An Energy Conservation Analysis of Ocean Drift in the CMIP5 Global Coupled Models*

WILL HOBBS

ARC Centre of Excellence for Climate System Science, IMAS, and Antarctic Climate and Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tasmania, Australia

MATTHEW D. PALMER

Met Office Hadley Centre, Exeter, United Kingdom

DIDIER MONSELESAN

CSIRO Oceans and Atmosphere, Hobart, Tasmania, Australia

(Manuscript received 10 July 2015, in final form 14 October 2015)

ABSTRACT

Climate model simulations of changes to Earth's energy budget are fundamental to improve understanding of both historical and future climate change. However, coupled models are prone to "drift" (i.e., they contain spurious unforced trends in state variables) due to incomplete spinup or nonclosure of the energy budget. This work assesses the globally integrated energy budgets of 25 models in phase 5 of CMIP (CMIP5). It is shown that for many of the models there is a significant disagreement between ocean heat content changes and net top-of-atmosphere radiation. The disagreement is largely time-constant and independent of forcing scenario. Furthermore, most of the nonconservation seems to occur as a result of energy leaks external to the ocean model realm. After drift correction, the time-varying energy budget is consistent at decadal time scales, and model responses to climate forcing are not sensitive to the magnitude of their drift. This demonstrates that, although drift terms can be significant, model output can be corrected post hoc without biasing results.

1. Introduction

Anthropogenic climate change is essentially an energy balance problem. The planet receives incoming shortwave solar radiation (SW_{in}), some of which is reflected back to space (SW_{out}), and the remainder is absorbed and reemitted as longwave thermal radiation (LW_{out}). A sustained nonzero sum of these three radiative components (netTOA) indicates disequilibrium (i.e., a net loss or gain of heat into the planet). Increased greenhouse gas concentrations impede LW_{out}, causing the tropospheric temperatures to increase and thus increase LW_{out} following the Stefan–Boltzmann law, in order to restore radiative equilibrium. It takes a considerable time period to restore equilibrium, and in the meantime there is a positive netTOA and the Earth system is absorbing heat. The observed 2006–13 netTOA was 0.4–0.6 $W m^{-2}$ averaged over the planet's surface (Roemmich et al. 2015). The ocean is by far Earth's largest heat sink, and it is estimated that >90% of this excess trapped energy is absorbed by the oceans (Rhein et al. 2013), increasing ocean heat content and leading to thermosteric sea level rise (Kuhlbrodt and Gregory 2012; Church et al. 2013).

Climate model simulations of changes of Earth's energy budget are fundamental to improving our understanding of both historical and future climate change. One of the essential characteristics of the climate system is the close correspondence between changes in the netTOA and global ocean heat content (OHC), which represents Earth's primary energy store (Palmer et al. 2011). Climate model simulations suggest that the ocean

* Supplemental information related to this paper is available at the Journals Online website: http://dx.doi.org/10.1175/JCLI-D-15-0477.s1.

Corresponding author address: Will Hobbs, ARC Centre of Excellence for Climate System Science, IMAS, Private Bag 129, University of Tasmania, Hobart, Tasmania 7001, Australia.
E-mail: whobbs@utas.edu.au

DOI: 10.1175/JCLI-D-15-0477.1

© 2016 American Meteorological Society

becomes the dominant term in Earth's energy budget on a time scale of about 12 months (Palmer and McNeall 2014).

A separate but related issue is the question of model "drift," where drift is a spurious trend in any state variable that generally persists throughout the length of a model simulation (Sen Gupta et al. 2013). Drift may occur in part because of insufficient spinup, when a model is still adjusting from its initialized state to its preferred equilibrium climate. Preindustrial control (piControl) simulations from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) are run using a common equilibrium solar forcing. Data resulting from the spinup procedure are rarely submitted to the CMIP5 archive by the modeling groups, although in practice spinup may take many centuries and even millennia for coupled models, and because of computational expense models may retain some drift due to spinup. Driftlike characteristics may also arise due to errors or missing terms in the model's energy budget, which we refer to here as energy leaks. As we will demonstrate, even relatively small energy leaks, when integrated over century scales, can cause significant spurious trends to appear. Since the ocean is the dominant inertial term in Earth's energy budget, these energy imbalances accumulate in the ocean, causing drift in ocean variables, especially temperature and its related quantity, ocean heat content. Although drifts can be accounted for by methods of varying degrees of complexity, there may be some ambiguity as to whether a control run trend is drift or a valid long-term mode of variability (Sen Gupta et al. 2013). Analysis of the global energy budget, and any associated energy leaks, may provide a more physically based insight into the most appropriate form of climate model drift correction, at least for globally integrated quantities; at regional scales there may be spurious redistributions of energy that are not evident in global integrals.

The purpose of the current work is to assess the representation of this first-order balance between netTOA and OHC in state-of-the-art climate model simulations conducted under CMIP5. We will discuss the implications for energy conservation in the models and the associated model drift characteristics (Sen Gupta et al. 2013), which must be corrected for in simulations with imposed external climate forcings.

2. Method

The model output are taken from the CMIP5 archive, and unless stated otherwise are from the piControl experiment. This simulation represents an estimate of unforced equilibrium climate and includes variability on a range of time scales arising from the emergent properties of the coupled climate system (e.g., ENSO). The 25 CMIP5 climate models included in this study are summarized in Table 1. All variables were obtained from the CMIP5 archive at monthly temporal resolution and were converted to annual means. To avoid any seasonal biasing of the results, a time weighting based on each model calendar month was used to calculate the annual means.

In addition to the piControl simulations, the sensitivity of model energy imbalances to the presence of external forcings was tested using data from the CMIP5 historical and representative concentration pathway 8.5 (RCP8.5) experiments (Meinshausen et al. 2011). The historical simulations were run by forcing the models with estimates of the observed anthropogenic (greenhouse gas, ozone depletion, and anthropogenic aerosols) and natural (solar and volcanic) forcing agents, and represents a simulation of climate for the period 1850–2005. The RCP8.5 experiments were run from the end of the historical simulations to the year 2300, using projected anthropogenic emissions under a "business as usual" scenario (i.e., no societal attempt to reduce greenhouse gas emissions over the twenty-first century). This represents the strongest forcing scenario in the CMIP5 suite. We used one ensemble member from each model for all the experiments.

The relationship between drift and model climate sensitivity was explored using output from the 1% yr^{-1} increasing CO_2 experiment (1pctCO2). The transient climate response (TCR; Cubasch et al. 2001) was estimated for each model by calculating the increase in global mean surface air temperature (CMIP5 variable name tasga) between the 20-yr mean centered on year 70 of the 1pctCO2 (at 70 yr is equivalent to a doubling of CO_2 at 1% yr^{-1}), and the piControl 20-yr mean immediately prior to the branch off of the 1pctCO2 experiment from the piControl mean. The difference was drift corrected using a first-order linear trend of the complete piControl simulation.

The analysis presented here compares the temporal evolution of three globally integrated energy budget variables. The netTOA was calculated by summing the model SW_{in}, SW_{out}, and LW_{out} at a grid point, and then calculating the areal integral. The net ocean surface heat flux (netSFC) was calculated by integrating the model output net surface heat flux (hfds in the CMIP5 variable nomenclature), where available. We did not attempt to calculate this flux from components because each model has a subtly different surface energy balance formulation, which becomes a complicating factor (and possible source of error) when dealing with large numbers of different models. OHC was calculated by integrating the

seawater potential temperature over the complete ocean volume, and multiplying by constant values of density ρ_o and specific heat capacity c_{po}. This implicitly assumes that the mass of the simulated oceans is fixed, which is not the case since almost all of the models have some ocean mass drift (see Fig. S1 in the supplemental material). To demonstrate that this assumption is valid, Fig. S2 in the supplemental material compares the time-mean piControl total ocean heat change [i.e., $c_{po}\delta(MT)/\delta t$, where M is the time-varying ocean mass and T is the global mean seawater temperature] with the component due to temperature change ($c_{po}M\delta T/\delta t$) and the component due to mass change only ($c_{po}T\delta M/\delta t$). For almost all of the models, the mass component is negligible, and we would expect the assumption of constant mass to be valid.

For the constant $\rho_o c_p$ we used a value of $4092 \pm 125\,\mathrm{kJ\,(K\,m^3)^{-1}}$, equivalent to $\rho_o = 1023\,\mathrm{kg\,m^{-3}}$ and $c_p = 4000\,\mathrm{J\,(K\,kg)^{-1}}$. This value was estimated by calculating ρ_o and c_p, using the same uniform distributions of temperature T, salinity S, and pressure P in the ranges $35° \geq T \geq 0°\mathrm{C}$, $40 \geq S \geq 0\,\mathrm{psu}$, and $5000 \geq P \geq 0\,\mathrm{dbar}$ ($1\,\mathrm{dbar} = 10^4\,\mathrm{Pa}$), and taking the extreme values of $\rho_o c_p$. This range results in an uncertainty of ocean heat content of just 3%.

3. Results

a. Control run energy imbalance

Figure 1 shows the evolution of the piControl annual-mean OHC (blue), the OHC induced the by time-integrated netTOA (assuming that all netTOA energy were absorbed by the ocean, red), and where available the OHC implied by the net heat flux into the ocean (orange); anomalies were calculated with respect to the initial value of each time series. Recall that for an equilibrium climate (i.e., one with no external climate forcing) we would expect all three quantities to fluctuate around a mean value of zero. This is clearly not the case for any of the models, all of which show significant net gains or losses of energy over the length of the piControl runs. For some of the models the magnitude of this imbalance is equal to or greater than the current observed energy imbalance of $0.5\,\mathrm{W\,m^{-2}}$ (black dashed lines; Roemmich et al. 2015). For most of the models shown, the gain or loss of energy is linear, implying a constant netTOA imbalance over time. IPSL-CM5B-LR has a significant curvature in the net energy gain, with the rate of gain decreasing over time. This curvature may be indicative of incomplete spinup (i.e., the model is still moving toward its preferred equilibrium state from its initialized state).

Recall that for an energy-conserving climate model we would expect the ocean to absorb almost all of the energy input to the system, and hence the OHC change should account for 90%–100% of the netTOA (Palmer and McNeall 2014). As Fig. 1 shows, there are often significant differences between the rate of OHC change and the netTOA. In most cases the rate of OHC change is less than that implied by the netTOA imbalance, although two models (GFDL CM3 and IPSL-CM5B-LR) have oceans that absorb more heat than is put into the system. In a few cases (CanESM2, GFDL-ESM2M, and MIROC-ESM) the OHC change has opposite sign of the netTOA. For the models that supplied surface heat flux, the agreement between OHC and netSFC is much better than between OHC and netTOA. In fact, the difference between $d\mathrm{OHC}/dt$ and netSFC is generally less than 10% of the total imbalance (Table 2, column 2), although there are some outliers with relatively poor agreement between netSFC and $d\mathrm{OHC}/dt$. This suggests that for most of the models, the majority of ocean temperature drift is due to surface flux forcings, and the greater part of difference between netTOA and $d\mathrm{OHC}/dt$ is explained by energy leaks in other components of the coupled models (e.g., atmosphere, land surface). The agreement between $d\mathrm{OHC}/dt$ and netSFC is not perfect, however, and there appear to be some energy leaks in the ocean models, albeit much smaller than elsewhere in the coupled models.

The time evolution of the nonconservation term (i.e., netTOA $-$ $d\mathrm{OHC}/dt$) is shown in Fig. 2. To allow easier comparison between the models, which have a large spread in their mean nonconservation term, the results are shown as anomalies. Although Fig. 1 shows that the nonconservation term is time-constant for most of the models, Fig. 2 shows some interesting time-varying behavior in a few of the models. Several of the models have a trend in the nonconservation term. In some cases this trend acts to decrease the magnitude of the nonconservation (ACCESS1.0, CCSM4, CSIRO Mk3.6.0, HadGEM2-ES, and IPSL-CM5B-LR); that is, energy balance improves over time. However, a similar number of models have nonconservation terms that grow over time (BCC_CSM1.1, IPSL-CM5A-LR, MIROC-ESM, and MIROC5). We should note that the trends are too large to be explained by uncertainty in the ocean heat content parameter ($\rho_o c_p$). It is also worth noting that compared to both the mean values and the variance of netTOA and $d\mathrm{OHC}/dt$ indicated in Fig. 1, the trend and variance of the nonconservation term are small, and constant to a first-order approximation.

Although there are significant time-mean nonconservation terms in the models, this nonconservation is relatively stable over time; Fig. 1 shows that the

TABLE 1. Summary of coupled models used in this study, including the approximate resolution of their ocean components, and the length of each model's control run. (Expansions of acronyms are available online at http://www.ametsoc.org/PubsAcronymList.)

Model	Modeling center	Ocean resolution (lon × lat)	Ocean model	Atmospheric model	piControl length (yr)
ACCESS1.0 (Bi et al. 2013)	CSIRO–BoM	1.0° × 0.6°	MOM, version 4.1	Global Atmosphere, version 1.0 (GA1.0)	500
BCC_CSM1.1 (Xin et al. 2013)	BCC	1.0° × 0.8°	MOM, version 4	BCC_AGCM, version 2.1	500
BCC_CSM1.1(m) (Xin et al. 2013)	BCC	1.0° × 0.8°	MOM, version4	BCC_AGCM, version 2.2	400
CanESM2 (Yang and Saenko 2012)	CCCma	1.4° × 0.9°	Fourth Generation Canadian Ocean Model (CanOM4)	CanAM4	996
CCSM4 (Gent et al. 2011)	NCAR	1.1° × 0.5°	POP, version 2	CAM, version 4	1051
CESM1(FASTCHEM) (Hurrell et al. 2013)	NCAR	1.1° × 0.5°	POP, version 2	Community Atmosphere Model with chemistry (CAM-CHEM)	222
CNRM-CM5 (Voldoire et al. 2013)	Centre National de Recherches Météorologiques (CNRM)–Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS)	1.0° × 0.6°	NEMO	ARPEGE, version 5.1	850
CSIRO Mk3.6.0 (Jeffrey et al. 2013)	CSIRO–Queensland Climate Change Centre of Excellence (QCCCE)	1.9° × 1.0°	MOM, version 2.2	AGCM, version 7.3.8	500
GFDL CM3 (Griffies et al. 2011)	NOAA/GFDL	1.0° × 0.9°	MOM, version 4.1	Atmospheric Model (AM), version 3 (AM3)	500
GFDL-ESM2G (Dunne et al. 2012)	NOAA/GFDL	1.0° × 0.9°	GOLD	AM, version 2.1	500
GFDL-ESM2M (Dunne et al. 2012)	NOAA/GFDL	1.0° × 0.9°	MOM, version 4.1	AM, version 2.1	500
GISS-E2-R (Schmidt et al. 2014)	NASA GISS	1.0° × 1.25°	Russell	ModelE, version 2	550
HadGEM2-CC (Martin et al. 2011)	Met Office Hadley Centre (MOHC)	1.0° × 0.8°	Met Office Unified Model (UKMO-UM)	UKMO-UM	240
HadGEM2-ES (Martin et al. 2011)	MOHC	1.0° × 0.8°	UKMO-UM	UKMO-UM	575
IPSL-CM5A-LR (Mignot and Bony 2013)	IPSL	2.0° × 1.2°	NEMO	LMDZ, version 5A (LMDZ5A)	1000
IPSL-CM5A-MR (Mignot and Bony 2013)	IPSL	2.0° × 1.2°	NEMO	LMDZ5A	500
IPSL-CM5B-LR (Mignot and Bony 2013)	IPSL	2.0° × 1.2°	NEMO	LMDZ, version 5B (LMDZ5B)	300
MIROC-ESM (Watanabe et al. 2011)	Model for Interdisciplinary Research on Climate (MIROC) contributors	1.4° × 0.9°	COCO3.4	MIROC-AGCM	630
MIROC5 (Watanabe et al. 2010)	MIROC contributors	1.4° × 0.8°	COCO4.5	FRCGC AGCM	670
MPI-ESM-LR (Jungclaus et al. 2013)	MPI for Meteorology (MPI-M)	1.4° × 0.8°	MPI-OM	ECHAM6	1000
MPI-ESM-MR (Jungclaus et al. 2013)	MPI-M	0.5° × 0.5°	MPI-OM	ECHAM6	1000

TABLE 1. (*Continued*)

Model	Modeling center	Ocean resolution (lon × lat)	Ocean model	Atmospheric model	piControl length (yr)
MPI-ESM-P (Jungclaus et al. 2013)	MPI-M	1.4° × 0.8°	MPI-OM	ECHAM6	1156
MRI-CGCM3 (Yukimoto et al. 2012)	Meteorological Research Institute (MRI)	1.0° × 0.5°	MRI Community Ocean Model, version 3 (MRI.COM3)	GSMUV	500
NorESM1-M (Bentsen et al. 2013)	Norwegian Climate Centre (NCC)	1.1° × 0.5°	Miami Isopycnic Coordinate Model–Hamburg Ocean Carbon Cycle Model (MICOM-HAMOCC)	Community Atmosphere Model–Oslo (CAM-Oslo)	500
NorESM1-ME (Bentsen et al. 2013)	NCC	1.1° × 0.5°	MICOM-HAMOCC	CAM-Oslo	252

nonconservation is dominated by a constant linear trend, and the variability of the nonconservation in Fig. 2 is small compared to the mean shown in Table 2. This means that after consideration of this bias, the time-varying energy balance seems to be reasonably well represented in the models (Palmer and McNeall 2014). Table 2 shows the regression coefficients between netTOA and ocean heat content change based on 10-yr averages. (At shorter time scales the netTOA and $d\mathrm{OHC}/dt$ may not be in equilibrium due to interannual redistributions of heat). Almost all of the models have regression coefficients that are greater than 0.8 (i.e., 80% of the ocean heat content change is explained by netTOA), close to what is expected in the real climate. This shows that after removal of the time-mean bias in netTOA, the models have a realistic energy balance. (Note that the time-mean netTOA bias is physically equivalent to the first-order linear drift in OHC).

b. *Nonconservation in forced simulations*

Typically, the effect of energy nonconservation is accounted for in forced simulations by calculating the nonconservation (i.e., nonzero term) in piControl, and removing that value from the forced simulation. This assumes that nonconservation is not sensitive to external forcings, and in this section we test that assumption. As we have shown in Table 2, after bias correction there is a good agreement between netTOA and $d\mathrm{OHC}/dt$ in the piControl run. However, it is possible that under different forcing conditions the energy imbalance characteristics may be amplified, bringing in to question the use of a control run for energy flux corrections. Because of the close relationship between energy imbalances and drift, this would also bring into question the use of control simulations for drift correction.

Figure 3 compares the time-mean piControl nonconservation term with the same quantities for two forced experiments: historical and RCP8.5. If the piControl run represents the imbalance under all forcing scenarios, we expect the values in Fig. 3 to lie on the one-to-one line, and this certainly holds for the historical simulation (Fig. 3a). The RCP8.5 is a much stronger radiative forcing than the historical simulation (indeed it is the strongest sustained forcing experiment in the CMIP5 design). Reassuringly, for the majority of the models the nonconservation term is the same in RCP8.5 as for their respective piControl runs, but a few models have small departures from the one-to-one relationship. This shows that for these models, the nonconservation term is sensitive to the imposed forcing scenario. This has implications for drift correction, since uncorrected nonconservation could potentially impact those models' responses to future forced climate change. For all three

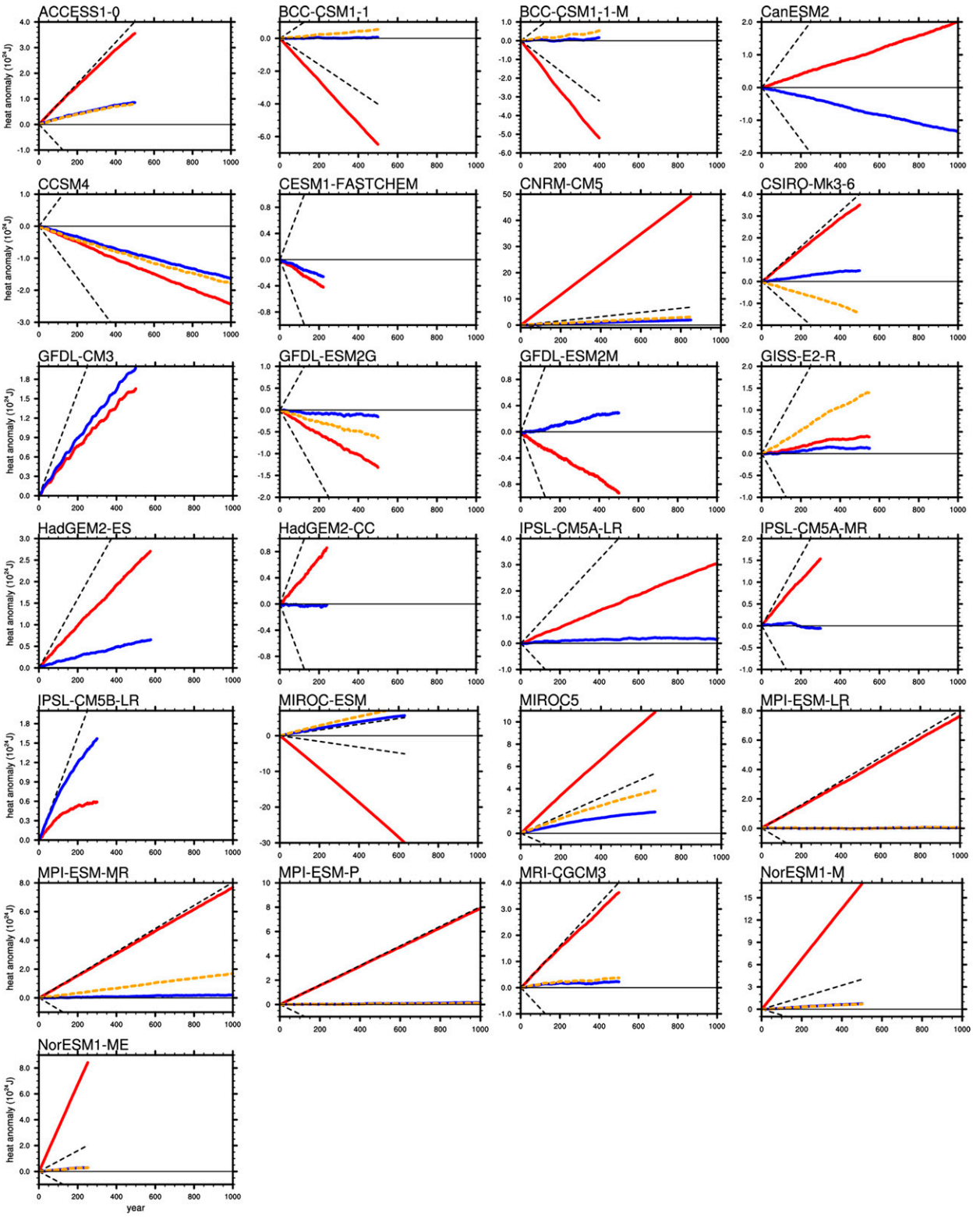

FIG. 1. Time plot by model of control run annual-mean heat content implied (10^{24} J) by netTOA (red line), OHC (blue line), and, where available, netSFC (orange line). To provide context, dashed black lines show the heat content implied by a netTOA of ± 0.5 W m^{-2}, the observed late-twentieth-century netTOA (Roemmich et al. 2015). Annual-mean flux contributions were calculated by integrating monthly mean fluxes.

TABLE 2. The second column shows the percentage of total control run nonconservation term (i.e., netTOA − dOHC/dt) that can be explained by energy leaks in the ocean component, for models where the netSFC variable is available. Column 3 shows the piControl mean nonconservation term. Column 4 shows the regression coefficients between the control run netTOA and change in ocean heat content (dOHC/dt), after a 10-yr moving average was applied to each time series. Uncertainty ranges are derived from the standard error of the residuals, as well as the uncertainty in using fixed values of c_p and ρ. Column 5 shows the maximum 10-yr mean nonconservation term in RCP8.5 that is not explained by the piControl mean [i.e., max(netTOA − dOHC/dt)$|_{\text{RCP8.5}}$ − mean(netTOA − dOHC/dt)$|_{\text{piControl}}$].

Model	netSFC − dOHC/dt as a percentage of netTOA − dOHC/dt	piControl mean netTOA − dOHC/dt (W m^{-2})	netTOA vs dOHC/dt regression	max(netTOA-dOHC/dt)$|_{\text{RCP8.5}}$ − mean(netTOA − dOHC/dt)$|_{\text{piControl}}$ (W m^{-2})
ACCESS1.0	−3	0.34	0.87 ± 0.10	0.14
BCC_CSM1.1	−7	−0.81	0.91 ± 0.13	−0.42
BCC_CSM1.1(m)	−7	−0.84	0.98 ± 0.17	−0.21
CanESM2	—	0.21	0.78 ± 0.09	0.12
CCSM4	1	−0.05	0.82 ± 0.09	—
CESM1(FASTCHEM)	—	−0.05	0.86 ± 0.14	—
CNRM-CM5	3	3.46	0.95 ± 0.13	0.40
CSIRO Mk3.6.0	−64	0.38	0.84 ± 0.10	—
GFDL CM3	—	−0.04	0.90 ± 0.10	−0.08
GFDL-ESM2G	43	−0.14	0.83 ± 0.11	−0.09
GFDL-ESM2M	—	−0.15	0.80 ± 0.12	−0.09
GISS-E2-R	645	0.03	0.88 ± 0.10	—
HadGEM2-CC	—	0.23	0.80 ± 0.13	—
HadGEM2-ES	—	0.22	0.81 ± 0.10	—
IPSL-CM5A-LR	—	0.18	0.92 ± 0.13	1.00
IPSL-CM5A-MR	—	0.33	0.85 ± 0.35	—
IPSL-CM5B-LR	—	−0.20	1.05 ± 0.25	—
MIROC-ESM	−7	−3.52	0.86 ± 0.11	−0.03
MIROC5	20	0.83	1.04 ± 0.18	0.17
MPI-ESM-LR	1	0.47	0.85 ± 0.09	—
MPI-ESM-MR	20	0.47	0.84 ± 0.07	—
MPI-ESM-P	−1	0.48	0.77 ± 0.10	—
MRI-CGCM3	4	0.43	0.94 ± 0.12	—
NorESM1-M	0	2.01	0.83 ± 0.10	0.10
NorESM1-ME	0	2.01	0.86 ± 0.17	—

experiments, the difference between the netSFC and dOHC/dt is at least an order of magnitude smaller than the netTOA − dOHC/dt term (Figs. 3c,d), with no discernible change under the historical and RCP8.5 forcings. This further supports the hypothesis that the most important energy leaks occur outside the ocean component in most of the models.

For the three models that show differences in their nonconservation under forced scenarios (BCC_CSM1.1, CNRM-CM5, and IPSL-CM5A-LR), the time evolution of the imbalance is shown in Fig. 4, concatenated across all three experiments (i.e., piControl–historical–RCP8.5). The change in energy imbalance under the historical forcing conditions is relatively modest compared to RCP8.5, with the only significant change being for IPSL-CM5A-LR, so it is perhaps not surprising that there is good agreement between the historical and piControl simulations. Under the stronger forcing conditions the nonconservation term starts changing in the early twenty-first century and does not stabilize for the length of the simulations. In all cases, whether positive or negative, the nonconservation term increases in magnitude. The maximum values of the nonconservation terms that are not accounted for by the respective piControl simulation are −0.22, 0.39, and 1.0 W m^{-2} for BCC_CSM1.1, CNRM-CM5, and IPSL-CM5A-LR, respectively (Table 2, column 5), which are significant compared to the increase in netTOA due to the RCP8.5 forcing. These extra nonconservation terms persist after the climate forcing reaches its maximum in 2200, and because of the length of the RCP8.5 simulation we cannot see whether or how long the nonconservation terms would return to their piControl values.

While it is beyond the scope of this paper to diagnose the models' energy budgets, it is interesting to consider possible sources of the change in nonconservation. A plausible and real response to strong climate forcing could be an increase in ocean stratification in regions of deepwater formation (e.g., the high-latitude North Atlantic) due to warming of the near-surface waters. This would reduce the downwelling of surface waters into the deep ocean and thereby reduce the deep ocean heat

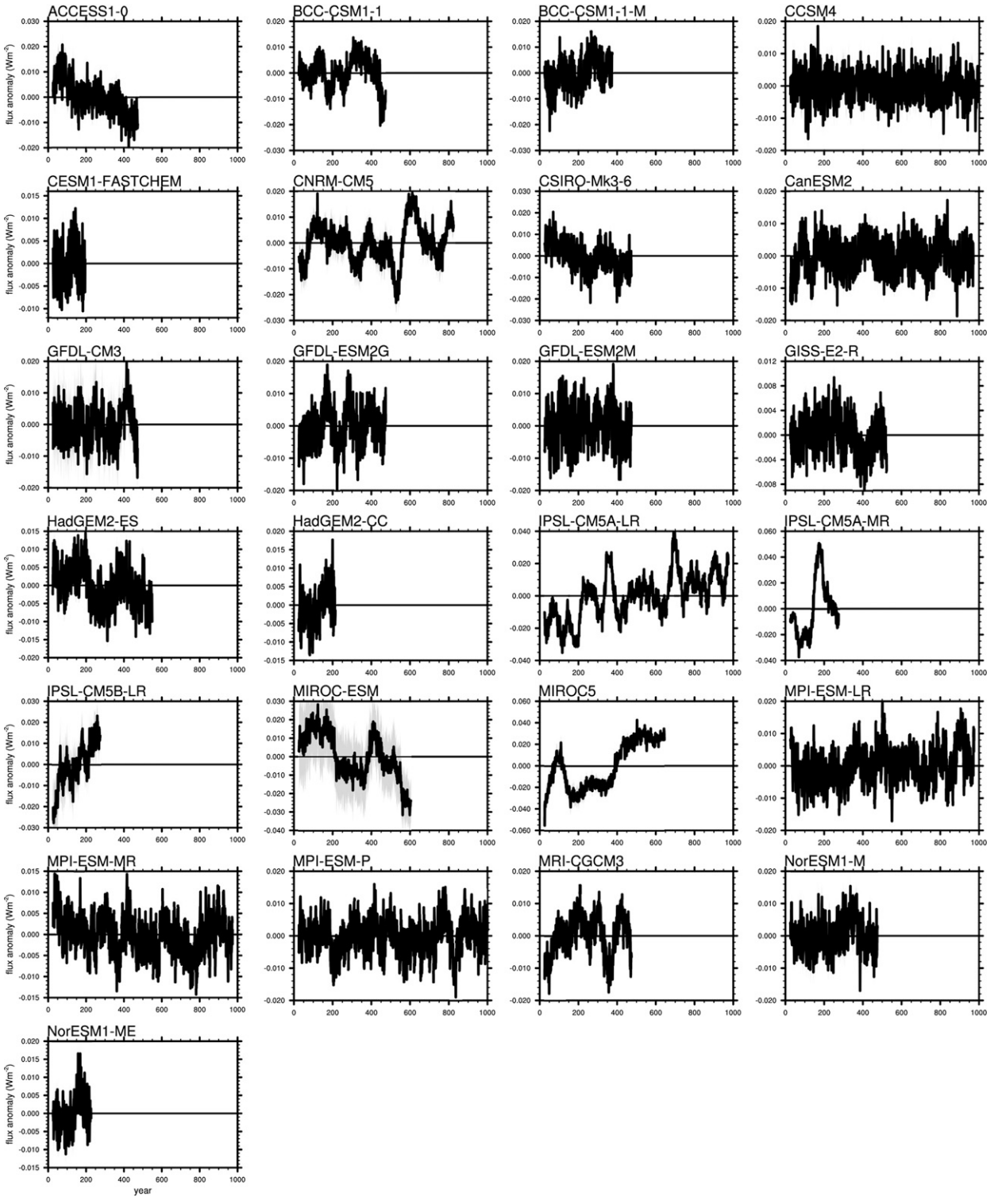

FIG. 2. Time plot by model of piControl nonconserved flux (i.e., netTOA − dOHC/dt, units are the equivalent W m^{-2} at top of atmosphere). Values are plotted as anomalies with respect to the time-mean nonconservation term, and a 50-yr running average has been applied. Gray shading shows the uncertainty due to using a fixed heat capacity term (i.e., $\rho_o c_p$); it is only visible for models with high mean nonconservation terms.

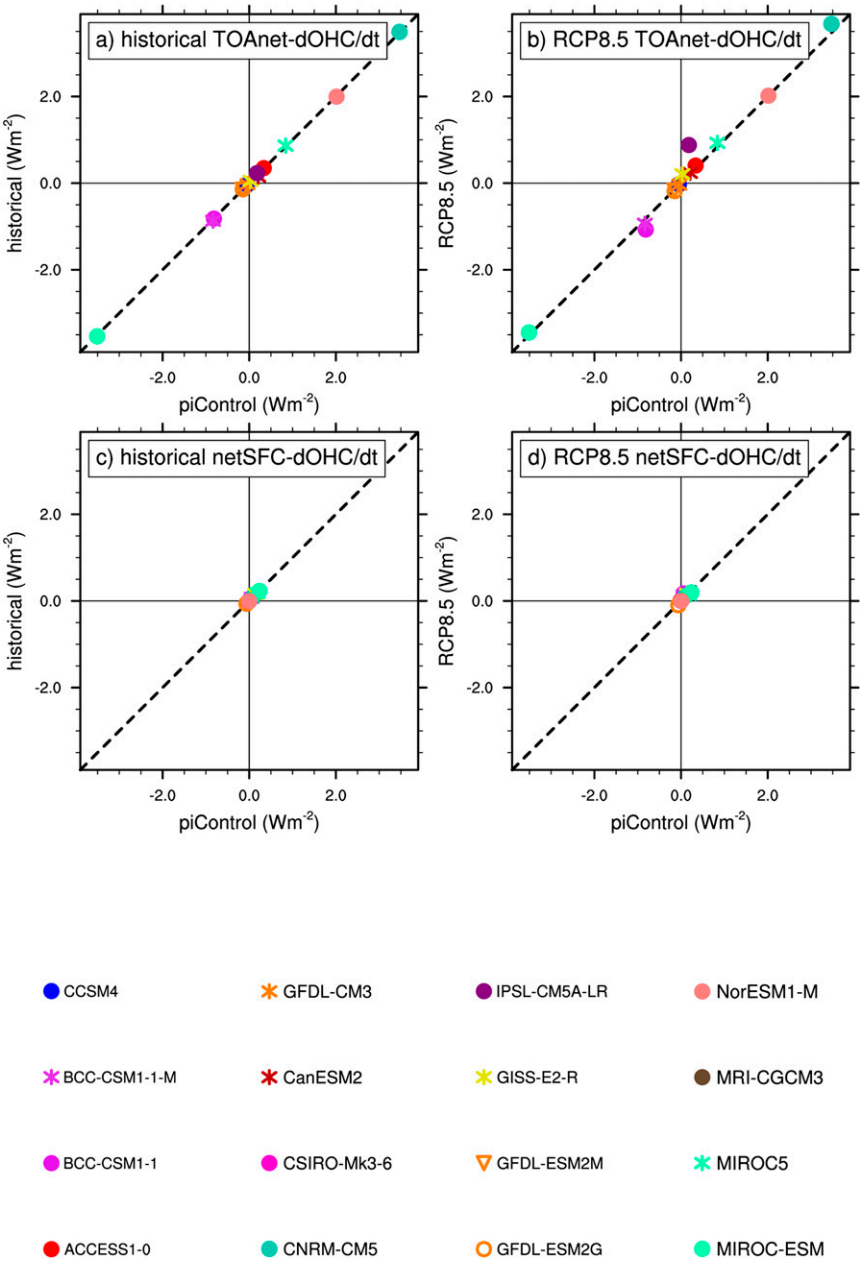

FIG. 3. Time-mean nonconservation terms (W m^{-2}) by model, for the piControl (x axes) and perturbed simulations (y axes) netSFC − dOHC/dt: (a) piControl against historical, (b) piControl against RCP8.5, (c) piControl against historical, and (d) piControl against RCP8.5. Dashed black lines show the 1:1 relationships.

uptake (Gregory 2000), so the response shown by IPSL-CM5A-LR and CNRM-CM5 could be physically reasonable. However, the other models all show a modest intensification of the nonconservation term under the RCP8.5 forcing (Table 2), whereas if the rate of ocean heat uptake were reducing across all models we would expect netTOA − dOHC/dt to increase in all the models, regardless of the sign of their uncorrected nonconservation term. It is also not clear under this scenario whether the climate response would absorb the heat elsewhere in the climate system or would act to increase outgoing longwave radiation (hence reducing netTOA). For example, an atmospheric absorption of approximately 0.4 W m^{-2} (as for CNRM-CM5) is equivalent to a rather drastic atmospheric warming of 1.2 K yr^{-1}, which is not the case in this particular simulation.

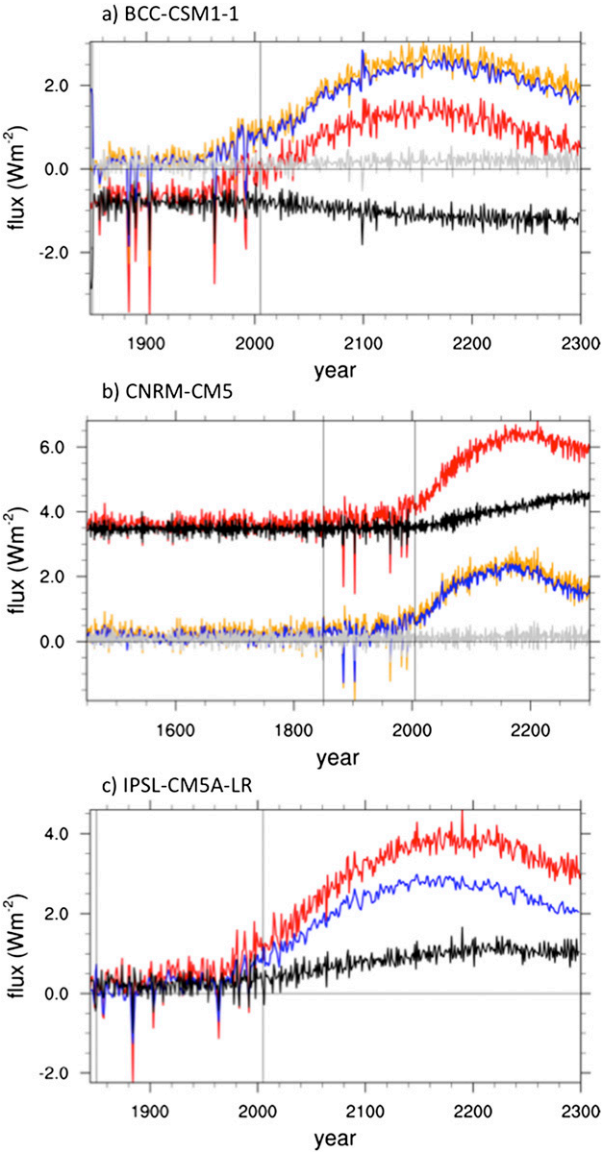

FIG. 4. Time series of annual-mean netTOA (red), *d*OHC/*dt* (blue), and nonconservation (black), concatenated across the piControl, historical, and RCP8.5 simulations, for (a) BCC_CSM1.1, (b) CNRM-CM5, and (c) IPSL-CM5A-LR. Where available, netSFC (orange) and netSFC − *d*OHC/*dt* (gray) are also plotted.

On the other hand, a known response to climate forcing is the intensification of the global hydrological cycle and increased atmospheric humidity. The hydrological cycle is a particular and problematic source of energy leaks in climate models; atmospheric moisture imbalances of the order of 0.01 Sv ($1\,\mathrm{Sv} \equiv 10^6\,\mathrm{m^3\,s^{-1}}$) can result in an excess latent heat term in the atmosphere of approximately $0.05\,\mathrm{W\,m^{-2}}$ (Liepert and Previdi 2012). A climate model is usually constructed by coupling separate model components (ocean, atmosphere, land surface, and cryosphere). Phase state changes can be particularly problematic for budgets when the cycle of phase changes occurs in different model components. For example water freezes as snow in the atmosphere's cloud scheme, releasing latent heat; this snow "falls" into the separate land surface model where it melts, absorbing latent heat. Not only must each component track these energy terms, but also a coupler must balance them. Under a forced intensification of the hydrological cycle, these energy leaks could well be magnified. This is just one example of the many transfers of mass and energy between different model components, and serves to highlight the difficulties of achieving a closed energy budget in a coupled framework.

Another question is whether the magnitude of energy imbalances in the models can impact the magnitude of forced response for variables that are indirectly related to the energy budget. Figure 5 compares the model piControl run energy biases with the TCR. None of the energy balance variables (netTOA, *d*OHC/*dt*, and the nonconservation term) seems to be a useful predictor for a model's TCR, and this was confirmed by performing a cross-correlation between TCR and the energy balance variables. Visual inspection of Figs. 5a and 5c might suggest that the strong netTOA biases in the CNRM-CM5 and MIROC-ESM models could mask relationships between netTOA or the nonconservation term and TCR. However, cross-correlations remain statistically insignificant even after removal of these models.

c. *Relationship with drift*

We have shown that after applying a simple linear correction for energy leaks, the CMIP5 models have a reasonable energy balance at the global scale. Furthermore, the majority of models in this study have energy leaks that are relatively insensitive to changes in the external forcing used. This is useful because it formally shows that control run data, in which any long-term nonzero energy imbalance is a bias, can be used to correct for the effects of nonconservation in perturbed simulations. The next question is this: What is the appropriate drift model to apply? The simplest approach is to calculate a linear trend over the length of a control run for the state variable of interest, and remove that trend from all experiments of interest. Applying this approach to the globally integrated ocean temperature is equivalent to removing the time-mean ocean heat content flux. However, as shown in Fig. 2 the energy leaks resulting in drift may not be constant (i.e., they may have trends or cyclical behavior), and previous studies suggest that linear drift correction is insufficient (Gleckler et al. 2012). Higher-order drift models (e.g.,

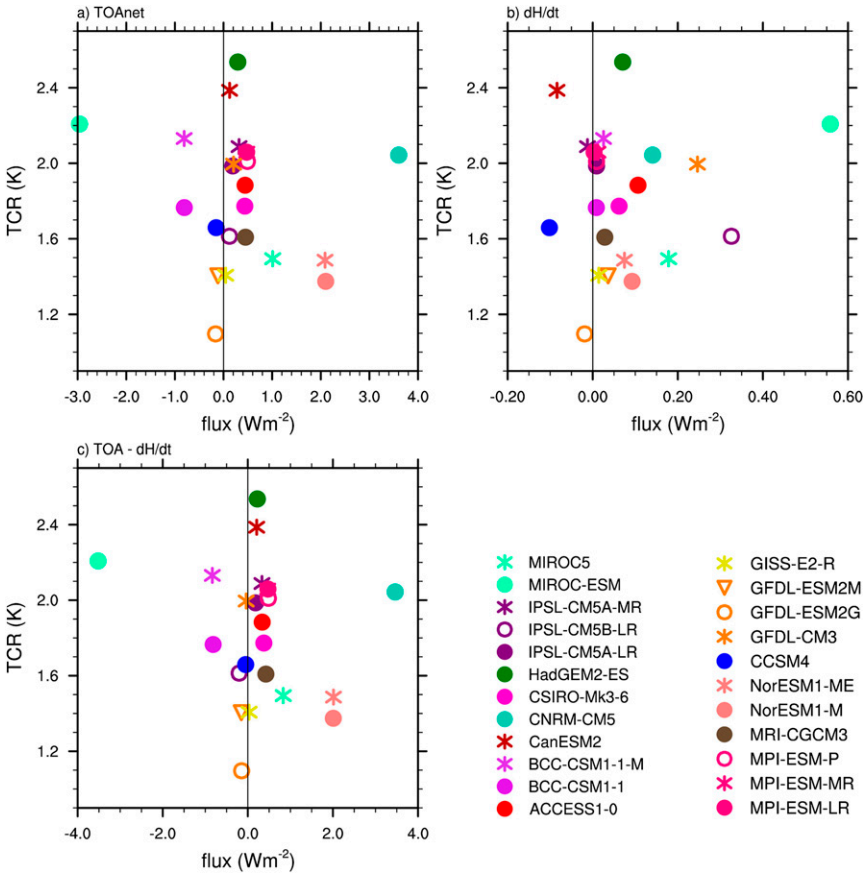

FIG. 5. The piControl time-mean flux terms (W m^{-2}; *x* axes) against TCR, where TCR is the change in decadal-mean surface air temperature following a doubling of atmospheric CO_2 after a 1% yr^{-1} increase.

quadratic or cubic trends) are sometimes employed (Gleckler et al. 2012; Sen Gupta et al. 2013), but in order to use these models to ''dedrift'' a perturbed simulation it is necessary to know the time period in the control simulation that parallels the forced simulation. (That is to say, CMIP5 historical simulations are initialized from a time step in the piControl simulation, referred to as the branch time; to remove a cubic or quadratic drift from the historical simulation one must know this branch time, which may or may not be supplied accurately as metadata.) Another approach to dedrifting control run data is to apply a long time scale high-pass filter (Palmer and McNeall 2014), but this is computationally intensive and not easily applied to perturbed simulations. Given the ambiguity in appropriate drift correction procedures, it is useful to consider the actual sensitivity to the choice of drift correction procedure.

Figure 6 shows the relationship between netTOA and ocean heat content change, after dedrifting using the different methods outlined above. As expected from Table 2, to a greater or lesser extent the models all show an approximately one-to-one relationship between netTOA and *d*OHC/*dt*. Importantly, none of the models shows a strong dependence on the method of drift correction; the results are largely the same regardless of the method used. Given its computational efficiency and ease of application to perturbed experiments, this suggests that a simple linear drift removal is an appropriate choice for global quantities, and more complex methods do not offer significant advantages. Furthermore, the simple linear drift correction does not perform appreciably worse than the more involved techniques used in Fig. 6. Where suspicious long-term variability is evident in the control run, researchers may be wise to perform the comparison that we make in Fig. 6 to guide choice, and here the concept of agreement between *d*OHC/*dt* and netTOA can be a useful guide. Whether the validity of the simple first-order correction holds true for regional variables is not assessed here, and we note that Sen Gupta et al. (2013) showed that there may be considerable spatial variations in model drift.

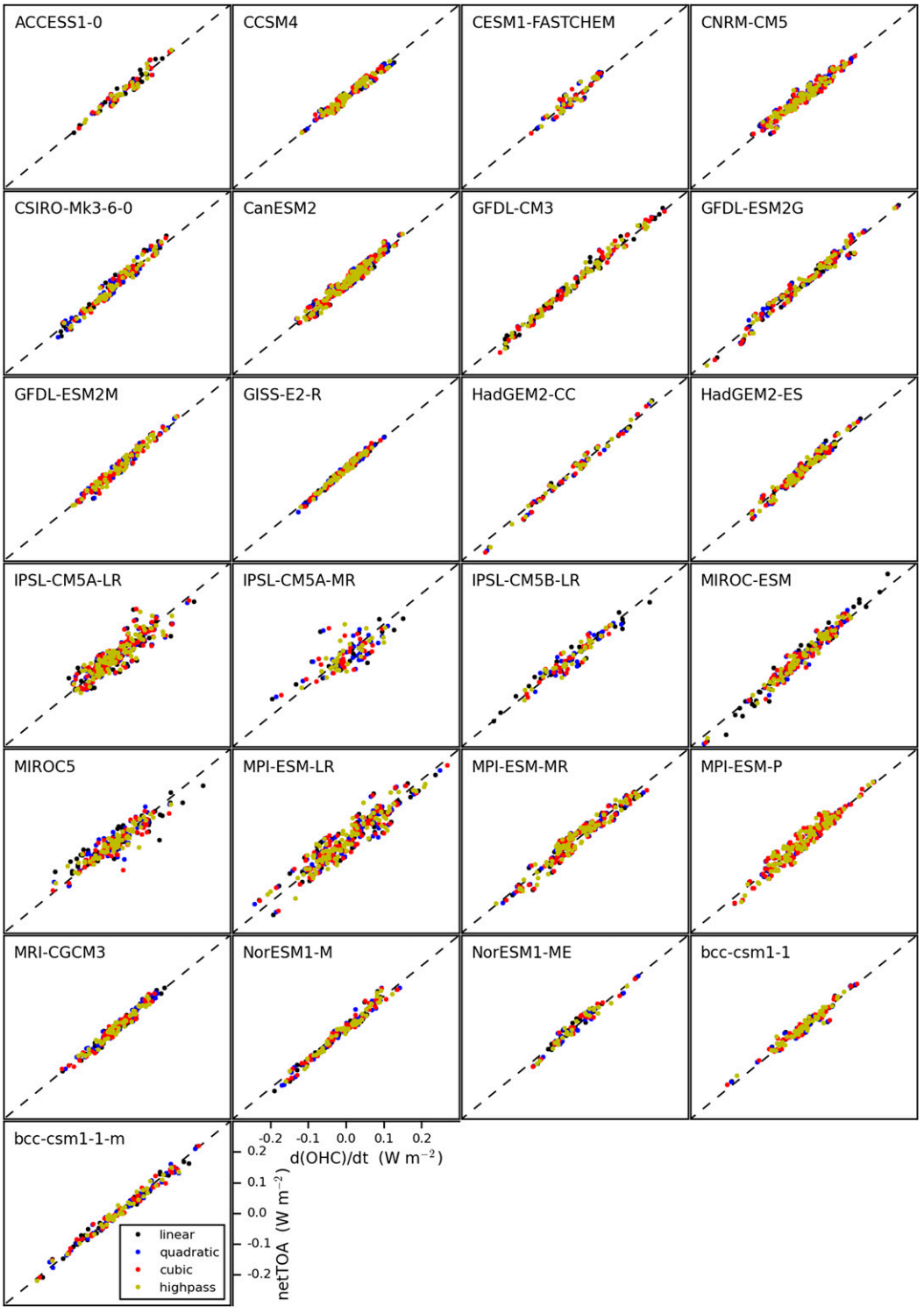

FIG. 6. Control run decadal-mean dOHC/dt (x axis) against netTOA (y axis), after drift correction using a linear trend (black markers), quadratic trend (blue), cubic trend (red), or a high pass filter with 100-yr cutoff (yellow).

4. Discussion and conclusions

We have presented an analysis of global energy balance and model drift in a range of coupled climate and Earth system models in the CMIP5 archive. Although we have considered only globally integrated quantities, the analysis demonstrates a number of insights that will be useful to researchers interpreting coupled model output. All the models analyzed have significant spurious top-of-atmosphere energy imbalances, and in some cases the imbalance is at least as large as the observed late-twentieth-century netTOA (i.e., the true climate forcing signal). In a model with a realistic energy balance, we would expect changes in ocean temperature to account for almost the entire energy imbalance, and indeed all the models analyzed have some spurious ocean heat content trend (i.e., drift). However, this rate of ocean heat content change is generally very different from the 90% of netTOA expected in the real climate, which implies significant nonconservation in the models. For most of the models that provided the net ocean surface flux variable, more than 80% of the nonconservation occurs between the ocean surface and the top of the atmosphere. This implies that

1) ocean model drift is largely (but not entirely) a physically valid response to surface heat flux, and
2) in most of the models, 80% or more of the energy leaks that lead to the nonconservation are outside the ocean, but a few of the models have a much greater proportion of drift due to internal ocean processes.

Consistent with previous work (Palmer and McNeall 2014), we show that on decadal time scales there is a close relationship between top-of-atmosphere energy balance and ocean heat content. This indicates that the models are energy conserving once their long-term nonconservation terms are accounted for. Furthermore, for most models the response to climate forcing does not seem to be affected by the magnitude of its energy bias or nonconservation term.

The nonconservation in these models is overwhelmingly characterized by a time-constant bias (i.e., the nonconservation is stationary). Some of the models have nonconservation terms with some organized variability (trends or multidecadal oscillations), with the magnitude of these nonstationarities being small compared to the mean (although for IPSL-CM5B-LR the nonstationarity is large enough to induce a ''curvature'' in the drift). As well as being constant in time, the energy biases are largely insensitive to changes in the models' forcing regimes (i.e., the nonconservation is the same for all experiments). This has a number of useful practical implications for the correction of model energy balances (e.g., in the calculation of equilibrium climate sensitivity) or for drift correction of global-mean state variables (e.g., ocean or surface air temperature). First, this result formally demonstrates that nonzero energy balances or drifts in a model's control run can be safely used to correct variables in the same model's perturbed experiments, as is commonly the case in studies requiring drift or energy balance correction. Second, it suggests that the simplest form of drift correction—the removal of a time-constant linear trend derived from the control run—is adequate, and that more complicated procedures add little improvement to the drift correction procedure. From a philosophical standpoint, we also argue that this approach reduces the risk of overfitting a drift correction term, since it may not be clear whether a significant curvature term in the control run time series due to drift or a valid mode of climate variability (Sen Gupta et al. 2013). However, this also raises the interesting question, which we have not addressed here, of why the transient energy balance is unaffected by time-mean energy leaks.

Finally, the comparison of globally integrated energy balance terms is a useful first-order diagnostic that can provide insight into model behavior as we have demonstrated here, and also as a first-pass check on postprocessed model output. Submission of formatted model output to the CMIP5 archive, and the upcoming CMIP6 (Meehl et al. 2014), represents an enormous effort by the participating modeling groups, and inevitably errors will be present in the archive, some of which may be subtle and hidden by the sheer number of output files. The expected good agreement between global netTOA and OHC changes (drift notwithstanding) provides a simple and effective quality assurance and can highlight errors in the postprocessed model output, such as missing or overlapping time periods. For CMIP6, all the core variables used in this study (global-mean ocean temperature, net heat flux into the ocean, and net top-of-atmosphere radiation) have been given the highest priority, and the archiving of model reference density and specific heat capacity has been recommended (Griffies et al. 2015). These steps should make global energy-balance analyses significantly easier, and we recommend that researchers employ these tools for understanding, model verification, and output quality control.

Acknowledgments. The authors are indebted to Dr. Jonathon Gregory, Dr. Alex Sen Gupta, and two anonymous reviewers for helpful comments in the preparation of this manuscript, and to Paola Petrelli for

help in managing the CMIP5 data. DM is partly supported financially by the Australian Climate Change Science Program. Much of the analysis took place during a visiting science fellowship undertaken by MDP that was sponsored by CSIRO and the University of Tasmania. MDP was supported by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101) and his work represents a contribution to the Natural Environment Research Council DEEP-C project NE/K005480/1. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. Much of the analysis was performed using the NCAR Command Language (NCL).

REFERENCES

Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation of the physical climate. *Geosci. Model Dev. Discuss.*, **6**, 687–720, doi:10.5194/gmd-6-687-2013.

Bi, D., and Coauthors, 2013: The ACCESS coupled model: Description, control climate and evaluation. *Aust. Meteor. Oceanogr. J.*, **63**, 41–64.

Church, J. A., and Coauthors, 2013: Sea level change. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1137–1216.

Cubasch, U., and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 525–582.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, doi:10.1175/JCLI-D-11-00560.1.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:10.1175/2011JCLI4083.1.

Gleckler, P. J., and Coauthors, 2012: Human-induced global ocean warming on multidecadal timescales. *Nat. Climate Change*, **2**, 524–529, doi:10.1038/nclimate1553.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect an time-dependent climate change. *Climate Dyn.*, **16**, 501–515, doi:10.1007/s003820000059.

Griffies, S. M., and Coauthors, 2011: The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations. *J. Climate*, **24**, 3520–3544, doi:10.1175/2011JCLI3964.1.

——, and Coauthors, 2015: Sampling the physical ocean in CMIP6 simulations. World Climate Research Panel, 68 pp. [Available online at http://www.clivar.org/documents/sampling-physical-ocean-cmip6-simulations.]

Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94**, 1339–1360, doi:10.1175/BAMS-D-12-00121.1.

Jeffrey, S., L. D. Rotstayn, M. A. Collier, S. M. Dravitzki, C. Hamalainen, C. Moeseneder, K.-K. Wong, and J. I. Syktus, 2013: Australia's CMIP5 submission using the CSIRO-Mk3.6 model. *Aust. Meteor. Oceanogr. J.*, **63**, 1–13.

Jungclaus, J. H., and Coauthors, 2013: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.*, **5**, 422–446, doi:10.1002/jame.20023.

Kuhlbrodt, T., and J. M. Gregory, 2012: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophys. Res. Lett.*, **39**, L18608, doi:10.1029/2012GL052952.

Liepert, B. G., and M. Previdi, 2012: Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models. *Environ. Res. Lett.*, **7**, 014006, doi:10.1088/1748-9326/7/1/014006.

Martin, G. M., and Coauthors, 2011: The HadGEM2 family of Met Office Unified Model climate configurations. *Geosci. Model Dev.*, **4**, 723–757, doi:10.5194/gmd-4-723-2011.

Meehl, G. A., R. Moss, K. E. Taylor, V. Eyring, R. J. Stouffer, S. Bony, and B. Stevens, 2014: Climate model intercomparisons: Preparing for the next phase. *Eos, Trans. Amer. Geophys. Union*, **95**, 77–78, doi:10.1002/2014EO090001.

Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241, doi:10.1007/s10584-011-0156-z.

Mignot, J., and S. Bony, 2013: Presentation and analysis of the IPSL and CNRM climate models used in CMIP5. *Climate Dyn.*, **40**, 2089, doi:10.1007/s00382-013-1720-1.

Palmer, M. D., and D. J. McNeall, 2014: Internal variability of Earth's energy budget simulated by CMIP5 climate models. *Environ. Res. Lett.*, **9**, 034016, doi:10.1088/1748-9326/9/3/034016.

——, ——, and N. J. Dunstone, 2011: Importance of the deep ocean for estimating decadal changes in Earth's radiation balance. *Geophys. Res. Lett.*, **38**, L13707, doi:10.1029/2011GL047835.

Rhein, M., and Coauthors, 2013: Observations: Ocean. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 255–315.

Roemmich, D., J. Church, J. Gilson, D. Monselesan, P. Sutton, and S. Wijffels, 2015: Unabated planetary warming and its ocean structure since 2006. *Nat. Climate Change*, **5**, 240–245, doi:10.1038/nclimate2513.

Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.*, **6**, 141–184, doi:10.1002/2013MS000265.

Sen Gupta, A., N. C. Jourdain, J. N. Brown, and D. Monselesan, 2013: Climate drift in the CMIP5 models. *J. Climate*, **26**, 8597–8615, doi:10.1175/JCLI-D-12-00521.1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Voldoire, A., and Coauthors, 2013: The CNRM-CM5.1 global climate model: Description and basic evaluation. *Climate Dyn.*, **40**, 2091–2121, doi:10.1007/s00382-011-1259-y.

Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.

Watanabe, S., and Coauthors, 2011: MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geosci. Model Dev.*, **4**, 845–872, doi:10.5194/gmd-4-845-2011.

Xin, X.-G., T.-W. Wu, and J. Zhang, 2013: Introduction of CMIP5 experiments carried out with the climate system models of Beijing Climate Center. *Adv. Climate Change Res.*, **4**, 41–49, doi:10.3724/SP.J.1248.2013.041.

Yang, D., and O. A. Saenko, 2012: Ocean heat transport and its projected change in CanESM2. *J. Climate*, **25**, 8148–8163, doi:10.1175/JCLI-D-11-00715.1.

Yukimoto, S., and Coauthors, 2012: A new global climate model of the Meteorological Research Institute: MRI-CGCM3—Model description and basic performance. *J. Meteor. Soc. Japan*, **90A**, 23–64, doi:10.2151/jmsj.2012-A02.