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 samples the standard error associated with drift in the control time series. We apply MCDC to an ensemble of global climate models from the Coupled Model Intercomparison Project Phase 6 (CMIP6). We find that drift correction partially addresses a problem related to drift: energy leakage. 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 drift: an alternative method would be to correct the bias before integrating the flux, but this alternative method amplifies the drift uncertainty. Assuming that drift is linear likely leads to an underestimation of drift uncertainty. Time series with weak trends may be especially susceptible to drift uncertainty: for historical thermosteric sea-level rise since the 1850s, the drift uncertainty can range from 3 to 24 mm, which is of comparable magnitude to the impact of omitting volcanic forcing in control simulations. Derived coefficients – such as the ocean's expansion efficiency of heat – can also be susceptible to drift uncertainty. 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 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., 2021), causing spurious trends in output variables such as thermosteric sea-level rise. Drift may also influence derived climate indices (Sanderson, 2020). 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 state (Stouffer, 2004; Stouffer et al., 2004; Eyring et al., 2016). Therefore, variables influenced by the deep ocean are especially susceptible to drift (Sen Gupta et al., 2013). The 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., 2021). Even when climate models are spun-up for long periods under equilibrium forcing, modelled variables may be inconsistent with a true quasi-equilibrium state – spurious trends remain (Hobbs et al., 2016). Of particular relevance here is the problem of *energy leakage* (Irving et al., 2021): 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 Vérard, 2018).

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 leakage (Hobbs et al., 2016; Irving et al., 2021). 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 of a state variable (Sen Gupta et al., 2013). This linear trend can then be subtracted from transient time series – such as time series produced by historical or projection 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. To do this, we propose a Monte Carlo drift correction (MCDC) framework (Sect. 3.2). We explore three alternative MCDC methods. Using MCDC, we produce drift-corrected time series of excess system energy (ΔE), excess ocean heat (ΔH), and thermosteric sea-level rise (ΔZ). Using the drift-corrected time series, we quantify the drift uncertainty associated with ΔE, ΔH, ΔZ during the historical period. We also quantify the drift uncertainty associated with the fraction of energy absorbed by the ocean (η) and the expansion efficiency of heat (ϵ). Our overarching aim is to *quantify drift uncertainty*. To support this aim, we ask two primary research questions:

1. Do the three alternative MCDC methods produce similar estimates of drift uncertainty? If not, which method is preferable?

2. What is the contribution of drift uncertainty to ΔE, ΔH, ΔZ, η, and ϵ?

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.

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

1. Net downward *top-of-atmosphere radiative flux* (E') is calculated as $E' = rsdt - rsut - rlut$, 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 (ΔE) is calculated by integrating E' cumulatively (Eq. A1). *Excess ocean heat* (ΔH) is calculated by integrating H' cumulatively (Eq. A2). We use the 1850s (1850–1859) as the reference period for $\Delta E = 0$, $\Delta H = 0$, and $\Delta Z = 0$. The *fraction of energy absorbed by the ocean* (η) is calculated as the linear regression coefficient of ΔH versus ΔE (Eq. A3). The *expansion efficiency of heat* (ϵ) is calculated as the linear regression coefficient of ΔZ versus ΔH (Eq. A4). The coefficients η and ϵ are both estimated using ordinary least squares (Appendix B).

Our MCDC framework uses the *control* simulations (Eyring et al., 2016). These control simulations follow equilibrium forcing representative of the year 1850. The time at which a historical simulation branches from a control simulation is called the *branch time*. We use the branch-time metadata to assign corresponding dates to each control simulation, so that the year 1850 of the control simulation corresponds to the start of the historical simulation.

We apply MCDC to the *historical* simulations (Eyring et al., 2016). 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). The projection

simulations are essentially continuations of the historical simulation, beginning where the historical simulation ends. Therefore, when processing the projection time series and applying MCDC, we prepend the historical period (1850–2014) before the projection period (2015–2100), so that the combined time series covers the period 1850–2100. For example, the cumulative integration that produces the ΔE SSP5-8.5 time series begins in 1850 (not 2015), and ΔE is subsequently referenced to the 1850s decadal mean. We subsequently remove the historical period from the projection time series before we calculate η and ϵ, to distinguish the projection simulations from one another and from the historical simulation.

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). Three of the figures presented in this manuscript show results for UKESM1 (Figs. 1, 2, and 4). Nevertheless, our analysis also includes the larger ensemble (Figs. 3, 5, and S1–S3; Tables 1 and S2–S6).

3 Drift correction methods

3.1 Conventional drift correction using a single estimate of drift

Sen Gupta et al. (2013) considered whether to use all or only part of the control time series. They recommended using the entire control time series to estimate the drift, "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, we should use the entire control time series.

Sen Gupta et al. (2013) also considered different drift correction methods, including linear, quadratic, and cubic drift correction approaches. Quadratic and cubic drift correction both risk 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. Among recent studies focusing on the earth's energy budget or sea-level change, a few consider the possibility that results may be sensitive to alternative linear, quadratic, and/or cubic models of drift (e.g. Sen Gupta et al., 2013; Hobbs et al., 2016; Jackson and Jevrejeva, 2016; Lyu et al., 2021; Hermans et al., 2021; Irving et al., 2021). However, most studies use only a single statistical model of drift: either linear (Jevrejeva et al., 2016; Palmer et al., 2018; Cuesta-Valero et al., 2021; Hamlington et al., 2021; Lambert et al., 2021), quadratic (e.g. Gleckler et al., 2016; Lyu et al., 2020; Harrison et al., 2021; Jevrejeva et al., 2021), or cubic (e.g. Irving et al., 2019). We observe that researchers generally select a statistical model

a priori, before analysing any data. They do not generally compare the *a posteriori* performance of alternative statistical models.

3.2 Monte Carlo drift correction (MCDC)

To quantify drift uncertainty, we propose a Monte Carlo drift correction (MCDC) framework. Following Sen Gupta et al. (2013), MCDC estimates the drift using the entire control time series. MCDC also samples the standard error associated with the estimated drift. MCDC proceeds as follows:

1. Select a statistical model of drift, such as a linear trend or quadratic polynomial. Alternative models of drift correspond to alternative MCDC methods: for example, a linear trend corresponds to linear-method MCDC (see below).

2. Using ordinary least squares, fit the statistical model to the entire control time series of a variable of interest (e.g. ΔE). This provides the best estimate of the drift, corresponding to a conventional drift correction approach.

3. Estimate the standard error associated with each parameter of the fitted model. To account for autocorrelation in the residuals, we use a heteroskedasticity and autocorrelation consistent covariance matrix (Newey and West, 1987).

4. For each parameter, randomly draw samples from a Gaussian distribution with mean equal to the best estimate (step 2) and standard deviation equal to the standard error (step 3). Use these samples to construct *drift samples*. Here, we draw 1500 drift samples, sufficient to quantify drift uncertainty using the 2nd–98th inter-percentile range.

5. For each scenario of interest (e.g. historical), subtract each drift sample from the time series to produce drift-corrected time series. Each drift sample will produce a different drift-corrected time series, enabling quantification of drift uncertainty.

Different statistical models of drift (step 1) correspond to different MCDC methods. In this paper, we apply three MCDC methods: integrated-bias-method MCDC, linear-method MCDC, and agnostic-method MCDC.

Integrated-bias-method MCDC can be applied to ΔE and ΔH, both of which are derived by cumulatively integrating a flux (E' or H'). MCDC is first applied to the flux by modelling the "drift" as a constant: in other words, we assume that the flux contains a constant bias. This application of MCDC to the flux could be referred to as *Monte Carlo bias correction*. Each bias-corrected E' (or H') time series is subsequently integrated cumulatively to produce an integrated-bias-corrected ΔE (or ΔH) time series.

Linear-method MCDC models drift in a state variable (e.g. ΔE, ΔH, or ΔZ) using a linear trend. In other words, we assume that the underlying drift is linear. For ΔE and ΔH, the assumption of linearity is consistent conceptually with the assumption that the underlying flux contains a constant bias. The assumption of linearity will likely contribute to an underestimation of drift uncertainty (Sect. 5.2).

Agnostic-method MCDC relaxes the assumption of linearity. In addition to sampling the uncertainty associated with the parameters of a given statistical model, agnostic-method MCDC also samples the uncertainty associated with the choice

between alternative statistical models: we assume that linear, quadratic, and cubic models of drift are equally valid. This corresponds to the common practice of selecting one of these alternative statistical models *a priori* (Sect. 3.1). For each of the three models of drift, we draw 500 drift samples. We then combine these samples, producing 1500 drift samples in total.

When applying a quadratic or cubic model of drift, the branch time must be known: which part of the control time series parallels the historical time series? For earlier generations of global climate model ensembles, this branch-time metadata may have been either unavailable or unreliable (Gleckler et al., 2016; Flynn and Mauritsen, 2020). To address this, Irving et al. (2019) proposed an alternative method to identify the branch time. Here, we assume that the branch-time metadata that accompany the CMIP6 data are reliable. When defining and fitting each statistical model of drift, we use the year 1850 as the reference year that corresponds to the origin.

4 Results

4.1 Uncorrected time series

An example of drift is illustrated in Fig. 1. For UKESM1's control simulation, the top-of-atmosphere radiative flux (E') is close to zero (Fig. 1a), consistent with quasi-equilibrium radiative balance at the top of the atmosphere. The UKESM1 team have achieved their stated goal of keeping E' close to zero when tuning the model (Sellar et al., 2019). Therefore, the drift in excess system energy (ΔE) 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 (ΔH; Fig. 1d). This illustrates the connection between bias and drift: a constant bias in a flux will drive a linear trend in the cumulatively-integrated flux (Hobbs et al., 2016). In this example, the drift in ΔH 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 ΔE–ΔH relationship should be approximately linear (Eq. A3). In contrast, for UKESM1's historical simulation, the drift in ΔH drives a strange ΔE–ΔH relationship (Fig. 1f): for most of the historical period, ΔE remains close to zero, but ΔH decreases. This reveals energy leakage. Furthermore, the inconsistency between ΔE and ΔH 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 sea-level rise signal during the 20th century. The drift will also contaminate future projections. Furthermore, the ΔH–ΔZ relationship (Fig. 1g) is inconsistent with Eq. (A4).

4.2 Excess system energy (ΔE)

Fig. 2 illustrates the application of MCDC to UKESM1's ΔE time series. For each MCDC method, we produce 1500 drift samples using the uncorrected control time series (Fig. 2a, d, and g). For UKESM1, these drift samples all indicate positive drift during the historical period. We subtract these drift samples from the uncorrected control time series to produce drift-corrected control time series (Fig. 2b, e, and h). In comparison to the uncorrected time series, the trends of these drift-corrected control time series are much closer to zero, as expected. We also subtract the drift samples from the uncorrected historical time series to produce drift-corrected historical time series (Fig. 2c, f, and i).

The spread among the drift samples and the drift-corrected time series indicates drift uncertainty. For UKESM1, the drift uncertainty is largest when integrated-bias-method MCDC is used (Fig. 2a–c). The linear-method drift uncertainty is much smaller (Fig. 2d–f). This demonstrates that integrated-bias-method MCDC and linear-method MCDC produce different estimates of drift uncertainty, even though these two methods have similar underlying assumptions – namely, a constant bias in E' drives a linear trend in ΔE. We ascribe the difference between the two methods to the size of the standard error: even if autocorrelation is accounted for, the standard error of the trend of ΔE is smaller than the standard error of the mean of E', because integrating cumulatively effectively averages over the substantial inter-annual variability in E'. Therefore, integrated-bias-method MCDC provides a much weaker constraint on the drift correction. To minimise drift uncertainty, it is preferable to use linear-method MCDC.

However, the underlying drift may be non-linear. If so, a linear model of drift may be inappropriate. Therefore, linear-method MCDC will likely underestimate drift uncertainty. By partially accounting for the possibility of non-linearity, agnostic-method MCDC should provide a more accurate estimate of drift uncertainty. For UKESM1, agnostic-method drift uncertainty is larger than linear-method drift uncertainty, yet smaller than integrated-bias drift uncertainty (Fig. 2; Table S2).

We extend our analysis to other global climate models within the CMIP6 ensemble by quantifying the drift uncertainty associated with ΔE during the historical period (2000s relative to 1850s; Fig. 3; Table S2). For all models, the linear-method drift uncertainty is substantially smaller than the integrated-bias-method drift uncertainty. The agnostic-method drift uncertainty is always larger than the linear-method drift uncertainty, and is often larger than the integrated-bias-method drift uncertainty. When averaged across the ensemble, agnostic-method MCDC produces the largest estimate of drift uncertainty for ΔE, with a median of 0.09 YJ (Table 1). The ensemble maximum drift uncertainty is 0.17 YJ, approximately twice as large as the ensemble median (Table 1; Fig. 3h).

Differences between global climate models also drive differences in ΔE (Fig. 3). We refer the inter-model range as *model uncertainty* (Tebaldi and Knutti, 2007; Knutti and Sedláček, 2013). For ΔE, the model uncertainty of 0.56 YJ is more than three times as large as the ensemble-maximum drift uncertainty (Table 1).

4.3 Excess ocean heat (ΔH)

For each CMIP6 model, the drift uncertainty associated with ΔH (Fig. S1; Table S3) is similar to the drift uncertainty associated with ΔE. The ensemble statistics are very similar to those of ΔE (Table 1). Therefore, we can draw very similar

conclusions for the ΔH results as we have for the ΔE results. This is not surprising because ΔH is closely related to ΔE: both consist of cumulatively-integrated fluxes, and ΔH should track ΔE closely if the fraction of excess energy absorbed by the ocean (η) is close to 1.0 (Sect. 4.5; Appendix A).

4.4 Thermosteric sea-level rise (ΔZ)

For ΔZ over the historical period (2000s relative to 1850s), the agnostic-method drift uncertainty is always larger than the linear-method drift uncertainty, as expected (Fig. S2; Table S4). (Integrated-bias-method MCDC is not applicable to ΔZ, which does not consist of a cumulatively-integrated flux.) The agnostic-method drift uncertainty ranges from 3 mm to 24 mm, with an ensemble median of 10 mm (Table 1). This is smaller than the model uncertainty of 64 mm (Table 1). On the other hand, the 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. Furthermore, for models with relatively large drift uncertainty, the drift uncertainty may influence the extent to which the historical time series agrees with estimates of ΔZ derived from observations, such as the reconstructions of Zanna et al. (2019) and Frederikse et al. (2020). For two CMIP6 models, the agnostic-method drift uncertainty is large enough to obscure the positive ΔZ signal during the historical period: the 2nd–98th percentile range includes negative values of ΔZ (Fig. S2a and n).

4.5 Fraction of excess energy absorbed by the ocean (η)

If energy is conserved and if the fraction of excess energy absorbed by the ocean (η) is constant, then excess ocean heat (ΔH) should be a linear function of excess system energy (ΔE): $\Delta H = \eta \Delta E$ (Eq. A3). However, as noted in Sect. 4.1 above, the uncorrected ΔE–ΔH relationship for UKESM1's historical time series is non-linear and reveals energy leakage (Fig. 1f). Can drift correction address this problem? After we apply agnostic-method MCDC, the trend-corrected ΔE–ΔH relationships are linear (Fig. 4a–b). Estimates of the coefficient η are positive and less than 1.0 (Fig. 4c). Therefore, for UKESM1, the drift-corrected ΔE–ΔH relationships are generally consistent with Eq. (A3).

When applying MCDC, we produce 1500 drift-corrected time series for each variable, scenario, model, and MCDC method. Therefore, we also calculate 1500 estimates of η, so we can quantify drift uncertainty. For all three MCDC methods, the drift uncertainty in η is relatively large during the historical period (Figs. 4c and 5): the relatively weak global warming signal in the ΔE and ΔH historical time series leaves them susceptible to drift uncertainty, which in turn influences the η estimates. The drift uncertainty in η is much smaller during the 21st century projection period (Figs. 4c and 5): global warming drives clear trends in the ΔE and ΔH projection time series, constraining the η estimates. We average across the projection simulations to summarise the drift uncertainty in η (Table S5; Table 1). For UKESM1, the agnostic-method drift uncertainty is approximately 0.01 (Table S5). Across the CMIP6 ensemble, the agnostic-method drift uncertainty ranges from 0.01 to 0.14, with a median of 0.06 (Table 1).

The η estimates differ slightly between the different projection simulations (Figs. 4c and 5). The differences arise from the response of the climate system to different forcing scenarios. We refer to the inter-scenario range as the *scenario uncer-*

tainty. For the UKESM1 projection simulations, the scenario uncertainty is approximately 0.01, similar to the drift uncertainty (Table S5). For most models in the CMIP6 ensemble, the scenario uncertainty is similarly small (Fig. 5; Table S5), with an ensemble median of 0.01 (Table 1): in general, η does not depend strongly on the scenario. However, for two models, the scenario uncertainty is 0.08 and 0.07 respectively (Table S5; Fig. 5c–d): for these models, η is sensitive to the choice of scenario.

Estimates of η vary between different models within the ensemble (Fig. 5). The model uncertainty of 0.17 is the largest source of uncertainty (Table 1). For most of the CMIP6 models, the η estimates lie in the range 0.9–1.0, consistent with our understanding of the climate system (Appendix A; von Schuckmann et al., 2020; Forster et al., 2021): even if the ocean were to absorb all excess energy entering the climate system, this would lead to a theoretical maximum of 1.0. However, some of the η estimates are larger than 1.0. For four models (25 % of the ensemble), the median η estimates are consistently greater than 1.0 for all projection scenarios (Fig. 5h, k, l, and m): even after drift correction, the energy balance of these models remains problematic.

4.6 Expansion efficiency of heat (ϵ)

As the ocean warms, it expands, leading to thermosteric sea-level rise. We expect the relationship between ΔH and ΔZ to be approximately linear: $\Delta Z = \epsilon \Delta H$ (Eq. A4). As noted in Sect. 4.1 above, the uncorrected data from UKESM1's historical simulation are inconsistent with Eq. (A4) (Fig. 1g). However, after we apply agnostic-method MCDC, the drift-corrected ΔH–ΔZ relationships are approximately linear (Fig. 4d–e). Drift correction successfully establishes meaningful ΔH–ΔZ relationships that can be interpreted using Eq. (A4).

However, in disagreement with Eq. (A4), the drift-corrected ΔH–ΔZ relationships during the historical period exhibit hysteresis-like behaviour (Fig. 4d). ΔH and ΔZ are referenced to the 1850s decadal mean, so the ΔH–ΔZ relationships begin near the origin. ΔH and ΔZ then become negative for much of the historical period, driving the relationships towards the lower-left of Fig. 4d. ΔH and ΔZ become positive later in the historical period, driving the relationships towards the upper-right. However, ΔH 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. It is possible that the ΔE–ΔH–ΔZ relationships may depend on the sign of the forcing (c.f. Bouttes et al., 2013). However, the hysteresis-like behaviour in Fig. 4d – and also in Fig. 4a – 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. As was the case for η, the drift uncertainty in ϵ is much larger during the historical period than the 21st century (Figs. 4f and S3) due to the comparatively weak forcing signal during the historical period. When drift uncertainty is considered, the historical period provides only a weak constraint on the value of ϵ.

For the UKESM1 projection simulations, the estimates of ϵ depend on the 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 agnostic-method drift uncertainty of 1 mm YJ^{-1} (Table S6), leading to distinct non-overlapping box plots for the projection scenarios

(Fig. 4f). For all models in the CMIP6 ensemble, the ϵ estimates are smallest for SSP1-2.6 and largest for SSP5-8.5 (Fig. S3). This suggests that the thermal expansion of the ocean is sensitive to the forcing scenario, and is not merely a linear function of ΔH (c.f. Wu et al., 2021). However, the strength of the dependence on scenario varies: across the ensemble, the scenario uncertainty varies from 4 to 10 mm YJ^{-1}, with a median of 7 mm YJ^{-1} (Table 1).

The agnostic-method drift uncertainty varies over a wider range, from 1 to 22 mm YJ^{-1}, with a median of 7 mm YJ^{-1} (Table 1). The linear-method drift uncertainty is much smaller, suggesting that linear-method MCDC underestimates the drift uncertainty. The model uncertainty is 12 mm YJ^{-1}. When analysing ϵ across the CMIP6 ensemble, drift uncertainty, scenario uncertainty, and model uncertainty all constitute substantial sources of uncertainty.

5 Discussion

5.1 Energy balance

Following Hobbs et al. (2016) and Irving et al. (2021), we have considered the energy budgets of global climate models. Inconsistent relationships between excess system energy (ΔE) and excess ocean heat (ΔH; 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. (2021), we have shown that drift correction can partially address this problem of energy leakage. Drift-corrected ΔE–ΔH relationships are approximately linear (Fig. 4a–b). For all sixteen models analysed here, the coefficient η – the fraction of excess energy absorbed by the ocean – is positive (Fig. 5). For most of these models, η is also less than 1.0, consistent with energy conservation. However, even after drift correction, the energy balance of several models remains suspect: energy leakage is revealed by η estimates greater than 1.0. For four models, the median η estimates are greater than 1.0 for all projection scenarios (Fig. 5).

When evaluating the performance of global climate models, the coefficient η 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

Sanderson (2020) previously explored the contribution of drift to uncertainty in global climate sensitivity indices. To do this, 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, using Monte Carlo drift correction (MCDC).

We have compared alternative methods of correcting drift in cumulatively-integrated fluxes (such as ΔE and ΔH). When applying *linear-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 flux is integrated cumulatively. Conceptually, these two alternative approaches are similar: a constant bias in flux (e.g. E') will drive a spurious trend in cumulatively-integrated flux (e.g. ΔE) (Hobbs et al., 2016). In practice, however, we find that these

two alternative methods produce differing results: the linear-method drift uncertainty is smaller than the integrated-bias-method drift uncertainty, because integrating cumulatively effectively averages over the substantial inter-annual variability, resulting in a smaller standard error. In other words, compared with integrated-bias-method MCDC, linear-method MCDC provides a much stronger constraint on the appropriate drift correction to apply. When correcting drift in an integrated variable, we recommend that it is generally preferable to integrate the variable first and then apply the drift correction afterward. Nevertheless, we recognise that there may be cases for which it makes sense to correct the bias before integrating: for example, when performing an analysis that involves both E' and ΔE, we may prefer to use the integrated-bias method for consistency. Researchers should clarify whether they correct drift before or after integration. Although we primarily refer to cumulative integration across time here, such clarity should also extend to other calculations: for example, Irving et al. (2019) clarify that they calculate spatial integrals before correcting drift.

When estimating drift uncertainty, our preferred method is agnostic-method MCDC. If the drift is non-linear, then linear-method MCDC will likely underestimate the drift uncertainty. Therefore, agnostic-method MCDC – which assigns equal prior probabilities to linear, quadratic, and cubic models of drift – should provide a more accurate estimate of drift uncertainty. For the variables analysed in this paper, we have found that agnostic-method drift uncertainty is larger than linear-method drift uncertainty, as expected. We emphasise that agnostic-method corresponds to an assumption that linear, quadratic, and cubic models of drift are all equally plausible. This assumption corresponds to the common practice of selecting a statistical model of drift *a priori*, before fitting the model to data (Sect. 3.1).

A possible further step would be to fit and compare alternative statistical models of drift *a posteriori*, using measures such as the Bayesian information criterion. We could then select the best statistical model(s) for a specific time series. When applying this "best-fit" approach, we could also consider additional statistical models of drift, including signal processing filters (e.g. Palmer et al., 2011). This best-fit approach would be a sensible way to correct drift. Application of the best-fit approach should lead to a reduction in drift uncertainty.

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. For the historical period, the agnostic-method drift uncertainty in ΔZ can be as large as 24 mm, of comparable magnitude to the impact of omitting volcanic forcing in control simulations (Gregory et al., 2013). Therefore, drift uncertainty deserves similar attention to that given to 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 reconstruction 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 climate models.

We further hypothesise that drift uncertainty may 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). Large variability will lead to a large standard error when fitting the statistical model of drift. 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.

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: energy leakage. Although energy leakage 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 η (the fraction of excess energy absorbed by the ocean): Is $\eta \leq 1$?

We draw four conclusions about drift correction:

1. When correcting drift in cumulatively-integrated fluxes (e.g. ΔE, the excess system energy), it is generally preferable to integrate the fluxes before correcting the drift, rather than correcting the bias before integrating the fluxes.

2. Linear-method MCDC may underestimate drift uncertainty due to non-linearity in the underlying drift. If we assume that linear, quadratic, and cubic statistical models of drift are equally plausible, then agnostic-method MCDC provides a more accurate estimate of drift uncertainty.

3. When analysing ΔZ (thermosteric sea-level rise) during the historical period, drift uncertainty is relatively large. For the period 1850s–2000s, the agnostic-method drift uncertainty ranges from 3 to 24 mm. This is comparable to the impact of omitting volcanic forcing in control simulations (Gregory et al., 2013).

4. When using 21st century projection scenarios to estimate ϵ (the expansion efficiency of heat), agnostic-method drift uncertainty of 1 to 22 mm YJ^{-1} constitutes an important source of uncertainty alongside scenario uncertainty and model uncertainty.

We propose two hypotheses – each with an accompanying question – to be tested in future work. First, 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? Second, 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.8219778 (Grandey, 2023).

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* (E') heats the earth's climate system (Melet and Meyssignac, 2015; von Schuckmann et al., 2020), increasing the *excess system energy* (ΔE). 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* (ΔH). 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 (E') cumulatively over time:

$$\text{Excess system energy} \equiv \Delta E \equiv \int_{t_0}^{t} E'(\tau)\mathrm{d}\tau \tag{A1}$$

where t_0 is the reference time. Expressed in global-mean units, E' is estimated to be 0.47 $\pm$ 0.1 W m^{-2} for the period 1971–2018 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 by integrating the global-total sea-surface heat flux (H') cumulatively over time:

$$\text{Excess ocean heat} \equiv \Delta H \equiv \int_{t_0}^{t} H'(\tau)\mathrm{d}\tau \tag{A2}$$

A constant 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 much smaller than E' – to focus on the relationship between E' and H'. The drift-corrected ΔH time series should be insensitive to the inclusion or exclusion of the geothermal heat flux: a constant

geothermal heat flux would essentially modify the bias in H' and the linear drift in ΔH, both of which are removed by drift correction.

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

$$\Delta H = \eta \Delta E \tag{A3}$$

For the period 1971–2018, η is estimated to be approximately 0.89–0.91 (von Schuckmann et al., 2020; Forster et al., 2021).

As the ocean warms, the water expands, causing thermosteric sea-level rise (Gregory et al., 2019). The thermal expansion of 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, 2015):

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

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

Appendix B: Estimating η and ϵ using ordinary least squares

Informed by Eq. (A3) and Eq. (A4), we estimate η 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 ΔE, ΔH, and ΔZ are all referenced to the 1850s decadal mean).

However, the drift-corrected ΔE–ΔH and ΔH–ΔZ relationships often exhibit hysteresis-like behaviour during the historical period (Figs. 4a and d, especially the "Max intercept" and "Min intercept" lines): when the time series transition from negative values to positive values, the ΔE–ΔH and ΔH–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 η 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. We thank Damien Irving , the anonymous referee, and the editor for their helpful comments. This work comprises EOS contribution number 547.

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.

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

Durack, P. J., Gleckler, P. J., Purkey, S. G., Johnson, G. C., Lyman, J. M., and Boyer, T. P.: 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.

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.

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.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., Zhou, B., and Zhou, B., Cambridge University Press, 2021.

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 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.

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.: Analysis Code for "Monte Carlo Drift Correction – Quantifying the Drift Uncertainty of Global Climate Models" (D22a-Mcdc), Zenodo, https://doi.org/10.5281/zenodo.8219778, 2023.

Gregory, J. M., Bi, D., Collier, M. A., Dix, M. R., Hirst, A., Hu, A., Huber, M., Knutti, R., Marsland, S. J., Meinshausen, M., Rashid, H. A., Rotstayn, L. D., Schurer, A., and Church, J. A.: 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, 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., 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.

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' 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.

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.

Irving, D., Hobbs, W., Church, J., and Zika, J.: A Mass and Energy Conservation Analysis of Drift in the CMIP6 Ensemble, J. Clim., pp. 3157–3170, https://doi.org/10.1175/JCLI-D-20-0281.1, 2021.

Irving, D. B., Wijffels, S., and Church, J. A.: Anthropogenic Aerosols, Greenhouse Gases, and the Uptake, Transport, and Storage of Excess Heat in the Climate System, Geophys. Res. Lett., 46, 4894–4903, https://doi.org/10.1029/2019GL082015, 2019.

Jackson, L. P. and Jevrejeva, S.: A Probabilistic Approach to 21st Century Regional Sea-Level Projections Using RCP and High-end Scenarios, Glob. Planet. Change, 146, 179–189, https://doi.org/10.1016/j.gloplacha.2016.10.006, 2016.

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.

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 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.

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, 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 Estimate the Earth Energy Imbalance, Front. Mar. Sci., 6, 432, https://doi.org/10.3389/fmars.2019.00432, 2019.

Newey, W. K. and West, K. D.: A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703, https://doi.org/10.2307/1913610, 1987.

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 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.

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.

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., Samir, K. C., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., 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., de 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.

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.

Wu, Q., Zhang, X., Church, J. A., Hu, J., and Gregory, J. M.: Evolving Patterns of Sterodynamic Sea-Level Rise under Mitigation Scenarios and Insights from Linear System Theory, Clim. Dyn., 57, 635–656, https://doi.org/10.1007/s00382-021-05727-7, 2021.

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global Reconstruction of Historical Ocean Heat Storage and Transport, Proc. Natl. Acad. Sci., 116, 1126–1131, https://doi.org/10.1073/pnas.1808838115, 2019.

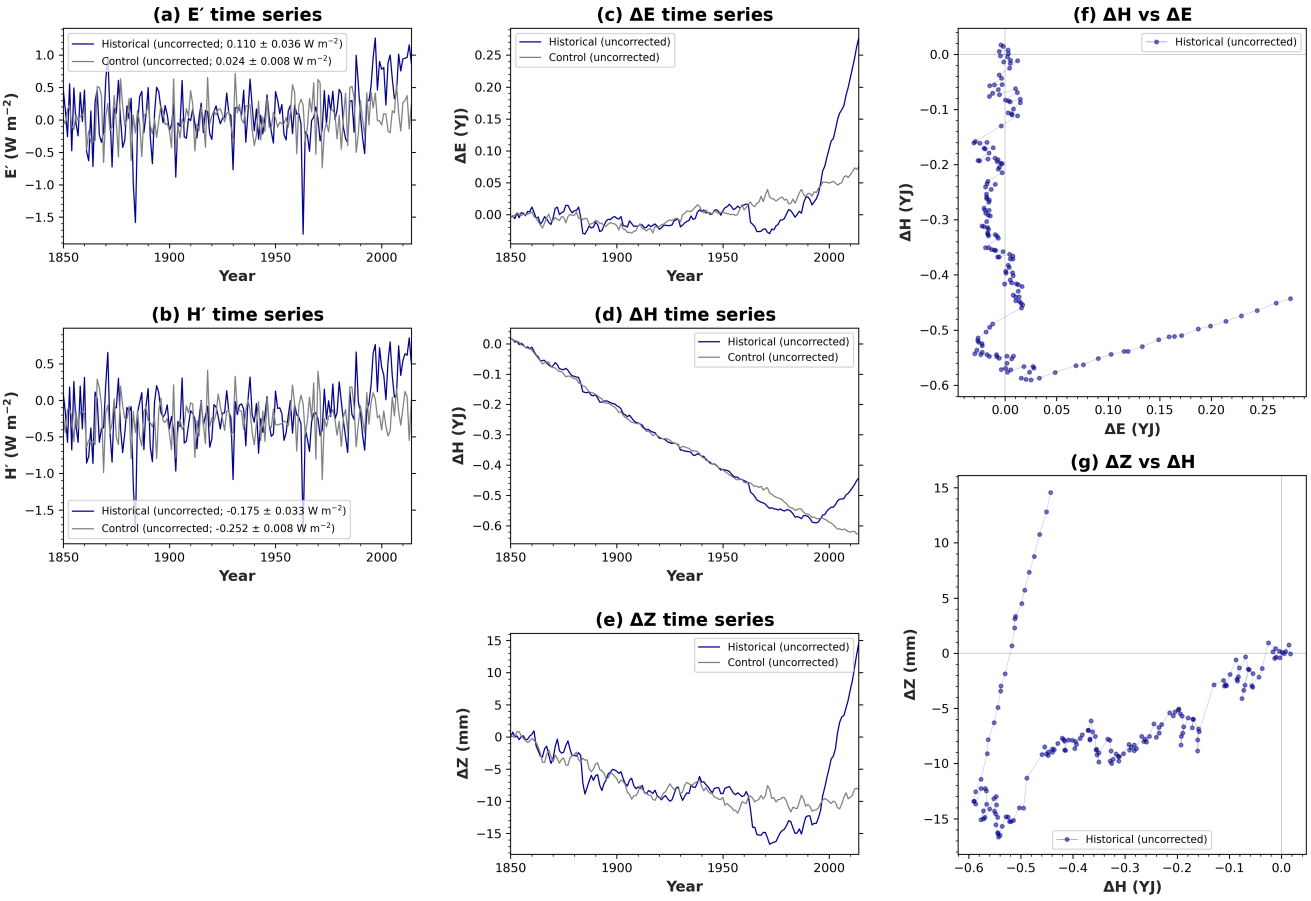

Figure 1. Uncorrected time series and bivariate relationships for the UK Earth System Model (UKESM1). These *uncorrected* time series have not yet been corrected for drift. Using data from the control and historical simulations, time series are shown for the following global variables: (a) top-of-atmosphere radiative flux (E'), (b) sea-surface heat flux (H'), (c) excess system energy (ΔE), (d) excess ocean heat (ΔH), and (e) thermosteric sea-level rise (ΔZ). Global-total E' and H' are expressed in global-mean units of W m^{-2}. In (a) and (b), the legend shows the mean of the entire time series and the standard error of the mean. In (a)–(e), only the segment of the control time series that corresponds to the historical period (1850–2014) is shown. (The full length of the entire control time series is 1100 years.) For the historical simulation, bivariate relationships are also shown: (f) ΔH versus ΔE and (g) ΔZ versus ΔH. Interpretation is offered in Sect. 4.1.

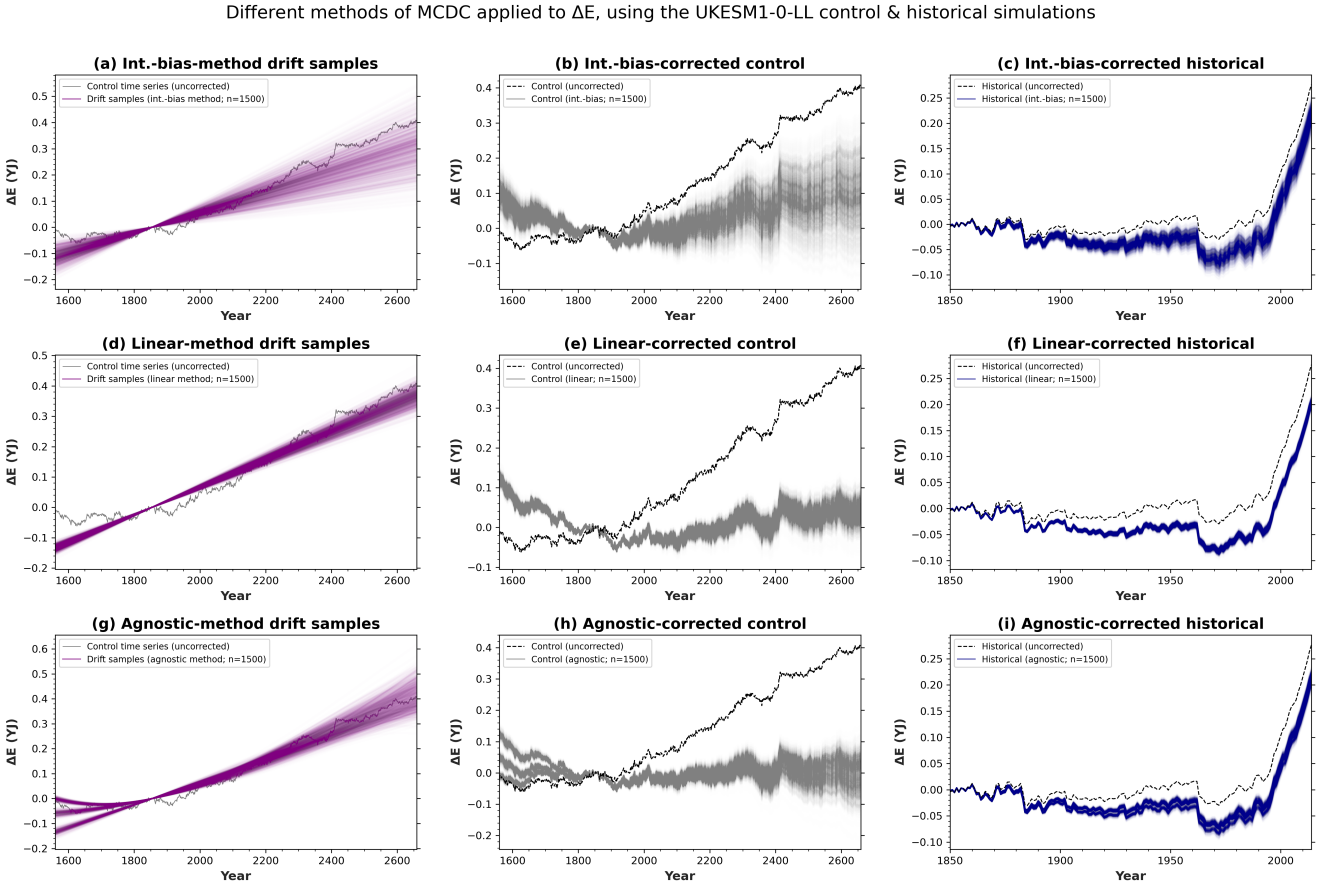

Figure 2. Different drift-correction methods applied to excess system energy (ΔE), using the UKESM1 control and historical simulations. The first row (a–c) shows integrated-bias-method MCDC results : (a) drift samples derived using the integrated-bias method , (b) integrated-bias-method drift-corrected control time series , and (c) integrated-bias-method drift-corrected historical time series , plotted alongside the uncorrected time series. The second row (d–f) shows corresponding linear-method MCDC results. The third row (g–i) shows corresponding agnostic-method MCDC results. These drift-correction methods are described in Sect. 3.2.

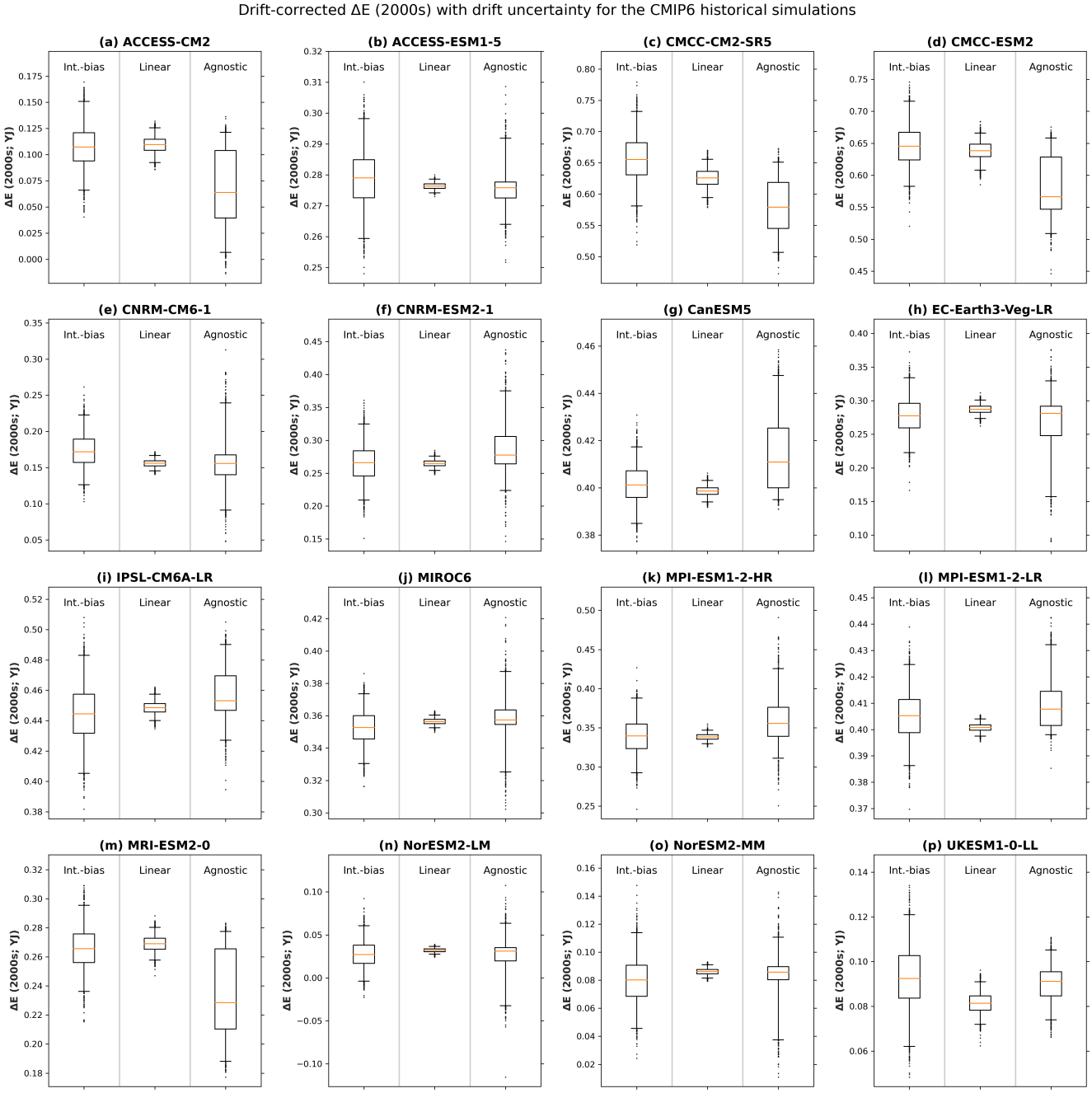

Figure 3. Drift-corrected excess system energy (ΔE) during the historical period and the corresponding drift uncertainty. We calculate ΔE as the decadal mean for the 2000s relative to the 1850s. Each panel shows results for one CMIP6 model (Table S1). Within each panel, each box plot shows results for one MCDC method. The central line shows the median, the box shows the inter-quartile range, the whiskers show the 2nd–98th inter-percentile range, and the dots show outliers beyond the range of the whiskers. Corresponding results for ΔH and ΔZ are shown in Figs. S1–S2.

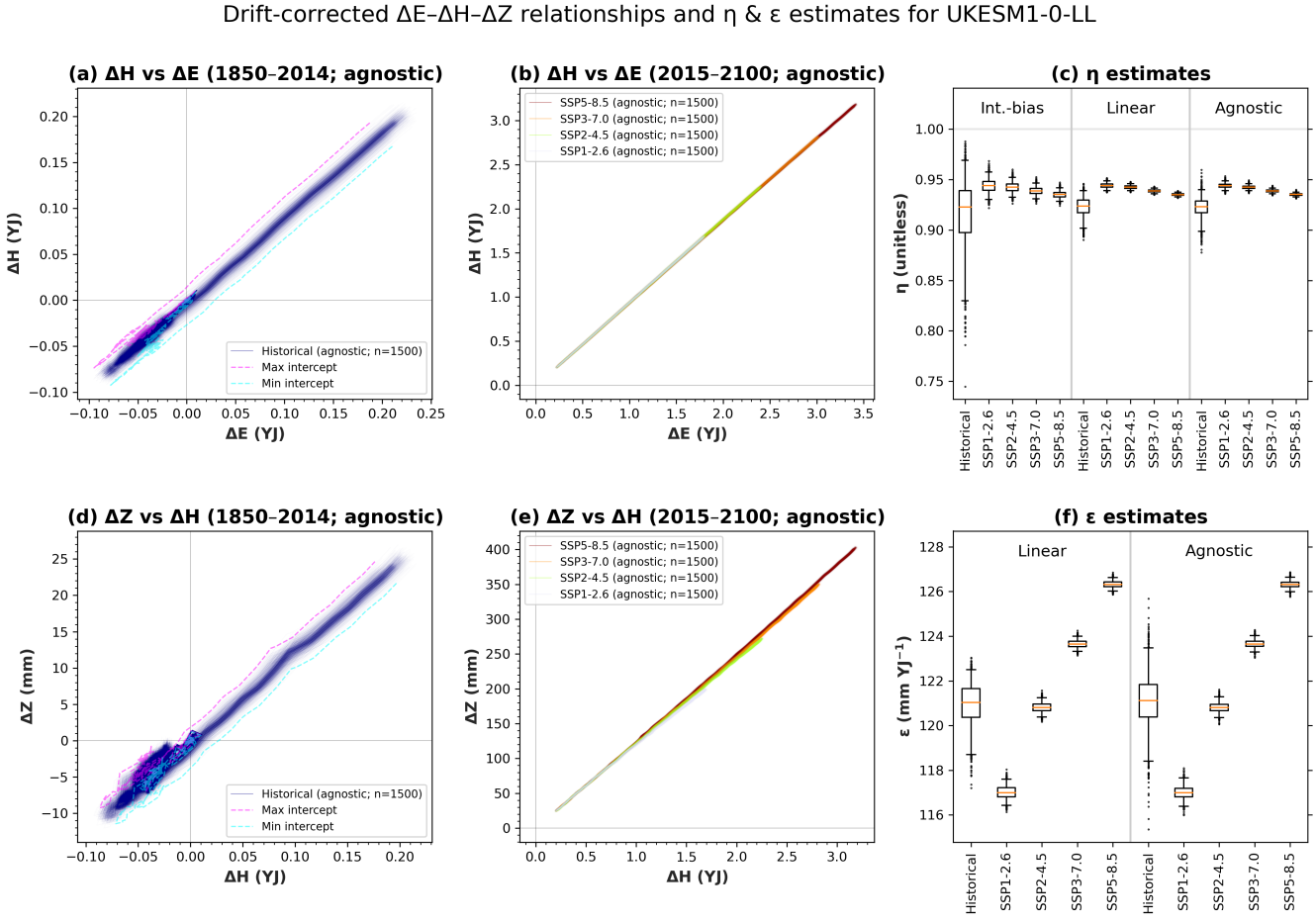

Figure 4. Drift-corrected bivariate relationships and regression coefficients (η and ϵ) for the UKESM1 simulations. (a–b) Agnostic-method drift-corrected excess ocean heat (ΔH) versus excess system energy (ΔE) for (a) the historical simulation and (b) the projection simulations. (c) Estimates of the fraction of excess energy absorbed by the ocean (η) and the corresponding drift uncertainty. The coefficient η is calculated as the linear regression coefficient of drift-corrected ΔH versus ΔE. For each combination of MCDC method and projection scenario, the central line shows the median, the box shows the inter-quartile range, the whiskers show the 2nd–98th inter-percentile range, and the dots show outliers beyond the range of the whiskers. (d–e) Agnostic-method thermosteric sea-level rise (ΔZ) versus ΔH for (d) the historical simulation and (e) the projection simulations. (f) Estimates of the expansion efficiency of heat (ϵ) and the corresponding drift uncertainty. The coefficient ϵ is calculated as the linear regression coefficient of drift-corrected ΔZ versus ΔH. Corresponding estimates of η and ϵ for other members of the CMIP6 ensemble are shown in Figs. 5 and S3.

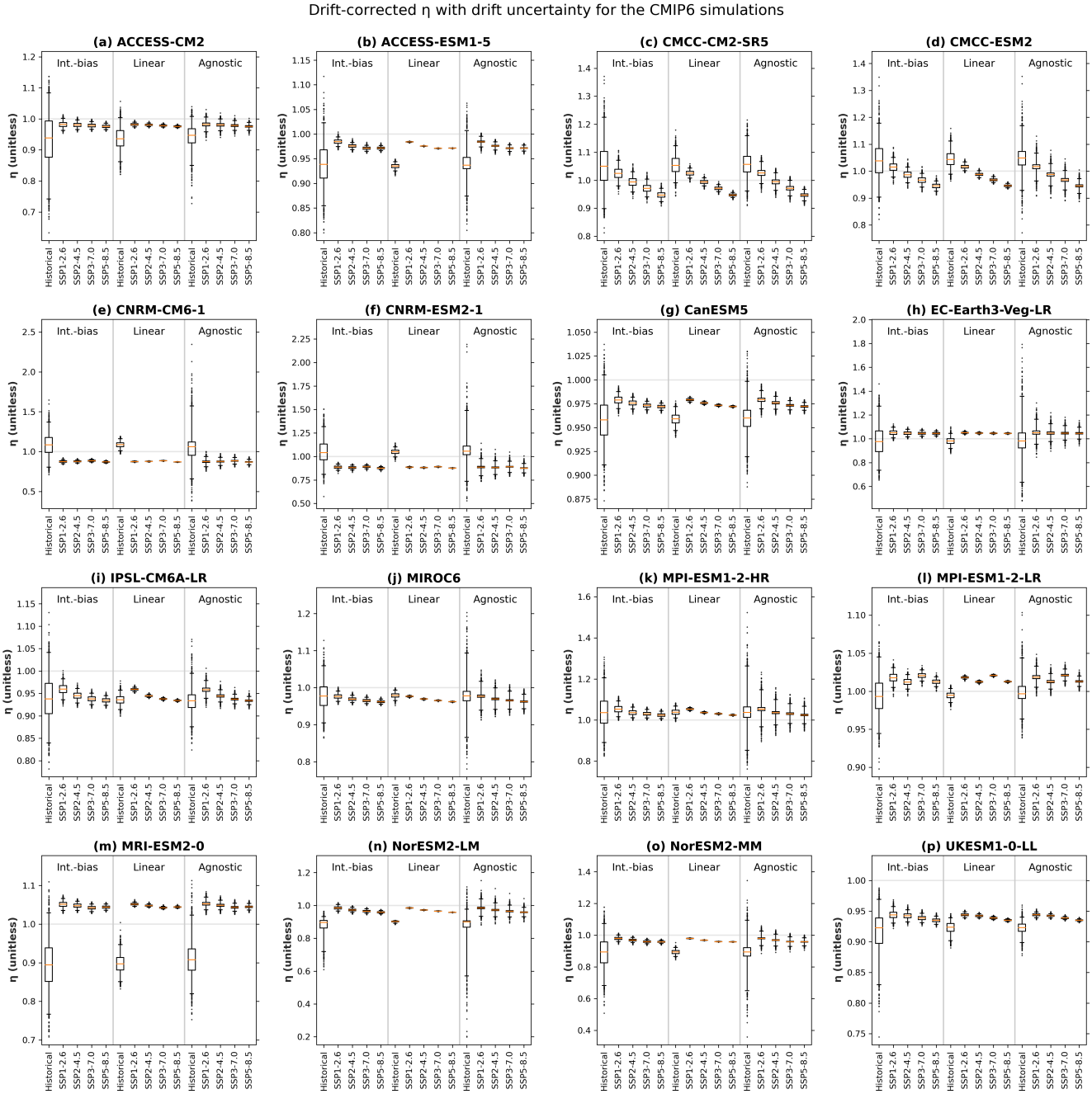

Figure 5. Drift-corrected estimates of the fraction of excess system energy absorbed by the ocean (η) and the corresponding drift uncertainty. Each panel shows results for one CMIP6 model. Within each panel, each box plot shows results for one combination of MCDC method and projection scenario. The central line shows the median, the box shows the inter-quartile range, the whiskers show the 2nd–98th inter-percentile range, and the dots show outliers beyond the range of the whiskers. The horizontal line at $\eta = 1.0$ shows the theoretical maximum value of η.

Table 1. CMIP6 ensemble median and range (minimum–maximum) for different sources of uncertainty. For each drift-correction method, *drift uncertainty* corresponds to the 2nd–98th inter-percentile range of the drift-corrected data. *Model uncertainty* corresponds to the inter-model range. *Scenario uncertainty* corresponds to the inter-scenario range. The statistics for η and ϵ are based on the 21st-century projection simulations (2015–2100). The ensemble statistics shown here correspond to the summary statistics shown in Tables S2–S6. For further details, see Tables S2–S6.

		ΔE (2000s; YJ)	ΔH (2000s; YJ)	ΔZ (2000s; mm)	η (unitless)	ϵ (mm YJ^{-1})
Drift uncertainty	Int.-bias	0.07 (0.03–0.15)	0.07 (0.03–0.16)		0.03 (0.01–0.07)	
	Linear	0.02 (0.00–0.06)	0.02 (0.00–0.06)	2 (0–8)	0.01 (0.00–0.03)	1 (0–3)
	Agnostic	0.09 (0.03–0.17)	0.08 (0.03–0.20)	10 (3–24)	0.06 (0.01–0.14)	7 (1–22)
Other uncertainty	Model	0.56	0.61	64	0.17 (0.16–0.18)	12 (11–13)
	Scenario				0.01 (0.01–0.08)	7 (4–10)