Climate Drift in the CMIP3 Models

ALEXANDER SEN GUPTA

Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia

LES C. MUIR AND JACLYN N. BROWN

Centre for Australian Weather and Climate Research, CSIRO Wealth from Oceans National Research Flagship, Hobart, Tasmania, Australia

STEVEN J. PHIPPS

Climate Change Research Centre, University of New South Wales, Sydney, New South Wales, Australia

PAUL J. DURACK, DIDIER MONSELESAN, AND SUSAN E. WIJFFELS

Centre for Australian Weather and Climate Research, CSIRO Wealth from Oceans National Research Flagship, Hobart, Tasmania, Australia

ABSTRACT

Even in the absence of external forcing, climate models often exhibit long-term trends that cannot be attributed to natural variability. This so-called climate drift arises for various reasons including the following: perturbations to the climate system on coupling component models together and deficiencies in model physics and numerics. When examining trends in historical or future climate simulations, it is important to know the error introduced by drift so that action can be taken where necessary. This study assesses the importance of drift for a number of climate properties at global and local scales. To illustrate this, the present paper focuses on simulated trends over the second half of the twentieth century. While drift in globally averaged surface properties is generally considerably smaller than observed and simulated twentieth-century trends, it can still introduce nontrivial errors in some models. Furthermore, errors become increasingly important at smaller scales. The direction of drift is not systematic across different models or variables, as such drift is considerably reduced in the multimodel mean. Despite drift being primarily associated with ocean adjustment, it is also apparent in atmospheric variables. For example, most models have local drift magnitudes in surface air and ocean temperatures that are typically between 15% and 35% of the twentieth-century simulation trend magnitudes for 1950–2000. Below depths of 1000–2000 m, drift dominates over any forced trend in most regions. As such steric sea level is strongly affected and for some models and regions the sea level trend direction is reversed. Thus depending on the application, drift may be negligible or may make up an important part of the simulated trend.

1. Introduction

Climate models are vital tools for helping us understand and attribute long–term changes in the global climate system. These models allow us to make physically plausible projections of how the ocean–atmosphere system might evolve in the future under given greenhouse gas emission scenarios. Models are not complete or perfect replicas of the real world, however. Many physical processes are only approximated or parameterized in the models while others are omitted entirely. This can lead to biases in the simulated climate. Here we focus on a particular problem inherent in coupled climate models that lead to spurious trends in climate simulations. This is commonly referred to as climate drift.

Climate drift is primarily associated with deficiencies in either the model representation of the real world or the procedure used to initialize the model. Over long time scales drift is primarily associated with slow adjustments

Corresponding author address: A. Sen Gupta, Climate Change Research Centre, University of New South Wales, Kensington, Sydney NSW 2052, Australia.
E-mail: a.sengupta@unsw.edu.au

TABLE 1. CMIP3 spinup information; data primarily from PCMDI model description and experiments table. Column 3 describes how the ocean-only model was initialized (and if available from what available climatology) prior to coupling. Column 4 is the number of years of coupled spinup prior to start of the preindustrial control output (PICNTRL) archived at PCMDI. Column 5 is the additional number of years from the start of the archived PICNTRL output prior to the start of 20C3M integrations (in some cases multiple 20C3M simulations are performed from different starting points in the PICNTRL simulation). Column 6: flux adjustment applied (if used): H = heat, F = freshwater, M = momentum. LEV82 is Levitus (1982); LEV94 is Levitus et al. (1994), Levitus and Boyer (1994); LEV98 is Conkright et al. (1999); and LEV is an unspecified Levitus database. PCMDI: http://www-pcmdi.llnl.gov/ipcc/time_correspondence_summary.htm; http://www-pcmdi.llnl.gov/ipcc/model_documentation/ipcc_model_documentation.php.

	Model	Initialized from	Coupled spinup (prior to PICNTRL) (yr)	Additional PICNTRL prior to 20C3M (yr)	Flux adjustment	Information source
1	BCCR-BCM2.0	LEV82	80	+0	—	PCMDI
2	CGCM3.1(T47)	LEV	250	+0	HF	(O. Saenko 2011, personal communication)
3	CGCM3.1(T63)	LEV	50	+0	HF	PCMDI
4	Centre National de Recherches Météorologiques Coupled Global Climate Model, version 3 (CNRM-CM3)	Older model initialized from LEV82	70	+110	—	PCMDI
5	CSIRO-Mk3.0	Individually spun up components	120	+0 to +20 (3 members)	—	PCMDI; (Collier et al. 2004)
6	CSIRO-Mk3.5	Individually spun up components	170	+0 to +40 (3 members)	—	PCMDI; (Gordon et al. 2002)
7	GFDL Climate Model version 2.0 (CM2.0)	LEV	300	+1 to +151 (3 members)	—	PCMDI. Some forcing changes in spinup prior to yr 200
8	GFDL-CM2.1	LEV	300	+1 to +81 (3 members)	—	PCMDI
9	GISS Atmosphere–Ocean Model (GISS-AOM)	LEV	200–250	0	—	PCMDI
10	GISS-EH	LEV	0	+120 to +160 (9 members)	—	PCMDI
11	GISS-ER	Older model version	200	+5 to +100 (9 members)	—	PCMDI
12	Flexible Global Ocean–Atmosphere–Land System Model gridpoint version 1.0 (FGOALS-g1.0)	Individually spun up components	100	+0 to +10 (3 members)	—	PCMDI
13	INGV Scale Interaction Experiment [SINTEX-G (SXG); ECHAM4]	Individually spun up components from LEV	0	+100	—	PCMDI (description unclear)
14	INM-CM3.0	LEV	280	+0	F	PCMDI. (forcing modified yr 200)
15	L'Institut Pierre-Simon Laplace Coupled Model, version 4 (IPSL CM4)	LEV82	330	+0	—	PCMDI.
16	MIROC3.2(hires)	Snapshot of previous version	109	+0	—	PCMDI.
17	MIROC3.2 medium-resolution version [MIROC3.2(medres)]	Snapshot of previous version	109/300	+0	—	PCMDI. (conflicting information)
18	Meteorological Institute of the University of Bonn, ECHO-G Model (MIUBECHOG)	Older model version from LEV	6	+0 to +200 (5 members)	HF	PCMDI

TABLE 1. (Continued)

	Model	Initialized from	Coupled spinup (prior to PICNTRL) (yr)	Additional PICNTRL prior to 20C3M (yr)	Flux adjustment	Information source
19	Max Planck Institute (MPI) ECHAM5	300 years of older version coupled simulation	144	+68 to +118 (3 members)	—	PCMDI
20	MRI CGCM2.3.2	LEV + 428 present-day forcing coupled spinup	450	+1 to +200 (5 members)	HFM	PCMDI.
21	National Center for Atmospheric Research (NCAR) Community Climate System Model, version 3 (CCSM3)	LEV98	300	+80 to +260 (9 members)	—	CCSM (spin-up history undocumented at PCMDI) (Large and Danabasoglu 2006)
22	NCAR Parallel Climate Model (PCM)	Individually spun up components	50	+100–120 (3 members)	—	PCMDI; (Washington et al. 2000) Information unclear
23	Third climate configuration of the Met Office (UKMO) Unified Model (HadCM3)	LEV94	300/360	+1 to +100/150 (2 members)		PCMDI; conflicting information provided
24	UKMO Hadley Centre Global Environmental Model version 1 (HadGEM1)	LEV98	85	+0		PCMDI.

in the simulated ocean that are independent of any external factors such as increased greenhouse gases. As a result, in trying to understand the long-term rate of change in a climate simulation arising from external forcing, we need to pay heed to both low-frequency natural variability and any spurious drift that exists in the model. Significant efforts by the climate modeling community have gone into reducing climate drift, which has meant that most climate models can now be successfully run without the need for unphysical flux adjustments. Nevertheless climate drift still persists. In this paper we quantify the size of drift relative to twentieth-century trends in climate models taking part in the Coupled Model Intercomparison Project phase 3 (CMIP3), which was used to inform the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4).

Climate drift tends to operate on two distinct time scales (previously referred to as "major drift" for large magnitude rapid drift and "minor drift" for slow long time scale drift, Cai and Gordon 1999). Large discontinuities in surface fluxes during the coupling of the various component models can cause rapid drift. The initial adjustment of the atmosphere, surface ocean, and sea ice to this perturbation is relatively fast, with a new equilibrium generally being achieved after a few years. A variety of techniques have been used to reduce this "coupling shock." For example, individual model components can be equilibrated using different boundary forcing combinations in sometimes quite elaborate multistage spinup procedures (e.g., Moore and Gordon 1994; Power 1995; Cai and Chu 1996; Large et al. 1997). However, some initial, albeit reduced adjustment to the coupling, generally persists.

A more pervasive problem relates to the millennial adjustment time scale of the deep ocean. This may be associated with various factors including 1) deficiencies in the model physics, 2) inaccuracies in the model formulation (which might for example lead to heat or salt/freshwater not being conserved), 3) the propagation of discontinuities associated with coupling shock through the ocean interior, and 4) a sparsity in the observational data used to initialize the model. Climate model simulations are often initialized from some observational dataset (Table 1). Given a perfect model and a perfect set of observations, the simulated climate system should be initialized in a dynamical balance. However, deficiencies in model physics mean that a model's dynamical balance will be different to that of the real world. The model will therefore drift. A lack of observational data, particularly in the deep ocean, and the need to interpolate this data will also mean that a dynamically consistent observationally based initial state is unlikely to exist. As such even with a perfect model some drift would still occur.
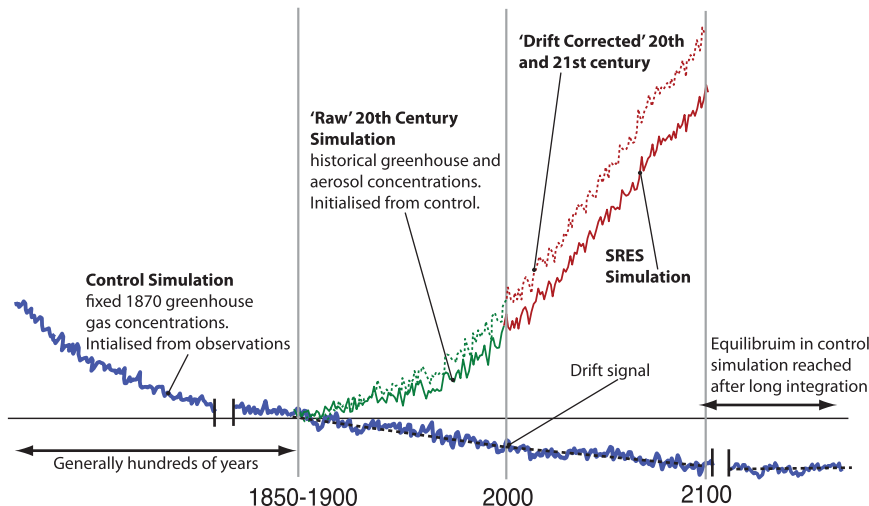
FIG. 1. Schematic of the temporal evolution of a simulated state variable (e.g., global temperature) during a preindustrial control simulation and subsequent simulations of the twentieth century and future scenarios.

Climate drift can, in principle, be alleviated via long model integrations. Such integrations are feasible and routinely done for low-resolution models (e.g., Phipps et al. 2011). However, to perform such simulations at the higher resolution used for climate projections is at present computationally prohibitive. Flux adjustments have also been widely used in the past whereby predetermined heat and/or freshwater adjustments are made over the duration of long climate simulations (Sausen et al. 1988). Such adjustments are a pragmatic partial solution to drift; however, they are also inherently nonphysical. It is therefore hard to envisage a satisfactory near-term solution to the issue of climate drift. As the models become more realistic, however, we would expect the problem to become less significant. This is already apparent. In the Coupled Model Intercomparison Project phase 2 (CMIP2), 10 of the 17 models employed ad hoc and nonphysical flux adjustments to reduce climate drift to maintain a relatively stable simulated climate state (Houghton et al. 2001; Räisänen 2001). Even with flux adjustment, however, drift was still evident (Covey et al. 2006). In CMIP3, only 6 of the 24 contributing models used flux adjustments. Yet despite the removal of flux adjustment the replication of the observed climate has improved considerably (Reichler and Kim 2008). This abandonment of flux adjustment can be partly attributed to improved and more physically consistent model parameterizations, increased resolution, and dynamical cores in the updated models. The shift was driven by a general discomfort with the use of physically untenable techniques. Previous work has demonstrated, for example, that the details of flux adjustment can have major impacts on transient simulations (Neelin and Dijkstra 1995; Tziperman 2000). Tziperman (2000), for instance, show that two equally plausible flux adjustment formulations can lead to a recovery or a sustained slowdown of the thermohaline circulation after an initial warming induced slowdown.

The quantification of drift requires the examination of control simulations in which forcing terms (e.g., solar irradiance, greenhouse gases) are maintained at fixed levels. Any long-term trend in these control simulations will be due to climate drift (and possibly low-frequency variability). Forced simulations that are initialized from these control simulations will therefore also contain a trend component that is spurious and associated with drift. In the CMIP3 models the twentieth-century hindcast simulations (20C3M, which extends from the late nineteenth century to ~2000) are initialized from a long preindustrial control simulation under constant late nineteenth-century boundary conditions. As the control simulation is, for most models, integrated beyond this branching point, a period of temporal overlap is available, which can in principle be used to identify, and where necessary remove, the drift from the forced simulation (Fig. 1). However, given a concurrent record of forced and control simulations (which is not guaranteed for all model/variable combinations in the CMIP3 repository) three interrelated complications exist. First, as with the identification of a forced trend, identifying the drift is hampered by inherent natural variability. When short time periods are analyzed or low-frequency natural variability exists, aliasing can produce spurious trends unrelated to either climate drift or external forcing.

Second, the trajectory of the drift is unknown. Long integrations (e.g., with models used for paleoclimate analysis) often show that for large-scale metrics (e.g., global temperature) there is an asymptotic approach to a final state that is approximately linear on sufficiently short time scales (see for example Cai and Gordon 1999, Fig. 3). However, at a regional scale, the form of the drift may be more complex as changes can propagate to a region via multiple pathways. Third, it is likely that there are nonlinear interactions between drift and the climate state. The fact that the rate of drift slows as the climate state approaches equilibrium demonstrates that drift is sensitive to the mean state. As the mean state of a forced simulation in general diverges from that of the corresponding control simulation, it seems inevitable that any drift in the control simulation will increasingly become a worse proxy for drift within the forced simulation with time.

A number of methods have been used to remove the drift from a forced simulation. Commonly a linear trend is calculated for an overlapping portion of the control simulation, either for some metric (e.g., global-average temperature) or on a grid point by grid point basis (Gregory and Lowe 2000; Meehl et al. 2007b; Sloyan and Kamenkovich 2007; Katsman et al. 2008; Santer et al. 2009; Sen Gupta et al. 2009; Downes et al. 2010). This trend from the control simulation (i.e., the drift) is then subtracted from the forced simulation trend. This technique assumes a linear drift that is coherent in the control and forced simulations and that low frequency variability is not significantly biasing the calculation of the linear trend. This method is equivalent to subtracting the control simulation from the forced simulation on a time step by time step basis and finding a linear trend from the resulting time series. An attempt to use the differenced time series to map out the temporal evolution of the forced response (as opposed to the long-term trend) would be somewhat misleading, however, as the subtraction of the control adds spurious variability to the resulting time series (e.g., Gregory et al. 2001; Sun and Hansen 2003). This can be easily demonstrated by subtracting two random time series. The variance of the differenced time series is equal to the sum of the variances of the constituent time series.

Another approach is to find the difference in a given climate variable between two time slices in both the forced and the control simulation. Subtraction of these differences provides an estimate of the forced signal. This method has the advantage of requiring less model output to calculate the forced change, but conversely is subject to more extreme aliasing as a result of natural variability.

More complex drift removal techniques may fit higher-order functions to the control output (e.g., Gregory et al.

2001, 2006; Ammann et al. 2007). While a linear drift seems justifiable on shorter time scales (as it is driven by slow ocean adjustment), the assumption becomes less tenable on longer time scales, as we might expect drift to diminish. As such for long simulations the use of higher-order drift removal is probably desirable (e.g., Gregory et al. 2006, uses a cubic polynomial drift to examine multicentennial sea level change). For shorter time scales, such higher-order drift techniques would be more likely to confuse drift and low-frequency natural variability.

Previous studies have examined the mechanisms driving drift in individual climate models (Rahmstorf 1995; Power 1995; Bryan 1998; Cai and Gordon 1999). In addition, Covey et al. (2006) performed a comprehensive assessment of drift for the previous generation of CMIP2+ models. They noted that while these models showed considerable improvement over those used in the previous intercomparison, the problem of drift, although improved, was still significant. To our knowledge no similar systematic assessment of the importance of drift has been conducted for the CMIP3 models. As such, our aim is to quantify the scale of the drift problem, in the context of simulated trends over the second half of the twentieth century, for the current generation of CMIP3 climate models, thereby updating the work of Covey et al. (2006). In particular we identify under what circumstances drift is important relative to the forced trends when presenting estimates of climate change. We demonstrate that drift remains an important factor, that must be taken into account not only in the analysis of ocean changes but also when estimating forced trends in the atmosphere.

In the remainder of this paper we provide a brief introduction to the CMIP3 models (section 2), discuss the method chosen for our drift estimation (section 3), and evaluate the size of the surface and interior ocean drift (section 4) for a variety of climate parameters. Finally a discussion and recommendations are provided in section 5.

2. CMIP3 models

In our examination of climate drift we use output from CMIP3, archived by the Program for Climate Model Diagnoses Intercomparison (PCMDI). CMIP3 is an initiative of the World Climate Research Programme (WCRP) to bring together output from an unprecedented array of 24 climate models used to inform the IPCC AR4 (Table 1; for details on the initiative see Meehl et al. 2007a). Output from two standard experiments is examined here: 1) a preindustrial control simulation (PICNTRL) incorporating a seasonally varying but annually unchanging forcing indicative of the late nineteenth century and 2) a climate simulation from the late nineteenth

century to ~2000 (20C3M) that incorporates observed greenhouse gas concentrations, with a subset of models also including variable solar radiation, volcanic aerosols, and anthropogenic ozone (along with other atmospheric forcing agents). We only examine the transient 20C3M simulations and no projection scenarios. In the 20C3M simulations the forced signal is still relatively weak (compared to future projection scenarios, which are subject to stronger radiative forcing) while the drift is relatively strong, as we might expect it to diminish with time. Examination of the twentieth century therefore provides a worst case estimate of the relative importance of climate drift to forced trend. For each model one or more 20C3M simulation is initialized from a snapshot of the corresponding PICNTRL simulation. There is considerable inconsistency with regard to the available variables, time spans, and numbers of realizations across the various experiments and models making up the CMIP3. As such analysis of different variables will generally involve slightly different subsets of models.

The CMIP3 repository contains a large set of nominally independent models encompassing a broad range of resolutions and incorporating a variety of different physical parameterizations (although component models and parameterizations are often implemented in more than one model; for more information see Sen Gupta et al. 2009, Table 1). For the ocean component, resolutions vary from the eddy-permitting Model for Interdisciplinary Research on Climate 3.2, high-resolution version [MIROC3.2(hires); $0.288° \times 0.198° \times 47$ levels] to the coarse-resolution Goddard Institute for Space Studies Model E-R (GISS-ER; $5° \times 4° \times 13$ levels). Most models employ a z-level vertical coordinate, although isopycnal, sigma, and hybrid schemes are also represented. They also implement some form of the Gent and McWilliams (1990) parameterization to account for the effect of unresolved eddy processes. As noted previously a major change from the CMIP2+ models is that the CMIP3 models—except for the INGV-ECHAM4, the global Hamburg Ocean Primitive Equation (ECHO-G) models, the Meteorological Research Institute Coupled General Circulation Model, version 2.3.2 (MRI CGCM2.3.2), the Institute of Numerical Mathematics Coupled Model, version 3.0 (INM-CM3.0), and the two Canadian Center for Climate Modelling and Analysis (CGCM3.1) models—do not use flux adjustments.

Table 1 shows information regarding model spinup times for the CMIP3 models. Spinup strategies employed by each modeling group and for each model are different. Some undertake a coupled spinup integration directly from an observed initial state, while others use a snapshot from a previous model realization or from separately spun up ocean and atmosphere model components. In comparison to CMIP2+ models (Covey et al. 2006, their Table 2) there has been a general reduction in spinup times (the time between coupling and the start of 20C3M simulations) in the CMIP3 generation of models. The median spinup time has gone from ~250 yr (Covey et al. 2006, their Table 2) to ~200 yr (CMIP3, Table 1). This is in part a consequence of improvements in model stability. However, it may also be symptomatic of the fact that the increase in model complexity and resolution and the need for multiple scenario experiments has overshadowed increases in computational power. In addition, timelines for inclusion in the CMIP–IPCC process often mean that modeling groups have insufficient time in which to perform long simulations. Table 1 contains various gaps and conflicting pieces of information, which are indicative of the lack of information provided by some of the modeling groups regarding the spinup procedures either to PCMDI or in the relevant model documentation.

3. Climate drift correction

Below we examine 1950–2000 trends from the 20C3M hindcasts for the CMIP3 models and investigate the bias introduced by climate drift. This period was selected as it has been the focus of recent new observational analysis (e.g. Durack and Wijffels 2010) and is only over the latter half of the twentieth century that there is strong evidence for an unambiguous anthropogenic signal in the observations (Solomon et al. 2007). As we are investigating relatively short time spans we make the common assumption that any drift will be approximately linear. This assumption would be less tenable if we were examining longer time scales over which drift might be expected to diminish over time. The short timespan also means that trend estimates, for both the PICNTRL and 20C3M simulations, are likely to be confounded by low-frequency natural variability. To reduce this effect we define the drift as a linear trend in the PICNTRL simulation over an extended period, where possible, 1900–2050 (which brackets 1950–2000). The time variable in PICNTRL experiments was offset, where required, such that the branch point for the PICNTRL and 20C3M simulations was labeled with the same year. Inspection of the control experiments suggests that drift tends to be approximately linear over this time scale. This time span is a subjective choice, but it provides a compromise between being sufficiently long to avoid some of the aliasing by natural variability yet sufficiently short to account for the fact that over long time scales the trajectory of the drift is likely to be nonlinear. Not all modeling groups provide sufficient PICNTRL output to
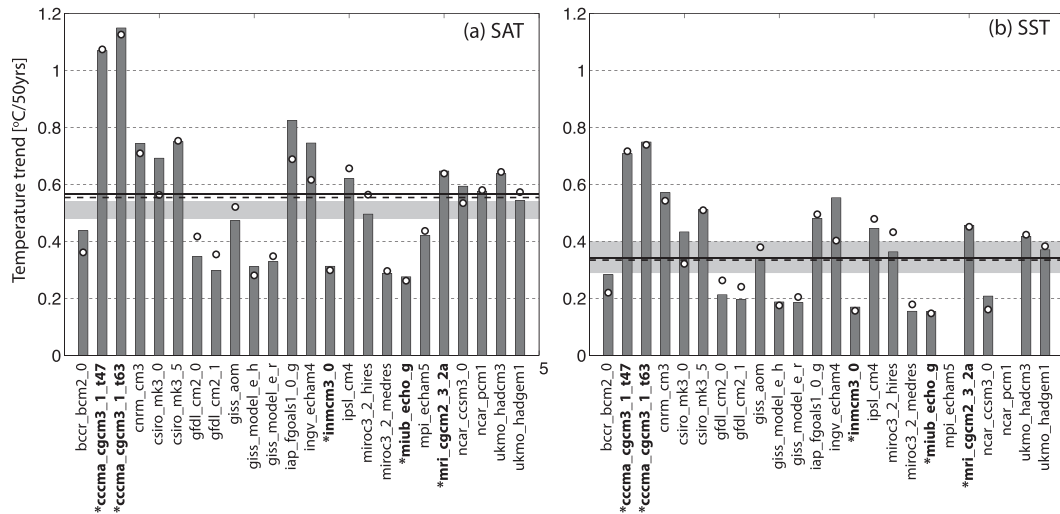
FIG. 2. Drift-corrected (bars) and raw (circles) 1950–2000 trend for (a) SAT and (b) SST. Horizontal lines indicate the multimodel means for drift-corrected (continuous) and uncorrected (dashed) SAT and SST. Horizontal gray bands indicate observational range based on two datasets in each case (see text for further details). Flux adjusted models are indicated in bold. Observational datasets include HadSST2, HadISST, HADCRUT3, and GISTEMP (see text).

meet this criterion, so the closest time period is then used. We do not apply the same procedure to the calculation of the forced 20C3M trend as the degree of forcing ramps up considerably with time and so the assumption of linearity would be less likely. However, where multiple twentieth-century realizations exist for a particular model (which is the case for a number of models), we averaged over all realizations to obtain trend estimates for that model. As low-frequency variability is not coherent across multiple realizations, such averaging helps to reduce the effect of aliasing.

An alternative trend estimation method was also performed whereby a multiparametric regression was performed on the 20C3M and PICNTRL that sought to reduce the effect of climate variability related to various climate drivers (e.g., El Niño–Southern Oscillation, Southern Annular Mode, and volcanic and solar forcing; method described in Durack and Wijffels 2010). While we do not present this here, these results were not substantively different to those shown here using the simple linear trend removal.

4. Results

a. Surface drift

Widely used metrics of global change are averaged surface air temperature (SAT) and sea surface temperature (SST). The observed linear trend of globally averaged SST over the last 50 years of the twentieth century is ~0.4 K $(50 \text{ yr})^{-1}$ based on the Second Hadley Centre

Sea Surface Temperature dataset (HadSST2) (Rayner et al. 2006) or ~0.3 K $(50 \text{ yr})^{-1}$ based on the interpolated HadISST1 (Rayner et al. 2003). Larger values are obtained for global air temperatures: ~0.48 K $(50 \text{ yr})^{-1}$ based on the Hadley Centre–Climate Research Unit Temperature Anomalies, version 3 (HADCRUT3) (Brohan et al. 2006) and ~0.54 K $(50 \text{ yr})^{-1}$ based on National Aeronautics and Space Administration (NASA) GISS surface temperature analysis (GISTEMP) (Hansen et al. 2010). The 50-yr trends from the available CMIP3 models show considerable spread (Fig. 2). However the multimodel mean changes [based on the raw 20C3M simulation output, 0.55 ± 0.22 K $(50 \text{ yr})^{-1}$ for SAT (black line) and 0.33 ± 0.2 K $(50 \text{ yr})^{-1}$ for SST (mean $\pm$ standard deviation) (gray line)] are consistent with the observational estimates. It is important to determine how much of the simulated warming is due to drift and how much is actually attributable to external forcing. For both SST and SAT the globally averaged drift ranges between about -0.16 to $+0.07$ K $(50 \text{ yr})^{-1}$. Consistent with results from the CMIP2 models (Covey et al. 2006), for individual models, the magnitude of the drift is considerably less than both observational estimates of the 1950–2000 trends and the corresponding 20C3M estimates. For most models the drift accounts for less than 20% of the signal, however, the drift magnitudes are often not negligible and should be accounted for in the final estimate of warming in each model. The globally averaged drift in SST and SAT is largest [>0.1 K $(50 \text{ yr})^{-1}$] for Istituto Nazionale di Geofisica e Vulcanologia (Italy) (INGV) ECHAM4, Commonwealth Scientific and
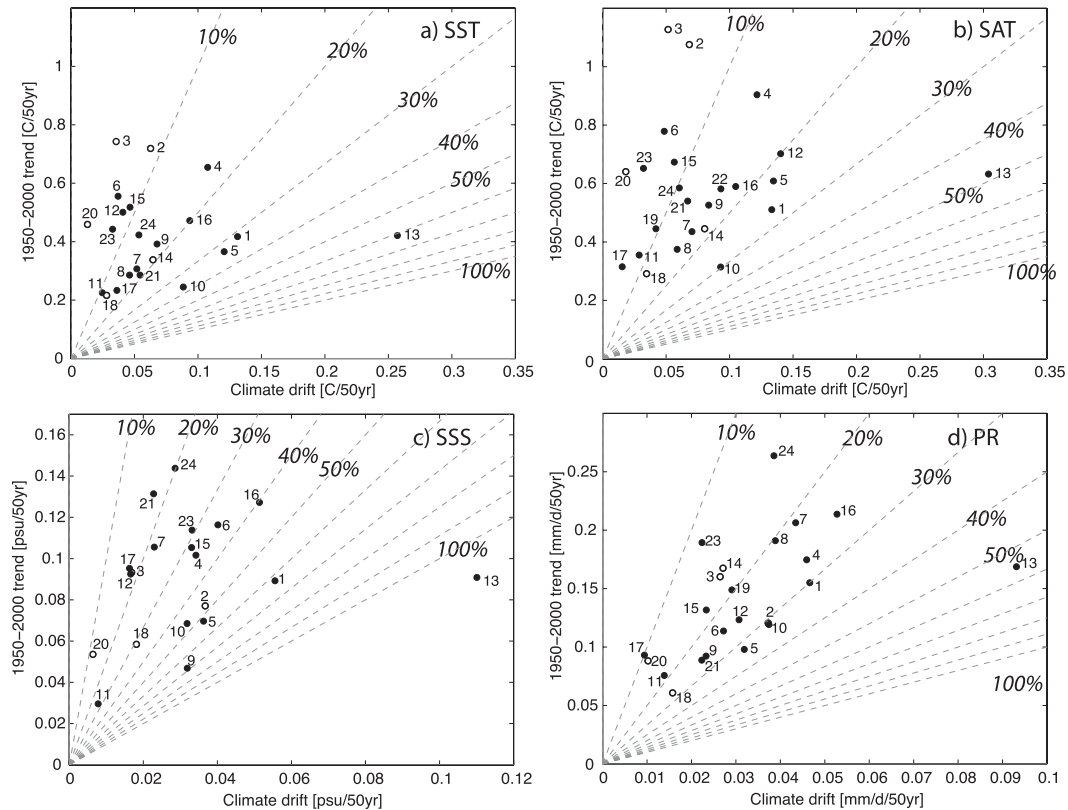
FIG. 3. Scatterplots of globally averaged drift magnitude vs globally averaged 20C3M trend magnitude (a) SST and (b) SAT [K (50 yr)$^{-1}$], (c) SSS [psu (50 yr)$^{-1}$], and (d) precipitation [mm (day)$^{-1}$ (50 yr)$^{-1}$]; 20C3M trend calculated from 1950–2000 period; drift calculated from 1900–2050 period in PICNTRL (or closest possible equivalent period; NB. only 100 years of PICNTRL available for INGV ECHAM4). 20C3M trends are averages of all available ensemble members. Open circles—flux adjusted. Lines indicate where climate drift is 10%, 20%, etc. of the 20C3M trend. Model numbers given in Table 1.

Industrial Research Organisation Mark version 3.0 (CSIRO Mk3.0), and the Institute of Atmospheric Physics (IAP) model (the latter model only for SAT). For ECHAM4 and CSIRO Mk3.0 this corresponds to a drift induced error in SST of over 30% in the raw 20C3M trends. Note that the ECHAM4 drift is based on only 100 years of PICNTRL that terminates before the start of the 20C3M simulation (a concurrent control period is unavailable). This seriously undermines the confidence that can be placed in this drift estimate. As expected, the flux adjusted models all have relatively small drifts (<10%). It is also apparent that the use of flux adjustment does not lead to consistently high or low estimates of 20C3M trends, as these models include both the fastest and slowest warming. Despite the drift making up a small but nontrivial fraction of the 1950–2000 forced trend for some models, the drift-corrected multimodel mean trends of 0.34 ± 0.21 K (50 yr)$^{-1}$ and 0.57 ± 0.24 K (50 yr)$^{-1}$ for SST and SAT, respectively, are statistically indistinguishable from the raw 20C3M

multimodel mean trends. This is because the drift is not systematic across the models and tends to cancel out in the model mean.

While climate drift is of negligible importance when considering the multimodel mean for large-scale surface properties, this is not necessarily the case when considering individual models or examining trends at regional or local scales, discussed below.

While forced trends in surface temperature are positive almost everywhere around the globe, this is not the case for the drift, nor is it the case for forced trends in other properties including salinity and precipitation. As such, a more appropriate metric for expressing the relative importance of the drift can be achieved by first taking the magnitude of the trend (for both 20C3M and PICNTRL) at each grid box before averaging globally (Fig. 3). This provides an average measure of the typical local error in the 20C3M trend if drift is unaccounted for—a very different measure to that shown in Fig. 2. For SST (Fig. 3a) the drift makes up ~20% or less of the

raw 20C3M trend in the majority of models. However, in four of the models the drift exceeds 30%, and in the case of ECHAM4 exceeds 60% of the 20C3M trend. Given the strong coupling between SST and SAT over the ocean, it is of little surprise that the scatter for SAT (Fig. 3b) follows a similar pattern to SST. In general, however, the SAT drift makes up a slightly lower proportion of the 20C3M trend. For salinity, drift magnitudes span 10%–70% of the 20C3M trend magnitude (excluding INGV ECHAM4 where the size of the drift actually exceeds that of the 20C3M trend), with a large proportion of models exceeding 30%. Despite using freshwater flux adjustment, two of the flux adjusted models still have drift magnitudes that exceed 30% [with CGCM3.1(T47) approaching a 50% error]. Drift in sea surface salinity (SSS) represents either a redistribution of salt within the ocean, a net flux of freshwater into or out of the ocean, or a failure to conserve either salt or freshwater.

Based on our examination of the literature, drift within atmospheric variables is rarely assessed in the examination of trends in transient climate simulations. For example in the IPCC AR4 report, only drift in SAT was accounted for. Other atmospheric variables were not subject to drift correction (H. Teng 2011, personal communication). A tight coupling between SST and the surface atmosphere means, however, that there is reason to expect drift in atmospheric properties and indeed we have shown this to be the case for SAT. In addition, to first order (i.e., assuming little change in the atmospheric circulation), changes in evaporation minus precipitation ($E - P$) are expected to scale in proportion to the changes in SAT and therefore over oceanic regions to SST (and the mean $E - P$ field; e.g., Held and Soden 2006; Covey et al. 2006; Romps 2011). Indeed, we find that most models have substantial drift magnitudes in precipitation of between 15% and 35% (Fig. 3d). The strong coupling between the ocean and atmospheric drift is evident in the high correlations ($r \sim 0.9$) between SST drift magnitudes and drift magnitudes across the models, in both SAT and precipitation.

b. Drift on regional scales

While the globally averaged drift magnitude provides an estimate of the typical size of local drift, the drift may be highly heterogeneous and some locations may have much larger drift magnitudes than others. Examination of the individual models shows that there are certain common regions where the magnitude of the drift is relatively large (Figs. 4b,e,h,k). While the magnitude of the drift may be coherent the sign of the drift in these regions is not (i.e., these regions may show either large positive or negative drift). Maps of multimodel-mean

drift magnitudes and the associated 20C3M trend magnitudes highlight some of the robust spatial structures in the drift (Fig. 4).

A number of studies pertaining to individual models find that drift in the ocean is sensitive to ocean convection and as a result drift magnitudes tend to be largest at high-latitude regions where deep convection occurs (Rahmstorf 1995; Cai and Chu 1996; Cai and Gordon 1999). Even when no discontinuity in surface fluxes occurs during model coupling, coupled feedbacks lead to instability and drift in the ocean convection zones (Rahmstorf 1995). A number of early studies noted a reduction of drift, in a variety of ocean and atmosphere variables, associated with the incorporation of the GM eddy parameterization (Gent and McWilliams 1990). This is a result of the parameterization's suppression of excessive convective activity (Boville and Gent 1998; Bryan 1998; Hirst et al. 2000). Large mid- to high-latitude drift magnitudes are common across the CMIP3 models (Figs. 4b,e,h). SST drift magnitudes generally reach a maximum in the midlatitude regions and in the vicinity of sea ice where strong convective activity takes place. This is particularly problematic, with regard to the estimation of forced trends, in the mid- to high-latitude Southern Ocean where the simulated 20C3M warming tends to be relatively weak (Fig. 4a) and so the drift makes up a large part of any trends in the 20C3M simulations. This weak warming trend persists under future projections (Sen Gupta et al. 2009), implying that the error associated with the drift will remain problematic when considering future projections. SST drift is of less importance at lower latitudes where the drift magnitude is small compared to the warming signal.

SAT drift (Fig. 4e) generally mirrors the pattern described for SST with a midlatitude enhancement in drift. The SAT and SST drift become decoupled, however, at high latitudes, particularly over the Arctic region and the Weddell Gyre where drift magnitude remains large in SAT, but is small in SST (Figs. 4b,e). This is likely due to the insulating effects of sea ice and the fact that a small change in sea ice cover can substantially change SAT, via modified air–sea heat exchange, while SST remains relatively unchanged close to the freezing point. Similarly, for the 20C3M raw trends (Figs. 4d,e), the polar amplification of temperatures, only significantly affects SAT and not SST.

The largest 20C3M trends in SSS are primarily related to freshening in the Arctic and the midlatitude northern Atlantic (Fig. 4h). The largest drifts are also evident at higher latitudes of the Northern Hemisphere. Although there is considerable intermodel spread, the largest drifts tend to occur in the northwestern North Atlantic, where SST drift magnitudes are also large. While this may be
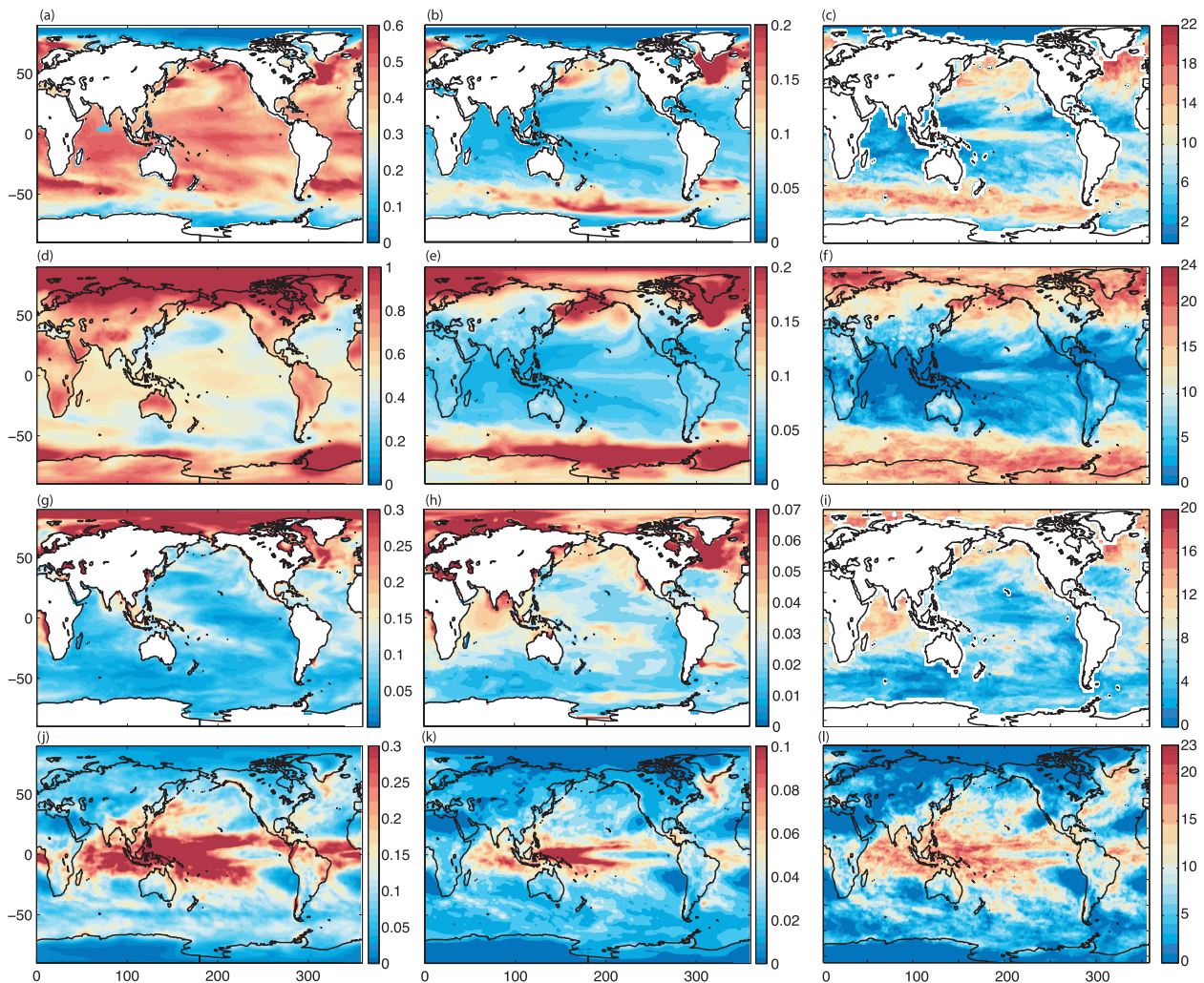
FIG. 4. (left) Multimodel mean 20C3M trend magnitude, (middle) multimodel mean drift magnitude, and (right) number of models in which the local drift magnitude exceeds the globally averaged drift magnitude for (a)–(c) SST [$^{\circ}$C (50 yr)$^{-1}$], (d)–(f) SAT [$^{\circ}$C (50 yr)$^{-1}$], (g)–(i) SSS [psu (50 yr)$^{-1}$], and (j)–(l) precipitation [mm (day)$^{-1}$(50 yr)$^{-1}$]. A total of 22, 24, 20, and 23 models where used in the analysis of SST, SAT, SSS, and precipitation, respectively.

related to oceanic processes alone, the SSS drift may stem in part from changes in the local water fluxes related to changes in SST. This is supported by the elevated precipitation drift magnitudes over this region, which (as noted previously) would scale with changes in SST assuming that the atmospheric circulation remains relatively unchanged.

While there are large intermodel differences in the simulated climate change trends across models, in general precipitation trends are consistent with an intensification of the hydrological cycle, with wet areas becoming wetter and dry areas becoming drier (Allen and Ingram 2002; Held and Soden 2006; Allan et al. 2010). Unlike surface temperature, the largest drifts in precipitation occur in the tropical regions (Fig. 4k). This is presumably a consequence of the

change in the hydrological cycle scaling not only with drift-related temperature changes (which are largest at midlatitudes, Fig. 4b), but also with the mean $E - P$ (Held and Soden 2006), which is greatly enhanced at tropical latitudes.

Figure 5 shows raw 20C3M trends and the associated drift over the high precipitation western tropical Pacific, for a selection of models. While the drift is generally smaller than the 1950–2000 20C3M trend, it becomes important in certain regions. For example in mottled regions the drift makes up at least 50% of the 20C3M trend. Such localized drift becomes important when conducting regional attribution studies or regional projections for individual islands—something that is becoming more prevalent as stakeholders require more policy-relevant, regional- and local-scale information.
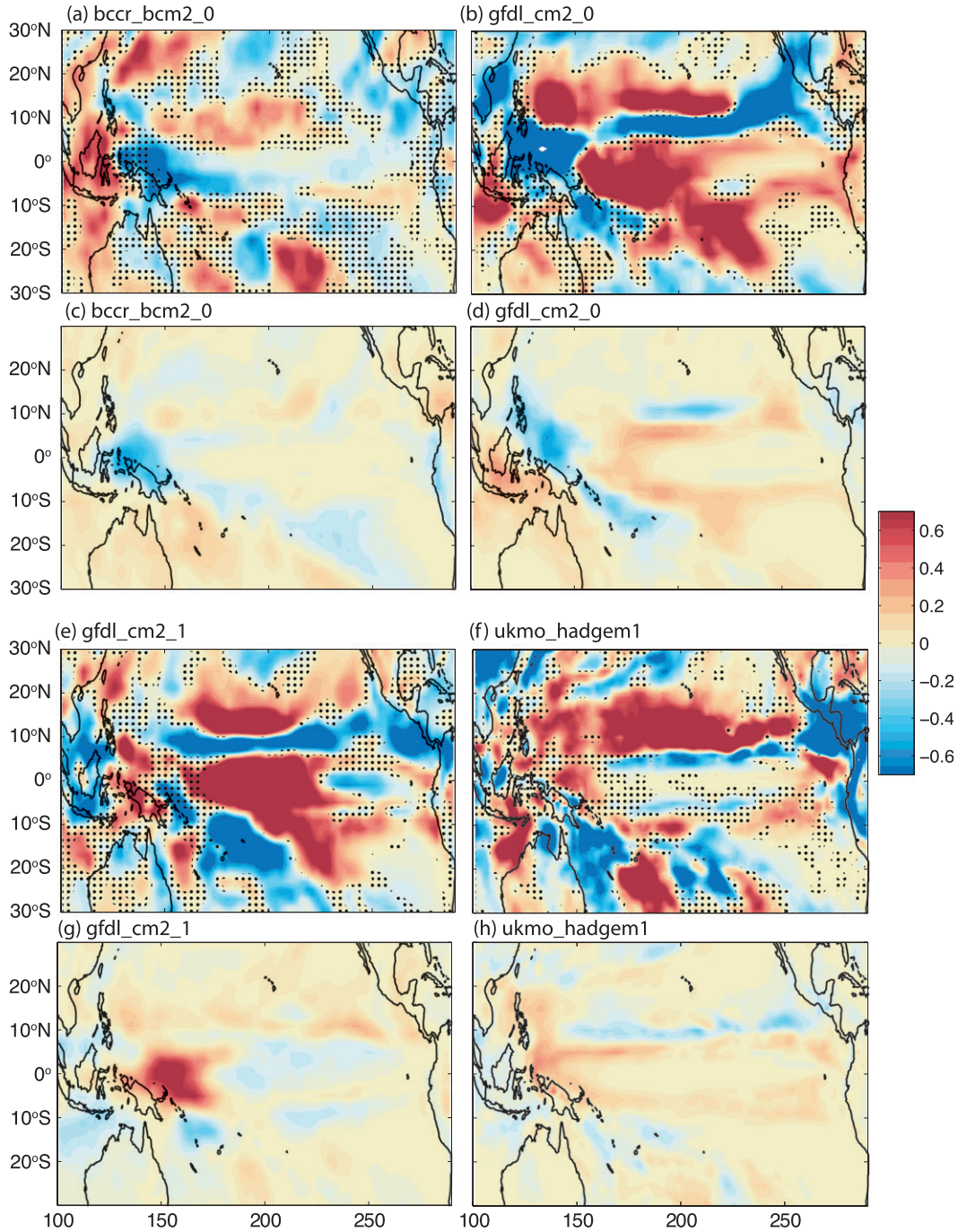
FIG. 5. Precipitation trends $[\text{mm (day)}^{-1} \text{ (50 yr)}^{-1}]$. (a),(b),(e),(f) "Raw" 1950–2000 20C3M trends and (c),(d),(g),(h) associated climate drift for a selection of models (listed at top of each panel) over the tropical Pacific. Mottling indicates regions where the magnitude of the drift or 50% or more of the 20C3M trend.

c. Drift in the ocean interior

It takes considerable time for surface temperature (or freshwater) anomalies resulting from increased anthropogenic forcing to be advected or mixed into the deep ocean. As a result, over the twentieth century, any warming signal, outside of the deep convection regions, is primarily constrained to the upper few hundred meters of the ocean (e.g., Levitus et al. 2005). However, the propagation of surface anomalies resulting from any coupling
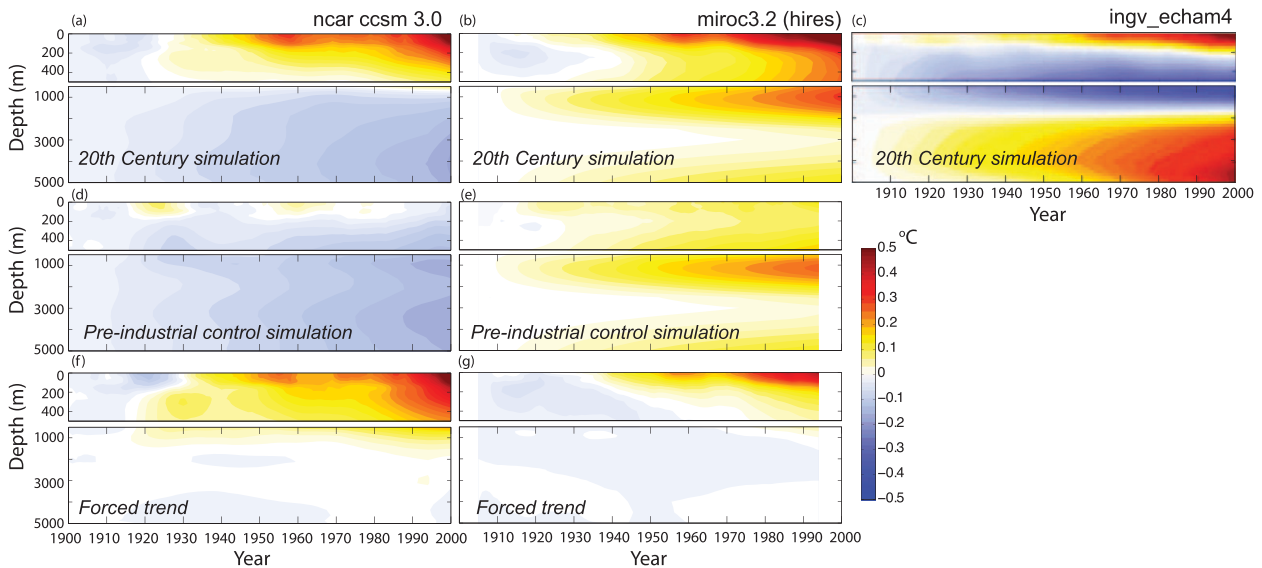
FIG. 6. Evolution of anomalous (relative to 1900) globally averaged ocean temperature at different depths for (left to right) three models: CCSM3.0, MIROC3.2(hires), and INGV ECHAM4 models. (top) 20C3M (1900–2000); (middle) concurrent period of PICNTRL; and (bottom) 20C3M minus PICNTRL (no concurrent PICNTRL available for ECHAM4 model). 10-yr running average applied.

shock over the spinup period and subsequent pre-industrial control and the growth of errors resulting from deficiencies in model physics, will have had a much longer time period (in most cases) over which to pervade the ocean interior. Consequently, we expect to see a comparatively strong drift in the deep ocean. The pre-dominance of drift over any climate change signal in the deep ocean was noted by Gleckler et al. (2006) when examining ocean heat content a subset of CMIP3 models. They found that the anthropogenic signal was generally confined to the upper 500 m over the 1850–2000 period.

The evolution of globally averaged 20C3M tempera-ture with depth is shown here for three CMIP3 models (Figs. 6a,b,c). As expected, the simulations demonstrate a surface-intensified warming becoming stronger over the century. However, significant model-dependent trends are also evident in the deeper ocean. These subsurface trends also exist in the concurrent PICNTRL simula-tions (Figs. 6d,e), indicating that they are not a result of any imposed forcing and are therefore spurious. In fact by simply subtracting the PICNTRL simulations from the 20C3M simulations, most of the signal deeper than ~500 m is removed (Figs. 6f,g). This clearly dem-onstrates a requirement for careful drift removal when investigating the subsurface ocean. The fact that sub-traction of the PICNTRL simulation so effectively re-moves the deep signal also indicates that, at least on these global scales, the drift component evident in the PICNTRL simulation exists relatively unmodified within the 20C3M simulation and nonlinear modulation of the drift in the forced experiment is relatively small. The

third model, INGV ECHAM4, shows quite dramatic deep spurious trends over the twentieth-century simu-lation (Fig. 6c). However, for this model no concurrent PICNTRL simulation is available. As such, we would be inclined to exclude this model from any analysis of the subsurface ocean and be wary of any conclusions drawn even at the surface.

Figures 7a and 7b show the globally averaged drift magnitude (i.e., the absolute value of the PICNTRL trends are taken prior to global averaging) for potential temperature and salinity with depth for each model and the multimodel mean. Significant drift in temperature and salinity occurs throughout the water column. The vertical structure of temperature drift is highly variable across the models. There is a weak tendency for drift to be larger in the upper 1500 m than at deeper levels, although it weakens again over the upper few tens of meters in many of the models. In contrast, the 20C3M warming trend (multimodel mean shown in red) is clearly in-tensified at the surface. After linear drift correction the forced trend is considerably smaller than the corre-sponding raw 20C3M trend at depth. This results in a sit-uation where forced trends dominate over the drift above ~1500 m while drift tends to dominate below this depth. This is similarly true for salinity, although the transition depth is shallower. This results from a systematic surface intensification in the drift magnitude for salinity across all the models.

The magnitude of the drift compared to the total 20C3M forced signal is quantified for the full set of CMIP3 models at two different depths (Fig. 8). At 100 m
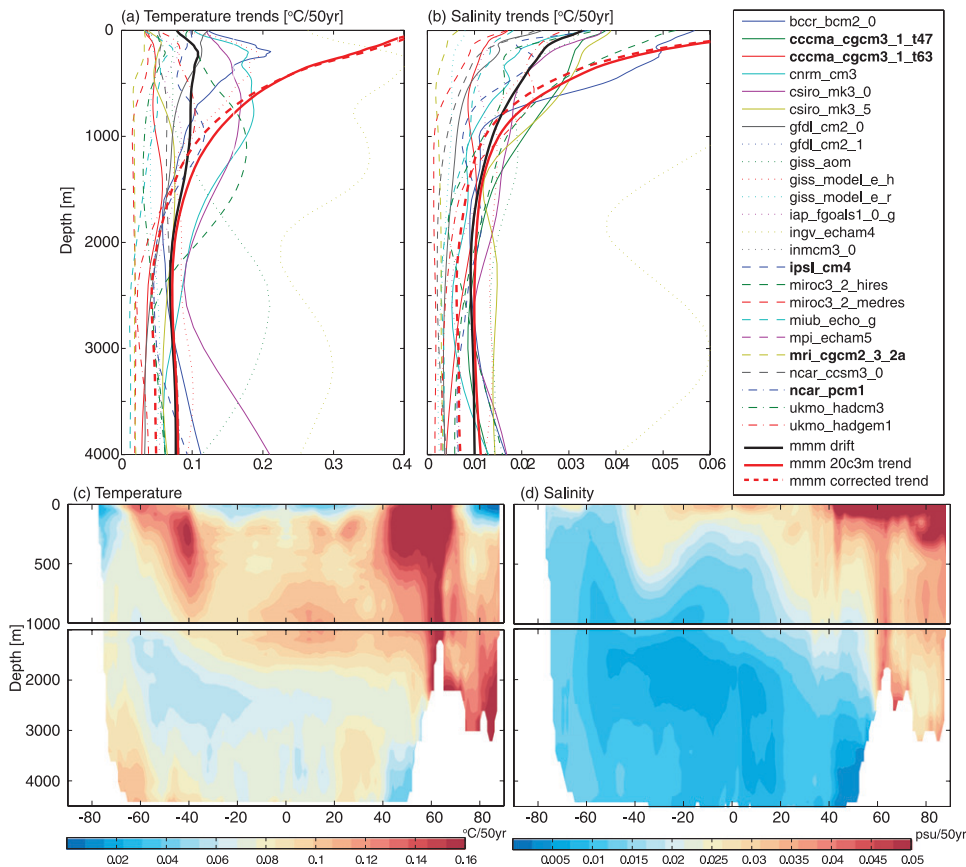
FIG. 7. (top) Globally averaged (left) temperature [K (50 yr)$^{-1}$] and (middle) salinity [psu (50 yr)$^{-1}$] drift magnitudes with depth. Shown are individual multimodel mean drift magnitude (thick black line) and multimodel mean 1950–2000 20C3M globally averaged trend magnitude (thick red line). Also shown are the drift-corrected 1950–2000 trends (thick red dashed line). (bottom) Multimodel mean zonally averaged drift magnitude with depth for (left) temperature and (right) salinity (calculated by taking gridpoint trend magnitudes and then zonally averaging). ECHAM4 is excluded in the calculation of multimodel means as no concurrent PICNTRL exits for drift calculation. Flux adjusted models are in bold.

the error introduced by the drift is mostly within 10% to 40% of the 20C3M signal for temperature and 20% to 70% for salinity. Even at this shallow depth the drift is considerably more important than at the surface (Figs. 3a,c). At 3000 m any trend in the forced experiment for both variables is almost entirely related to drift (i.e., all points sit close to the 100% line). It is again apparent, particularly in the case of salinity, that the near-surface drift exceeds the deep drift (suggesting that atmosphere–ocean freshwater fluxes are playing an amplifying role).

Some consistent patterns of drift can be found across the models (Fig. 7, lower panels). For temperature, the enhancement of drift in the upper part of the water column is located primarily in the regions of deep convection and along the sea ice edge, particularly in the Northern Atlantic (but also the north eastern Pacific and around the Southern Ocean, see also Fig. 4b). At most

latitudes the drift magnitude at the surface decreases again. A possible explanation for this is that surface drift may be subject to damping by the atmosphere, so that heat from areas of positive drift is transferred to areas of negative drift.

The surface intensification of the salinity drift described above is evident at all latitudes but is strongest in the North Atlantic and Arctic Oceans. Little drift amplification is evident around Antarctica. This pattern of enhanced Arctic salinity change is also present in the trend pattern in the 20C3M simulations (Figs. 4g,h). This suggests that there are amplifying feedback processes acting in the Arctic that are less important in the Antarctic, whether the system is being driven by change associated with drift or external forcing. Such feedbacks may be related to the presence of extensive multiyear ice in the Arctic and the fact that Arctic temperatures are
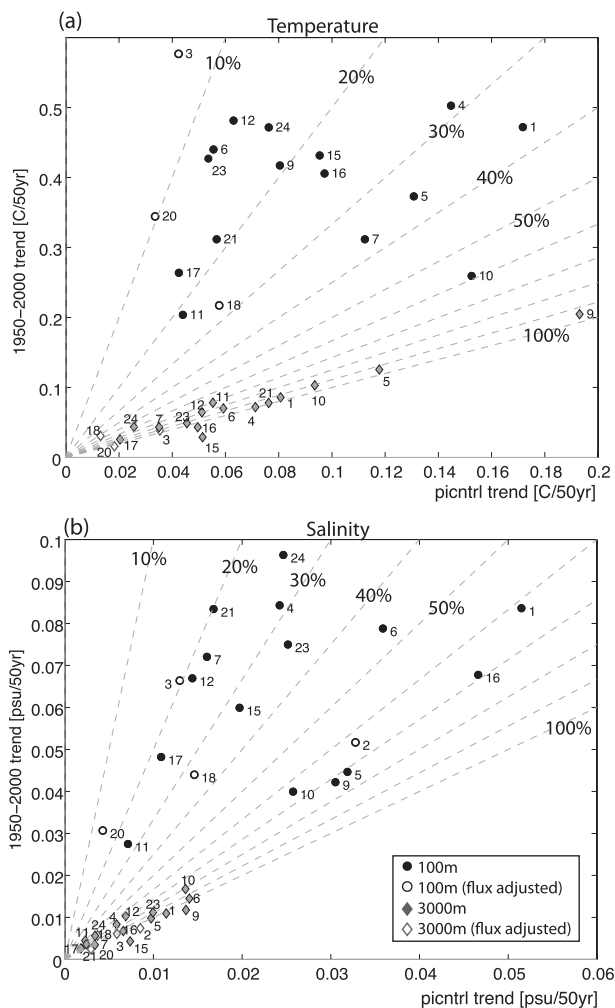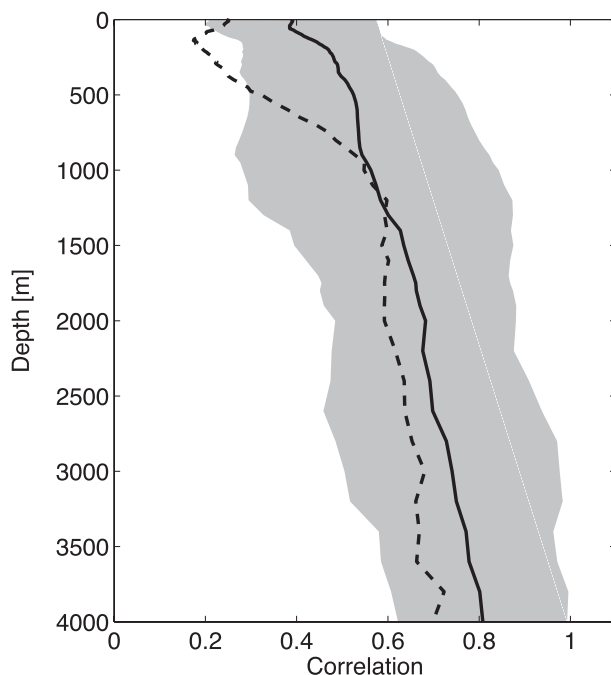
FIG. 9. Correlation between gridpoint drifts in temperature and salinity for different depths, averaged across all models (thick line). The spread in the correlations, represented by an intermodel standard deviation is shown in gray. The 90% significance level is also shown (dashed line). This has been calculated by cross-correlating temperature and salinity fields from all possible combinations of models (at each depth level) and examining the resulting correlation probability distribution function. The large correlations required for 90% significance, for increasing depth, is a result of trends in temperature and salinity becoming much more spatially homogenous at greater depths. This increased spatial autocorrelation in effect dramatically reduces the number of degrees of freedom when computing confidence levels.

FIG. 8. Scatterplots of PICNTRL (1900–2050) vs 20C3M globally averaged trend magnitude (1950–2000) globally averaged trend at 100 and 3000 m for (a) potential temperature and (b) salinity. Lines indicate where climate drift is 10%, 20%, etc. of the 20C3M trend. The ECHAM4 model is off the scale shown with drift in excess of $0.3°C \text{ yr}^{-1}$. Model numbers given in Table 1.

generally warmer and close to the ice melting point (Serreze and Francis 2006). As noted for SST drift, the drift magnitudes in temperature around Antarctica and in the Arctic Ocean are small (Figs. 7 and 4a,b), as upper-ocean temperatures are insensitive to any changes in sea ice. A further factor that may account for the enhanced surface salinity drift away from regions of sea ice is that drift in the SST will drive changes to the atmospheric hydrological cycle (via coupled changes to lower tropospheric temperatures), thus causing spurious drifts in salinity that would be independent of any direct ocean derived salinity drift.

It is interesting to note that there are positive correlations between temperature and salinity gridpoint drifts at most depths across all the models (Fig. 9). This

suggests that the changes in temperature and salinity are partially density compensating. This becomes particularly evident in the deep ocean where correlations are very large (although care must be taken in assigning confidence to the correlation values, as the spatial pattern of temperature and salinity drift have high degrees of spatial autocorrelation). Covariance of temperature and salinity will act to partially compensate the effect drift has on density and thus on steric sea level rise and circulation. Drift in these properties still persists however. Sen Gupta et al. (2009) has shown that, for some CMIP3 models, the twenty-first-century projected changes in the overturning circulation of the Southern Ocean are significantly modified by drift, although in most cases the drift is relatively small compared to the forced change. For 100-yr changes under the Special Report on Emissions Scenarios (SRES) A2 emissions scenario, they find that drift was more important in the Antarctic and abyssal overturning cells, related to the formation and

northward export of bottom waters, than in the wind driven Deacon cell. A large drift in the Atlantic overturning circulation has also been noted in one of the climate models (Fichefet et al. 2003).

d. Implications for steric sea level rise

As discussed above, drift in temperature and salinity dominates 20C3M trends throughout most of the subsurface ocean. In the calculation of steric sea level rise, a given temperature or salinity change will generally have less effect at depth than near the surface. As the amount of expansion for a given change in temperature or salinity is itself a function of temperature, salinity, and pressure (in particular warmer water expands more than colder water for the same increase in heat content), the changes in temperature near the warm surface ocean have a proportionally larger influence on steric sea level rise than temperature changes in the cold deeper ocean (at least away from the well mixed high-latitude regions). Nevertheless, given that the global warming signal over the twentieth century is predominantly limited to the top few hundred meters, in most regions, while ocean drift extends through the entire water column, drift still introduces considerable bias into both regional and global sea level rise.

The CMIP3 models show a broad range of estimates for steric sea level rise over 1950–2000 (Fig. 10a). The spread in the raw 20C3M estimates is considerable (standard deviation $\sim$0.76 mm yr^{-1} with a multimodel mean of 0.45 mm yr^{-1}). In addition a number of the models indicate a lowering of sea level over the period. For the drift-corrected sea level rise (i.e., by using drift corrected temperature and salinity) values become considerably more consistent (standard deviation $\sim$0.36 mm yr^{-1}) and all models now indicate a rise in sea level. While considerable intermodel variability still exists the drift-corrected multimodel mean ($\sim$0.59 mm yr^{-1}) is consistent with the Domingues et al. (2008) observational estimate (0.52 $\pm$ 0.08 mm yr^{-1}, for 0–700 m, 1950–2003). Figure 10a shows raw 20C3M trends and drift-corrected estimates of forced trend for steric sea level rise, including multiple ensemble members where available; ensemble members for a given model are generally initialized from the same PICNTRL experiment but from different points in time, usually separated by multiple years (Table 1). Nevertheless the drift, which is derived from different time periods from a single PICNTRL simulation, is very similar across ensemble members, suggesting that the linear drift approximation is valid and that natural variability is not having a major effect on the drift estimates. Figure 10b shows a scatter of the raw 20C3M trend magnitudes versus drift magnitudes. The drift-related error varies considerably across the

models from less than 10% to over 200% for the ECHAM4 model (see previous discussion of this model).

As with surface drift, subsurface drift in temperature and salinity is spatially heterogeneous and so can result in a larger bias on regional scales. This is particularly important for assessing twentieth-century regional changes, where the steric component of sea level rise is a major component of the total (e.g., Domingues et al. 2008). Figure 11 shows both the raw 20C3M and drift-corrected 1950–2000 trends for three models (calculated from the surface to the bottom). A few models (e.g., MRI-CGCM2.3.2) have a well-equilibrated pre-industrial control throughout the ocean and so are essentially untroubled by drift. However, most models are significantly affected in certain regions. In fact for many models and regions the sign of the sea level trend is changed by the spurious drift. For instance in the CSIRO Mk3.0 model the steric sea level anomaly over much of the tropics and midlatitudes, estimated from the raw 20C3M temperature and salinity, changes sign once the drift is taken into account.

5. Discussion

Despite major advances in the fidelity of coupled climate models in their reproduction of the observed climate system (Randall et al. 2007; Reichler and Kim 2008) spurious trends in model simulations, known as climate drift, still persist, independent of any external forcing. Despite the best efforts of modeling groups, climate drift is an issue that is likely to persist for some time to come. As such, the appropriate level of importance must be given to this problem depending on the application at hand. In some instances drift is of primary importance and cannot be ignored. For example, in the deep ocean or for depth-integrated properties drift may dominate over any externally forced signal. In some applications, however, climate drift has a relatively minor effect and can be safely ignored. This is often the case when dealing with multimodel means of the surface climate. Drift appears not to be systematic with regards to its sign and tends to cancel out where a large number of models are considered. In addition the relative importance of drift will generally diminish into the future as the forced trend becomes larger (at least for some time) and the drift (at least in principle) should diminish.

a. Drift issues for examining climate model output

Below we raise a number of points that should be considered by those examining model output in the context of forced trends:
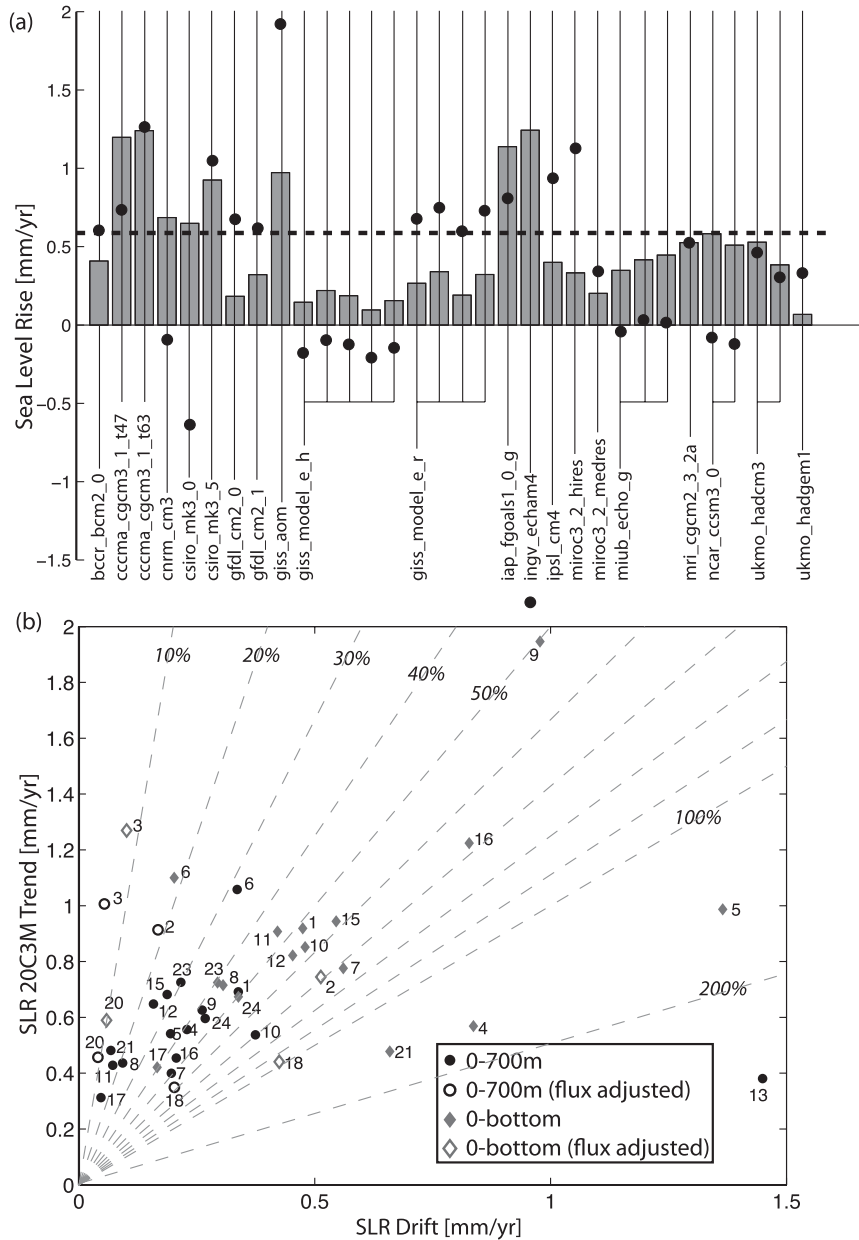
FIG. 10. Steric sea level rise trends (1950–2000). (top) Raw (circles) and drift-corrected (bars) globally averaged steric sea level rise for individual models and drift-corrected multimodel mean (dashed line). (bottom) Scatterplots of PICNTRL (1900–2050) vs 20C3M (1950–2000) globally averaged steric SLR trend magnitude calculated for surface to 700 m and for the full water column. Steric SLR drift is computed by differencing trends computed from raw and drift-corrected temperature and salinity.

Below ~(1–2) km, the drift generally dominates over any forced trend. Any study examining subsurface processes or depth-integrated properties like steric sea level rise, must pay careful attention to how drift is treated. The drift in sea level can be large enough to reverse the sign of the forced change both regionally and in some models for the global average. Conclusions drawn from

such studies may be sensitive to the method by which the drift is corrected for.

Globally averaged drift for SST and SAT is for all models substantially smaller than forced trends in the twentieth century. For this reason when considering such globally averaged variables, drift is often considered of relatively minor importance, especially compared to the
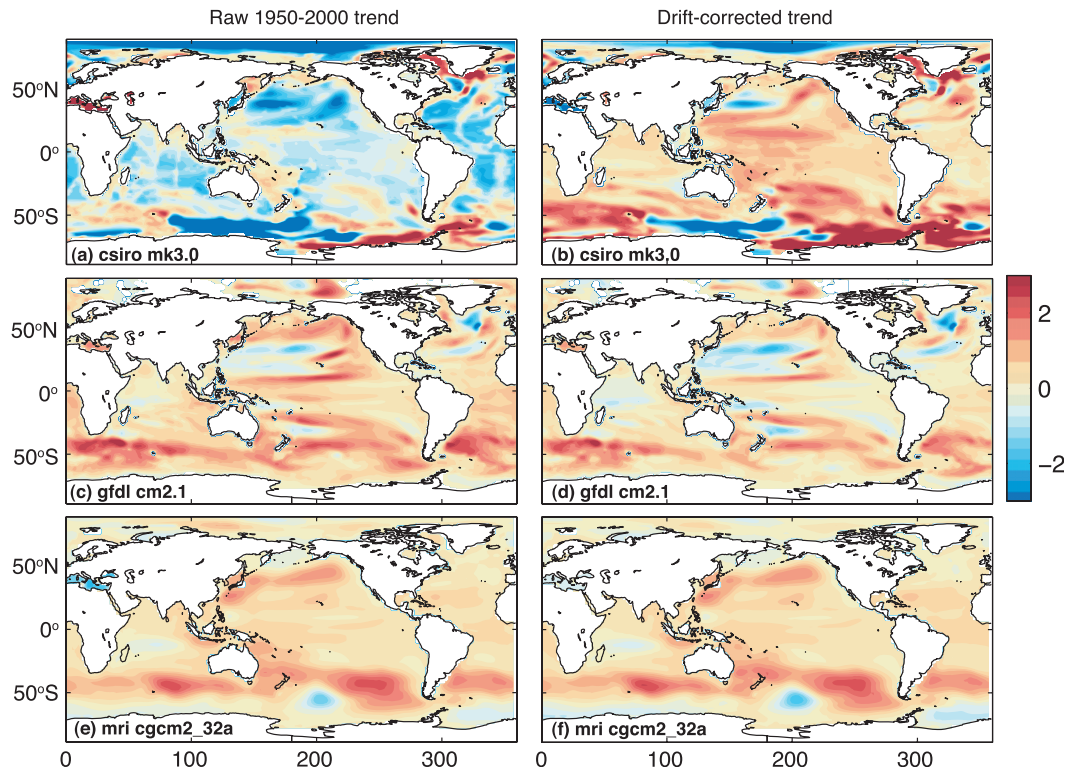
FIG. 11. (left) Raw 20C3M and (right) drift-corrected steric sea level rise for the full water column (mm yr^{-1}) for (top to bottom) three of the CMIP3 models. Drift-corrected values are calculated using drift-corrected temperature and salinity.

uncertainty in the model spread. This is in part due to the fact that the forced trend is positive almost everywhere, while the sign of the drift varies regionally. The drift in this case will on average make up ~10% of the forced trend (and not exceed 30% for SST or 20% for SAT, see Fig. 2). In addition, the sign of drift does not appear to be systematic across the models. Thus, when using an unweighted multimodel mean drift has a much reduced impact. This appears to be generally true for all the variables considered here.

As surface drift is spatially heterogeneous, the regional importance of drift for individual models can be much larger than the global figures suggest. We have presented a number of analyses showing globally averaged drift magnitudes versus 20C3M trend magnitudes (Figs. 3, 8, and 10) for various variables. These plots provide a measure of the average error that would be incurred in estimating the forced 1950–2000 trend for a particular location and model, if no drift correction were applied. For example a typical error in calculating a regional forced SST trend in the Bjerknes Center for Climate Research (BCCR) Bergen Climate Model, version 2.0 (BCM2.0), CSIRO Mk3.0, and GISS-EH models without accounting for drift would be 30% to

40% (Fig. 3a). This is an average value and so larger (and smaller) errors would be expected at different locations.

Studies examining ocean fields from coupled climate models routinely take climate drift into account through some form of correction. As far as we are aware, this is not generally the case for the analysis of atmospheric fields. However, surface ocean drift will necessarily propagate to the atmosphere via air–sea coupling. This is evident in the strong correlations that exist between globally averaged SST drift magnitude and the drift magnitude of SAT and precipitation across the models. As such, drift in atmospheric properties, for example, SAT and precipitation, can make up a significant proportion of 20C3M trends. As an example, for precipitation (Fig. 3d) in 13 out of 21 models the error incurred by ignoring drift at a given location typically exceeds 20% for 1950–2000. Consideration of regional drift is particularly important as there is an increasing effort to use regional and local-scale information from individual climate models to inform regional impact studies.

Spurious trends in temperature and salinity suggests that density would also show substantial drift, although

some degree of density compensation tends to occur in the models (Fig. 9). This would in turn drive dynamical changes to the ocean via changes in stratification, the overturning circulation, and geostrophic flow where there are spatial differences in density drift. It has already been noted that the sensitivity to stratification changes at high latitudes is an important factor in amplifying drift in these regions (Rahmstorf 1995; Cai and Gordon 1999). While we have not explicitly examined drift in circulation, Sen Gupta et al. (2009) have shown that, for some models, the projected changes in the overturning circulation of the Southern Ocean over the coming century are significantly modified by drift, although in most cases the drift is relatively small compared to the future forced change. Preliminary analysis also suggests that drift in the tropical Pacific circulation may be important in selected models.

As evidenced by the fact that drift diminishes with time, climate drift is sensitive to the mean state of a model. The mean state can change because of external forcing or because of natural variability. This implies that drift in the preindustrial control will not be a perfect proxy for the drift within a transient simulation. While we offer no solution to this problem, it is important to recognize that this introduces some degree of uncertainty into any drift-corrected forced trend estimate.

Long spinup simulations can greatly reduce the rate of climate drift. However, this comes at a cost. A long integration necessarily means that the climate state has more time to diverge from the initial "observed" state. This has implications for the evaluation of climate models; that is, assessing their realism in simulating the observed system. It is often assumed that a "good" model is simply one that can adequately reproduce a realistic mean state. Such an assessment is often used to select models or even weight models to provide a best estimate of future projections (see Knutti 2010 for a review). However, historical simulations and subsequent projections are branched from spinup integrations of very different lengths. As such, a physically realistic model that has been integrated for a long period of time (to reduce any drift) may exhibit a poorer mean state than a less physically realistic model that has had only a short spinup time and so retains a strong memory of the observational data used in the model initialization. Knowledge of the rate of a model's drift during the spinup phase may in itself be a useful indicator of model realism, as a realistic model would have a final state that is close to the observationally derived initial state (assuming the initialized observed fields are dynamically consistent). The relative importance of a stable climate versus a realistic mean state must be carefully considered.

b. Improving the assessment of climate drift in the future

While the WCRP's Working Group on Coupled Modelling (WGCM) provide detailed guidelines with respect to historical, projection, and sensitivity experiments, little guidance is provided with respect to the control simulation, other than that the control should be extended for sufficient duration to span any historical and projection simulations (Taylor et al. 2011). With the next IPCC round looming, simulation strategies for CMIP5 are set and simulations well underway. Nevertheless, some steps can still be taken to aid with the analysis of model output in the context of climate drift, and certain steps can be considered for future sets of simulations:

- Length of spinup simulations should not be neglected in the attempt to produce an increasing number of projections. The lack of a stable climate is a considerable hindrance in the primary objective of the CMIP5 climate models—making future projections. It is encouraging that all four of the Geophysical Fluid Dynamics Laboratory (GFDL) climate models that will provide output for CMIP5 have been spun up in coupled configuration for ~2000 yr (S. Griffies 2011, personal communication). Similarly the new generation of Canadian Centre for Climate Modelling and Analysis (CCCma) models has spun up for ~800 yr (O. Saenko 2011, personal communication).
- Modeling groups should make available the longest possible period of preindustrial control that extends prior to the start of forced simulations. This greatly aids in understanding the temporal evolution of the drift and allows a better separation between drift and low-frequency natural variability. As a minimum a preindustrial control simulation for a period running concurrently to any forced simulations should be made available (while this has usually been the case for CMIP3, it is not universally so).
- All variables archived as part of the forced simulations should also be provided for the control simulations. As shown above ocean drift is also inherent in the dynamical ocean fields and propagates to the rest of the climate system.
- Where possible multiple control simulations should be provided. As with forced experiments, this would be a powerful way to extract a robust trend when considerable low-frequency variability exists. Three modeling groups provided multiple control simulations (that ran for greater than 100 yr) as part of CMIP3.
- Spinup procedures and experimental design should be fully documented. This information was often lacking at the PCMDI repository for some of the CMIP3 models.

- In the absence of a clear direction forward to alleviate climate drift in the near term, it seems important to keep open the question of flux adjustment within climate models that suffer from considerable drift. Flux adjustments are nonphysical and therefore inherently undesirable. They may also fundamentally alter the evolution of a transient climate response (Neelin and Dijkstra 1995; Tziperman 2000). Nevertheless, flux adjustment can alleviate climate drift, at least in surface temperature, which is also nonphysical and inherently undesirable.

REFERENCES

Allan, R. P., B. J. Soden, V. O. John, W. Ingram, and P. Good, 2010: Current changes in tropical precipitation. *Environ. Res. Lett.,* **5,** 025205, doi:10.1088/1748-9326/5/2/025205.

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature,* **419,** 224–232.

Ammann, C. M., F. Joos, D. S. Schimel, B. L. Otto-Bliesner, and R. A. Tomas, 2007: Solar influence on climate during the past millennium: Results from transient simulations with the NCAR Climate System Model. *Proc. Natl. Acad. Sci. USA,* **104,** 3713–3718.

Boville, B. A., and P. R. Gent, 1998: The NCAR Climate System Model, version one. *J. Climate,* **11,** 1115–1130.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.,* **111,** D12106, doi:10.1029/2005JD006548.

Bryan, F. O., 1998: Climate drift in a multicentury integration of the NCAR climate system model. *J. Climate,* **11,** 1455–1471.

Cai, W., and P. C. Chu, 1996: Ocean climate drift and interdecadal oscillation due to a change in thermal damping. *J. Climate,* **9,** 2821–2833.

——, and H. B. Gordon, 1999: Southern high-latitude ocean climate drift in a coupled model. *J. Climate,* **12,** 132–146.

Collier, M. A., A. C. Hirst, M. R. Dix, R. P. Cechet, H. B. Gordon, E. A. Kowalczyk, S. P. O'Farrell, and L. D. Rotstayn, 2004: CMIP experimental results from the CSIRO Mk3 climate system model: Comparison with reanalysis and observations. CSIRO Atmospheric Research Tech. Paper 69, 73 pp.

Conkright, M. E., and Coauthors, 1999: World ocean database 1998: Documentation and quality control version 2.0. Ocean Climate Laboratory National Oceanographic Data Center Internal Rep. 14, 117 pp.

Covey, C., P. J. Gleckler, T. J. Phillips, and D. C. Bader, 2006: Secular trends and climate drift in coupled ocean–atmosphere general circulation models. *J. Geophys. Res.,* **111,** D03107, doi:10.1029/2005JD006009.

Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature,* **453,** 1090–1093.

Downes, S. M., N. L. Bindoff, and S. R. Rintoul, 2010: Changes in the subduction of Southern Ocean water masses at the end of the twenty-first century in eight IPCC models. *J. Climate,* **23,** 6526–6541.

Durack, P. J., and S. E. Wijffels, 2010: Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *J. Climate,* **23,** 4342–4362.

Fichefet, T., C. Poncin, H. Goosse, P. Huybrechts, I. Janssens, and H. Le Treut, 2003: Implications of changes in freshwater flux from the Greenland ice sheet for the climate of the 21st century. *Geophys. Res. Lett.,* **30,** 1911, doi:10.1029/2003GL017826.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.,* **20,** 150–155.

Gleckler, P. J., K. AchutaRao, J. M. Gregory, B. D. Santer, K. E. Taylor, and T. M. L. Wigley, 2006: Krakatoa lives: The effect of volcanic eruptions on ocean heat content and thermal expansion. *Geophys. Res. Lett.,* **33,** L17702, doi:10.1029/2006GL026771.

Gordon, H. B., and Coauthors, 2002: The CSIRO Mk3 Climate System Model. CSIRO Atmospheric Research Tech. Paper 60, 130 pp.

Gregory, J. M., and J. A. Lowe, 2000: Predictions of global and regional sea-level rise using AOGCMs with and without flux adjustment. *Geophys. Res. Lett.,* **27,** 3069–3072.

——, and Coauthors, 2001: Comparison of results from several AOGCMs for global and regional sea level change 1900–2100. *Climate Dyn.,* **18,** 225–240.

——, J. A. Lowe, and S. F. B. Tett, 2006: Simulated global-mean sea level changes over the last half millennium. *J. Climate,* **19,** 4576–4591.

Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.,* **48,** RG4004, doi:10.1029/2010RG000345.

Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate,* **19,** 5686–5699.

Hirst, A. C., S. P. O'Farrell, and H. B. Gordon, 2000: Comparison of a coupled ocean–atmosphere model with and without oceanic eddy-induced advection. Part I: Ocean spinup and control integrations. *J. Climate,* **13,** 139–163.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. Van der Linden, X. Dai, K. Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis.* Cambridge University Press, 892 pp.

Katsman, C. A., W. Hazeleger, S. S. Drijfhout, G. J. Oldenborgh, and G. Burgers, 2008: Climate scenarios of sea level rise for the northeast Atlantic Ocean: A study including the effects of ocean dynamics and gravity changes induced by ice melt. *Climatic Change,* **91,** 351–374.

Knutti, R., 2010: The end of model democracy? *Climatic Change,* **102,** 395–404.

Large, W. G., and G. Danabasoglu, 2006: Attribution and impacts of upper-ocean biases in CCSM3. *J. Climate,* **19,** 2325–2346.

——, ——, S. C. Doney, and J. C. McWilliams, 1997: Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology. *J. Phys. Oceanogr.,* **27,** 2418–2447.

Levitus, S., 1982: Climatological Atlas of the World Ocean. NOAA Prof. Paper 13, 173 pp. and 17 microfiche.

——, and T. P. Boyer, 1994: *Temperature.* Vol. 4, *World Ocean Atlas 1994,* NOAA Atlas NESDIS 4, 117 pp.

——, R. Burgett, and T. P. Boyer, 1994: *Salinity.* Vol. 3, *World Ocean Atlas 1994,* NOAA Atlas NESDIS 3, 99 pp.

——, J. Antonov, and T. Boyer, 2005: Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.,* **32,** L02604, doi:10.1029/2004GL021592.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007a: The WCRP CMIP3 multimodel dataset. *Bull. Amer. Meteor. Soc.,* **88,** 1383–1394.

——, and Coauthors, 2007b: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–845.

Moore, A. M., and H. B. Gordon, 1994: An investigation of climate drift in a coupled atmosphere-ocean-sea ice model. *Climate Dyn.,* **10,** 81–95.

Neelin, J. D., and H. A. Dijkstra, 1995: Ocean–atmosphere interaction and the tropical climatology. Part I: The dangers of flux correction. *J. Climate,* **8,** 1325–1342.

Phipps, S. J., L. D. Rotstayn, H. B. Gordon, J. L. Roberts, A. C. Hirst, and W. F. Budd, 2011: The CSIRO Mk3L climate system model version 1.0—Part 1: Description and evaluation. *Geosci. Model Dev.,* **4,** 483–509.

Power, S. B., 1995: Climate drift in a global ocean general circulation model. *J. Phys. Oceanogr.,* **25,** 1025–1036.

Rahmstorf, S., 1995: Climate drift in an ocean model coupled to a simple, perfectly matched atmosphere. *Climate Dyn.,* **11,** 447–458.

Räisänen, J., 2001: CO$_2$-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability. *J. Climate,* **14,** 2088–2104.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 589–662.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.,* **108,** 4407, doi:10.1029/2002JD002670.

——, P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. J. Ansell, and S. F. B. Tett, 2006: Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *J. Climate,* **19,** 446–469.

Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.,* **89,** 303–311.

Romps, D. M., 2011: Response of tropical precipitation to global warming. *J. Atmos. Sci.,* **68,** 123–138.

Santer, B. D., and Coauthors, 2009: Incorporating model quality information in climate change detection and attribution studies. *Proc. Natl. Acad. Sci. USA,* **106,** doi:10.1073/pnas.0901736106.

Sausen, R., K. Barthel, and K. Hasselmann, 1988: Coupled ocean-atmosphere models with flux correction. *Climate Dyn.,* **2,** 145–163.

Sen Gupta, A., A. Santoso, A. S. Taschetto, C. C. Ummenhofer, J. Trevena, and M. H. England, 2009: Projected changes to the Southern Hemisphere ocean and sea ice in the IPCC AR4 climate models. *J. Climate,* **22,** 3047–3078.

Serreze, M. C., and J. A. Francis, 2006: The Arctic amplification debate. *Climatic Change,* **76,** 241–264.

Sloyan, B. M., and I. V. Kamenkovich, 2007: Simulation of Subantarctic Mode and Antarctic Intermediate Waters in climate models. *J. Climate,* **20,** 5061–5080.

Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis.* Cambridge University Press, 996 pp.

Sun, S., and J. E. Hansen, 2003: Climate simulations for 1951–2050 with a coupled atmosphere–ocean model. *J. Climate,* **16,** 2807–2826.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2011: A summary of the CMIP5 experiment design. World Climate Research Programme Rep., 33 pp. [Available online at http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf.]

Tziperman, E., 2000: Uncertainties in thermohaline circulation response to greenhouse warming. *Geophys. Res. Lett.,* **27,** 3077–3080.