapore, and National Environment Agency, Singapore under the National Sea Level Programme Funding Initiative (Award No. USS-IF-2020-3). We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

References

Bouttes, N., Gregory, J. M., and Lowe, J. A.: The Reversibility of Sea Level Rise, J. Clim., 26, 2502–2513, https://doi.org/10.1175/JCLI-D-12-00285.1, 2013.

525 Brunetti, M. and Vérard, C.: How to Reduce Long-Term Drift in Present-Day and Deep-Time Simulations?, Clim. Dyn., 50, 4425–4436, https://doi.org/10.1007/s00382-017-3883-7, 2018.

Choudhury, D., Sen Gupta, A., Sharma, A., Mehrotra, R., and Sivakumar, B.: An Assessment of Drift Correction Alternatives for CMIP5 Decadal Predictions, J. Geophys. Res. Atmospheres, 122, https://doi.org/10.1002/2017JD026900, 2017.

Cuesta-Valero, F. J., García-García, A., Beltrami, H., and Finnis, J.: First Assessment of the Earth Heat Inventory within CMIP5 Historical
530 Simulations, Earth Syst. Dyn., 12, 581–600, https://doi.org/10.5194/esd-12-581-2021, 2021.

Davies, J. H. and Davies, D. R.: Earth's Surface Heat Flux, Solid Earth, 1, 5–24, https://doi.org/10.5194/se-1-5-2010, 2010.

Durack, P., Gleckler, P., Purkey, S., Johnson, G., Lyman, J., and Boywe, T.: Ocean Warming: From the Surface to the Deep in Observations and Models, Oceanography, 31, 41–51, https://doi.org/10.5670/oceanog.2018.227, 2018.

Durack, P. J., Gleckler, P. J., Landerer, F. W., and Taylor, K. E.: Quantifying Underestimates of Long-Term Upper-Ocean Warming, Nat.
535 Clim. Change, 4, 999–1005, https://doi.org/10.1038/nclimate2389, 2014.

Eisenhauer, J. G.: Regression through the Origin, Teach. Stat., 25, 76–80, https://doi.org/10.1111/1467-9639.00136, 2003.

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, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

540 Flynn, C. M. and Mauritsen, T.: On the Climate Sensitivity and Historical Warming Evolution in Recent Coupled Model Ensembles, Atmospheric Chem. Phys., 20, 7829–7842, https://doi.org/10.5194/acp-20-7829-2020, 2020.

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y.: Ocean, Cryosphere and Sea Level Change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment
545 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.

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., Dangendorf, S., Hogarth, P., Zanna, L., Cheng, L., and Wu, Y.-H.: The Causes of Sea-Level Rise since 1900, Nature, 584, 393–397, https://doi.org/10.1038/s41586-020-2591-3, 2020.

550 Gleckler, P. J., Durack, P. J., Stouffer, R. J., Johnson, G. C., and Forest, C. E.: Industrial-Era Global Ocean Heat Uptake Doubles in Recent Decades, Nat. Clim. Change, 6, 394–398, https://doi.org/10.1038/nclimate2915, 2016.

Good, P., Sellar, A., Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Kuhlbrodt, T., and Walton, J.: MOHC UKESM1.0-LL Model Output Prepared for CMIP6 ScenarioMIP, https://doi.org/10.22033/ESGF/CMIP6.1567, 2019.

Grandey, B. S.: D22a-Mcdc: Analysis Code for "Monte Carlo Drift Correction – Quantifying the Drift Uncertainty of Global Climate
555 Models", Zenodo, https://doi.org/10.5281/zenodo.7488335, 2022.

Gregory, J., Bi, D., Collier, M., Dix, M., Hirst, A., Hu, A., Huber, M., Knutti, R., Marsland, S., Meinshausen, M., Rashid, H., Rotstayn, L., Schurer, A., and Church, J.: Climate Models without Preindustrial Volcanic Forcing Underestimate Historical Ocean Thermal Expansion, Geophys. Res. Lett., 40, 1600–1604, https://doi.org/10.1002/grl.50339, 2013.

Gregory, J. M., Griffies, S. M., Hughes, C. W., Lowe, J. A., Church, J. A., Fukimori, I., Gomez, N., Kopp, R. E., Landerer, F., Cozannet,
560   G. L., Ponte, R. M., Stammer, D., Tamisiea, M. E., and van de Wal, R. S. W.: Concepts and Terminology for Sea Level: Mean, Variability
       and Change, Both Local and Global, Surv. Geophys., 40, 1251–1289, https://doi.org/10.1007/s10712-019-09525-z, 2019.

Griffies, S. M. and Greatbatch, R. J.: Physical Processes That Impact the Evolution of Global Mean Sea Level in Ocean Climate Models,
       Ocean Model., 51, 37–72, https://doi.org/10.1016/j.ocemod.2012.04.003, 2012.

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., Chassignet, E. P., Curchitser, E., Deshayes, J.,
565   Drange, H., Fox-Kemper, B., Gleckler, P. J., Gregory, J. M., Haak, H., Hallberg, R. W., Heimbach, P., Hewitt, H. T., Holland, D. M.,
       Ilyina, T., Jungclaus, J. H., Komuro, Y., Krasting, J. P., Large, W. G., Marsland, S. J., Masina, S., McDougall, T. J., Nurser, A. J. G., Orr,
       J. C., Pirani, A., Qiao, F., Stouffer, R. J., Taylor, K. E., Treguier, A. M., Tsujino, H., Uotila, P., Valdivieso, M., Wang, Q., Winton, M.,
       and Yeager, S. G.: OMIP Contribution to CMIP6: Experimental and Diagnostic Protocol for the Physical Component of the Ocean Model
       Intercomparison Project, Geosci. Model Dev., 9, 3231–3296, https://doi.org/10.5194/gmd-9-3231-2016, 2016.

570   Hamlington, B. D., Frederikse, T., Thompson, P. R., Willis, J. K., Nerem, R. S., and Fasullo, J. T.: Past, Present, and Future Pacific Sea-Level
       Change, Earths Future, 9, https://doi.org/10.1029/2020EF001839, 2021.

Harrison, B. J., Daron, J. D., Palmer, M. D., and Weeks, J. H.: Future Sea-Level Rise Projections for Tide Gauge Locations in South Asia,
       Environ. Res. Commun., 3, 115 003, https://doi.org/10.1088/2515-7620/ac2e6e, 2021.

Hausfather, Z., Marvel, K., Schmidt, G. A., Nielsen-Gammon, J. W., and Zelinka, M.: Climate Simulations: Recognize the 'Hot Model'
575   Problem, Nature, 605, 26–29, https://doi.org/10.1038/d41586-022-01192-2, 2022.

Hermans, T. H. J., Gregory, J. M., Palmer, M. D., Ringer, M. A., Katsman, C. A., and Slangen, A. B. A.: Projecting Global Mean Sea-Level
       Change Using CMIP6 Models, Geophys. Res. Lett., 48, https://doi.org/10.1029/2020GL092064, 2021.

Hobbs, W., Palmer, M. D., and Monselesan, D.: An Energy Conservation Analysis of Ocean Drift in the CMIP5 Global Coupled Models, J.
       Clim., 29, 1639–1653, https://doi.org/10.1175/JCLI-D-15-0477.1, 2016.

580   Hossain, M. M., Garg, N., Anwar, A. H. M. F., Prakash, M., and Bari, M.: Intercomparison of Drift Correction Alternatives for CMIP5
       Decadal Precipitation, Int. J. Climatol., 42, 1015–1037, https://doi.org/10.1002/joc.7287, 2021.

Huber, P. J.: Robust Estimation of a Location Parameter, Ann. Math. Statist., 35, 73–101, https://doi.org/10.1214/aoms/1177703732, 1964.

Huber, P. J.: Robust Statistics, Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1981.

Irving, D., Hobbs, W., Church, J., and Zika, J.: A Mass and Energy Conservation Analysis of Drift in the CMIP6 Ensemble, J. Clim., pp.
585   1–43, https://doi.org/10.1175/JCLI-D-20-0281.1, 2020.

Jevrejeva, S., Jackson, L. P., Riva, R. E. M., Grinsted, A., and Moore, J. C.: Coastal Sea Level Rise with Warming above 2 °C, Proc. Natl.
       Acad. Sci., 113, 13 342–13 347, https://doi.org/10.1073/pnas.1605312113, 2016.

Jevrejeva, S., Palanisamy, H., and Jackson, L. P.: Global Mean Thermosteric Sea Level Projections by 2100 in CMIP6 Climate Models,
       Environ. Res. Lett., 16, 014 028, https://doi.org/10.1088/1748-9326/abceea, 2021.

590   Knutti, R. and Sedláček, J.: Robustness and Uncertainties in the New CMIP5 Climate Model Projections, Nat. Clim. Change, 3, 369–373,
       https://doi.org/10.1038/nclimate1716, 2013.

Kuhlbrodt, T. and Gregory, J. M.: Ocean Heat Uptake and Its Consequences for the Magnitude of Sea Level Rise and Climate Change,
       Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL052952, 2012.

Lambert, E., Le Bars, D., Goelzer, H., and van de Wal, R. S.: Correlations Between Sea-Level Components Are Driven by Regional Climate
595   Change, Earths Future, 9, https://doi.org/10.1029/2020EF001825, 2021.

Lucarini, V. and Ragone, F.: Energetics Of Climate Models: Net Energy Balance And Meridional Enthalpy Transport, Rev. Geophys., 49, RG1001, https://doi.org/10.1029/2009RG000323, 2011.

Lyu, K., Zhang, X., and Church, J. A.: Regional Dynamic Sea Level Simulated in the CMIP5 and CMIP6 Models: Mean Biases, Future Projections, and Their Linkages, J. Clim., 33, 6377–6398, https://doi.org/10.1175/JCLI-D-19-1029.1, 2020.

600 Lyu, K., Zhang, X., and Church, J. A.: Projected Ocean Warming Constrained by the Ocean Observational Record, Nat. Clim. Change, 11, 834–839, https://doi.org/10.1038/s41558-021-01151-1, 2021.

Mayer, M., Haimberger, L., Edwards, J. M., and Hyder, P.: Toward Consistent Diagnostics of the Coupled Atmosphere and Ocean Energy Budgets, J. Clim., 30, 9225–9246, https://doi.org/10.1175/JCLI-D-17-0137.1, 2017.

Melet, A. and Meyssignac, B.: Explaining the Spread in Global Mean Thermosteric Sea Level Rise in CMIP5 Climate Models, J. Clim., 28,
605 9918–9940, https://doi.org/10.1175/JCLI-D-15-0200.1, 2015.

Meyssignac, B., Boyer, T., Zhao, Z., Hakuba, M. Z., Landerer, F. W., Stammer, D., Köhl, A., Kato, S., L'Ecuyer, T., Ablain, M., Abraham, J. P., Blazquez, A., Cazenave, A., Church, J. A., Cowley, R., Cheng, L., Domingues, C. M., Giglio, D., Gouretski, V., Ishii, M., Johnson, G. C., Killick, R. E., Legler, D., Llovel, W., Lyman, J., Palmer, M. D., Piotrowicz, S., Purkey, S. G., Roemmich, D., Roca, R., Savita, A., von Schuckmann, K., Speich, S., Stephens, G., Wang, G., Wijffels, S. E., and Zilberman, N.: Measuring Global Ocean Heat Content to
610 Estimate the Earth Energy Imbalance, Front. Mar. Sci., 6, 432, https://doi.org/10.3389/fmars.2019.00432, 2019.

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, Geosci. Model Dev., 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

Paeth, H., Li, J., Pollinger, F., Müller, W. A., Pohlmann, H., Feldmann, H., and Panitz, H.-J.: An Effective Drift Correction for Dynamical
615 Downscaling of Decadal Global Climate Predictions, Clim. Dyn., 52, 1343–1357, https://doi.org/10.1007/s00382-018-4195-2, 2019.

Palmer, M. D. and McNeall, D. J.: Internal Variability of Earth's Energy Budget Simulated by CMIP5 Climate Models, Environ. Res. Lett., 9, 034 016, https://doi.org/10.1088/1748-9326/9/3/034016, 2014.

Palmer, M. D., McNeall, D. J., and Dunstone, N. J.: Importance of the Deep Ocean for Estimating Decadal Changes in Earth's Radiation Balance, Geophys. Res. Lett., 38, https://doi.org/10.1029/2011GL047835, 2011.

620 Palmer, M. D., Harris, G. R., and Gregory, J. M.: Extending CMIP5 Projections of Global Mean Temperature Change and Sea Level Rise Due to Thermal Expansion Using a Physically-Based Emulator, Environ. Res. Lett., 13, 084 003, https://doi.org/10.1088/1748-9326/aad2e4, 2018.

Palmer, M. D., Gregory, J. M., Bagge, M., Calvert, D., Hagedoorn, J. M., Howard, T., Klemann, V., Lowe, J. A., Roberts, C. D., Slangen, A. B. A., and Spada, G.: Exploring the Drivers of Global and Local Sea-Level Change Over the 21st Century and Beyond, Earths Future, 8,
625 https://doi.org/10.1029/2019EF001413, 2020.

Piecuch, C. G. and Ponte, R. M.: Mechanisms of Global-Mean Steric Sea Level Change, J. Clim., 27, 824–834, https://doi.org/10.1175/JCLI-D-13-00373.1, 2014.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., Kc, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P.,
630 Humpenöder, F., Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The Shared Socioeconomic Pathways and Their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview, Glob. Environ. Change, 42, 153–168, https://doi.org/10.1016/j.gloenvcha.2016.05.009, 2017.

Richter, K., Meyssignac, B., Slangen, A. B. A., Melet, A., Church, J. A., Fettweis, X., Marzeion, B., Agosta, C., Ligtenberg, S. R. M., Spada, G., Palmer, M. D., Roberts, C. D., and Champollion, N.: Detecting a Forced Signal in Satellite-Era Sea-Level Change, Environ. Res. Lett., 15, 094 079, https://doi.org/10.1088/1748-9326/ab986e, 2020.

Russell, G. L., Gornitz, V., and Miller, J. R.: Regional Sea Level Changes Projected by the NASA/GISS Atmosphere-Ocean Model, Clim. Dyn., 16, 789–797, https://doi.org/10.1007/s003820000090, 2000.

Sanderson, B.: Relating Climate Sensitivity Indices to Projection Uncertainty, Earth Syst. Dyn., 11, 721–735, https://doi.org/10.5194/esd-11-721-2020, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, J. Adv. Model. Earth Syst., 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Sen Gupta, A., Jourdain, N. C., Brown, J. N., and Monselesan, D.: Climate Drift in the CMIP5 Models, J. Clim., 26, 8597–8615, https://doi.org/10.1175/JCLI-D-12-00521.1, 2013.

Slangen, A. B. A., Meyssignac, B., Agosta, C., Champollion, N., Church, J. A., Fettweis, X., Ligtenberg, S. R. M., Marzeion, B., Melet, A., Palmer, M. D., Richter, K., Roberts, C. D., and Spada, G.: Evaluating Model Simulations of Twentieth-Century Sea Level Rise. Part I: Global Mean Sea Level Change, J. Clim., 30, 8539–8563, https://doi.org/10.1175/JCLI-D-17-0110.1, 2017.

Stouffer, R. J.: Time Scales of Climate Response, J. Clim., 17, 209–217, https://doi.org/10.1175/1520-0442(2004)017<0209:TSOCR>2.0.CO;2, 2004.

Stouffer, R. J., Weaver, A. J., and Eby, M.: A Method for Obtaining Pre-Twentieth Century Initial Conditions for Use in Climate Change Studies, Clim. Dyn., 23, 327–339, https://doi.org/10.1007/s00382-004-0446-5, 2004.

Tang, Y., Rumbold, S., Ellis, R., Kelley, D., Mulcahy, J., Sellar, A., Walton, J., and Jones, C.: MOHC UKESM1.0-LL Model Output Prepared for CMIP6 CMIP, https://doi.org/10.22033/ESGF/CMIP6.1569, 2019.

Tebaldi, C. and Knutti, R.: The Use of the Multi-Model Ensemble in Probabilistic Climate Projections, Phil. Trans. R. Soc. A, 365, 2053–2075, https://doi.org/10.1098/rsta.2007.2076, 2007.

von Schuckmann, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., Aich, V., Adusumilli, S., Beltrami, H., Boyer, T., Cuesta-Valero, F. J., Desbruyères, D., Domingues, C., García-García, A., Gentine, P., Gilson, J., Gorfer, M., Haimberger, L., Ishii, M., Johnson, G. C., Killick, R., King, B. A., Kirchengast, G., Kolodziejczyk, N., Lyman, J., Marzeion, B., Mayer, M., Monier, M., Monselesan, D. P., Purkey, S., Roemmich, D., Schweiger, A., Seneviratne, S. I., Shepherd, A., Slater, D. A., Steiner, A. K., Straneo, F., Timmermans, M.-L., and Wijffels, S. E.: Heat Stored in the Earth System: Where Does the Energy Go?, Earth Syst. Sci. Data, 12, 2013–2041, https://doi.org/10.5194/essd-12-2013-2020, 2020.

Wild, M.: The Global Energy Balance as Represented in CMIP6 Climate Models, Clim. Dyn., 55, 553–577, https://doi.org/10.1007/s00382-020-05282-7, 2020.

Uncorrected time series and $\int R - \int H - \Delta Z$ relationships for the UKESM1−0−LL control & historical simulations, demonstrating the problem of drift
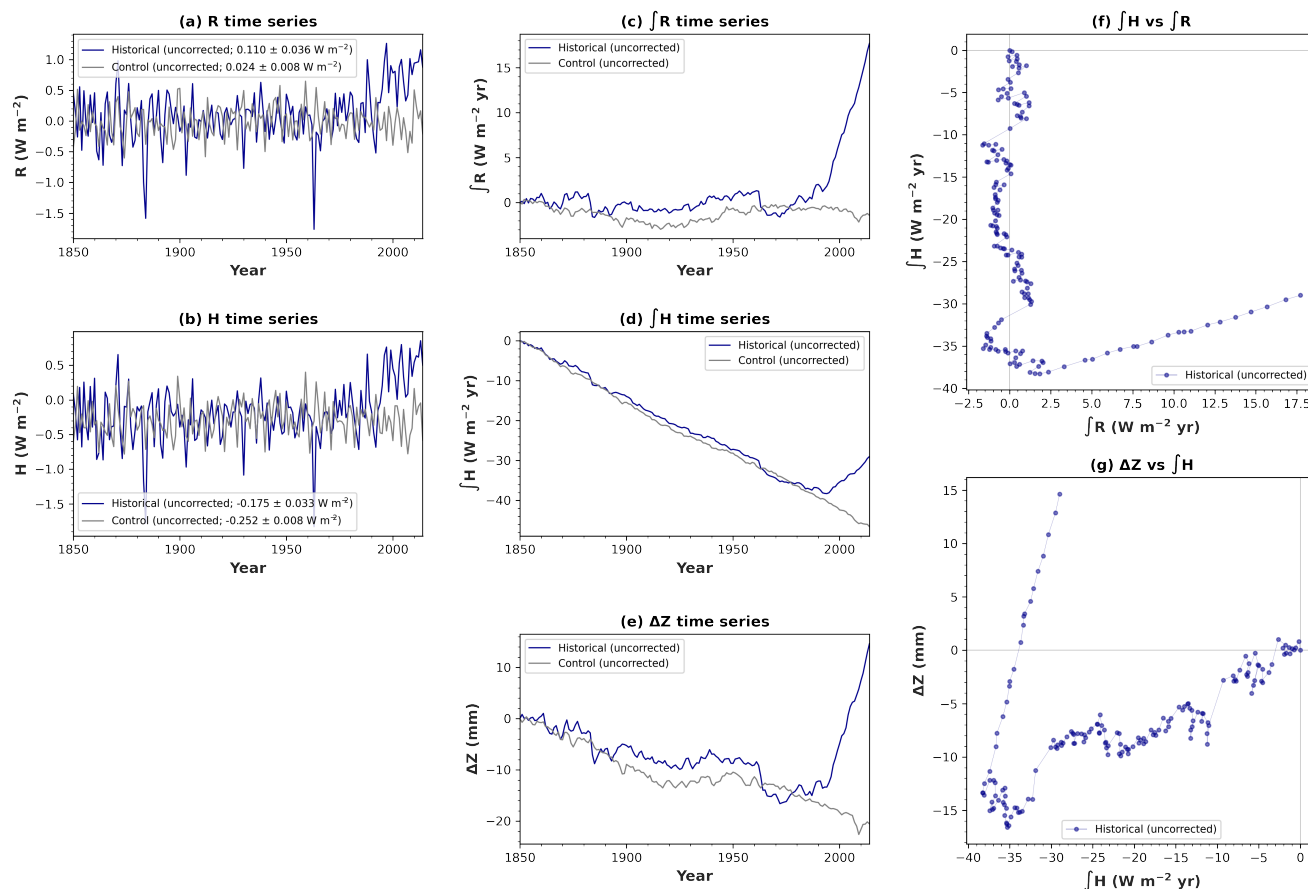
Figure 1. Uncorrected time series and bivariate relationships for the UK Earth System Model (UKESM1). These *uncorrected* time series have not yet been corrected for drift. Using data from the control and historical simulations, time series are shown for the following global variables: (a) top-of-atmosphere radiative flux (R), (b) sea-surface heat flux (H), (c) excess system energy ($\int R$), (d) excess ocean heat ($\int H$), and (e) thermosteric sea-level rise (ΔZ). Global-total R and H are expressed in global-mean units of W m^{-2}; global-total $\int R$ and $\int H$ are expressed in global-mean units of W m^{-2} yr. In (a) and (b), the legend shows the mean of the entire time series and the standard error of the mean. In (a)–(e), the control time series are arbitrarily shifted to start in the year 1850, for ease of comparison with the historical time series. The length of the control time series is 1100 years, but only the first 165 years are shown (in contrast to Fig. 2). For the historical simulation, bivariate relationships are also shown: (f) $\int H$ versus $\int R$ and (g) ΔZ versus $\int H$. Interpretation is offered in Sect. 4.1.

Figure 2. Trend-method Monte Carlo drift correction (MCDC). The first row (a–d) shows results for excess system energy ($\int R$). (a) Five hundred trend samples are produced by sampling 500 150 yr segments of the uncorrected control time series. (b) The trend samples are displayed as a histogram. (c) Each trend sample is subtracted from the uncorrected control time series, producing 500 trend-corrected control time series. (d) Each trend sample (derived from the control time series) is also subtracted from the uncorrected historical time series, producing 500 trend-corrected historical time series. The second row (e–h) shows results for excess ocean heat ($\int H$). The third row (i–l) shows results for thermosteric sea-level rise (ΔZ). Trend-method MCDC is described in more detail in Sect. 3.2.
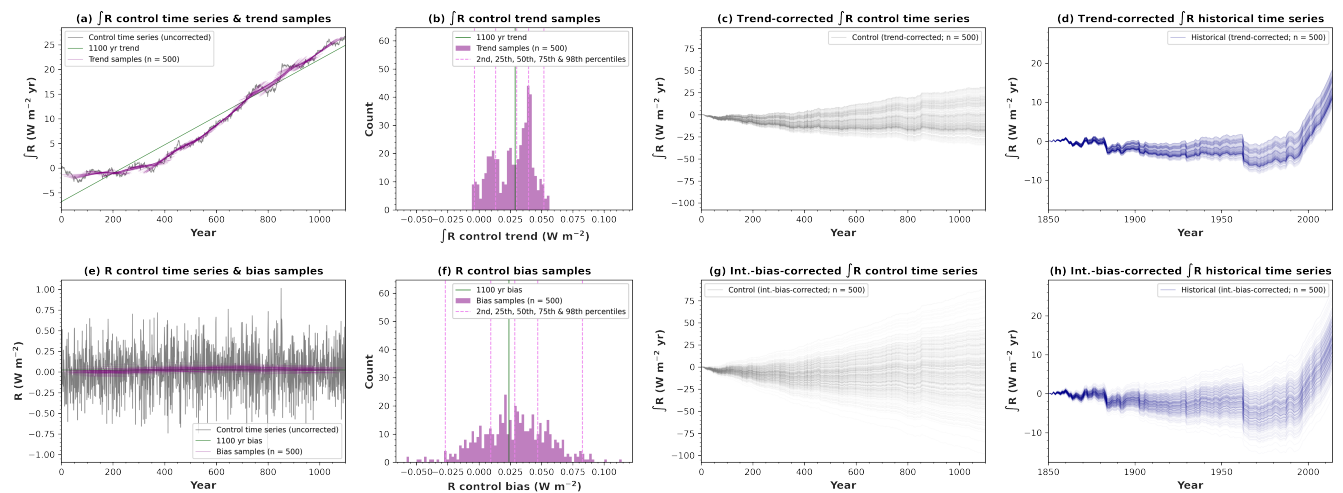
Figure 3. Integrated-bias-method MCDC (second row) compared with trend-method MCDC (first row), applied to excess system energy ($\int R$). The first row (a–d) shows trend-method MCDC results. (Although these trend-method MCDC results have already been shown in Fig. 2a–d above, they are reproduced here – with expanded axis ranges in (b)–(d) – for ease of comparison with the integrated-bias-method MCDC results.) The second row (e–h) shows integrated-bias-method MCDC results. (e) Five hundred bias samples are produced by sampling 500 150 yr segments of the uncorrected control time series of top-of-atmosphere radiative flux (R). (f) The bias samples are displayed as a histogram. (g) Each bias sample is subtracted from the uncorrected R control time series, before integrating the bias-corrected R control time series cumulatively to produce 500 integrated-bias-corrected $\int R$ control time series. (h) Each bias sample (derived from the R control time series) is also subtracted from the uncorrected R historical time series, before integrating cumulatively to produce 500 integrated-bias-corrected $\int R$ historical time series. Integrated-bias-method MCDC is described in more detail in Sect. 3.3.
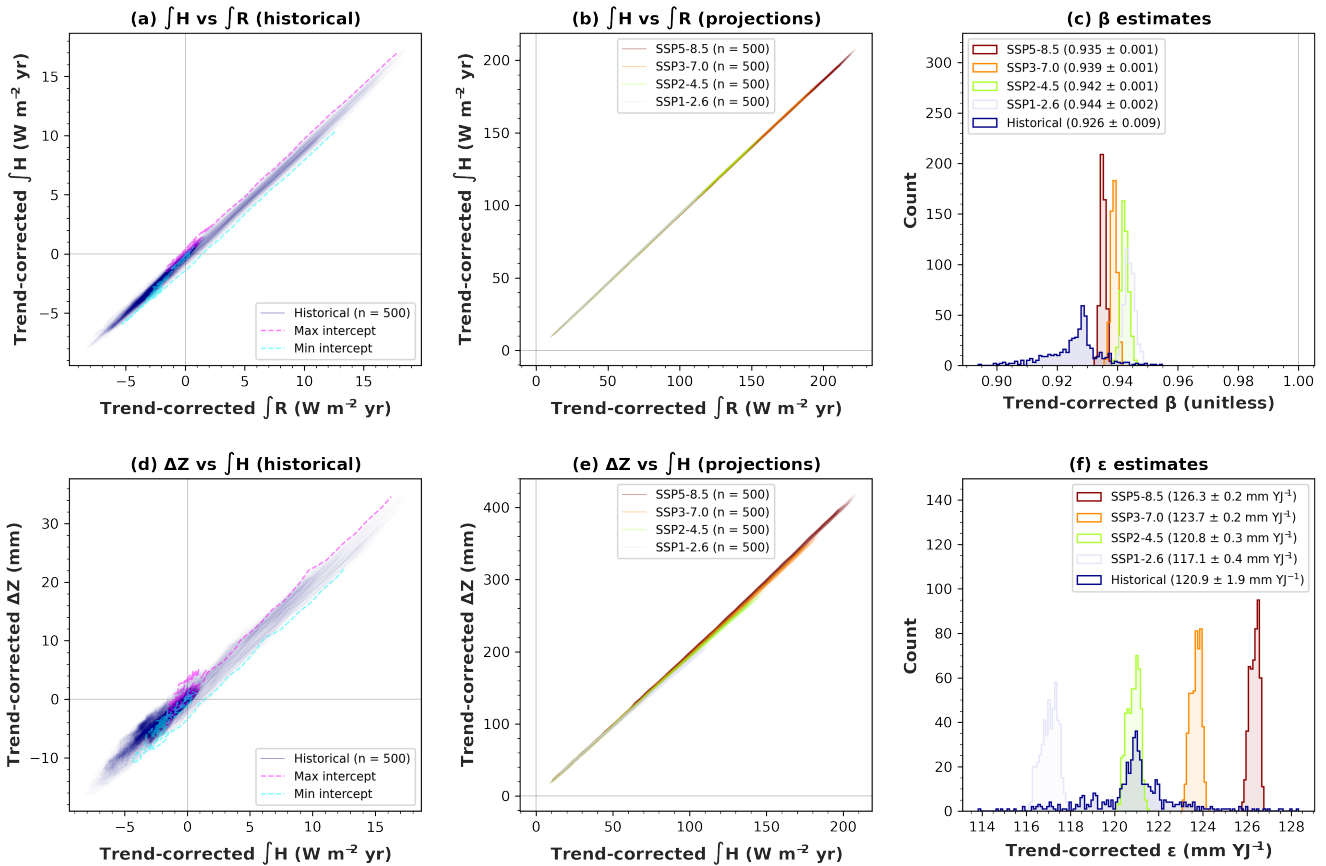
Figure 4. Trend-corrected bivariate relationships and regression coefficients. (a) Trend-corrected excess ocean heat ($\int H$) versus trend-corrected excess system energy ($\int R$) for the historical simulation. The legend shows the number of MCDC samples ($n = 500$). "Max intercept" refers to the line with the maximum y-intercept; "Min intercept" refers to the line with the minimum y-intercept. (b) Trend-corrected $\int H$ versus trend-corrected $\int R$ for the four projection simulations. (c) Trend-corrected estimates of the fraction of excess energy absorbed by the ocean (β; Eq. A3), calculated as the linear regression coefficient of trend-corrected $\int H$ versus trend-corrected $\int R$ using ordinary least squares (Appendix B). The legend shows the mean and standard deviation (*not* standard error) for each scenario. (d) Trend-corrected thermosteric sea-level rise (ΔZ) versus trend-corrected $\int H$ for the historical simulation. (e) Trend-corrected ΔZ versus trend-corrected $\int H$ for the four projection simulations. (f) Trend-corrected estimates of the expansion efficiency of heat (ϵ; Eq. A4), calculated as the linear regression coefficient of trend-corrected ΔZ versus trend-corrected $\int H$. (See Appendix A for a brief explanation of the relationship between the units of ϵ and the units of $\int H$).
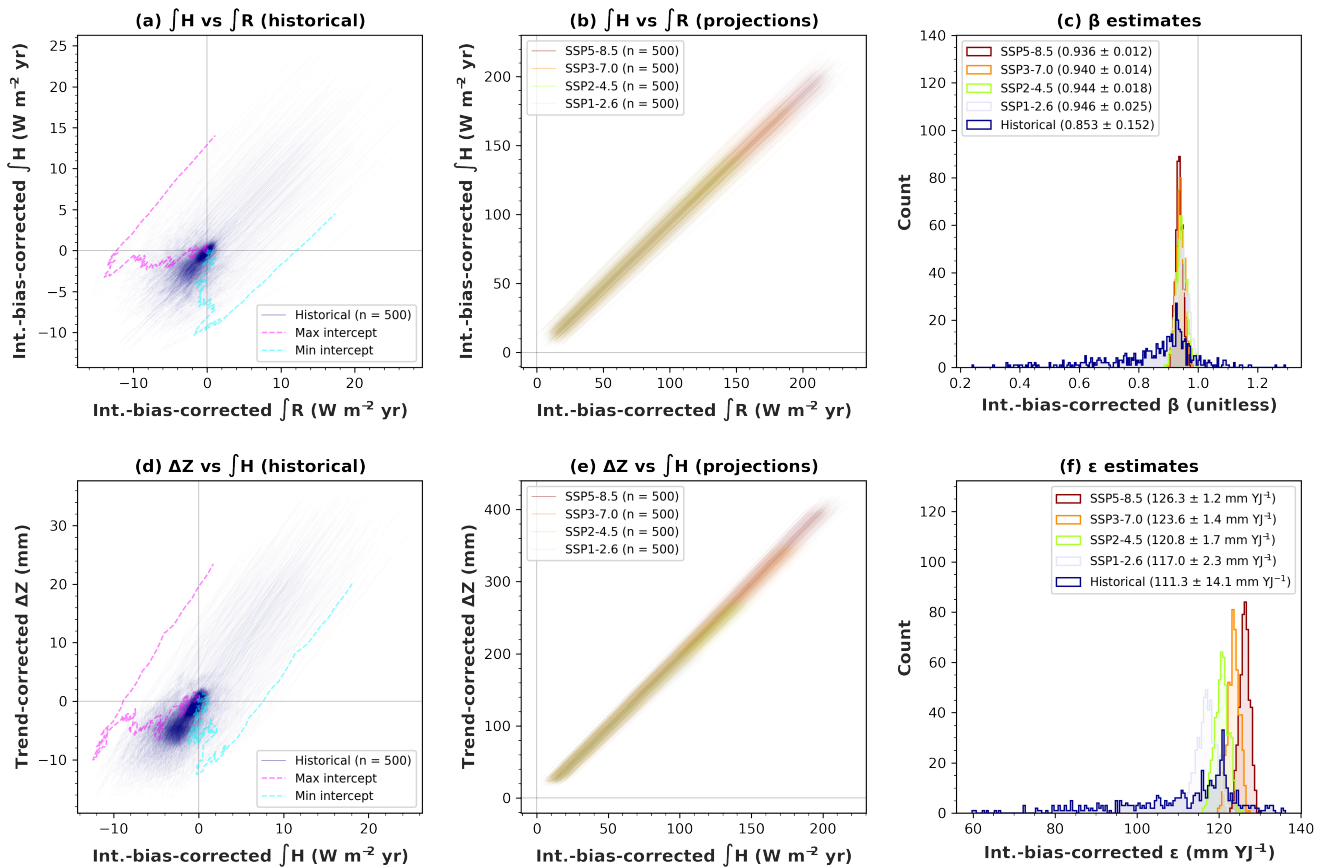
Figure 5. Integrated-bias-corrected bivariate relationships and regression coefficients. (a) Integrated-bias-corrected excess ocean heat ($\int H$) versus integrated-bias-corrected excess system energy ($\int R$) for the historical simulation. (b) Integrated-bias-corrected $\int H$ versus integrated-bias-corrected $\int R$ for the four projection simulations. (c) Integrated-bias-corrected estimates of the fraction of excess energy absorbed by the ocean (β), calculated as the linear regression coefficient of integrated-bias-corrected $\int H$ versus integrated-bias-corrected $\int R$. The legend shows the mean and standard deviation for each scenario. (d) Trend-corrected thermosteric sea-level rise (ΔZ) versus integrated-bias-corrected $\int H$ for the historical simulation. (Integrated-bias-method MCDC cannot be applied to ΔZ, so trend-method MCDC is used instead.) (e) Trend-corrected ΔZ versus integrated-bias-corrected $\int H$ for the four projection simulations. (f) Integrated-bias-corrected estimates of the expansion efficiency of heat (ϵ), calculated as the linear regression coefficient of trend-corrected ΔZ versus integrated-bias-corrected $\int H$.
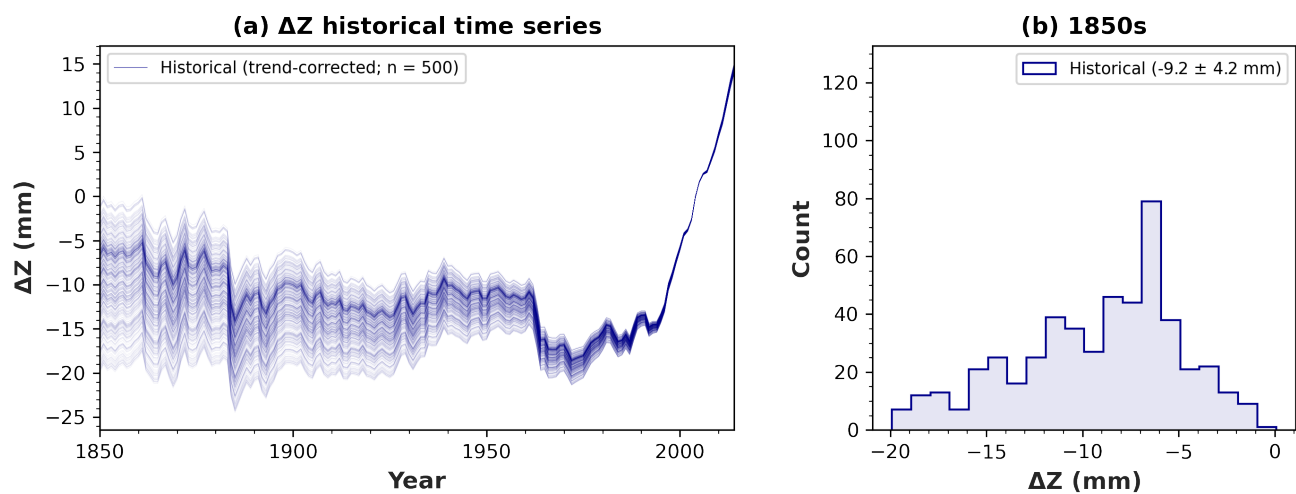
Figure 6. Trend-corrected thermosteric sea-level rise (ΔZ) for the historical simulation. (a) Historical time series of ΔZ. The legend shows the number of MCDC samples ($n = 500$). (b) Histogram of ΔZ averaged across 1850–1859. The legend shows the mean and standard deviation (*not* standard error). Results are shown relative to 1995–2014.
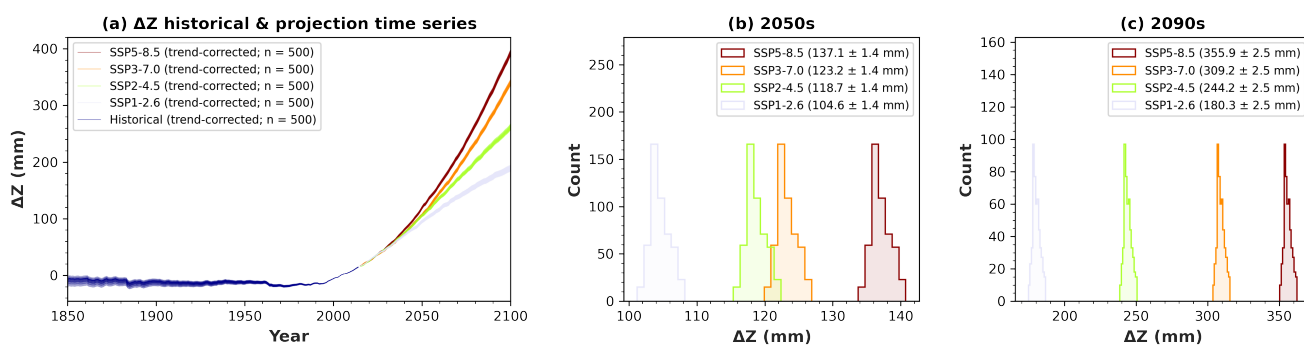
Figure 7. Trend-corrected thermosteric sea-level rise (ΔZ) for the projection simulations. (a) Time series of ΔZ. (b) Histogram of ΔZ averaged across 2050–2059. The legend shows the mean and standard deviation for each scenario. (c) Histogram of ΔZ averaged across 2090–2099. Results are shown relative to 1995–2014.

Table 1. Sources of uncertainty in β (fraction of excess energy absorbed by the ocean), ϵ (expansion efficiency of heat), and ΔZ (thermosteric sea-level rise, relative to 1995–2014), for the CMIP6 ensemble. *Drift uncertainty* is derived from the 2nd–98th inter-percentile range of the drift-corrected data. *Scenario uncertainty* is derived from the inter-scenario range. *Model uncertainty* is derived from the inter-model range. For β and ϵ, drift uncertainty is calculated using both trend-method MCDC and integrated-bias-method MCDC. Scenario uncertainty and model uncertainty are relatively insensitive to the choice of drift correction method, so these sources of uncertainty are shown for trend-method MCDC only. When calculating the uncertainty in β, ϵ, and ΔZ for future decades, the four projection scenarios are used (but not the historical scenario). When calculating the uncertainty in ΔZ for the 1850s, the single historical scenario is used instead, hence scenario uncertainty cannot be calculated for the 1850s. Expanded results – showing the uncertainties for different models and scenarios – are shown in Table S2. The expanded results (Table S2) have been averaged across models and scenarios to produce Table 1.

Source of uncertainty	Uncertainty in β (unitless)		Uncertainty in ϵ (mm YJ^{-1})		Uncertainty in ΔZ (mm)		
	Trend-method	Int.-bias-method	Trend-method	Int.-bias-method	1850s	2050s	2090s
Drift uncertainty	0.01	0.11	1.4	9.6	15	5	9
Scenario uncertainty	0.02		7.6			23	138
Model uncertainty	0.17		12.2		69	45	98